This is to certify that the
dissertation entitled

BUILDING A STRATEGIC THEORY OF DECEPTION

presented by

BRIAN HAHN KIM

has been accepted towards fulfillment
of the requirements for the

Ph. D. degree in PSYCHOLOGY

J Kevin Ford
Major Professor's Signature

October 26, 2006
Date

MSU is an Affirmative Action/Equal Opportunity Institution

BUILDING A THEORY OF STRATEGIC DECEPTION

By

Brian Hahn Kim

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

2006

ABSTRACT

BUILDING A STRATEGIC THEORY OF DECEPTION

By

Brian Hahn Kim

Although research on deception, lying, faking, and related concepts has often examined factors motivating the use of deception and factors affecting people's attempts to detect specific signs of deception, little has been stated about the theoretical concept of deception, particularly regarding the psychological constructs that mediate the process from motivation to the behavioral performance of deception. In this paper, existing concepts and theoretical propositions related to deception are integrated and then organized into a unified theoretical framework. More specifically, a set of cognitive deception strategies were formulated as a mechanism to explain how motivation leads to the production of certain acts of deception, which might later be examined in others' attempts to detect the deception. An exploratory, verbal protocol study was conducted to examine the cognitions of people performing deception on a selection test battery in a job applicant simulation, using the think-aloud procedure. Qualitative and quantitative analyses of these cognitions showed that the general set of theorized deception strategies were used by people, though some strategies were more prevalent than others. Refinements to the theoretical framework of deception introduced here and practical applications of the theory, particularly in relation to psychological measures of test faking, are then discussed.

ACKNOWLEDGEMENTS

Certain people deserve recognition for their assistance, which enabled this project to be completed. I am truly indebted to Ben Saxe, Lydia Leonard, Liesel Wilfinger, John Klavanian, and Joseph Butterwick for collecting data with me. I give special recognition to Lydia, in particular, for also transcribing data.

I am especially grateful to Alyssa Friede and Stephanie Drzakowski for serving as expert coders, to Fred Oswald and Zack Hambrick for allowing me to use their laboratories, and to Neal Schmitt for providing me with critical research funds. If aspects of this research can be applauded, I credit my dissertation committee (Kevin Ford, Ann Marie Ryan, Rick DeShon, Neal Schmitt, and Fred Oswald) for providing invaluable guidance, and also to Steve Kozlowski for teaching me to build a theory.

I also thank my peers for providing general encouragement and advice during my graduate experience. In particular, Mike Gillespie, Aaron Schmidt, Christine Scheu, Jen (Carr) Gillespie, Darin Wiechmann, Brad Chambers, Ernest Park, Jaclyn Nowakowski, Lauren Ramsay, Chi Dang, Dan Watola, Dustin Jundt, and Alyssa Friede were always willing to provide help when I needed to lean on them. I also thank Brendan Baird, Kimdy Le, Amani El-Alayli, Guihyun Park, Sonia Ghumman, and Tony Boyce for being other sources of friendship that helped to make graduate life and East Lansing bearable, at the very least. Given that any graduate education will have its "ups and downs," I am

fortunate enough to say that I had a wonderful experience and received a great amount of general support by faculty and peers in the Michigan State I/O program.

I would be remiss if I did not make a special acknowledgement for my two faculty mentors. Thank you Kevin for helping me through this massive project and for always pushing me to "be theoretical." Thank you Neal for mentoring me from day one, providing both guidance and the freedom to explore things on my own. I am sure this is really only the beginning of my education, but trust that the training I received will prepare me for whatever lies ahead.

Finally, I offer my deepest thanks to my sister, Janice, and best friend, Smriti. Thank you for not only for providing research assistance on this particular project as coders, but for all wonderful moments in life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

All human beings rely on various forms of communication to develop personal and societal relationships. However, the meaningfulness of communicated messages cannot be determined without first assessing their veracity to some degree. People must have a sense of whether the information that they encounter was transmitted honestly or deceptively as one indicator of the information's accuracy, with the consequences of failing to determine veracity ranging from mundane to ominous.

Examples of deception that affect people's everyday lives are ubiquitous, at all levels of interaction (Galasínksi, 2000). Children lie to their parents about performing their chores. Job applicants exaggerate their qualifications. Corporations like Enron and WorldCom fabricate financial reports (CNN, 2002; McLean & Elkind, 2003). Politicians use deception to obtain office and to push certain agendas, including wars (e.g., Duelfer, 2004 and Van Natta, 2006). In a poignant example, U.S. meat packing employees have lied on enough reports of microbial tests related to the deadly E. coli bacterium that one federal investigator for the Department of Agriculture's Food Safety and Inspection Service stated, "We give no serious validity to company-generated records. ... There's a lot of falsification going on" (Schlosser, 2004, p. 216). At the same time, not all instances of deception produce negative outcomes. Social norms are often supported by "little white lies" that enable smooth social interactions, as when people are expected to give glib, positive responses to daily questions (DePaulo, Kashy, Kirkendol, & Wyer, 1996; Miller & Stiff, 1993). A classic example of socially accepted deceit is lying to conceal plans for a "surprise" birthday party. In professional settings, authority figures (e.g.,

parents, store managers, or presidents) may be expected to use deception to present a happy or confident demeanor to motivate their subordinates.

Overall, deception encompasses many types of processes and behaviors that range from harmless practical jokes to dire military strategies (Hyman, 1989), having the potential to produce both positive and negative outcomes. Encountering and performing deception, at least to some degree, is an inevitable process rooted in the human experience (Galasínski, 2000; Goffman, 1959; Levin & Zickar, 2002; Miller & Stiff, 1993). In one study, people told lies in approximately 25% of their daily social interactions (DePaulo et al., 1996). Thu s, there is a basic scientific need to understand the pervasive psychological phenomenon of deception, the process of intentionally communicating a lie or other misrepresentation.

In this paper, I attempt to map the core psychological processes underlying deception. I begin by reviewing a broad array of related concepts and integrating them to formulate a general definition of deception. I then review and critique existing theories of relevant phenomena to build a theory of strategic deception that was specifically designed to fill a major theoretical gap in the literature by describing the motivational, cognitive, and behavioral events that lead to the production of a deceptive message, and eventual interpretation of that message by a second party. While this theory is proposed primarily for the purpose of description, I do address its relation to a few practical applications concerning the detection, prevention, and control of deception.

An exploratory empirical investigation of the deceptive process, with a focus on the theorized set of strategies, was conducted using verbal protocol data collected from people instructed to describe their thoughts while creating deceptive test responses, under

varying conditions. Qualitative and quantitative analysis helped to identify key concepts and to provide some initial confirmation that the further development and study of deception strategies can explain key processes. Insights about potential modifications to and elaborations of the theory are then discussed, as well as practical implications for applied psychological endeavors.

Relevant Concepts

Lying and lies. No act of deception is more clearly defined than lying, where "lies" are the message outcomes of the act. One early definition stated that lying is "the intentional deliverance of a conscious untruth for the purpose of deception" (Larson, 1932, p. 3). More recently, Smith (2004) defined lying as "any form of behavior the function of which is to provide others with false information or to deprive them of true information" (p. 14).[1] The Merriam-Webster Online Dictionary (at www.m-w.com) defines a lie as 1) "an assertion of something known or believed by the speaker to be untrue with intent to deceive" or "an untrue or inaccurate statement that may or may not be believed true by the speaker," and as 2) "something that misleads or deceives"; it defines lying as 1) "to make an untrue statement with intent to deceive" or 2) "to create a false or misleading impression." With one exception these and numerous alternative definitions have a common stipulation: the liar must perceive the message communicated to be untrue (Bok, 1978; Ekman, 1985; Frankfurt, 2005; Galasínksi, 2000; Kashy & DePaulo, 1996; Miller & Stiff, 1993).

Typically, the term lying is not used to describe subtle misrepresentations of an object, like slight exaggerations of otherwise true information. Also, definitions of lying

[1] Adopting an evolutionary perspective, Smith focused on the concept of "function", but the word "intention" is similar enough to be used as a substitute in broader contexts.

tend not to include nonverbal communications (Hopper & Bell, 1984) and situations in which true information fails to be communicated (Galasínski, 2000; Levin & Zickar, 2002). Because selective presentation and "acts of omission" do tend to be viewed as deception, Miller and Stiff (1993, p. 23) conclude that "all lies would be instances of deceptive communication, but not all instances of deceptive communication would involve lying."

Fraud. Certain kinds of criminal acts are tied directly to deception. Perjury is essentially synonymous with lying, but applies specifically to legal contexts. Blanco (2001), a claims investigator, defines fraudulent as "of or pertaining to the deceiving or swindling of someone out of money or possessions; taking without permission, knowledge, or authorization; or distorting or misleading for personal gain" (p. 213), and Merriam-Webster Online defines fraud as "deceit" used to make someone part with something of value and as pretending to be something untrue or an impostor. Forgery seems to be a specific type of fraud. "Every person who, with the intent to defraud, knowing that he or she has no authority to do so, signs the name of another person or of a fictitious person to any of the items listed in subdivision (d) is guilty of forgery," according to the California Penal Code (Blanco, 2001; p. 228).

Clearly, there is a close conceptual link between the concepts of fraud and deception, though fraud tends to place a greater emphasis on the outcome of the deception as being related to the acquisition of other people's possessions. As with lying, the intent to deceive is again a critical characteristic that must be attributed to some action defined as forgery. Simply signing a false name on a check and then disposing of it

does not pass for forgery. Rather, the forged signature must be used to trick another person.

Integrity. Integrity often refers to being honest and to adhering to a set of standards (e.g., American Psychological Association's Ethics Code, 2002). Whether the referent is a person or thing (e.g., data), an object having integrity is viewed as reliably genuine or true. Regarding the use of psychological "integrity" tests in organizational settings, assessments have covered a range of ethical and moral behavior (e.g., theft), as well as dishonesty and trustworthiness (Sackett, Burris, & Callahan, 1989). When defined in this way, deception/honesty refers to just one aspect of integrity.

Faking (tests). Despite the topic's popularity over decades, clear definitions or detailed descriptions of "faking" remain elusive. In the colloquial speech, to fake or be fake can refer to social interactions, athletic moves in sports (including "feints"), counterfeit items (e.g., jewelry and art), costumes (e.g., beard on a shopping mall Santa Claus), and sexual activities. Based on early psychological research on lying (e.g., Hartshorne & May, 1928; Hathaway & McKinley, 1951), Meehl and Hathaway (1946) often used the terms "lying" and "faking" in conjunction and virtually treated them synonymously despite developing separate scales to detect obvious and subtle forms of these phenomena separately on the Minnesota Multiphasic Personality Inventory (MMPI). In fact, they distinguished the broader term "plus-getting" from "faking good" (and "defensiveness" from "faking bad") by claiming that the latter occurred only when test responses were "*extremes*, where such deliberate deception seems likely" (Meehl & Hathaway, 1946, p. 533), again equating the terms of lying and faking.

Though this definition and subsequent ones (e.g., McFarland, Ryan, & Ellis, 2002) clearly have strong links to lying, many current definitions of faking were apparently uncoupled at some point from the same concerns, and expanded to include behaviors that also cause self-reported information to be inaccurate (i.e., biases and errors) without relating to deception. Examples of these two factors are directly implied, simultaneously, in Paulhus' (1991) statement about psychological scales: "A content scale is resistant to faking if, under fake-good instructions, the SDR scale exceeds the cutoff point but the content scale does not change" (p. 20). The statement implies that faking resistance is defined completely by the *operational* instructions provided rather than by the construct(s) underlying the scale and that resistance is influenced by any factor affecting the scale's validity.

Other definitions of faking concepts seem to fall somewhere in between pure deceit and general measurement error, such as those specifying that faking is the process of raising (or lowering) one's scores on a measure in a socially desirable manner (Becker & Colquitt, 1992; Kluger, Reilly, & Russell, 1991; McFarland & Ryan, 2000; Ones, Viswesvaran, & Reiss, 1996; Smith & Ellingson, 2002), without actually explaining much about the process. Similarly, Ones et al. (1996) view faking as "claiming unlikely virtues, denying common faults and unpopular attitudes, exaggerating personal strengths, good impression, self-enhancement...." (p. 660). Overall, there are similarities between faking and lying, but one apparent distinction is that faking refers primarily to distortions of *self*-presentation, rather than to misrepresentations of external objects.

Socially desirable responding (SDR). The concept of socially desirable responding has sparked researchers' interests in multiple scientific fields, being studied in

multiple contexts, including behavior in social situations (Leary & Kowalski, 1990) and response sets/biases/styles on psychological tests (Ones et al., 1996; Paulhus, 1991, 2002). Yet, since it is typically described as being synonymous with or subsumed by "faking," its definitions have also been ambiguous. One reason for this may be that researchers have relied almost exclusively on empirical methods to understand and measure SDR (Paulhus, 1991; 2002). Furthermore, measures are often designed to detect SDR as it occurs when measuring other constructs (e.g., the Lie and K scales are for use with the MMPI). As such, Paulhus (1991, p. 21) warns that SDR scales "may have little application outside the inventory" with which they were designed to be used. Regardless of the reasons for their development, research has clearly demonstrated that existing measures of SDR tend to have low intercorrelations (Paulhus, 1991).

In their review, Zerbe and Paulhus (1987) defined SDR as "the tendency of individuals to present themselves favorably with respect to current social norms and standards" (p. 250). However, Paulhus (2002) later modified the definition to emphasize that SDR should include only "overly positive" presentations that "indicate a departure from reality" (p. 50). While the more recent definition clearly overlaps with lying, the former one represents the majority of research to date. A central problem with Zerbe and Paulhus's (1987) definition, and similar ones, is that self-presentation effects can be performed for many reasons, including nonconscious ones, leading to the study of both truthful (e.g., truthfully reporting positive characteristics or being persuasive) and dishonest response processes (see Paulhus, 1991, for a review of popular conceptions and measures of SDR).

While SDR has typically been viewed as a "contaminant" and source of measurement error (Paulhus, 1991), some now regard it as a stable individual difference characteristic related to aspects of personality[2] (Christiansen, Goffin, Johnston & Rothstein, 1994; McCrae & Costa, 1983; Smith & Ellingson, 2002; Zerbe & Paulhus, 1987). Despite claims and correlational evidence, it is not clear this view is justified or even useful. Currently, we do not know if personality is related to the honest aspect, deceptive aspect, or both aspects of SDR. We also do not know how empirical findings (even meta-analytic ones) have been influenced by differences in measures that show theoretical (i.e., not surface level) differences. For instance, some scales measure the general tendency to engage in SDR across time and situations, while others measure SDR as it occurs simultaneously on a content measure of interest (e.g., personality). Finally, the correlational findings support an alternative proposition that SDR is partially *determined* by personality, rather than that SDR represents an aspect of personality.

In any case, contemporary researchers often treat SDR as having two general facets: impression management and self-deception, based largely on Paulhus' (1984, 1986) factor-analytic work on college student responses to common measures (Zerbe & Paulhus, 1987). *Impression management* (IM) refers to conscious attempts to make oneself appear favorably on psychological tests or in social interactions, based on other people's expectations and standards (Paulhus, 1984). Thus, IM involves deception *only when* the favorable impressions created are simultaneously (perceived to be) untrue, meaning that the impressions are misrepresentations of oneself.

[2] Ones et al. (1996) produced meta-analytic correlations of .37 with Emotional Stability, .20 with Conscientiousness, .15 with Agreeableness, .06 with Extraversion, and .00 with Openness to experience.

Nondeceptive forms of IM may be actions like repeatedly emphasizing one's true strengths so that the target focuses on them. That favorable impressions can be true may explain why IM tends to be viewed as a more socially accepted form of "deception" (McCornack, Levine, Solowczuk, Torres, & Campbell, 1992). People are expected to engage in IM in many situations to produce appropriate, "face-maintaining" actions that demonstrate social competence. Even so, impressions that are very discrepant from the truth generally seem to be regarded negatively, much like deception.

Based on their literature review, Leary and Kowalski (1990) viewed IM as a function of 1) the maximal cost-reward ratio in social situations, 2) attempts to maintain self-esteem, and 3) attempts to create an identity. They then formulated a "2-component" model of IM. The *impression motivation* component, as the name implies, concerns a desire to manage impressions and appear differently to observers. The *impression construction* component relates to the set of cognitions about one's self-concept, possible images/impressions, role constraints, and information about observers that help one decide *how* to communicate a impression (e.g., verbally or nonverbally). As will be described in further detail below, these facets of IM help to explain the general process of deception, but are not exclusive to Leary and Kowalski's model.

Self-deception, the other major, presumed dimension of SDR, clearly exists in some forms of human behavior (Sackeim & Gur, 1978). People can successfully perceive magic tricks and visual illusions (e.g., motion pictures), despite knowing that they are victims of flaws and limitations in their sensory and perceptual faculties (Zackay & Bentwich, 1997). Sometimes clinical patients appear to hold genuinely contradictory beliefs. Paulhus viewed self-deception as "the unconscious tendency to see oneself in a

favorable light" and to deny a psychologically threatening thought "that manifests itself in socially desirable, positively biased self-descriptions that the respondent actually believes are accurate" (Zerbe & Paulhus, 1987, p. 253). As its name implies, self-deception is focused inwardly and based on a set of internal expectations and standards that are assessed often and automatically (Paulhus, 1984; Sackeim & Gur, 1978). (Delusions and fantasies are sometimes grouped with self-deception, but delusions do not have to be self-induced, and fantasies do not have to be perceived as "real.")

It would be convenient if Paulhus' conception of self-deception was analogous to descriptions of deception in social interactions. Unfortunately, such a direct generalization appears to be rarely justified given that (social) deception is most often defined as an intentional process requiring a "consciousness of falsity" (Ford, King, & Hollender, 1988). While a person may attempt to ignore, forget, "erase," or avoid a natural thought or behavior, it is difficult to explain how a mind can intentionally trick itself into believing something, while also knowing the truth (Greenwald, 1997; Smith, 2004). Consequently, the *intention* criterion for defining lies and related concepts implies that self-deception may be conceptually impossible.

One resolution to this possible contradiction would be to view self-deception as resulting from a dual mind. General theories of dual cognitive processes do exist (e.g., Bem, 1967; Feldman Barrett, Tugade, & Engle, 2004; Jones & Pittman, 1982), but usually refer to an interaction between conscious and nonconscious processes, offering limited explanations of self-deception. Although little empirical evidence exists to settle this issue, that "abandoning the assumption of personal unity [within an individual's

knowledge system] seems a drastic step" (Greenwald, 1997, p. 55) has lead many to treat self-deception as conceptually distinct from social deception.

Sackeim and Gur (1978) provided a thorough treatment of the conceptual dilemmas involved in ascribing definitions of social deception to self-deception and concluded that self-deception is similar to social deception. However, this conclusion seems to be based on their decision to completely disregard the importance of intention as a defining aspect of deception, based on the existence of rare colloquial phrases like "I unintentionally deceived Jones" (see p. 150). As most of the literature since that time has underscored the importance of intention (Buller & Burgoon, 1996; Handel, 1985; Hyman, 1989; LaFrenière, 1988; Levin & Zickar, 2002; Lee, 2004; Thompson, 1986; Vasek, 1986; Vrij, 2000), I contend that self-deception represents a complex psychological process that is distinct from social deception.

One other problem with treating self and social deception analogously is that self-deception tends to be defined more broadly. For example, Agassi (1997) considered self-deception to be any situation in which a person espouses contradictory beliefs. Defining self-deception as a behavior instead of a process results in alternative explanations being ignored. For instance, people with low cognitive ability, low self-awareness, little desire for social conformity, or general avoidance tendencies will also tend to state contradictory beliefs, but not because they are being dishonest.

Ultimately, one should recognize that SDR, whether treated as impression management and/or self-deception, is *not* synonymous with deception or lies; SDR can be defined simply as a pattern of responding that reflects social standards, rather than as a person characteristic related to deceit. If a test is explicitly designed to measure social

values and related constructs like social conformity or an awareness of social norms, SDR responses will also be honest (cf. Zerbe & Paulhus, 1987; Schmitt & Oswald, 2006). SDR will also overlap with true patterns of responding when a person's natural characteristics or values happen to be socially desirable. Thus, SDR is only guaranteed to reveal deception when a socially desirable response cannot possibly be true.

Selective presentation. Levin and Zickar (2002, p. 256) formally defined *presentation behaviors* as "structuring the representation of truth." Miller and Stiff (1993, p. 18) argue that "selectivity and oversimplification are usually not considered deceptive unless the message recipient has reason to suspect the message source of duplicity" because "virtually all communication exchanges are marked by the omission of information." As such, selective presentation seems to be deceptive only when the omission of information results in some object being *misrepresented*, preventing the recipient of communication from understanding the truth.

Response distortion. Literally, response distortion is the alteration of true signals in some way. When related to deception, the concept is usually specified in more detail and defined as a person's intentional attempt to appear more positive with respect to some standard (Barrick & Mount, 1996). Turner, Edgley, and Olmstead (1975) defined information distortion as a situation in which "the actor gives information which misrepresents that which he believes to be accurate or true" (p. 71). Others have treated response distortion as synonymous with faking and as strongly related to lies, expedient behaviors, selective presentation, and socially desirable responding (Levin & Zickar, 2002; Ones et al. 1996; Potosky & Bobko, 1997; Schmitt & Kunce, 2002).

Tactical deception. Tactical deception is a term developed from observations of nonhuman primates and is defined as "acts from the normal repertoire of the individual, used at low frequency and in contexts different from those in which [a primate] uses the high frequency (honest) version of the act, such that another familiar individual is likely to misinterpret what the acts signify, to the advantage of the actor" (Byrne & Whiten, 1985, p. 672). It relies heavily on statistical deviance criteria (Miles, 1986) because the primates studied could not convey their true intentions, as humans do in self-reports. Still, some definitions and measurement methods (e.g., lie scales) of human deception also focus on statistical deviance.

Deception as personality. Based mostly on empirical findings associated with SDR, a number of researchers have proposed that measures of faking are partially measures of personality (Ashton, Lee, & Son, 2000; Smith & Ellingson, 2002). Such claims seem to reflect the fact that there are a variety of definitions of SDR and faking (as mentioned above) more than that research has demonstrated clear links between personality constructs and acts of lying and misrepresentation. While aspects of personality may predict who is motivated or able to deceive in certain situations, it is debatable whether there exists an individual difference characteristic that creates an enduring tendency to fake across situations.

Some clinical theories state or imply that deception is an aspect of disordered personality. The Diagnostic and Statistical Manual (DSM-IV-R) lists "deceitfulness, as indicated by repeated lying, uses of aliases, or conning others for personal profit or pleasure" (American Psychiatric Association, 2000, p. 706) as one of seven criteria used to diagnose as having antisocial personality disorder, with estimates of the disorder's

prevalence in clinical samples ranging from 3% to 30%. Some evidence supports the notion that certain people perform deception on a regular basis. In a sample of Turkish adults, 66% of those diagnosed with conduct disorder and 68% of those diagnosed antisocial personality disorder demonstrated lying symptoms, compared to 36% of the adults in the normal sample (Doğan, Önder, Doğan, & Akyüz, 2004). Though not serving as a diagnostic criterion (Ford et al., 1988), lying has also been associated with other rare disorders like *psuedologia phantastica*, where individuals regularly tell lies or exaggerate the truth in the absence of clearly identifiable, external motivators (Larson, 1932; Myslobodsky, 1997; Selling, 1942; Wiersma, 1933). For such "pathological liars," lying may serve as an intrinsically satisfying outcome, rather than as a means to obtain some external reward (Ford et al., 1988). Also, some view clinical levels of narcissism as the result of self-deception (Paulhus, 2003). Though this claim seems reasonable in some cases, theoretical issues concerning the concept of self-deception (mentioned above) limit the credibility of this belief without further explanation.

Unfortunately, little empirical work has been conducted specifically to address these issues. At present, it then seems prudent to assume that the motivation to perform deception could result both from situational factors and from a person's natural disposition. It is also important to recognize that, though deception may be a "factor" of disordered personality, models of personality *disorders* are related only indirectly to traditional models of normal personality (Davis & Millon, 1999).

Expedient behaviors: bullshit and bluffing. Frankfurt (2005) defined the concept of "bullshit," or "humbug," to in such a way as to distinguish it from lying and deception: to bullshit is to make statements beyond one's knowledge while recognizing that some of

the statements are untrue. Although it is generally difficult to claim something as irrevocably true/untrue from a philosophical perspective (e.g., Bradac, Friedman, & Giles 1986), this concept is meant to capture a lack of concern or an "indifference to how things really are" (Frankfurt, 2005, p. 34), rather than any situation in which information perceived to be untrue is communicated (i.e., lying). Levin and Zickar (2002) re-labeled this concept as *expedient behaviors* "because truth and falsehood are no longer the basis for presentation... information presented is whatever is expedient at attaining the end" (p. 257), unlike other concepts such as "falsification" (i.e., lying) and (selective) "presentation."

Despite these arguments, my own view is that expedient behaviors are not distinct from other deceptive behaviors. They seem to represent a compound construct that describes the ratio of truths to falsehoods occurring within a single communicative interaction. Yet, virtually all forms of deception involve some degree of honest communication. Furthermore, the defining characteristic of expedient behaviors, indifference, is very similar to the existing, better-understood concept of conscientiousness. In engaging in expedient behavior, a person implicitly accepts that lies will be communicated. That lies will be communicated among truth in an unfocused or careless manner does not mean that it is distinct from deception. Rather, it appears to be a form of low-quality lying by people lacking conscientiousness or some other motivation to lie with care. This concept is still important to consider, however, in the sense that it implies that there will be individual differences in the way deceivers lie.

Similar terminology can be found in card games involving deception. In fact, a game called Bullshit, or Cheat, requires players to create false demeanors that reveal

neither truth nor dishonesty, much like bluffing in Poker. Bluffing is the sending of signals (e.g., betting in greater or lesser amounts) that suggest one has a better or worse set of cards than is actually the case (Sklansky, 1999). In general, to bluff is 1) to deter or frighten by pretense or a mere show of strength, 2) to deceive, and 3) to feign (Merriam-Webster Online). Thus, terms like bullshit hold meanings very similar to deception.

Concepts Distinct From Deception

Cognitive errors. A vast amount of psychological research (e.g., Bassili & Krosnick, 2000; Gilovich, 1991; Murphy & Balzer, 1989; Nisbett & Ross, 1980; Rynes, Colbert, & Brown, 2002; Schwarz, 1999; Shafir & Tversky, 1995; Tversky & Kahneman, 1974) has uncovered various types of errors in sensation, perception, and processing (e.g., fundamental attribution error, halo rating error, anchor and adjustment heuristic) that lead people to state and/or believe false statements. Whether these errors are isolated accidents, the result of natural cognitive heuristics, or indicators of low general intelligence, they explain why people say things that are untrue. To ignore these ubiquitous factors in the study of deception would lead many people to be accused of lying, despite their lack of a true intent to deceive. The mere act of stating something untrue is not a sufficient justification for defining deception (Galasínski, 2000); a person must believe that his/her message is inaccurate and untrue, regardless of whether it actually is or not.

In addition, the tendency of weak intentions *not* to become behavior may also explain why people "say one thing, but do another," without relying on deception concepts. Research has shown that intentions typically account for less than 30% of the variance in behavior (Armitage & Conner, 2001; Gollwitzer, 1999; Sutton, 1998).

Therefore, people who make mistakes when honestly estimating their intentions or who legitimately change their mind because of situational changes can endorse behaviors at two different times that are contradictory. In most cases, however, people would not be held accountable for errors in their predictions of future events unless the predictions were stated with an extremely high degree of certainty, they claimed to (or could be expected to) have the requisite knowledge to understand relevant causal factors, and/or they had control over the predicted events. When one or more of these things is missing from the context, people are more likely to be viewed as having been foolish than as dishonest. For obvious reasons, clear examples of these types of errors are made by weatherpersons making forecasts that turn out to be wrong; they are rarely viewed as liars.

At the same time, I am not arguing that the prediction of future events never involves deception. People can selectively add phrases like "I intend to...," "I plan to...," or "I think... will happen" to their predictions to remove accountability, should the statement turn out to be false. They can also add phrases in a manipulative manner to misrepresent their true level of confidence in the prediction, phrases like "I definitely...," "... without a doubt," etc. Therefore, it seems that failed predictions are interpreted partly on the basis of how they are communicated. The general point that I wish to establish is that messages can be complex. Deceivers can manipulate not only the direct content of a message, but also "meta-content," about things like one's conviction that the message is true or false.

Lastly, typical social communication norms can lead people to communicate messages that are not objectively true, in the absence of deceptive intent. One norm

encourages the omission of extraneous information to create "synoptic" messages (Miller & Stiff, 1993). When people are asked about the events of their day, they typically censor uninteresting activities. This kind of omission is distinct from selective presentation and other forms of deceit because the information excluded is not necessary for the target of communication to develop an accurate representation of the truth. In summary, there are many factors that cause honest people to make statements that can be viewed as incorrect and not true.

Regulation and intent to change. People can act inconsistently over time and still be honest. As mentioned above, the strength of people's stated intentions is one factor affecting whether they will be viewed as liars. Changes in the situation are another factor. Yet, another factor is any attempt (or failure) to self-regulate. It is not uncommon for people to break New Year's resolutions, for example. People who try to maintain their resolution, but give up, may be considered "weak-willed" as opposed to deceivers. In another example, people may automatically believe in negative stereotypes, but actively suppress the stereotypes (Crandall & Eshleman, 2003). If they genuinely desire to eradicate such beliefs, people could be viewed as making an "honest" attempt to change, though natural stereotypes may manifest themselves at times.

Spinning. While lies and blatant exaggeration constitute deception, the act of "spinning" seems to reflect a general type of persuasive communication. Spinning is giving "accounts [to] explain an offense to an audience by justifying it, excusing it, refusing to accept responsibility for it, or conceding to having committed it" (Folkes & Wang, 2003). Although deception is often used to accomplish the same goals as spinning, it is not defined by the avoidance of taking responsibility and does not include

honest attempts to justify an offense. Thus, deception refers to only a portion of the concept of spinning.

Sarcasm and humor. Sarcasm and humor appear to be related to deception (T. R. Levine, *personal communication*, August, 2, 2005), though somewhat ambiguously. In both cases, people may intentionally communicate messages that are false. Yet, the defining characteristic of sarcasm/humor is that the target of communication understands that messages received are literally false. That is, the source of the sarcasm sends two signals to the target in the forms of content and tone (or facial expression), where tone honestly reveals that the content is false (Vrij, 2000). So, one could view tone as providing a kind of meta-content for the exact purpose of revealing that the message content is untrue. Similarly, humorous jokes are often told without any real intention to make the audience believe the information to be factual. (Exceptions to this may be "practical jokes" and pranks that rely on deception to embarrass a person.) As a result, sarcasm and humor generally fail to capture the nature of what it means to deceive.

Silence and partial disclosure. Unlike selective presentation, an explicit refusal to state the complete truth is an honest form of communication, as is common practice with consent forms in psychology experiments using deception or in courts where people plead the 5th Amendment. So, an honest admission that partial disclosure was provided is not an act of deception (Ekman, 2001). Furthermore, an example given earlier showed that it is not always necessary to provide a comprehensive, detailed account of some object or event to allow the communication target to develop an accurate representation.

Formal and understood roles. A consideration of roles reveals one issue that theorists may find difficult to resolve. Particularly when discussing personality traits, the

distinction between honesty and dishonesty depends on whether one believes that expressions of relatively stable individual difference characteristics will manifest themselves similarly across different times and situations. If this assumption is true, people who express themselves differently across situations with respect to some trait could be viewed as being deceptive. A person who is bubbly and vivacious at home, but solemn at work, could be seen as fake in (at least) one of the contexts. The opposing view allows people to have honest expressions that differ across settings (Janis & Gilmore, 1965). Although the issue of how personality and other characteristics should manifest themselves in varied situations is beyond the scope of this paper, it does affect how deceptive behaviors are defined fundamentally.

Role theory provides one framework that resolves this issue sufficiently enough so as to allow deception to be defined clearly, even if the underlying theory is imperfect. The major assumption is that situations place role demands on people that alter prototypical behaviors (Leary & Kowalski, 1990). People who strive to fulfill their roles should then be expected to act differently depending on the time and place. Using an exaggerated example to demonstrate this point clearly, people who kiss their spouse before leaving home are not considered fake if they do not kiss their boss before leaving work. Role theory, however, does not mean that all behaviors are genuine; role expressions can themselves be faked when people override their natural tendencies in specific settings.

One domain directly relevant to the discussion of roles is acting (Buller & Burgoon, 1994; Dawson, 1980; DePaulo & Rosenthal, 1979). Even though actors may use deceptive techniques (e.g., wearing a wig) to alter the audience's perception, the

larger context of acting is not deceptive because the audience knows that actors are playing their roles. For instance, audiences know from the film credits that Ben Kingsley was not actually Gandhi (in the movie "Gandhi"), even if believing that Kinglsey accurately portrayed Gandhi. One exception to this rule is that "actors" in real situations (e.g., spies and undercover law enforcement agents) do intend to fool their targets and do not reveal this intention during the act of deception.

A Definition of Deception

Although the discussion above reveals that deception encompasses various concepts, a definition and theory of deception should show "coherent psychological properties" (Hyman, 1989, p. 134). Hyman (1989) surrendered to the fragments of theory and empirical work available at the time and asserted that "no psychology of deception exists in this strong sense... no one has managed so far to demonstrate that a single, coherent framework can meaningfully account for the psychological issues involved in various types of deception" (p. 134).[3] While the study of deception is no easy task, the review above strongly implies that a perceptible set of commonalities tie together concepts like lying, faking, and so forth. First, a person must perceive something as being true, about the world or about oneself, before he or she can distort, hide, or mask it (Frankfurt, 2005). Second, deception is always characterized by an intention to have the target of communication accept information as true. "An intention is a representation of a future course of action to be performed. It is not simply an expectation or prediction of future actions but a proactive commitment to bringing them about" (Bandura, 2001, p. 6).

[3] Upon concluding his review, however, Hyman expressed an "optimistic" attitude about "the possibility of a coherent psychology of deception in the strong sense" (p. 151).

Miller (1983) defined deceptive communication as "message distortion resulting from deliberate falsification or omission of information by a communicator with the intent of stimulating in another, or others, a belief that the communicator himself or herself does not believe" (p. 92-93). Galasínski (2000) provided a somewhat broader definition: "[communicated] false beliefs intended to be induced in the target" (p. 19). Based on these definitions and the literature reviewed, I define deception as *the communicative process of intentionally presenting an inaccurate representation of some object.* For the most part, this definition is consistent with preexisting ones and may simply be considered an operational one that underlies the building of theory in this paper.

Before proceeding, I explain the five aspects of this definition. First, deception is a form of persuasive communication (Wilson, 1997). Whether transmitted verbally, physically, or through some other medium (e.g., electronically), deception describes how a signal is sent by one person ("source" or "agent") to alter another person's ("target") perception (Whiten & Byrne, 1988a). Also, the source and target may be different entities or one and the same. The target may also be nonhuman, like an animal or automated machine (e.g., polygraph or computerized test). Second, deception is a process involving motivation, cognition, behavior, and social interaction, with no single aspect adequately describing deception. Third, deceptive acts are intentional, not merely acts of presenting incorrect information that might result from random error, biases, or low intelligence. Deception might also involve true statements, made either by accident (i.e., by a poor deceiver) or when "double bluffing" (tricking the target into believing that correct information is actually false). A focus on intention also implies that deception should be

distinguished from its outcomes; deception occurs whenever a person engages in it, even if the target is not fooled. Fourth, deception encompasses any form of intentional misrepresentation, ranging from selective silence to subtle exaggerations, to outright lies. (In a simple example, failing to say the word "not" when communicating a message can drastically alter its meaning, even if all other words are true.) Since people may not "know" truth in an absolute sense, misrepresentations can be made on a relative scale and include probabilistic messages. The notion of misrepresentation (Galasínski, 2000) as a form of deception is evident in the legal domain where people must tell "the *whole* truth and nothing but the truth." Fifth, a deceptive message or signal response can refer to any *object* (DePaulo et al., 1996). Lies may refer to a person's thoughts, feelings, behavior, or personality, as well as to nonliving objects (e.g., "The sun is blue").

The definition above can be used to understand a range of communicative behaviors across different contexts, and provides the initial step towards creating a unified, psychological theory of deception. In the following sections, I review formal theories and propositions about when and how deception is thought to occur. Literature is drawn from multiple domains including studies of nonhuman primates, warfare, daily social interactions, computer security, test faking by job applicants, and criminal lie detection methods. Throughout this review, it will become evident that researchers have tended to focus on some aspects of the general deception process and not others, necessitating the development of additional theory.

Before describing the general process of deception, however, it is helpful to establish some dimensions that differentiate certain types of deceptive actions from others. While formal taxonomies of lie content do exist in the literature (e.g., DePaulo &

Rosenthal, 1979; Hopper & Bell, 1984; Metts, 1989; Turner et al., 1975), my main intention here is to recognize just two important factors that have a moderating effect on general deceptive processes discussed across scientific domains.

Dimensions of Deception

Dichotomous versus categorical/continuous. The truthfulness (or deceitfulness) of any communicated message will have a certain form. To borrow from the language of psychometrics, deceptive responses can be dichotomous or categorical/continuous. Many simple messages have a dichotomous form (e.g., yes-no responses), such that one answer indicates the truth and the only alternative indicates a lie. The dichotomous lie is particularly useful in criminal interrogation because detecting or preventing a person's deception will reveal the truth.

More complex messages are associated with multiple possible responses that make it difficult to ascertain the truth even if the target knows the information being received is false. Such messages also provide more opportunities for the deceiver to construct a false or distorted message. For example, a college applicant who lies about his or her ethnicity to increase the chance of being admitted may be able to select from a set number of alternatives (i.e., categorical message), or even be allowed to write in a category (e.g., "Other Ethnicity"). Categorical/continuous messages, however, may also be more difficult to construct than dichotomous messages (Walczyk, Schwartz, Clifton, Adams, Wei, & Zha, 2005). Deceptive responses to open ended questions usually must conform to social communication norms and meet informational demands.

Offensive versus defensive. Researchers have drawn a strong distinction between two types of lies, based on their function (e.g., Ekman & Friesen, 1969; Galasinski, 2000;

Hopper & Bell, 1984; LaFrenière, 1988; Lee, 2004; Metts, 1989; Turner et al., 1975). Some lies seem to be offensive acts (also called commissive, active, and overt) that deliver false or distorted information to the target. Though typified by outright lies, offensive acts can include subtle or indirect attempts to actively influence the target's perception, such as exaggerations. Acts of denial when accused of some action would also fall under this category (Inbau, Reid, Buckley, & Jayne, 2004).[4]

In contrast, defensive (or omissive) acts of deception involve the withholding of information from the target. The decision to present true information selectively is one such act that still leads a target to develop a distorted view of some object (Buller & Burgoon, 1994; Hopper & Bell, 1984; Miller & Stiff, 1993). Egregious cases of defensive deception involve selective presentation when the norms of communication or social interaction clearly dictate more disclosure, as when people are asked to describe something completely and accurately (Elliott & Culver, 1992, provide one example). Another type of defensive action is the inhibition of natural responses. The classic example of this is the inhibition of perspiration to avoid revealing nervousness, either through mental concentration or by physical means (e.g., wiping it away). This act is defensive because the source is preventing the communication of true information that would naturally be sent to the target and allow the target to understand the truth. Lastly, it should be emphasized that both offensive and defensive acts can constitute deception, and that both do not need to occur simultaneously.

[4] Claims for distinguishing lies based on whether they increase a person's image in the positive direction versus whether they involve denying negative attributions about one's image have only received marginal empirical support (Paulhus, 1984).

The Process of Deception

At a general level, any rational attempt to deceive is caused by a discrepancy between the truth and a desired state, a subsequent decision to act deceptively, the selection of one or more strategies that would cause a target to perceive the discrepancy as reduced or nonexistent, and the execution (or performance) of the chosen strategy to generate a deceptive message. Beyond this process, a series of subsequent processes largely determine whether the deception is ultimately effective for attaining some reward. Upon receiving the deceptive message, the target of deception may or may not attempt to evaluate its credibility. If the target does not evaluate the message, for whatever reason, the source will be considered successful. If the target does attempt to evaluate the message, it may still fail to detect deception, again resulting in the deceiver being rewarded. If, however, deception is detected, then a different set of negative outcomes usually result.

Motivation

Given that any intentional behavior is enacted for a purpose, deception is goal-directed (Ford et al., 1988; Frankfurt, 2005; Mitchell & Daniels, 2002), with the general higher-order goal being to make a target believe something perceived to be untrue (Ekman, 1985; Lewicki, 1983; Shulsky, 2002; Zuckerman, DePaulo, & Rosenthal, 1981). The end goal of any deception-related action is probably related to the attainment of some kind of personal, social, or material reward. Thus, psychological theories of motivation should be used to help explain who engages in deception, and the means by which people produce deceptive messages.

Typically, deception is performed when a person's true state is discrepant from a desired state. By reducing or eliminating the *appearance* of the discrepancy, as perceived by a target (labeled "perceived discrepancy" from here on), the source can claim to have achieved the desired state and receive a contingent reward (Ekman, 2001; Leary & Kowalski, 1990). For example, job applicants who embellish their qualifications probably do so because they do not feel qualified (McFarland & Ryan, 2000). It is also widely believed that the students who are most likely to cheat on an exam are the ones who have a low ability, low motivation, or limited time to study (Schab, 1991). Reasons for cheating may also stem from competition, when others increase the level at which tangible and social rewards are obtained (Anderman, Griesinger, & Westerfield, 1998).

Conversely, students who are knowledgeable and confident about their test preparation are probably less likely to use deception, particularly if it requires more effort than acting honestly or entails a high risk of punishment. In support of this general notion, Levitt and Dubner (2005) inferred that teachers tended to help their students on state exams (by giving test answers) more often when the students produced initially low scores. Consistent with this notion, Jones and Pittman's (1982) descriptions of "an ingratiator's dilemma" and "self-promoter's paradox" posit that undesirable or unqualified people are most likely to resort to flattery or boasting, respectively. More generally, Burton (1963) found honest behavior and IQ to be correlated (r = .22).

Discrepancies can involve comparisons between any number of features associated with a person, object, or situation (see Medin, Goldstone, & Gentner, 1993). Higgins (1987) proposes that psychological states of discomfort (i.e., negative affective

states) become stronger as more features are mismatched and as the magnitude of those mismatches increases. Various psychological theories (e.g., goal-orientation, self-regulation, and self-discrepancy) help to explain how people feel and act when facing different kinds of discrepancies, but most propose that the psychological discomfort produces motivation to reduce the discrepancy because people naturally seek consistency (Cooper & Fazio, 1984; Elliot & Devine, 1994; Higgins, 1987; Higgins, Roney, Crowe, & Hymes, 1994; Higgins & Tykocinski, 1992). With deception, however, the aim is to reduce the discrepancy perceived by the target, rather than to transform one's true state into the desired one.

Specific discrepancies motivating people to deceive may vary by context, referring to things such as the remaining progress needed to accomplish some task (e.g., Carver & Scheier, 1990), inconsistencies in one's beliefs (e.g., Festinger, 1957) or self-concepts (Bandura & Locke, 2003; Higgins, 1987), or the degree of fair treatment received (e.g., Robinson & Bennett, 1997). In addition, individuals probably vary in their propensity and ability to assess discrepancies (Higgins, 1987). I provide just one example using goal orientation concepts (Dweck, 1986) to show how motivation might be used to predict differences in discrepancy assessment. Students with a strong learning goal orientation (i.e., attempting to master new material) should be relatively unconcerned with the discrepancy between their performance and others' performance. If they are given a class test scored normatively, then they should be expected to be uninterested in cheating since high scores achieved in that way would not demonstrate mastery. In contrast, students with a strong performance goal orientation (i.e., seeking to gain the

favorable judgment of others) should find cheating appealing because it accomplishes their goals, assuming that the act of cheating is not discovered.

At the same time, motivation for deception may also be determined by higher-level judgments, beyond task discrepancies. People with a strong sense of integrity may refuse even to consider using deception because it would create a new discrepancy between their actions and their values and self-concept. Similarly, social values and pride may override motivational rewards offered for discrepancy reduction through deception. For example, some military groups in history only resorted to deception in warfare when they viewed themselves as weaker or inferior to their opponents, and other groups avoided the use of deceptive tactics because they viewed themselves as superior (Handel, 1989). Similarly, one study showed that certain naval cadets refused to engage in self-presentation because they held negative views of social conformity by people of high status (Gergen & Taylor, 1969). Thus, recognizing that a discrepancy was small actually motivated people towards honesty.

In addition to rewards, deception is also associated with motivation to avoid detection and associated punishment. Without the threat of punishment, deceivers may seek to reduce discrepancies in extreme ways and create deceptive messages that are implausible. General research findings support this notion. In fact, many measures of faking are aimed at identifying people who fail to produce realistic messages, making extravagant claims like "I never lie or steal." However, "certain research indicates that faking in real life consists more of a subtle slanting of responses than one of uninhibited dissimulation... than just to 'look as good as possible'" (Doll, 1971, p. 9; Smith & Ellingson, 2002). When detection and punishment are concerns, effective deceivers must

plan and monitor their actions in such a way as to maximize their chance of obtaining rewards while minimizing the chance of being punished. Therefore, discrepancy reduction is expected to occur in more moderate degrees and through more sophisticated or subtle means.

It is axiomatic that people will attempt to avoid large or severe punishments. However, their actions are also determined by perceptions of the *likelihood* of punishment. A deceiver who perceives a low probability of being detected engaging in deception may engage in riskier behaviors, even if the punishment is severe. This idea is consistent with the consensus that research has failed to demonstrate the effectiveness of capital punishment as a deterrent of murder (e.g., Bailey & Peterson, 1999; Fagan, 2005). From a theoretical standpoint, Vroom's (1964) expectancy theory provides an appropriate framework in which to view the various rewards and punishments that exist in many situations (Leary & Kowalski, 1990; Schlenker, 1980). According to VIE theory, a motivational force results from the valence (value) of an outcome, the instrumentality (functional usefulness as a means to achieve other valued outcomes) of a particular level of performance, and expectancy (perceived probability) that an action will lead to the outcome (Mitchell & Daniels, 2002).

Without denying the importance of outcome valence or instrumentality, Figure 1 illustrates a hypothesis of the basic process determining people's motivation to deceive based on their expectancies for rewards and punishments. Regardless of whether people engage in specific, rational calculations of probabilities or make subjective clinical judgments, the expected value of receiving a reward through deceptive means interacts with the expected value of receiving punishment, assuming the target attempts to and/or

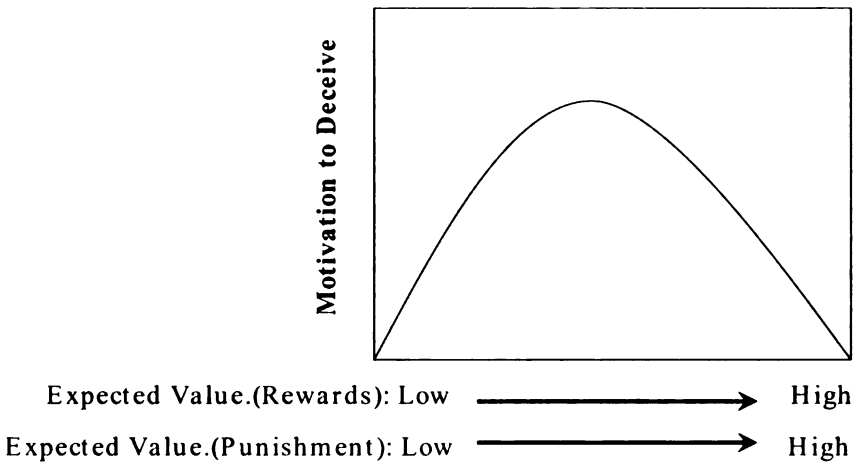
can detect deception.[5] The motivation to deceive should be low when the reward expectancy is low, but should increase as the reward expectancy increases. However, this motivation will begin to decrease as the threat of punishment increases. As depicted in Figure 1, the threat of punishment has a relatively small effect until a person is actually likely to use deception.

Simple models of deception, such as those used to describe job applicants' faked responses, have assumed that the motivation to deceive is positively, linearly related to one's behavior. In the real world, however, rewards tend to be distributed dichotomously (i.e., people either get the job or they get nothing) rather than continuously and linearly, especially when the stakes are high (Snell, Sydell, & Lueke, 1999). Consider a job applicant who produces faked responses on a personnel selection test that is used to hire 2 out of 50 applicants. In this case, "there is no reward for increasing rank order incrementally, yet there is a substantial reward for increasing rank order in a quantum leap to the top of the pool" (Levin & Zickar, 2002, p. 261). As a result, it seems more accurate to model the motivation to deceive in relation to the expected distribution of reward creating that motivation, assuming people are rational.

Figure 2 depicts a modified version of the theoretical relationship in Figure 1 that might be more accurate in real-world contexts. First, there is an assumption that a person's *motivation* to deceive will show some correspondence with deceptive *behavior* (Armitage & Conner, 2001), to keep the model simple. Thus, Figure 2 substitutes "deceptive performance" for "motivation to deceive." The difference emphasized is that deceptive performance occurs in a stepwise fashion, as suggested by Levin and Zickar,

[5] This example presupposes there is a reason to deceive, meaning that the expected value of obtaining a reward is higher through deceptive means than honest means.

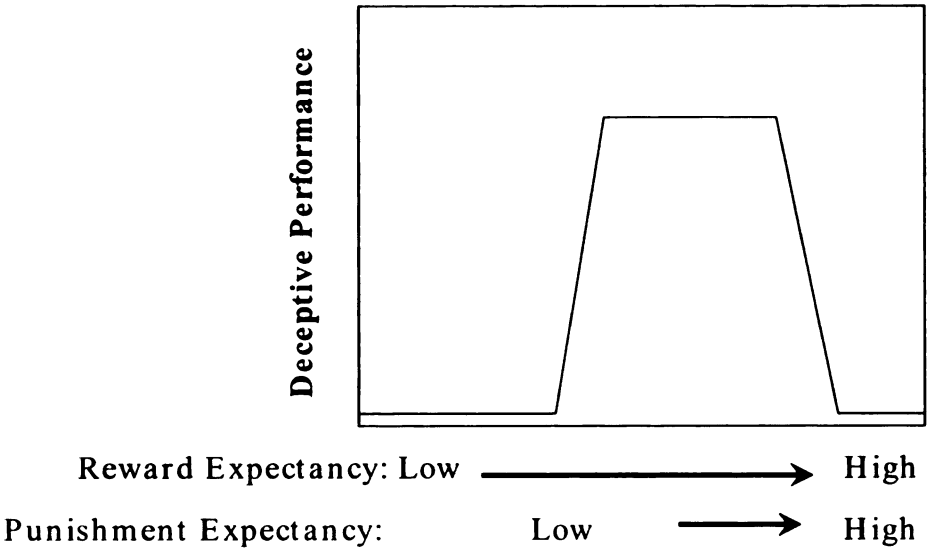
Figure 1. Motivation to Deceive Based on Expect Values of Rewards and Punishment.



such that people will only engage in deception when the likelihood of obtaining a reward reaches an acceptable level; below that level, the person would not even bother to try. The use of deceptive performance then remains relatively constant so as to ensure that the dichotomously distributed reward will be received, and remains high as long as the threat of punishment remains low. For the job applicant example, Levin and Zickar (2002, p. 261) explain, "When a single candidate is to be selected for a key position, particularly one high in power, prestige, or compensation, that position offers high relative rewards to a job seeker. The motivation and rewards for the applicant to increase the rank order of presented qualifications is high." Thus, although the motivation to deceive may naturally fall on a continuous distribution of probabilities, the deception-based performance observed in the real world may occur in an all-or-none fashion.

It should also be stated that, for the sake of simplicity, Figure 2 describes situations when people do not take into account the possibility that others competing for the same reward may engage in deception as well. That is, people assume that performing deception will directly increase the probability of reward attainment (e.g., gaining employment). In reality, however, deceivers are likely to be competing with other deceivers, as well as honest people (cf. Snell et al., 1999). When people are aware of that possibility, reward and punishment expectancies will change and necessitate a more complex model. It is likely that the threshold for reducing deception with punishment will shift upwards (to the extent possible) even when the threat of punishment is high, when the reward cannot be obtained unless one presents his/her current state as being radically better than other deceivers, as well as better than honest people.

Figure 2. Deceptive Performance Based on Expectancies for Rewards and Punishment.



It is also important to point out that punishments will only have an effect when they are noticed. Research has shown that explicit warning statements about punishments can dissuade people from using deception at least on paper-and-pencil tests. Data from a large-scale survey showed that honor codes were associated with reduced rates of cheating on college exams (McCabe & Bowers, 1994). For job applicants taking personality and cognitive ability tests, warning statements (e.g., "Responses are subject to verification... attempts to falsify information will be grounds for dismissal") also reduced score inflation, presumably resulting from deception (Kluger & Colella, 1993; Vasilopoulus, Cucina, & McElreath, 2005).

Kluger and Collella (1993) administered a customized biographical data, or biodata, instrument to 429 applicants for nurse's assistant positions, half of whom were warned not to fake because a "special scoring system" would detect dishonest responses and reduce false scores. They found that the warned group tended to produce less extreme average scores on Likert-type scales than the unwarned group, with a small effect size (average $|d| = .17$). This effect was found only for items that were transparent based on social desirability ratings, supporting the hypothesis that people were faking more often on items for which the best answer was apparent.

In one of the few studies testing multiple motivation conditions, Doll (1971) placed 300 naval aviation candidates in a simulated selection situation. Every participant first completed a biodata inventory "as honestly as possible" and then pretended they were applying for a research position and answered the inventory again according to one of the following instruction sets: 1) "subtle-interview" – providing answers that might need to be defended in an interview, 2) "subtle-lie score" – providing answers that are not

exaggerated enough to trigger a lie score, and 3) "unsubtle" – answering to look as good as possible. Across different types of response formats, item-level signs of faking were much more frequent when people were in the unsubtle condition versus the two subtle conditions.

Schrader and Osburn (1977) extended Doll's work and examined whether subtle versus unsubtle faking attempts interacted with how much specific information about the situation was known. In that study, participants were instructed to act like applicants for either a general managerial job or a specific sales representative position, where detailed job information was provided in the latter case. Although the study results did not confirm the bulk of their predictions, data did indicate that people using subtle faking approaches to obtain the sales position had biodata scores that were, on average, barely larger than honest participants, as compared to the other three cell conditions.

Dwight and Donovan (2003) reviewed the 15 studies on warning effects in the literature and discovered that their mean-weighted effect was small, $d = .23$. Recognizing the considerable variance around this estimate (ranging from no effect to $d = .90$), the authors conducted an experimental simulation in which 190 undergraduates applied to a fictitious university, where accepted applicants would receive $75. Applying consisted of taking personality surveys and other measures, with people assigned to different conditions: unwarned, warned about being detected through special questions, warned about penalties for faking (i.e., invalidated scores), and warned about both detection and penalties. Results showed that warnings appear to be most effective for reducing score inflation when mentioning both the detection method used and penalties incurred (*d*s ranging from .38 to .75 depending on the measurement scale). All of the warnings were

associated with reduced scores on lie scales comprised of "bogus items" (about fictitious events or characteristics), with *d*s ranging from .27 to .41 (Dwight & Donovan, 2003). Furthermore, more people having high bogus scale scores were able to increase their rank-order within their group when unwarned.

Ultimately, it seems that the effect of a warning statement on deception probably depends mostly on what is being warned. If the threat of detection is realistic, the punishment is severe, and the reward expectancy is low (i.e., not worth the trouble), any rational person would avoid using deception. Conversely, people should be more likely to use deception when rewards expectancy is high, the threat of detection is unrealistic, and punishment is weak and worth the risk. Kluger and Colella's study supports this idea since it used a large reward and a moderately realistic threat, but only a moderate punishment (i.e., scores were just reduced, not invalidated), and produced a smaller effect than other studies with stronger punishments.

Before concluding, it is noted that punishments and warning statements are not purely motivational factors on deception; they can also have informational value on deceptive decision making processes. For instance, Hughes, Dunn, and Baxter (1956) found that when hiring managers in an insurance company altered the records of favored applicants, the managers only raised credentials to a minimum degree rather than maximally (e.g., an answer would be raised to just $7000, to bump a person up into the $7000-9999 scoring bracket). In this way, knowledge about the detection method provided in warnings may inform people about how much they should use deception to achieve rewards before being detected, as suggested in Figure 2.

Heckhausen (1987, 1991; Heckhausen & Gollwitzer, 1986) and Gollwitzer's (1990) rational action (or Rubicon) theory helps to bridge this gap between motivational and cognitive models of action (Mitchell & Daniels, 2002). In brief, rational action theory proposes that any motivational decision to commit to a goal and act intentionally is followed by the development of an action plan or strategy that enables behavioral progress toward that goal. Various cognitive processes help to arrange the multiple, hierarchical goals that people have so that actions directed toward subgoals will eventually help to accomplish superordinate ones (DeShon & Gillespie, 2005). Following this planning step, people monitor their goal progress and adjust their behaviors when necessary. (This latter point adds complexity that is not central to the theory developed in this paper.) Thus, it seems that punishment and warnings can reduce the likelihood that a person will use deception, but can also be processed cognitively by deceivers who wish to develop optimal faking strategies that maximize their reward expectancy while minimizing their punishment expectancy.

Cognitive and Behavioral Models of Deception

Occasionally, researchers state some beliefs and assumptions about prior knowledge or skills needed to execute a deceptive plan (e.g., Buller & Burgoon, 1994; DePaulo & Rosenthal, 1979; McFarland, Ryan, & Kriska, 2003; Snell et al., 1999). For the most part, however, theorists have specified little about the cognitive and behavioral processes that enable people to perform deception, either successfully or unsuccessfully. Deception has instead been treated as a "black box" that just requires some kind of motivation and "effortful processing" to produce a deceptive message, or faked behavior.

Hyman (1989) reviewed the use of deception by magicians, psychics, and swindlers since the late 1800s and discovered a reliance on "conjurer" models. These models produced visual illusions by exploiting unconscious inferences in perception, including what are regarded today as schemas, priming, suggestion, the manipulation of passive attention, and more. After the wave of behaviorism washed over these "mentalistic models," specific theories of deception-related acts remained dormant until the cognitive revolution decades later (Hyman, 1989).

Activation-Decision-Construction Model. Recently, Walczyk, Roper, Seemann, and Humphrey (2003) proposed one cognitive framework for the detection of "person-to-person [lies] about life events" (but primarily for criminal lie detection) that rests on three main assumptions. First, semantic long-term memory enables people to organize information about how others view the world, and episodic memory enables them to organize autobiographical information based on similarity and temporal contiguity. Second, long-term memory primes other processes in a cognitive network, once activated. Third, the retrieval of true information from long-term and working memory is easier and occurs in a relatively short amount of time (approximately 400ms or less), as compared to the retrieval or fabrication of false information. Based on these cognitive principles, the researchers formulated their Activation-Decision-Construction Model (ADCM).

The ADCM appears to resemble rational action theory and Leary and Kowalski's (1990) two impression management components quite closely. The activation component refers to attentional and awareness processes that cause a person to recognize the possibility of using deception in a given situation. That is, the person perceives

something to be true and is aware that the truth may be replaced with a lie. Once aware of a deception opportunity, the person then must engage in goal-relevant processing to decide whether he or she will attempt to use deception. Once this decision is made, the person then begins the construction of a lie. With the "truth" in active memory, the decision to lie will enact an inhibition process that prevents the deceiver from expressing the truth. The truth is then replaced with a lie constructed based on knowledge and inferences about the target (e.g., intelligence level and degree of suspicion held). Social-cognitive processes also monitor and censor implausible or inconsistent information from being used to construct the lie, before it is eventually expressed as a behavior.

Because the construction component accounts for most of the cognitive processes that would be defined as the actual process of lying, it is regrettable to note that the ADCM provides only a minimal description of this component. Walcyzk and colleagues were instead focused on improving *lie detection* by measuring response times, based on the premise that lies take longer to express than truths. One can also question the degree to which most people are able to construct high quality lies. Limited by their abilities and knowledge and by situational constraints (e.g., the amount of time allowed for the deceiver to respond), people may not be able to engage in complex monitoring and censoring processes. In summary, the ADCM provides a good starting point for developing a theory of deception and has received general support in initial experiments (Walczyk et al., 2003; Walczyk et al., 2005), but leaves critical questions unanswered.

Applicant faking model. McFarland (2000; McFarland & Ryan, 2000) developed a broad model to explain faking behavior in job applicants, based on Ajzen's (1991) theory of planned behavior. In the model, perceived behavioral control (i.e., efficacy),

subjective norms, beliefs, and attitudes toward faking predict an intention to fake. Subjective norms consisted of perceptions about how "important others" view faking (e.g., parents and friends), whereas attitudes referred to the test-taker's beliefs. These three predictors of the intention to fake were hypothesized to be moderated by situational factors related to motivation, such as warnings not to fake and the valence of a reward (e.g., a job, or doing well on a test). The intention to fake then interacted with the "ability to fake" to produce "faking behavior," operationally defined as impression management scores and difference scores between honest and faked answers. Faking behaviors then affected the test's validity.

This model provided a unique contribution over previous ones, especially with its explicit mention of a "faking ability" variable. It was proposed that this ability consists of self-monitoring, knowledge of the constructs measured by the faked test, and the transparency of test items (McFarland & Ryan, 2000). For the most part, McFarland's initial test of the model showed that these variables were related, though not always in hypothesized ways. However, the model contains a number of conceptual ambiguities that limit its usefulness. For one thing, it is puzzling that the valence of a reward is regarded as a moderator of intention to fake rather than as a direct cause, whereas norms and attitudes toward faking are viewed as causes of faking. Although McFarland's model is mathematically equivalent to one in which these relationships are reversed, such an alternative model would seem to make more sense, per VIE theory. That is, reward valence and perceived behavioral control should determine the intention to fake directly. Also ambiguous was that one of the two operational definitions for faking behavior really represented an outcome of faking behavior (i.e., mathematical difference in scores

between a person's honest and faked answers). McFarland noted this key limitation of the model (p. 122): "the model of faking behavior presented in this study does not address the *process* of faking. That is, the thought processes and strategies people use when faking was not examined."

Interactional Model of faking. Snell et al. (1999) provided a simpler, narrower model of job applicant faking. Their Interactional Model proposes that "successful faking" (a variable never defined theoretically) is the outcome of an interaction between one's ability to fake and one's motivation to fake. As the model was used to predict motivational outcomes (i.e., who fakes and by how much), this interactive portion of the model contributes little to theories of faking beyond the application of VIE theory. However, the model does specify that the ability to fake is determined by dispositional factors (e.g., general mental ability and emotional intelligence), experiential factors (i.e., knowledge of the behavior to be faked and desirable responses, and contextual factors), and test characteristics[6] (including the item format, type, and scoring methods). The motivation to fake is determined by demographic variables, dispositional factors (e.g., impression management tendencies, integrity and morality, Machiavellianism, manipulativeness, and locus of control), and perceptions of 1) other's faking behavior, 2) other's attitudes toward faking, 3) fairness (and the use of faking to remedy unfairness), 4) own attitudes toward faking, 5) expectations for success (i.e., expectancy), and 6) outcome valence.

Information Manipulation Theory. McCornack's (1992) Information Manipulation Theory (IMT) was intended to explain how people distort communications

[6] If generalized to more interactive social contexts, test characteristics affect one's ability to deceive in much the same way as mode of communication does. Thus, an oral versus a written test would be analogous to lying verbally versus through a written message.

for deceptive purposes. Based on Grice's (1989) theory of conversational sense-making, IMT posits that four "maxims" of verbal communication that can be monitored and altered (McCornack et al., 1992). In conversation, individuals expect information to be provided in sufficient *amounts* of information (i.e., without missing key information), and to have *veracity* (or quality, regarding truthfulness and accuracy), *relevance* to the overall conversation, and *clarity*. IMT further proposes that deceivers actively manipulate (i.e., alter, reduce, or omit) information to violate these maxims to create a lie. As such, IMT also implies that people who notice or perceive communicated messages to be in violation of any maxim will be more likely to suspect that deception is occurring or to detect it (Levine et al., 2000).

In an early empirical test of IMT, McCornack et al. (1992) had more than 1,000 undergraduate students rate their perceptions of honesty and of "communicative competence" (i.e., a type of social skill/competence) for messages embedded in 40 communication scenarios, where some messages were manipulated in some manner and other messages were honest. The results showed that perceptions of message honesty and of communicative competence were positively correlated with the maxims, though manipulations of message amount were perceived as more competent than manipulations of the other maxims. A few other studies (e.g., Jacobs, Dawson, & Brashers, 1996) have replicated and expanded support for IMT, with one study demonstrating cross-cultural differences, such that students in Hong Kong only perceived violations of quality and relevance as deceptive (Yeung, Levine, & Nishiyama, 1999).

Overall, IMT provides valuable insights about the immediate goals people might have for generating lies. However, it does not actually describe the causal processes

involved in manipulating the amount, veracity, relevance, or clarity of information. Also, empirical tests of IMT have focused on whether people perceive messages in violation of the maxims as deceptive, not whether deceivers actually alter these message characteristics. In fact, such research implies that the logic of IMT is actually backward. IMT claims that people violate maxims to create deceptive messages, but fails to acknowledge that skilled deceivers probably aim to fabricate messages that do *not* violate any maxims, while still altering the truth.[7] Thus, IMT provides a categorization system for judging the appearance of a message as deceptive, but does not explicates the process of how deception is accomplished.

Interpersonal Deception Theory. Buller and Burgoon's (1996) interpersonal deception theory (IDT) describes how deception occurs in a dynamic, socially-constructed environment. IDT also uses Grices' maxims as standards by which targets can judge the credibility of a statement, as well as standards that deceivers aim to meet when producing their false messages. IDT considers deceptive messages to be comprised of three components: a central message (i.e., verbal content), ancillary messages that support the credibility of the message and protect the deceiver (i.e., meta-messages), and inadvertent behaviors (i.e., leakage). These deceptions are caused, in part, by the motivational factors of arousal, negative affect (including guilt), cognitive effort, and attempted control. However, situational factors interact with the person variables to determine the complexity and timing of communication, including transmission and reception of deception.

[7] A counter argument might be that it is impossible to distort the truth without violating these maxims. This point certainly holds true for veracity, because veracity is essentially synonymous truth/deception and, therefore, tautological. However, it might be possible to produce deception without violating the other maxims.

Linguistically, *central messages* are produced by strategically reducing the amount and clarity of information (Buller, Burgoon, Buslig, & Roiger, 1996). They can also be made to refer to irrelevant things or be given improper specificity (e.g., everyone vs. a few people), qualified, objectified, and altered in style (e.g., past versus present tense). Ancillary messages are created by controlling behaviors to meet conversational expectations, such as appearing pleasant and relaxed. Upon receiving a deceptive message, the target may evaluate the message for suspiciousness, but this process also depends on contextual factors like the target's propensity to assume messages are true, the deceiver's reputation for credibility, priming, and more. Once a deceiver transmits a message, feedback is used to regulate behavioral attempts to discount suspicion and to suppress additional leakage.

McCornack (1997, p. 91) began a review of deception by remarking: "Despite nearly 25 years of research, deceptive communication remains a scholarly domain devoid of viable theory." He then critiqued IMT and IDT, the two theories spawning the most dominant lines of deceptive communications research. Among the damning criticisms of his own theory, McCornack admits that "IMT is not actually a *theory* at all" (p. 91) since it has no testable propositions and poorly specifies the processes of message production, message characteristics (i.e., formation), and message reception. IDT, on the other hand, was viewed as being too broad and detailed, lacking integration across the various concepts. Others have harshly criticized IDT as a non-theory because its propositions tend to be purely descriptive and not explanatory (DePaulo, Ansfield, & Bell, 1996; Stiff, 1996). IDT also confounds general acts of persuasion with deception (Burgoon & Buller, 1996; Stiff, 1996) and strives to address many processes that are irrelevant to acts of

communicating false information, partly because of its emphasis of deception as a broad dynamic interaction. Furthermore, McCornack asserted that both IMT and IDT both lacked even a minimally sufficient explanation of *how* people create strategic messages.

Other frameworks and propositions. This last section addresses predictions about deception that do not arise out of formal theories. One classic proposition that underlies much of the research to date is Ekman and Friesen's (1969) proposal that deceivers always need to make concerted attempts to prevent signs of leakage that reveal one's true response and indicate that deception is occurring (e.g., concealing feelings of guilt, anxiety, and shame caused by lying). The authors specifically predicted that leakage would be observable primarily through channels that are monitored and controlled least. Thus, leg movements should reveal anxiety more than hand movements, and hand movements should reveal more than the facial signals. Although no specific behavior is directly linked to deception (Ekman, 2001), studies have supported generalized leakage effects, even when people are highly motivated to control their behaviors (DePaulo, Lanier, & Davis, 1983).

Hopper and Bell (1984, p. 289) state: "In actuality, deception often involves enacting a role rather than making a statement." This comment suggests that people conceptualize roles to adopt and then attempt to act out those roles. How roles are conceptualized in deceptive situations is not clear, but the process probably depends on one's knowledge of the situation and of possible roles that would lead to goal attainment. One implication of this proposition is that people can create false messages by attempting to fulfill a certain role, as actors might do. Kroger and Turnbull (1975) conducted one of the few tests of this belief by asking college students to produce responses to the MMPI

by creating profiles of responses similar to that of individuals who occupy certain roles. They found that students were able to create realistic profiles when knowledgeable about the role to be mimicked.

Cohen, Lambert, Preston, Berry, Stewart, and Thomas (2005) conducted an exceptionally broad sweep of the literature on deception, including the topics of warfare, cognition, street cons, magic tricks and illusions, influence tactics, and computer security. They supported Ekman and Friesen's (1969) distinction between acts that present a target with false information ("simulations") and acts that withhold or conceal information ("concealments"), and proposed that the latter are useful in almost all instances of deception. Concealment by individuals might entail performing actions in private or distracting the target while deceptive actions are performed. At the level of groups, deception requires the same tactics, but also internal security, where all members of the group work to keep their plans secret. In Cohen et al.'s framework of deception, a goal leads the source to engage in complex cognitive processes that search and test the "space of possible deception methods" by simulating how the target would react to each method under different circumstances, using a "state of mind" (see "theory of mind" below). As the focus of their theory involved deception in computer systems, specific processing algorithms were suggested, assuming that deceptive strategies could be executed perfectly.

In conclusion, theoretical models of deception-related phenomena have grown in complexity in recent years, but they tend to provide detailed explanations of just one or two aspects of the overall process. Unexpectedly, the least amount of attention has been paid to the cognitive and behavioral processes mediating the effects of motivation on

message outcomes (i.e., lies), the processes that actually *are* deception. As a result, researchers have sometimes made rather large and unjustified assumptions about what happens when a deceptive message is produced.

A Review of Deception Strategies and Tactics in the Literature

In contrast to broad process-oriented theories of deception, various writers in the scientific domain and beyond have described different kinds of specific actions, termed "tactics" or techniques, that cause targets to be deceived, or that enhance other deceptive actions. Although they can reveal some of the assumptions underlying beliefs about deception, tactic research generally tends to yield unwieldy lists of heterogeneous sets of narrowly defined activities. Heuer (2002, p. 33-34) commented, "...there is little worth saying about deception in general. Virtually everything one might conclude about deception – whether it is pervasive or uncommon, a major threat or not, how best to deal with it, even whether deception requires secrecy to be successful, all that and more – depends upon what kind of deception one is talking about."

Despite this, there are some very apparent similarities between the themes found in work on tactical research across different disciplines. In a classic literary text, Sun Tzu (1991) claimed that "All warfare is deception" and provided military advice, some of which involved the alteration of one's appearance or of the environment's appearance (e.g., placing underbrush over a pit) and use of spies to feed opponents false information. Likewise, modern analysts of warfare tactics (e.g., Dunnigan & Nofi, 1995; Godson & Wirtz, 2002; Handel, 1989) have asserted that deception often depends on keeping true intentions and deceptive plans concealed. To provide an example, Dewar's (1989) warfare "techniques" are presented in Table 1. Research on nonhuman primates has also

Table 1

Dewar's Techniques of Deceptive Warfare.

Techniques	Description
Distraction	Diverting the enemy's attention towards a false plan, away from a true plan/intention
Lure	Provide the enemy with a seemingly beneficial opportunity as a trap
Repetitive Process	Perform something repeatedly until it appears true and natural
Double Bluff	Revealing true information to the enemy, knowing the enemy will expect to receive deceptive, false information
Unintentional Mistake	Allow the enemy to interpret an intentional act as a "mistake" to be exploited
Piece of Bad Luck	Allow the enemy to interpret an intentional act as a lucky opportunity (or unlucky opportunity for its opponent)
Substitution	Encourage the enemy to regard information as false, after it has become true
Disguise	Dressing up in the enemies uniforms or dressing in camouflage to move unnoticed in an environment (e.g., forest or desert)

produced lists of deceptive behaviors that rely on not revealing one's true intentions (to the extent that intentions can be inferred from nonhumans), which is typically accomplished by keeping the target distracted while the deception is performed (Savage-Rumbaugh & McDonald, 1988; Whiten & Byrne, 1988a, 1988b).

Though focused on the acquisition and maintenance of power rather than on pure "impression management," theories (e.g., Jones & Pittman, 1982) and studies of impression management (IM) tactics (e.g., Ellis, West, Ryan, & DeShon, 2002; Fletcher 1990; McFarland et al., 2003) have also found the concealment of one's true motives to be a necessary or beneficial component. In an early theory of self-presentation behaviors, Jones and Pittman (1982) described two concepts that seem to overlap with basic definitions of lying, whenever a person's true state is discrepant from the image presented. *Self-promotion* is seeking to increase others' perceptions of one's competence (rather than likeability and attractiveness) generally or with respect to a specific task, and *exemplification* is seeking to project integrity and moral worthiness.

Jones and Pittman (1982) argued that their theory applies more to "selective disclosure and omissions, matters of emphasis and toning rather than of deceit and simulation" (p. 233). IM tactics (offensive and defensive) do seem to refer to both honest and dishonest behaviors, as does the concept of impression management (as explained above). However, Bratton and Kacmar's (2004) concept of "extreme careerism" (the propensity to gain career related rewards through nonperformance based activity) shows greater overlap with deception. Extreme careerists are believed to use "negative IM" tactics, actions aimed at improving one's image at the expense of another. These tactics include blaming others for one's own mistakes, taking credit for another's work,

discrediting the personal qualities of a competitor, intimidating others, and negative projection (i.e., exaggerating obstacles to avoid responsibility for failures and self-handicapping). Except for intimidation, these negative tactics closely resemble actions that are traditionally regarded as forms of deception, though the tactics are context specific.

Buller and Burgoon (1994) synthesized the extant research on deceptive communications and identified a set of "relational message categories" based on their consequences (Table 2). Strategic behaviors are purposive and tend to be short, not disclosing, less continuous, and aimed at suppressing leakage. Nonstrategic messages are leaked and more pronounced when deceivers are facing suspicion and asked probing questions. They include blinking, higher voice pitch, hesitation and speech errors, less smiling, less nodding, word repetitions, irregular tempo, irregular behaviors, and a lack of spontaneity. Both behaviors were thought to be a function of the communication channel (i.e., verbal cues or body movements), motivation, planning and practice, personality, and communication skill (i.e., public self-consciousness). Empirical research (including a meta-analysis by Zuckerman & Driver, 1985) has consistently supported all of the categories, but provided mixed support for aspects of categories 2, 3, and 4.

With regard to deception on tests, Kluger and Colella (1993) suggested that fakers adopt one of three tactics when faced with nontransparent test items using a Likert scale. Nontransparent items do not reveal which answer choices are "best" simply based on their content or structure, unlike regular Likert items for which the best answers are usually one of the extreme values (e.g., 1 or 5, on a 5-point scale). For nontrasparent items, fakers might select an intermediate response option that is safe, neither good nor

Table 2

Buller and Burgoon's (1994) Relational Message Categories.

Message Category	Description
Strategic	
1. Uncertainty and vagueness	Predominantly linguistic behaviors used to create ambiguous, opaque, or intentionally mixed messages
2. Nonimmediacy, reticence, and withdrawal	Verbal and nonverbal means used to distance oneself from others, to disaffiliate, and to close off scrutiny or probing communication
3. Disassociation	Predominantly linguistic behaviors used to distance oneself from responsibility for one's own statements and actions and to imply dependence on the actions of others. This corresponds in some respects to an external locus of control.
4. Image- and relationship-protecting behavior	Verbal and nonverbal behaviors used to make oneself appear sincere and trustworthy and to sustain the self-presentation one has created.
Nonstrategic	
5. Revealing (arousal and nervousness)	Exhibition of nonverbal cues that betray one's heightened state of physiological arousal. The arousal associated with "duping delight" – glee at successfully misleading others – also fits here.
6.Revealing negative affect	Exhibition of verbal and nonverbal behaviors that leak unpleasant feelings possibly associated with guilt and embarrassment
7. Incompetent communication performance	Verbal and nonverbal behavior patterns such as exteme and nonnormative behavior, awkward conversation, and discrepancies between channels that yield an awkward conversation, substandard communicative performance. These decrements in communicative capability may be the results of the cognitive complexity of the deception task or of excessive motivation to succeed, leading to a forced, stilted, halting, or incongruous presentation

Note. These descriptions are taken directly from Buller & Burgoon (1994, p. 204).

bad. Another tactic is to guess an extreme value, just as if the item were transparent. This strategy often provides a 50-50 chance of getting a maximal item score. The third tactic is to abandon the deceptive approach and respond honestly. For transparent items, they propose that respondents answer in a socially desirable manner, answer in a job desirable manner (endorsing qualities valued by an employer), commit "outright fraud" by copying answers from empirical keys, or distorting their responses unconsciously (i.e., self-deceptions).

In contrast to typical strategies used by experimental subjects instructed to "fake good," some alternative strategies proposed included answering randomly or carelessly and answering according to one's perception of a social role. The latter strategy was believed to be more useful when people were motivated to evade detection, via lie scales. Viewing faking behaviors more broadly than as responses to individual items, Kroger and Turnbull (1975) proposed fakers generate realistic profiles of responses by selecting and "enacting a specific social role." However, it is not clear whether this type of response process is naturally occurring or only found when induced or suggested experimentally.

A Taxonomy of Deception Strategies

Based on the literature reviewed, it is clear that there remains a critical gap in general theories of deception regarding the process by which a deceptive message is actually constructed. Aimed at the performance of specific behaviors, the work on deception-related tactics seems to offer one way of bridging theories of motivational factors with the vast amount of empirical research on efforts to detect deception based on message outcomes (detailed further in subsequent sections). As yet, these literatures have not been integrated to form a unified and coherent theory. However, there are at least a

couple major issues preventing the direct importation of existing taxonomies of tactics into deception theories.

For one thing, behavioral tactics tend to be very context-specific, providing little insight about the underlying, core constructs that enable persuasive or deceptive communications to be produced. Another dilemma is that tactics (e.g., those by Buller & Burgoon, 1994) tend to be defined in terms of their consequences. This confounds the meaning of identified concepts because the exact same behavior may be defined as a tactic if it leads to success, but ignored if it does not. These factors have resulted in a variety of behaviors being labeled as "tactics." As discussed, some tactics in the literature involve the use of deception, while others involve honesty (e.g., self-promotion) and/or persuasion (e.g., ingratiation). Still other behaviors serve supportive functions (e.g., distracting a target so deception can be performed), and others represent accidental "leakage" (e.g., nonstrategic behaviors in Table 2), rather than behaviors directed at an intentional higher-order goal. Thus, tactics seem to cover behaviors that result from different goals and psychological mechanisms.

To circumvent such issues, I believe that an approach centered on potential deception strategies would better explain deception processes with general psychological concepts of motivation and behavior. Unlike behavioral tactics, strategies refer to broad processes driven by higher-order goals, including the planning and execution of a deceptive act, or even its adaptation in dynamic interactions. Also, strategies reside at a more general level of abstraction and are flexible enough to accommodate new actions that might be observed in the future. Because no prior work on deception strategies

exists, their existence is an empirical issue. This leads to the first research question (RQ) posed in this study:

RQ1: What kinds of cognitive strategies do people use to produce a faked message?

Although this question is investigated here, I used an approach that was primarily deductive, or perhaps better labeled "abductive" (see Haig, 2005), by using logic and insights from existing research to create six types of theoretical deception strategies. These strategies are fairly broad and are intended to subsume the total set of actions performed by people in both deceptive social interactions and written communications. The first four strategies can typically be described as offensive, using the terminology established above. Though the strategies are described in general terms below, specific tactics/techniques/behaviors are provided as examples of the strategies (in Table 3), across a range of situations.

In discussing dissimulation behaviors by experimental participants, Rosenzweig (1933, p. 348) notes that "The inference is that if the subject knows what it is that may be revealed by his behavior, his need for inviolacy is able to operate directly in behalf of his self-respect. Forewarned if forearmed." Consistent with this notion, the offensive strategies fall within a loose hierarchy based on how much information the deceiver possesses about the message that should be produced and the context in which it will be communicated. Generally, having greater, higher-quality knowledge about a situation allows for the use of more effective and efficient strategies, as will become evident in the strategy descriptions below.

The last two strategies can be viewed as defensive because they do not involve sending fabricated information to the target. The offensive and defensive strategies, however, are not opposing and may be used in combination. Also, all strategies may be combined with truthful communications, unintentionally or intentionally, to make a deceptive message appear more plausible or deemphasized (Buller & Burgoon, 1996; Shulsky, 2002). Unlike previous work, the use of honest communications is kept distinct from the deception strategies.

Finally, all strategies, whether pertaining to verbal or nonverbal communication, can be described as operating according to the fundamental principles of addition and subtraction. That is, any truthful message can be augmented through pure fabrication, exaggeration, or connections to other things that are irrelevant or the opposite (Galsínski, 2000). Any truth can also be reduced or omitted in scope or magnitude. In both cases, the statement will become less true and thus, more deceptive. These rules have been noted in past work implicitly and explicitly, but not always with respect to the *kinds* of information, behaviors, signals, etc. that are being manipulated.

Offensive Strategies

Reproduction. The first offensive strategy, reproduction, is the most straightforward and conceptually simplistic. As such, it may also be a strategy that is easily prevented. A reproduction strategy is planning to create a desired message response based on some existing model (i.e., person, object, one's past behaviors and deceptions, etc.) as a substitute for a true response. A clear example of how this strategy is implemented comes from students' use of cheat sheets and answer keys obtained beforehand when cheating, or the copying a classmate's answers. This strategy also includes the copying of a

Table 3

Examples of Strategies in Multiple Contexts.

	Example Context			
Strategy	General Communications	Job Interviews	Personality Tests (multiple choice, empirical)	Cognitive Ability Tests (multiple choice, objective)
Reproduction	Overhearing some idea and telling other people the exact same idea, as if it were one's own idea	In a group interview, repeating what another person says	Obtaining and using an official answer sheet	Obtaining and using an official answer sheet
	Calling in sick for work and using the same excuse as another person, in the same situation	Obtaining and using another person's answers to the same type of interview	Obtaining another person's answers before or during the test, and reproducing them	Obtaining another person's answers before or during the test, and reproducing them
	Impersonation			
	Forging a name on a document, based on a real signature			
	Plagiarism – reproducing another's paper			
Generalization	Calling in sick to work and using a ready-made excuse matched to the situation	Learning a list of structured interview questions beforehand, and formulating a faked answer to use in the interview	Using advice from a coaching program or test-taking guide and applying it to the test	Using advice from a coaching program or test-taking guide and applying it to the test

Table 3 (*cont.*)

Generalization (*cont.*)	Confirming false information; denying truthful information	Using general advice from a help resource as a guide to create a faked answer	Learning about the "important" dimensions in a personality test and answering only exaggerating values for those items	Secretly using notes or a textbook to help calculate an answer during a "closed book" exam, where the notes do not reveal the correct answer
	Displaying an emotional expression based on social norms and cues		Learning that rating extreme answers (e.g., "Agree Strongly") is generally beneficial	
	Taking credit for another person's work			
Extrapolation /Interpolation	Using empathy and the target's reaction to select and display a faked emotional response	Imagining what a qualified candidate would say do and say and acting in accordance with those beliefs	Imagining a person of a particular personality type and guessing how the person would respond	Imagining a person of a particular personality type and guessing how the person would respond
	Making up an excuse for missing a party, based on whether the friend seems angry or not	Using interviewer's comments to determine what questions the interviewer cares about and elaborating false answers to those questions	Using the questions to guess what dimensions of personality are being measured, and then reasoning the best answer choice	Using information from surrounding questions and answers to create an answer to a new test question

Table 3 (cont.)

Extrapolation /Interpolation (*cont.*)	Using past experiences to imagine new events that could have resulted from them	Feigning laughter whenever the interviewer tells a joke	Inferring that raising (or lowering) every response by a certain degree is beneficial after starting the test	
	Being accused and denying having performed an action, despite not knowing what the specific accusation is			
	Asking a classmate about his/her paper, then using the ideas to write a new one			
Conditional Randomization	Giving purely fabricated information made up on the spot (e.g., a story, making up a dance)	Telling a false story without knowing whether it is an appropriate response	Guessing randomly when not allowed to guess	Guessing randomly when not allowed to guess
	Calling in sick to work and giving a random reason without any regard to whether it fits the situation or is plausible		Random guessing after one's "true" instinctual response is ignored, suppressed, or removed from consideration	Identifying answer choices that are clearly wrong and then guessing an answer that is between (or beyond) the range of the incorrect choices
	Saying bullshit			
	Forging a name, by altering one's natural handwriting patterns			

signature during forgery or the copying of text during plagiarism. In social interactions, a person may recreate the idea or movement of another person and claim it as one's own.

Generally, this strategy requires little creativity and cognitive effort, relative to other strategies. Functionally, the most important requirement of this strategy is having access to a model that accurately represents the desirable "true" response. (At times when the model cannot be accessed during the deceptive attempt, a good memory may also be needed.) One final point is that reproduction is not limited to situations in which a person has or acquires full information about the model, or in which the model is perfect.

Generalization. The second strategy is the application of *preexisting* knowledge about specific deception tactics and desirable responses to fabricate a deceptive message in a given situation. In contrast to reproduction strategies for which the deceiver attempts to reproduce a model directly, knowledge strategies require the deceiver to make some kind of generalization about whether previously used strategies or responses can be applied appropriately to the current situation. This strategy, however, does not require the deceiver to generate new tactics (as do later strategies); the knowledge perceived to be relevant by the deceiver must have been acquired prior to the deceptive attempt and then used to determine subsequent actions.

Perhaps being too obvious to warrant mention, one general tactic that can be used in virtually any situation warranting the production of a dichotomous lie is to confirm verbally information known to be false (e.g., saying "Yes, that's true" when asked about information) or to deny the truth with a verbal negation. Another general tactic is to exaggerate any truthful piece of information by increasing the degree or frequency of some referent, such as hours worked, one's IQ score, and how often one cleans the

kitchen. More specific tactics are relevant only to certain situations and require cognitive decision-making processes. For example, a person might lie about missing a party by using a lie commonly used to explain absence from work (e.g., "I couldn't make it to the party because I was feeling ill"). In other daily communications, people can attempt to generalize behavioral signals like emotional expressions.

Generalization strategies are common in academic and work selection contexts because some tactics are taught explicitly. Test coaching (Sackett, Burris, & Ryan, 1989), for instance, can teach people specific behaviors to perform when taking particular tests, though it should be recognized that not all forms of coaching encourage the use of deception. In such cases, test takers do not know exactly what answer needs to be reproduced, but they do have a base of knowledge about techniques they can use to produce a specific kind of faked answer, based on contextual cues detected. As an example, Cullen and Sackett (2003) coached students to always select non-extreme answer choices on a situational judgment test, a strategy favored by one popular scoring format for such tests.

More informally, job applicants can find "tips" from books and websites (e.g., www.cabincrewjobs.com) about how they should act and what information they need to convey during an employment interview. (If, however, a person was to memorize an actual list of phrases to give to predetermined questions, this would be an example of the previous strategy.) A periodical (Furchgott, June 1998) provided tips on "How to Lie on Your Resume" based on published research and information provided by pre-employment checking organizations (i.e., the knowledge was based on actual deception research findings). Tips included applying for employment at organizations that do not

conduct background checks or use professional services, claiming previous employment with a company that went out of business (or that had the employer die), explaining away employment gaps with a need to take "time out for family reasons," and pretending to be one's own previous employer by using a post office box or having a friend act as a previous employer. Sackett et al. (1989) also discussed the possibility that people obtain knowledge relevant to deception through "grapevine coaching," the dissemination of information through subsequent individuals within some social network.

Extrapolation / Interpolation (E/I). Moving beyond readily available models and tactics, deceivers in unfamiliar situations will need to assess their situation and extrapolate, or infer, the desired, deceptive response that should be produced in order to obtain the reward. In daily conversation, the deceiver may be asked a question for which the desirable answer is not apparent, precluding the use of previous strategies. In this case, the deceiver may use other cues, such as previous questions or the target's facial expressions, to decide what the target would perceive to be a favorable response. This situation is analogous to one in which a person attempts to fake test answers without knowing anything about the test beforehand and without having been coached. Once the deceiver reasons what response should be produced, he or she must use interpolation to reduce the perceived discrepancy between the truth and desired states, where interpolation is the process of inserting values in a gap to connect existing values. Thus, this strategy suggests that people do not merely alter some preexisting information (per Grice's maxims). Instead, they create new information to fill in a gap (Ekman & Friesen, 1969), as when creating an original excuse after being accused of an action.

Some clear examples of E/I strategies come from the acting domain, particularly when actors must portray historical figures. In some cases, the actor can simply watch an old video clip of a figure and reproduce each movement and statement; this would represent the first strategy. In other cases, the actor may acquire knowledge about the figure's behavioral tendencies (e.g., the figure usually told a joke when feeling nervous); this would represent the second strategy. Often, however, actors must use known information about the figure to infer the basic thought processes and the values guiding the figure's actions, and then interpolate how that person would have reacted in other situations. (Of course, less devoted actors may not bother to make such "educated guesses.") Similarly, it has been suggested that people fake personality tests by imagining or reasoning how others would respond to them (Kroger & Turnbull, 1975; Levin & Zickar, 2002).

E/I strategies appear to be particularly useful when deceivers have little prior knowledge about their situation (cf. Kroger & Turnbull, 1975). Although they may still involve basic reasoning and decision making processes, it seems that deceivers might be more successful at using E/I strategies when they have adaptive and creative skills enabling them to produce interpolated information that is plausible and relevant (Handel, 1989). More general skills related to general self-awareness and empathy toward others may further help a deceiver make necessary inferences.

Conditional randomization. The fourth strategy of conditionally random responding is used when the deceiver aims to distort the communication of a true response, but has virtually no idea about what the faked response should look like, and does not infer specific behavioral goals based on contextual information (cf. McCornack,

1997). Regardless of whether contextual information is unavailable or ignored, this strategy produces a random behavior after a decision is made to misrepresent the truth.

One example of this may occur when people attempt to forge signatures, without knowing what the true signature looks like. In this case, the forger could write the forged name in his true handwriting style, just as he would write any other word. Yet, if the forger believes that his true handwriting style would lead to detection, he can use conditional randomization strategy and attempt to alter his writing style (e.g., using his non-dominant hand) in a random manner that has no relationship to the desired response (i.e., the victim's true signature). Similarly, a person may fake a psychological multiple-choice test simply by eliminating the true response from consideration and then guessing one of the alternatives.

Concerning deceit in general communications, a deceiver might try to impress her romantic date by claiming that she is a good salsa dancer, though she knows nothing about that style of dance. If the date asks her to demonstrate her dancing skills, she may recognize that whatever dance moves she is naturally inclined to perform will probably not appear to be salsa and, therefore, attempt to make random movements. Though this example is admittedly a bit contrived, it describes the manner in which this strategy operates.

At a glance, this strategy seems to have limited effectiveness (i.e., produce responses that reduce image discrepancy) in most situations. Because the deceiver lacks relevant contextual knowledge, messages produced with this strategy may be nonsensical or implausible, and actually increase the chance of detection. Yet, conditional randomization may actually be very effective when a person is also concerned about

concealing true responses (Kroger & Turnbull, 1975), since the message sent to the deceiver has no systematic relation to the truth (e.g., lies are not the exact opposite of the truth). Attempts to detect this form of deception may then be quite difficult, since observed responses are random and have no meaning. Thus, this strategy may be of greater value when the deceiver is focused on defensive strategies (i.e., concealing true representations that would surely prevent the attainment of a reward) and cares little about the specific message communicated.

For any of the four offensive strategies described, their immediate outcomes will be some kind of message transmitted to the target, verbally or nonverbally. In some contexts, messages may serve to make something or someone appear more positively. In others, messages may serve to create a negative impression, still reducing image discrepancy (i.e., the desired state is more negative than the current state). For instance, deceivers using the IM tactic of intimidation work to create false negative impressions that seem threatening, as a means of manipulating their targets. In other contexts, cross-cultural differences suggest that people in collectivist cultures may act to appear socially undesirable to increase the relative status of others (Kim, 1985; Xin, 2004). Thus, the specific messages produced by strategies may take on different forms, but the underlying psychological processes are the same.

Defensive Strategies

Selective presentation. The use of selective presentation requires deceivers to make cognitive evaluations about whether the communication of different pieces of truthful information will facilitate reward attainment, by way of perceived discrepancy reduction. Although selective presentation might involve the concealment of true

responses, I use the term to connote the source's unwillingness to *volunteer* information necessary for the target to develop an accurate representation of the object of communication. So, messages produced by this strategy can be complete omissions, vague or ambiguous appearances, or misleading portions of truth (Buller et al., 1996).

This strategy cannot be used when people are asked directly about the specific information being withheld, whether during a lie detector or a multiple-choice test (not allowing blank answers). People who are asked if they know or did _____ and then falsely answer "no" would be creating an offensive lie. Indirect, open-ended questions, on the other hand, allow this strategy to operate. For instance, people asked, "What did you do this weekend?" have the discretion to present whatever combination of truths they deem to be important for the target to understand the true answer. If a person spent 99% of the time eating, sleeping, and watching television and only 1% exercising, the statement "I exercised" would technically be true, but it would also be a gross misrepresentation of the truth, even by subjective standards.

Suppression. Unlike selective presentation, suppression involves inhibiting the otherwise automatic expression of natural thoughts, behaviors, and emotions (Baumeister, Bratslavsky, Muraven, & Tice, 1998; Crandall & Eshleman, 2003; Muraven & Baumeister, 2000; Vohs, Baumeister, & Ciarocco, 2005; Wegner, Erbe, & Zanakos, 1993). This is thought to be accomplished either when the deceiver seeks out distracter thoughts or when thoughts/behaviors are monitored directly and censored (Handel, 1989; Wegner et al., 1993). In practical matters, it seems that suppression is typically used to prevent nonverbal meta-messages from being leaked, so the target does not realize deception is occurring. For example, suppressing perspiration or twitching associated

with nervousness may do little to reduce any perceived discrepancies related to reward attainment, but can prevent the target from recognizing an offensive deceptive attempt. Suppression can also refer to the concealment of objects (e.g., covering up physical evidence with a blanket, or telling screaming children to go to their room before guests arrive to communicate happiness in the household).

However, suppression can also refer to the content of verbal communications. People may remove or prevent pieces of a message from being sent to the target. The key distinction between suppression and selective presentation in such cases is admittedly subtle, but suppression only applies to information that would be *automatically* transmitted or available to the target, as when radio or television program signals with obscenities are transmitted, but then censored with silence.

For both offensive and defensive strategies, the current theoretical framework does not, at this time, provide any expectations about their prevalence. Although I expect each strategy to exist, I offer no suppositions about the strategies' frequencies, in naturalistic settings or in specific contexts. This leads to the second research question of this study:

RQ2: How prevalent are different deception strategies in a given setting?

This initial investigation of the existence and prevalence of different strategies was limited by certain features of the experimental design (i.e., deception on paper-and-pencil psychological tests), but does provide the first step toward establishing general patterns of deceptive communication.

If certain strategies are more prevalent than others, it would seem wise to ask whether the use of various strategies is related to differences in the goals that deceivers

have or focus on, given the motivational theories reviewed above. Thus, a third research question was formulated:

RQ3: What is the relationship between strategy use and deceivers goals?

Also, it was suggested that certain strategies or their specific tactics might be more effective than others for different purposes (e.g., perceived image discrepancy and not being detected). However, it is not clear whether certain strategies will be more effective (in helping deceivers accomplish their respective goals) for the reasons proposed, or whether they will differ in effectiveness at all.

Conceptual Layers of Deception Outcomes

Although they differ, both offensive and defensive strategies can be understood as components of a unified process of deception from a construct-oriented perspective. Traditional models of psychological measurement assume that observable behaviors (or psychological test scores) reflect surface-level manifestations of true latent constructs, theoretical concepts of thoughts and feelings that cannot be measured directly (Bollen, 2002; Edwards & Bagozzi, 2000). In the case of honesty, people produce messages that indicate their true standing on a latent construct after being influenced by a set of biases and errors (Murphy & DeShon, 2000). Rosenzweig (1933; 1934), for example, states that research subjects will have a natural, true reaction to an experimental manipulation, but may provide self-reports of their internal states that are distorted to reflect more positive self-conceptions (i.e., "What a man thinks he would like to be"; 1934, p. 401).

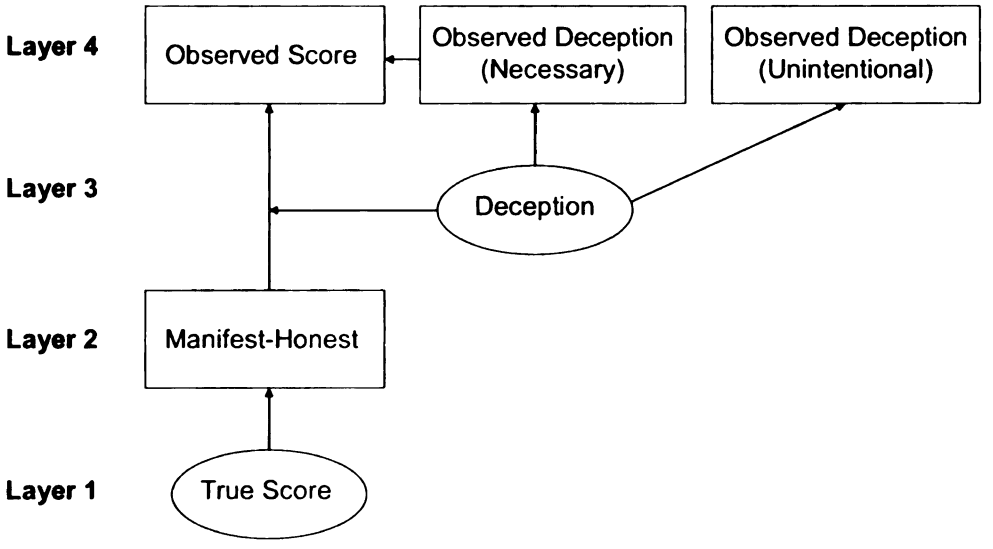
Figure 3 depicts a basic model of latent constructs (in ovals) and observed behaviors (in boxes) as they operate during the misrepresentation of some personal attribute. The immediate feature to notice is that the variables are placed at different

conceptual levels, or what I term "layers," representing the order in which they occur. As stated above, Layer 1 represents a person's latent attribute and Layer 2 represents an observable manifestation of (or measured score for) that attribute, which is subject to biases and errors. At Layer 3, the performance of deception occurs by means of latent psychological constructs. What these constructs are has not been specified in the literature, but the proposed strategic theory of deception asserts that cognitive decision making processes lead to the choice and execution of a set of behaviors or tactics to produce a deceptive message.

Layer 4 shows three types of observable behaviors that might result from the simplest case of deception. First, "observed score" represents the deceptive message intended to be sent to the target, often crafted to have appearance of truth. In the ideal case, the source would be able to fashion a deceptive message that seems to be true in every way (i.e., perfectly mimicking an honest person's manifest-honest score) without producing any indication that deception was performed. Thus, observed scores would be the only variable at Layer 4. Most theories, however, posit that people always produce some kind of observable behaviors (i.e., leakage) indicating that deception was performed, and that these byproducts can potentially be detected by targets.

Interestingly, since Layer 3 tends to be treated as a "black box" with unknown processes occurring, leakage is almost always assumed to be some kind of behavior that covaries with the performance of deception, without directly reflecting the process of deception. For instance, a shifting gaze and perspiration were two commonly studied leakage behaviors in early detection research (described in detail in the next section). Yet, no one ever claimed that these behaviors were necessary for creating a deceptive

Figure 3. Constructs Involved in the Process of Deceptive Responding By Level.



message.

By contrast, the focus on constructs in this paper reveals the importance of specifying what happens conceptually at Layer 3. It is plausible that cognitions and behaviors that enable the process of deception will produce observable signs (i.e., behaviors). Unlike accidental leakage, however, these behaviors are *necessary* components of deception and performed intentionally, even if the target of communication may be aware of them. Thus, any behaviors necessary for performing deception (ranging from physical movements to speech to neurological signals) are labeled "observed deception (necessary)" in Figure 3. Unintentional byproducts or covariates (i.e., 3rd variables) of deception like perspiration, which might even be controllable by skilled deceivers, are labeled "observed deception (unintentional)."

By using this model to organize the extant literature on attempts to detect deception, it becomes apparent that the many techniques have been focused on different layers, with no method accounting for factors across the entire process of deception. Early attempts at detecting deception (e.g., Hartshorne & May, 1928; Liberty, Lunneborg, & Atkinson, 1964) varied in their level of sophistication, but typically focused on identifying people who claimed to do statistically deviant or impossible things. As such, they were aimed at comparisons of deceivers' Layer 4 observed scores to honest people's Layer 2 scores. Layer 3 was assumed to exist, without explanation. Other attempts to measure and detect deception in communications and social psychology have focused on identifying unintentional leakage behaviors, physical and verbal, through comparisons with Layer 2 observed scores.

More traditional lie detection studies of polygraphs and interrogation techniques are somewhat ambiguous in meaning. Originally, physiological measures of heart rate, perspiration, and other behaviors were supposed signs of observed deception (unintentional), assumed to result directly from Layer 3. Over the last two decades, however, Lykken (e.g., Lykken, 1998) and others have argued convincingly that such behaviors are better conceptualized as signs of guilt, which is a likely outcome of deception, but not a part of deception itself.

Alternatively, a few approaches to detect lies using response latencies (e.g., Walczyk et al., 2005) could be viewed in different ways, since the authors of studies have not always been explicit about their justifications. If deceivers take longer to respond because they are nervous or unsure about how to respond to a target, delays would represent observed deception (unintentional). If, on the other hand, deceivers responded slower because they were engaging in cognitive processes (perhaps strategic ones) that produced the deceptive messages, then delays would indicate observed deception (necessary).

Few studies to date have described Layer 3 conceptually or attempted to measure it (observed deception – necessary). One approach has been to use multiple-choice "bogus" items that ask people about their experiences with fictional events (Stokes & Toth, 1996). Since the true answer is always "none" or "never," people *must* select a higher answer if they intend to lie at all. In this case, the behavior necessary to perform deception is also the one that leads to its immediate detection. By comparison, "unlikely virtues" scales detect people who want to lie, but who unintentionally produce signs of this goal (i.e., selecting highly unlikely and unrealistic events); selecting unlikely virtues

is not necessary to perform the deception in most situations. In more modern approaches, researchers have begun to use item-response theory to model latent scores at Layer 3 (e.g., Levin & Zickar, 2002; Stark, Chernyshenko, Chan, Lee, & Drasgow, 2001; Zickar, Gibby, & Robie, 2004; Zickar & Robie, 1999). Other work using neurological imagining techniques attempt to measure very proximal behaviors (in the brain) at Layer 3 that directly result from deception, though these are still observed deception (necessary).

Performance of Deception: Deception Capability and Strategy Execution

Strategies enable deceivers to determine which specific set of actions need to be performed to produce a message that will eventually lead to reward attainment. The broader process of deception then involves some degree of decision making (e.g., which strategy to use in light of personal and situational factors) and performance. Because both of these processes tend to exhibit variance between persons, including specifically when faking tests (McFarland & Ryan, 2000), it is sensible to explore the possibility that individuals differ in their capability to deceive. This view is also encouraged by research on possible prerequisites in cognitive processing that enable deception to be performed by nonhuman primates (e.g., Premack, 1988; Whiten & Byrne, 1988a, 1988b) and young children (e.g., Banerjee & Yuill, 1999; Broomfield, Robinson, & Robinson, 2002; LaFrenière, 1988; Pezdek et al., 2004; Vasek, 1986).

Snell et al. (1999) and McFarland (2000) have provided unique models, both pertaining to job applicant faking, that explicitly account for individual differences in faking ability, but researchers in other domains have often implied that certain skills are important (e.g., Buller & Burgoon, 1994; Dunnigan & Nofi, 1995; Handel, 1989; LaFrenière, 1988), including both innate abilities and acquired skills or knowledge. I list

just a few of the proposed aspects of deception "capability" (i.e., term not connoting exclusively innate characteristics) here: cognitive ability and general awareness (Kashy & DePaulo, 1996), knowledge of the context and target (Buller & Burgoon, 1994; Godson & Wirtz, 2002; Shulsky, 2002; Snell et al., 1999), physical, vocal, and emotional control (Buller & Burgoon, 1996), verbal fluency (Buller & Burgoon, 1994, 1996), and social adroitness (Kashy & DePaulo, 1996). It was also discussed earlier that personality variables correlate with deception-related concepts.

Overall, little research has attempted to measure this supposed deception capability, much less identify its determinants, beyond examining correlations between individuals' characteristics and whether their lies were detected. Among the few findings that exist, Bond and DePaulo (2005) produced meta-analytic evidence suggesting that lies tend to be detected less often when planned, though others have suggested that too much planning may lead to the creation of esoteric messages not resembling normal communication (Handel, 1989). Similarly, some have proposed that practicing strategy execution is beneficial because it increases behavioral automaticity (Dewar, 1989; Wilson, 1997) and decreases response times (Walczyk et al., 2003). To my knowledge, just two studies (Dawson, 1980 and DePaulo & Rosenthal, 1979) have examined differences in the global ability to act and pretend (i.e., Stanislavsky "method" acting and general "hamming behaviors, respectively), but they produced inconclusive findings and have not been replicated.

In this study, I examine relationships between the proposed deception strategies by deceivers and potential correlate of deception capability. As no validated measure of

capability exists, the following research question is proposed in reference to an experimental measure based on self-reported data:

RQ4: What are the relationships between the proposed deception strategies and a self-reported measure of deception capability?

Despite the propositions developed in this section, research has shown that many people are biased toward assuming that information given to them is true, in a phenomenon labeled "the truth bias" (Hilton, Fein, & Miller, 1993; McCornack & Levine, 1990; McCornack & Parks, 1986; Levine, Park, & McCornack, 1999; Taylor, Gittes, O'Neal, & Brown, 1994). The truth bias is beneficial in most social relationships because it establishes the basis for conversation (Galasínski, 2000); it may also support self-confirmation biases (Vrij, 2000). In any case, the existence of a truth bias directly implies that deceivers do not always need to bother creating and applying elaborate strategies with care, or make an effort to conceal their deceptive intentions (cf. Langer, Blank, & Chanowitz, 1978; Langer & Newman, 1979). (This is not to say that targets cannot reject messages on other grounds.) Therefore, deceivers may also demonstrate satisficing, performing just sufficiently well to achieve a goal, which may require only minimal skill.

Deception Effectiveness: Message Quality and Detection

As of yet, there really is no standard method for evaluating the quality of deceptive messages based on their features. When differences between lies are examined, a distal outcome typically provides the distinguishing criterion, and the criterion is most often whether a target detects the deception. Thus, the outcome of detection (or no detection) serves as a proxy for message quality, though it is influenced by other factors

as well (e.g., the truth bias). However, the ultimate outcome of *deception effectiveness* is determined by whether the target is influenced by the deceptive message in such a way and enough so as to perceive little discrepancy between the source's presentation and desired state.

Detection research can generally be classified according to three types of mechanisms: polygraphs and machines, unaided human judgment, and psychological tests. Apart from differences in the communication medium typically used within each domain (i.e., physiological, visual and oral, and written, respectively), research on each mechanism has provided a unique set of insights into how deception can be detected and/or measured (on a continuum). Overall, however, empirical findings have tended to be disappointing and have led many to conclude that deception is very difficult to detect accurately in laboratory and real-world settings by novices, experts, and lay people alike (Hyman, 1989; Miller & Stiff, 1993).

Criminal lie detection and polygraphs. Informal methods of detecting deception have existed since the advent of communication; "we are all human lie detectors" (Lykken, 1998, p. 23). Since the first published account of a lie detector devices (i.e., Benussi, 1914) a variety of equipment and corresponding interrogation procedures have been used to detect lies, with the most being related in some way to polygraphs that measure physiological (e.g, skin conductance, respiration, heart rate) response patterns. Researchers originally believed that these responses resided in Layer 2 (of Figure 3) and that they were expressed automatically whenever people engaged in deception. Early studies attempted to identify unique patterns in the responses of lying versus honest subjects, but it soon became apparent that both groups produced similar patterns,

especially when honest people were simply nervous (Lykken, 1998; Waid & Orne, 1981). Consequently, polygraphers began to interpret differences in response patterns within each person, comparing physiological data when the interrogated subject is asked to make true ("control") and deceptive statements (Lykken, 1998; Waid & Orne, 1981).

To circumvent conceptual ambiguities plaguing earlier methods, Lykken (1959) developed a popular procedure called The Guilty Knowledge Test (GKT). As the name implies, the GKT aims to measure direct signs of *guilt,* rather than more subtle signals associated with acts of lying. However, "guilt" was conceptualized as a cognitive awareness that occurs when people encounter evidence of their lies/crimes as opposed to neutral information (Lykken, 1960), not as an emotional response (Lykken, 1998). MacLaren's (2001) meta-analysis of laboratory studies using multiple-choice GKTs estimated the average sensitivity (i.e., a function of the percentage of people correctly classified as guilty or innocent) to be 76% for correctly detecting "innocent" people, as compared to the 20% rate expected by chance, and to be 83% for detecting "guilty" people who were asked to deny having encountered some piece of information beforehand. Another meta-analysis (Ben-Shakur & Elaad, 2003) estimated validity of the GKT to be .55 across 80 studies. This estimate rose to .65 when including only experiments in which participants performed mock crimes before providing "guilty" information. Together, these data suggest that the GKT has some validity.

However, the results of two initial field studies of the operational effectiveness of the GKT by Israeli police provided less optimistic results (Elaad 1990; Elaad, Ginton, & Jungman, 1992). Archival records of actual GKT results were compared to records of criminal confessions to estimate detection accuracy rates. While the percentage of people

judged to be innocent by the GKT was close to the percentage found to be innocent based on confessions (i.e., they did not confess), the detection of people who gave confessions was only correct for about half of sample. Thus, the GKT appears to be biased toward determining innocence.

This finding, however, could be attributed to the use of a fairly flawed criterion, confessions, as most people are probably not motivated to make them, particularly if their GKT results turned out to be negative. Other explanations for this finding offered by the researchers include the possibility that criminals do not attend to certain details of their crime, that details are forgotten after the crime, or that the facts of the crime used as the "guilty knowledge" were obtained from the criminals themselves at some time prior to their interrogation. In the end, the GKT has demonstrated some degree of validity as a scientific measure of "guilt," but it also produces more false negative errors than would be tolerated in many field settings.

In recent years, more sophisticated and technologically advanced approaches (e.g., using fMRI and PET equipment) have been used to measure neurological response patterns like event-related potentials (Farwell & Donchin, 1991; Rosenfeld & Bessinger, 1990; Rosenfeld, Cantwell, Nasman,Wojdac, Ivanov, & Mazzeiri, 1988), in addition to or in place of traditional physiological signals. Under the assumption that neurological signals are more proximal and automatic indicators of the psychological process of lying, these studies suggest that more refined methods of detection will improve accuracy rates. Additional studies are still needed to support this claim, but initial work seems to be promising (Seymour, Seifert, Shafto, & Mosmann, 2000).

Other studies have investigated simpler, cost-effective approaches, using computers to measure response time in honest and deceptive responders. In one such study, Kluger et al. (1991) found no relationship between response latencies for people who answered biodata questions honestly and for people pretending to be job applicants. In contrast, Walczyk et al. (2003) found that a sample of college undergraduates responded more slowly on average to a set of personal, factual questions when lying, regardless of whether the questions were dichotomous (i.e., Yes-No) or categorical. Their results also suggested that liars with good verbal skills will take longer to respond because they tend to construct lies that are more elaborate. This effect was also demonstrated when people only answered some questions dishonestly and others honestly, though the response times were generally slower under those circumstances.

In a study applying response time measures to the GKT, Seymour et al. (2000) found that the GKT was able to differentiate people on the basis of their response times, even when people were instructed to beat the test by consciously manipulating their response latency. Participants' responses to GKT questions when presented with guilty knowledge were generally around 600ms for innocent people and longer than 850ms for guilty people. (Response times for the guilty group were around 650ms when presented with irrelevant, control stimuli.) These findings also suggest that costly neurological measures may not be necessary to achieve acceptable rates of detection.

Nonetheless, the Employee Polygraph Protection Act of 1988, allowing testing in employment settings only under strict conditions, exemplifies the current scientific consensus on the efficacy of polygraph procedures (cf. Committee to Review Scientific Evidence on the Polygraph, 2003; Sackett et al., 1989). Practically, the largest cause for

concern in the U.S. is the occurrence of false positive errors (i.e., accusing innocent people on the basis of polygraph results). Early research supported this notion by showing that innocent people could produce the same patterns as guilty ones, but for unintended reasons. It has also been shown that "guilty" laboratory subjects can produce innocent-looking response patterns, as when they forget the guilty knowledge (Waid, Orne, & Orne, 1978). At least some research (Elaad, 1990; Elaad et al., 1992) suggests, on the other hand, that the GKT may have a greater tendency to produce false negatives. Either way, the accuracy rate of such tests leaves much room for improvement, especially when accounting for the fact that the chance probability of correctly accusing someone as being guilty is 50%.

Human lie detectors. Empirical evidence supporting humans' natural abilities as accurate lie detectors is mixed, but it generally suggests that accuracy rates are just above that expected by chance (Bond & DePaulo, 2005; Zuckerman et al., 1981). Although deception seems to require more arousal and effortful processing (Zuckerman et al., 1981), research over decades has failed to support the common belief that lies can be detected invariably through specific behavioral cues (e.g., facial signals or pausing in speech; Ekman, 2001; Inbau et al., 2004). Among the behavioral cues frequently studied, speech cues (for content and delivery) seem to be the most indicative of verbal lies (Bond & DePaulo, 2005; Zuckerman et al., 1981), but they still produce unimpressive detection rates.

In a renowned study, Ekman and O'Sullivan (1991) compared the lie detection prowess of different kinds of "professional lie catchers." As part of the study design, 5 videotaped interviewees told lies and 5 told truths, but all interviewees were selected for

use as stimuli in the study because they produced facial "microexpressions" (expressions lasting less than 1/25th of a second) that allowed them to be classified into their respective groups. In the study, participants belonging to various occupational groups watched 1-minute video clips of the 10 interviewees and rated the perceived truthfulness of each interviewee. Among, judges, psychiatrists, special interest group members, police, federal polygraphers, and a control group of undergraduates, only Secret Service agents (SSA) consistently detected lies at a level above chance (53% of the SSA correctly classified 70-100% of the clips). For SSA, age was negatively correlated with accuracy, as was confidence in one's performance on the task. Interestingly, the results also showed that no group explicitly reported using microexpressions to detect liars.

Despite the findings in this and similar studies using professional interrogators (providing both supportive and contradictory results, see Inbau et al., 2004), broad literature reviews have offered compelling data that people's capability to detect lies tends to barely exceeds chance levels (50%). Estimates of the average detection rate have fallen between 50-70% (e.g., DePaulo et al., 1981; Kraut, 1980; Mattson, Ryan, Allen, & Miller, 2000; Vrij, 2000). In the largest meta-analysis to date, Bond and DePaulo (2005) found the average rate of detection accuracy across 200 studies of human detection attempts to be $d = .40$, where d is the "mean difference between the percentage of truth judgments to truths and the percentage of truth judgments to lies divided by a percentage standard deviation" (p. 9). This result can be broken down as an estimate of detection rates in studies that require subjects to make dichotomous decisions (i.e., lying versus honest) and studies measuring accuracy as a d score. For the former subsample, sampling error accounted for much of the variability in detection rates and the mean sample-

weighted detection rate was 53.42% (C.I.: 53.28 to 53.57), equivalent to $d = .42$. For the latter subsample, the mean sample-weighted rate was $d = .37$ (C.I: .33 to .41), an estimate that is about half the size of ones found in prior reviews. Overall, the data suggest conclusions similar to ones made for polygraph research: humans can perform lie detection at levels somewhat better than chance, but make many classification errors.

In contrast, some of the most optimistic claims for the efficacy of lie detectors comes from analyses of linguistic message styles (including types and patterns of words) in conversation and writing. In their review of the literature, Pennebaker, Mehl, and Niederhoffer (2003) concluded that deceivers avoid ownership of their statements and use first person singular pronouns less often when telling stories. They also claimed that studies support the notion that deceivers display leakage resulting from guilt and anxiety in the form of more negative emotion words, whereas honest people make more qualifications in their statements, using the words "but, except, without, and exclude," and produce statements of greater cognitive complexity. At the same time, contradictory evidence also exists, suggesting that people who experience shame or guilt tend to use the "projective second person" more often in their writing (Tangney, 1992). So, the linguistic cues researched are not used exclusively by deceivers and can result from other psychological and situational factors. It may also be the case that *skilled* deceivers are able to avoid producing Layer 4 responses that have these features. Nonetheless, linguistic cues offer one interesting and potentially viable method for measuring deception.

Faking on psychological tests. With an early focus on diagnostic clinical measures and a later focus on job applicant responses in personnel selection, much

research has examined test faking (Barrick & Mount, 1996; Dunnette, McCartney, Carlson, & Kirchner, 1962; Ones et al., 1996). A typical study (e.g., Dannenbaum & Lanyon, 1993; Sackett et al., 1989; Viswesvaran & Ones, 1999) involves the administration of a psychological test to participants under varied instruction sets, where some people respond honestly and some "fake good" (strive to obtain as high a test score as possible). Alternate conditions may instruct people to respond as if one were applying for a job or to "fake bad" (strive to obtain low scores), as when pretending to have a clinical disorder. Many studies also include direct and indirect measures of faking itself, particularly when the focus is on preventing or correcting test scores for the effect of faking. Examples of direct measures include the use of "lie scales" assessing the endorsement of extremely rare events and bogus items, which are similar to lie scales, but refer to impossible events (Anderson, Warner, & Spencer, 1984; Panone, 1984). More indirect or "subtle" measures include socially desirable responding (SDR) scales that tend to assess general tendencies to select socially valued responses (see review by Paulhus, 1991).

One of the most popular issues of study and debate is whether people actually fake when given the opportunity. In many cases, faking on psychological tests is very similar to cheating on academic exams because both concern people who attempt to alter their appearance by artificially inflating (or deflating) their tests scores, though academic exams may have more objective questions with uniquely correct answers. Generally, survey research suggests that the incidence of cheating is quite high. Schab's (1991) large-scale surveys of high school students over a twenty-year span suggest that the majority of students have cheated on tests (68% in 1989) and plagiarized work (70%).

Additionally, responses to the item "To succeed in business requires some dishonesty" revealed that approximately 40% of the students have favorable attitudes toward the use of deception. A survey of high school Girl Scouts (Harris, 1990) showed that 65% claimed they would cheat on an "important exam." Similarly, 76% of the students in a sample of Who's Who Among American High School Students admitted to having cheated at one time, and 92% of the admitted cheaters reported having never been caught (Newberger, 2003).

Field and laboratory have shown that at least some tests are susceptible to faking (e.g., Alliger & Dwight, 2000; Cunningham, Wong, & Barbee, 1994; Doll, 1971; Ones et al., 1996; Ones & Viswesvaran, 1998), but the differences in scores between fakers and honest responders tend to vary considerably between studies (Kluger & Colella, 1993; McFarland & Ryan, 2000). In both types of studies, people are often instructed to engage in deception (e.g., "fake good") without any regard to negative consequences for being detected (Smith & Ellingson, 2002), suggesting that research estimates of the frequency or magnitude of faking are inflated. At the same time, rewards for faking in studies (e.g., McFarland and Ryan, 2000, gave $15 to the top 15% of high scoring test takers) tend to be much smaller than those offered in real life, such as job employment, suggesting a counterargument.

Perhaps the most convincing evidence to date comes from a meta-analysis of 51 studies showing that standardized mean differences between groups told to fake good and to be honest ranged from .48 to .65 for the Big Five personality traits and was 1.06 for SDR scales (Viswesvaran & Ones, 1999). Some of these effect sizes increased by a factor of about 1.5 (for Emotional Stability and Conscientiousness) when examining

within-person effects, where people responded to the same test honestly and deceptively; effect sizes ranged from .47 to .93 for the Big Five and was equal to 2.26 for SDR scales. Data on fake bad manipulations were found in fewer studies, but the meta-analysis also showed that the same standardized group mean differences on test scores were around -1.90 with Emotional Stability and -1.0 with SDR scales. As with fake good manipulations, the effects of faking bad tended to be larger when examined within-persons, ranging from -.91 to -3.66.

Given that people can and do fake their scores in at least some situations, the valid measurement of other psychological characteristics requires that faking effects are minimized or eliminated. As a result, some studies have specifically examined whether indices (presumably of) the degree to which people fake (e.g., SDR and lie scales) can be used to correct the scores on psychological measures of interest, sometimes called the "content" or "substance" measures. Because the issues involved with "faking corrections" are broader and more numerous than can be covered adequately here, I refer the reader to a recent treatment of the issue by Schmitt and Oswald (2006), and only point out a couple of important issues.

The first issue is that many researchers who focus on statistical corrections tend to ignore the psychological processes underlying people's honest and deceptive responses. As such, debates have been typically focused on the criterion-related validity of content tests, at the expense of considerations of construct validity (see American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999; Society for Industrial and Organizational Psychology, 2003). Due to its conceptual nature, deception almost invariably threatens construct

validity. In a very simple example, if a 25-year encounters a question about age and answers "4 years old," there can be no debate that the content of the answer has little, if any, substantive meaning. Despite the obvious value of construct validity, just a few studies have examined it in relation to faking (e.g., Cellar, Miller, Doverspike, & Klawsky, 1996; Ellingson et al., 1999; Schmit & Ryan, 1993; Stark et al., 2001).[8] Regarding the vast research on criterion-related validity, empirical findings suggest that measures of faking tend to be weakly correlated with personality tests and their validation criteria, meaning that corrections are rarely useful in practice (Ellingson, Sackett, & Hough, 1999; Hough, 1998; Ones et al., 1996), especially given the influence of other situational factors, such as the selection ratio for job applicants (Schmitt & Oswald, 2006).

A second issue is that the construct validity of the faking measures may be questioned. Typically based on empirical research, many popular faking, lie, and SDR scales were designed to discriminate liars from truth-tellers based on whether liars selected statistically rare test responses (Crowne & Marlowe, 1960; Kroger & Turnbull, 1975). The problem with this method of detection is that measures often include statistically rare, socially valued events that can actually be true. Despite the encouragement of some to only use *realistically unattainable* responses as indicators of deception (e.g., Paulhus, 1994, 2002), scales often contain at least some items (e.g., "I believe there is never an excuse for lying" or "I am always ready to start afresh" from the *Unlikely Virtues Scale* of the IPIP; Goldberg, 1999) that could be endorsed by people honestly. When honest people with stellar responses are measured for unlikely virtues

[8] Smith & Ellingson (2002) recently suggested that faking in large scale selection procedures does not affect construct validity, in terms of a test's factor structure, but they did not establish, with direct evidence, that people in their sample were faking.

and other rare events, they will be assessed as fakers (i.e., "false positives"), despite being the exact people the target (e.g., employing organization) wants to reward (Elliot, Lawty-Jones, & Jackson, 1996). This "Mother Teresa Catch," as I label it, may be a very serious concern when the target can only reward the very best, high-scoring respondents (e.g., has a low selection-ratio). Therefore, the logic underlying criteria of statistical deviance is sound, but limited in its ability to help targets (e.g., test developers) accomplish their goals of detecting faking and deception. (Of course, this is less of a concern when targets are focused more on avoiding false negatives.)

Given the concerns about construct validity in these first two issues, an interesting notion arises. Since deception *should* decrease construct validity based on simple theories, the repeated null findings found for the efficacy of faking corrections might actually suggest that the nomological network is poorly specified. This may lead one to ask whether 1) faking measures lack adequate criterion-related validity, 2) predictor measures lack adequate validity, or 3) both faking and predictor measures lack adequate validity. Although a wealth of research supports the validity of personality predictors (commonly studied in faking research) for predicting job performance criteria, correlations found are far from perfect (e.g., Barrick & Mount, 1991; Barrick Mount & Judge, 2001), meaning that there may not be much error-free variance in personality scales to correct, with a poor measure of faking.

Across the three general bodies of work on detection, it is safe to say that most detection approaches to date have produced unremarkable results. Researchers have investigated a wide range of possible indicators of leakage, but found few specific observable behaviors that have any kind of consistent relationship with deception-related

processes. In some cases (e.g., the use of SDR scales, linguistic cues), it is not even clear that deception is being assessed well. In fact, much of the work on lie detection seems to explain how receivers generally perceive messages as truthful or deceptive (e.g., Bond et al., 1992; Levine et al., 2000) more than how they identify deceptive messages. Since few have delineated the constructs that are involved in producing a deceptive message at Layer 3, the vast research on the different methods of detection could be described as operating somewhat blindly. Current detection methods attempt to detect some signs of deception without always specifying what those signs should look like, apart from behaviors that represent unintended leakage (observed deception-unintentional at Layer 4).

For this particular study, the major implication of this literature is that there exist few accurate methods of measuring deception directly and the quality of messages produced by the process. The use of bogus items that reference impossible events appears to be the most clear conceptually and is used in this study, along with a standard SDR measure as an additional reference marker. The more distal outcome of effectiveness, however, has been studied quite extensively and is the final factor (within a single deceptive communication process) that affects whether the rewards and/or punishments (contingent upon perceived image discrepancy) are given to the source.

Given this specification of an ultimate outcome, it is then pertinent to ask whether the strategies have any effect on deception effectiveness. That is, does using a particular strategy, or set of strategies, actually benefit the deceiver? Although effects like the truth bias suggest that many deceptive actions are unnecessary in certain situations, I ask this question from a theoretical standpoint:

RQ5: What is the relationship between the type of strategy used and deception effectiveness?

A General Framework for Predicting Deception Effectiveness

The depiction in Figure 4 outlines the basic conceptual process of deception proposed here, based on the theories reviewed and strategies generated, for the communication of a single deceptive message. Since deception involves an intentional act of communication, it must be directed by some goal. In most cases, this goal is a reward (tangible, social, personal, etc.) contingent upon the delivery of a deceptive message that causes a target to perceive the source as meeting some standard. If the source is truly at the standard already, there is no need for him or her to enter into the deception process. If the standard is higher than a person's actual state, then it automatically creates an image discrepancy, and the motivation to reduce the target's perceptions of this discrepancy is what drives the rest of the communication process, starting with a decision to deceive. (Note that the higher "standard" can either mean that a person needs to appear better or worse, in a socially desirable sense.)

The decision to deceive may be all-or-none or of some magnitude (i.e., lying on half of the questions asked) depending on how it is conceptualized and the form of communication to be produced. Also affecting this decision are the source's characteristics and values (Snell et al., 1999). Some people may refuse to use deception as a means of discrepancy reduction, perhaps attempting to genuinely change. A number of contextual factors may also affect the decision to deceive. In some cases, the environment may be so tightly controlled that it is not even possible to deceive, or the source at least perceives that deception is impossible to perform. The major contextual

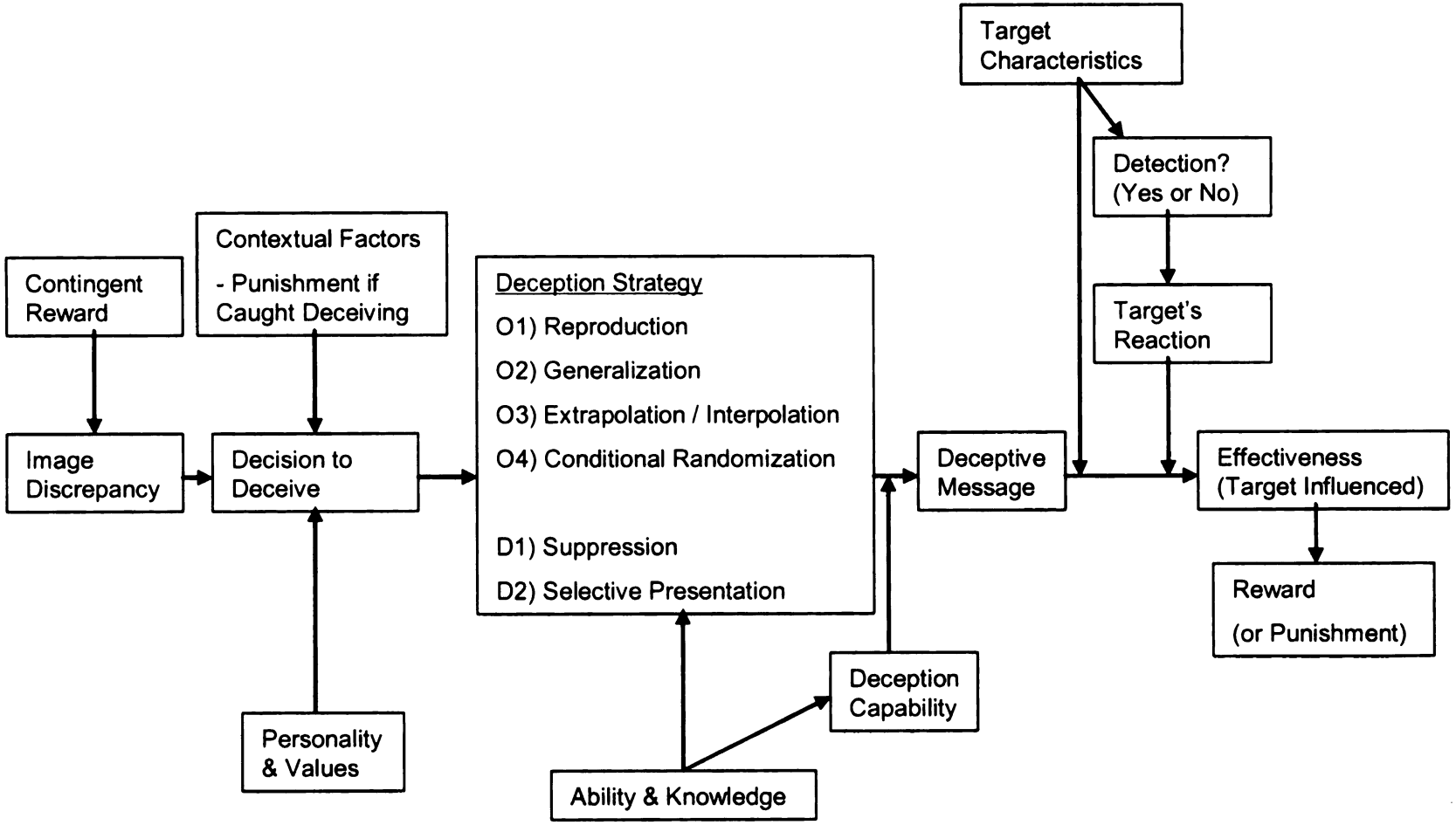
factor focused on in Figure 4 is the presence of a punishment that is contingent upon the source's deception being detected. This punishment may be known to the source through an explicit warning, for example.

Once the decision to deceive is made, the source must strategically plan, even if poorly or quickly, some manner of creating and delivering a deceptive message, including which behavioral tactics to use. I speculate that the selection, development, and implementation of one or more strategies will be determined by general cognitive ability and knowledge, at least in part. Knowledge may be general or specific to the context in which deception is performed (e.g., who the target will be, what the target desires, the structure and format of messages that will be acceptable to the target, etc.). The strategies are then executed as behaviors to produce a message outcome (e.g., verbal, written, or physical), and the quality of the message is partly a function of the deceivers' capabilities.

Finally, the target receives the message and interprets it. The interpretation is, of course, affected by the target's personal characteristics. If the target accepts the message blindly or automatically, then the message's effectiveness can be evaluated based on whether it reduces the originally perceived image discrepancy and results in a reward. If, however, the target attempts to assess the message in an attempt to detect possible deception, then the process becomes more complex.

Regardless of whether the target attempts detection due to his/her personal characteristics (e.g., paranoia or astuteness) or some situational demand (i.e., target is told to look for deception), the attempt may be successful or unsuccessful. A number of different scenarios can be developed to describe how targets react after detecting

Figure 4. Conceptual Depiction of Deception Process.



deception or overlooking it. For the moment, it is sufficient to assume that the target will have some kind of reaction that alters the influence of the deceptive message (i.e., its effectiveness). When deception is detected, punishments of some type are often conferred, though the target may also choose to pardon the offense.

Conclusion

My view is that deception-related behaviors are guided by the same set of underlying principles and follow the same general process, regardless of whether they are statements on a polygraph test, exaggerated answers on a multiple-choice exam, stories told on romantic dates, physical impersonations, or other forms of communication. By integrating literature on different concepts from different fields, I have attempted to point out similarities and derive a general definition of deception as well as a description of the deceptive process using psychological constructs.

One impediment to a simple integration of literatures and development of a coherent unified theoretical framework is the lack of attention paid to the cognitions and behaviors that comprise deception. Most of the work on deception has described its antecedents and consequences, while a separate body of literature has studied specific behavioral tactics as applied to very specific contexts. As a result, I used existing research and theoretical insights to generate a set of broad strategies that explain how motivation is translated into behavioral outcomes in the form of deceptive messages. The theory is by no means perfect and is purposefully left at an intermediate level of abstraction, rather than classifying a potentially infinite number of behavioral tactics. Yet, it expounds the need for a greater focus on psychological constructs and causal processes related to cognition, motivation, and behavior.

Empirical Investigation

The study conducted here was designed to examine some of the basic components of the deception process proposed above, as well as to explore their relationships with each other and relevant outcomes. Both qualitative and quantitative self-report data were collected to triangulate on different phenomena, primarily in an exploratory manner, though data were intended to provide some initial confirmation of the proposed theoretical framework. Since prior research has not explored the use of deception strategies, I applied the theoretical framework to describe a very focused and well-studied type of deception, psychological test faking, which allows experimenters to control key aspects of the communication.

Multiple-choice tests were used because they captured a very structured type of communication, while still allowing responses to be distorted in a deceptive manner. Also, multiple-choice questions automatically solicit information such that offensive strategies must be used, although respondents still maintained the discretion to skip questions, due to ethical concerns. In this study, participants were also told not to go backwards in the test and revise previous answers, thereby focusing the examination on sequential and relatively static communications.

Informed by existing research, the context was controlled even further to provide participants with specific, meaningful goals. People responded to the test while pretending to be job applicants seeking employment. By manipulating the types of relevant goals, it was hoped that strategies could be understood more clearly, particularly with respect to RQ3 (regarding relationships between strategies and goals). In conjunction with these goals, item responses were viewed as deceptive messages.

However, test scores were treated as proxy measures of deception effectiveness, under the assumption that a target (i.e., hiring organization) could make selection decisions and administer rewards (i.e., job offers) based directly on the observed test scores.

The five research questions posed throughout this paper provide the main foci of this study. Generally, the approach was to collect a vast amount of data and analyze relationships to determine what kinds of deception strategies exist, how prevalent strategies might be, whether the selection or execution of different strategies is related to the goals people have, whether strategies are related to any general skills or capabilities for performing deception, and whether strategies differ in their effectiveness. It was hoped that answers to these questions, when viewed together, would help to describe the broader process of deception.

A final research question was posed for additional exploration of situational factors that affect the above relationships. Different types of measures were presented to study participants in case different goals became salient, different strategies were used, or different strategies seemed to be more effective. All questions were still in the multiple-choice format, but each measured different constructs (at the scale and subscale/dimension level). Thus, research question 6 generally asks:

RQ6: What is the relationship between the type of strategy used and the type of communication (i.e., question format)?

In general, measures were chosen because they were commonly studied with respect to deception and used in applied psychology, but also because of their accessibility (Goldberg et al., 2006). Personality tests and biographical data were the two content measures of main interest. They tend to complement each other since personality

refers to general tendencies while biodata often refers to discrete (and sometimes verifiable) experiences (Nickels, 1994). A situational judgment measure was also included to explore a more complex question format. The situational judgment measure is described in the Methods section. So, I only point out that it was included in this study based on the assumption that it was less transparent than the other measures. The reasons for this assumption were threefold: 1) researchers have experienced difficulties trying to explain what exactly situational judgment items measure (McDaniel, Buhn-Finnegan, Morgeson, Campion, & Braverman, 2001; Schmitt & Chan, in press), 2) items require responses to two related, but distinct questions, and 3) the multiple-choice response options sometimes resemble a forced-choice test in that options may assess different constructs and may appear to be relatively equal, since empirical scoring keys are used. Finally, two measures of faking are investigated. It was proposed earlier that bogus items may be relatively accurate measures of deception. However, SDR measures are most common in the research literature. As a result, both types were studied.

METHOD

Sample

Study participants were college students in psychology courses who were recruited through a departmental "subject" pool. They received course credit for their participation. After scheduling an experimental session on the Internet, they were administered a brief online survey in Part I of the study. Usable data were collected from 192 people, but only 142 of these participants also attended their scheduled session and completed a paper survey in Part II. As expected, a number of participants did not provide usable verbal protocol data in Part II. The manipulation checks were also used to exclude people who did not fake their answers or who did not pretend like they were a job applicant. In addition, one record was lost due to a computer malfunction that corrupted an audio file, and three records were simply inaudible. In the end, 127 participants provided usable data for both Parts I and II of the study. Females comprised 68.5% of the sample. Sixty-four percent of the sample was 18 or 19 years of age, and 95% was 21 years or younger; approximately half of the sample was first-year students. The ethnic breakdown of the sample was as follows: 79.5% Caucasian, 10.2% African-American, 3.1% Asian, 3.1% multi-racial, and 4.1% other.

Procedure

The study consisted of two parts: an online survey and a job applicant simulation. The Part I survey consisted of a pretest of personality, an experimental measure of deception skill and efficacy, and demographic questions. Between 1 and 32 days later, participants attended a 75-minute experimental session. The average time between parts I and II was 11 days, but 80% of the participants waited between 5 and 20 days. In Part II,

people were told to pretend they were job applicants and respond to a test battery comprised of biographical data, personality, situational judgment, impression management, and mock interview questions, in that order. While taking the test, they were prompted at certain points to vocalize every thought in their head using the "think-aloud" procedure described by Ericsson & Simon (1993) and explained below.

Each participant was run in a separate session so that candid verbal reports would be solicited in a confidential environment (Kuusela & Paul, 2000). Individual sessions also enhanced the quality of audio recordings of the verbal reports collected during think-alouds. At the same time, this design decision required numerous resources and necessitated the involvement of multiple experimenters, including myself and five undergraduate research assistants, because each subject needed to be monitored constantly in case he/she failed to think aloud. Each undergraduate was trained to run experiments according to the experimental protocol in Appendix A and to deal with potential questions and problems. To insure consistency of the experimental procedure, assistants were required to observe me run one session and to run two sessions under my supervision, at a minimum.

The experimental procedure is summarized here (see Appendix A for further details). Following a general introduction and the procuring of informed consent, experimenters activated a Sony ICD-SX digital recorder. Next, participants were taught to think-aloud with instructions adapted from Ericsson & Simon (1993) and practical recommendations in van Someren, Barnard, and Sandberg (1994). Participants practiced providing verbal reports, or protocols, with five questions and received encouragement and feedback when necessary (per Ericsson & Simon, 1993). Then, participants were

told to select a job they would want and to imagine applying for it when being tested throughout the experiment. They were also given instructions specific to their randomly assigned condition, that directed them to respond honestly or in a deceitful manner (see below and Appendix A).

Following the condition-specific instructions, all participants were told that the hypothetical hiring organization was looking for applicants who are "hardworking," "responsible," "cooperative," and "helpful in a team," but not necessarily "outgoing" or "extraverted." The test battery was then administered, with certain questions prompting participants to say their thoughts out loud (Appendix B). Upon completion of the test, the digital recorder was stopped, and participants completed a final series of questions: a deception training item, experience with personnel selection instruments, knowledge of "ideal others," perception of subjective norms regarding lying, attitudes toward lying, and manipulation check items. Prior to departure, participants were asked for permission to publish their verbal protocol anonymously and were given a debriefing that included a request not to discuss the study and its manipulations with potential participants.

One important note is that a concerted effort was made to ease the cognitive burden place on participants since they had to focus on thinking aloud, pretend to be job applicants, and take a test battery while faking (for the relevant conditions). The instructions were repeated during the introduction as many times as was necessary and participants were given a reminder sheet of the major instructions (Appendix C). Visual cues next to each item requiring participants to think-aloud made the procedure simple and intuitive. Also, experimenters paused the test at specified points (after the questions

9, 26, 42, 66, 92, 102, and 140) to repeat the major instructions (see fourth page of Appendix A), though they performed minimal to no interactions in between these pauses.

Think-aloud and verbal protocol (VP). Verbal protocol data were collected as a means of assessing the types of cognitions that people produce when creating deceptive test responses. Although the think-aloud method for obtaining VP has been used infrequently in applied psychology, it allows for the direct measurement of task-based processes (Ford, Schmitt, Schechtman, Hults, & Doherty, 1989) and nondiscrete internal cognitions (e.g., Ball, Langholtz, Auble, & Sopchak, 1998; Barber & Roehling, 1993; Williamson, Ranyard, & Cuthbert, 2000). Neuroimaging techniques now being used to explore deception-related cognitions (Henig, 2006) appear to be the only alternative methods for direct measurement. Here, concurrent verbal reports were collected during each "task" in this study, with a task being the generation of a response to a single test item. As such, the tasks were very short in duration and numerous. Had retrospective reports been collected at the end of the test, the verbal protocol would be greatly influenced by memory effects (Nisbett & Wilson, 1977) and probably be less detailed (Kuusela & Paul, 2000).

Unlike the highly reactive method of introspection that tends to produce invalid self-reports of what a person was thinking or performing (Nisbett & Wilson, 1977), think-aloud procedures ask participants to state literally every thought that occurs to them spontaneously (i.e., "pops in their head"). This instruction discourages people from interpreting the meaning of their cognitions and from inferring the occurrence/meaning of other cognitions/behaviors that did not actually occur (Ericsson, 2002). Though the thoughts that can be accessed in this manner are limited to ones in consciousness, studies

(e.g., Barber & Roehling, 1993; Cable & Graham, 2000; Isenberg, 1986; Sonnentag, 1998) have demonstrated that rich and valid data can be obtained with this method (Ericsson, 2002). As such, normal patterns of test responding should not be drastically altered by requiring people to vocalize their thoughts.

Robie, Brown, and Beaty (2006) conducted what appears to be one of few, if any, verbal protocol studies examining test faking. Twelve people thought-aloud while completing a personality survey in the simulated context of applying for a job, with the chance to win a cash prize. Using the verbal protocols obtained, Robie et al. classified respondents as honest (n = 9), slight (n = 2), or extreme fakers (n = 1). Unfortunately, the small sample size of the study and the lack of other, more objective criteria for identifying fakers means that the conclusions reached required a substantial degree of inference on the part of the coder.[9] This potential problem is further exacerbated by the base rate of faking that was observed, with just one extreme faker. By contrast, the current study describes many more subjects, as they respond to different types of measures, under three instructions to deceive: to fake maximally, to fake in a realistic manner, and to respond like an ideal applicant.

Before proceeding, I explain a few aspects of the specific think-aloud procedure used here. Based on a pilot study, 5 seconds was found to be a suitable length of time to wait before prodding participants to keep saying their thoughts aloud, per standard procedures (e.g., Linderholm & van den Broek, 2002; Sonnentag, 1998). To reduce the number of cognitive processes that would be analyzed and to insure that verbal reports would not become convoluted, participants were not permitted to revise previous answers. Think-alouds were required for just 73 of the 140 items in the mock selection

[9] The study provided no mention of a coding process. It is not clear whether multiple raters were involved.

test battery in an attempt to keep people focused on the job applicant simulation and to eliciting normal test taking and faking behaviors (Kuusela & Paul, 2000), as well as to prevent rambling. In the hope that the verbal reports would reveal clear information about deception, I attempted to collect data for enough items within each test scale to analyze differences between them. Five items for each biodata subscale, 7 items for each personality dimension, and 10 items within each socially desirable responding subscale were initially chosen at random to request participants to think aloud (see Appendix B for the exact item numbers within the test); all situational judgment items were think-alouds. After these items were selected, I substituted items when two were very similar, but one seemed clearer and likely to produce more interpretable verbal protocols. I also made substitutions when certain items seemed more interesting and potentially able to elicit insightful cognitions. These decisions were somewhat arbitrary, but not unjustified given the exploratory nature of this investigation.

Finally, some of the verbal protocols were excluded from coding or analyses based on decisions described below. Here, I only emphasize that protocol studies typically use small samples of approximately 10-20 or participants (e.g., Cable & Graham, 2000; Sonnentag, 1998) or find large portions of the data to be unusable (e.g., Linderholm & van den Broek, 2002, discarded 30% of their participants and Harrison, McLaughlin, & Coalter, 1996, discarded 50%). The main reasons for this trend seems to be that not every participant finds thinking-aloud easy, produces complex thoughts, and/or provides informative descriptions of their cognitions.

Experimental manipulations. All participants were told to take the test battery as if they were applying for a job. Some researchers have suggested that fakers can produce

higher scores or, at least, more interpretable scores when focused on faking for the purpose of obtaining a specific job (e.g., Bridgman & Hollenbeck, 1961; Elliot et al., 1996; Schrader & Osburn, 1977). So, each participant was asked to select a specific job that he/she would like to apply for and to write it down, as an additional reminder. Participants were allowed to choose their simulated job as a way of mildly enhancing their intrinsic motivation to engage in the applicant simulation and perform well on the test. This discretion also enabled participants to choose jobs that they would find at least somewhat familiar, even if only through general information and stereotypes.

Because all test battery questions could be viewed as measuring something desired by the organization, participants were given specific goals in the form of direct recommendations to enhance certain characteristics (listed earlier and shown in Appendices A and C). This design feature was expected to create differential patterns of test responding so that goal-directed faking could be distinguished from other faking response processes (e.g., general socially desirable responding or use of a generalization rule like picking the highest numerical answer for all positively worded questions). Differential response patterns were made even more obvious by including questions about artistic appreciation and culture that were rarely regarded as being job relevant and were not mentioned in the experimental instructions. It should also be noted that despite the variance in jobs selected by participants (Appendix E), no job was wholly incompatible with the instructions about the specific qualities desired by the hiring organization. Overall, there were few instances (on specific questions) when people disagreed with the job desirable information provided in the instructions.

Participants were randomly assigned to each condition (in Part II of the study) in a somewhat indirect manner to increase the likelihood that usable data would be collected in all conditions, given that not every participant was expected to provide complete VPs. With a target of 35 participants per condition, the first 120 participants to complete Part I of the study were randomly assigned to one of the four conditions in Part II. If a person failed to attend the scheduled session or provided inadequate or inaudible VPs (despite being reminded to speak up), that person's condition assignment was recycled and given to the next participant signed up after the original 120. Because the initial condition assignments were ordered randomly and because there was no systematic pattern regarding data found to be unusable, recycling conditions in this way did not create new bias. After the first 120 (presumably) usable records were collected, 20 additional sessions were assigned in the same manner as the first set of conditions. All of the data collected were then reevaluated for inclusion in the analyses.

Now, I summarize the conditions and usable data available in each. The thirty-three people in the honest (H) condition were told to pretend they were in a job applicant setting, but to answer everything accurately and honestly, "for purposes of this experiment." These instructions initially created some slight confusion for a few participants. Repeating the instructions when necessary appeared to eliminate any further confusion. Thirty-three participants in the fake good (FG) condition were told to "produce responses that would result in the best scores," to "not worry about your real answers," and to "fake the best" answers. Thirty-one participants in the fake realistic (FR) condition were also told to "produce the best scores" to be hired and to not be concerned with their honest answers. In addition, they were told that a computer would

evaluate their patterns of responses by looking for ones that appeared to be "too good, abnormally high, or unrealistic," and told that they would "lose [their] chance to get the job" if caught. The latter information was crafted to be consistent with the warning research by Dwight and Donovan (2003) that suggests warning statements are effective when mentioning both the detection method and penalty. In the initial instructions and later repetitions, it was emphasized that the participant should try to get the best score without looking "too unrealistic." Lastly, thirty-one participants in the ideal applicant (IA) condition were told to "look like as good an applicant as possible," while not worrying about their real answers. It was also emphasized that they should produce responses to "look like an actual person who would definitely get the job." In this way, it was hoped that people would focus on the instructions more literally and not answer questions like a super-human being or ideal person; they were instructed to look like the ideal "applicant" given normal limitations of human beings.

Variants of the three faking conditions have been popular in past research (including many of the studies cited in the Introduction). The ideal applicant condition does require an assumption that the participants would not actually be stellar applicants naturally, and potential limitations of this feature are discussed later. The main reason for including the ideal applicant manipulation was that it seemed most likely, as compared with the other two faking manipulations, to frame the task as one where role-taking would be beneficial. Role-taking and the creation of ideal profiles have been studied (e.g., Kroger and Turnbull) in the past, and were suggested examples of the E/I strategy (in Table 3).

Special practice item. As stated, one of the practice think-aloud problems was presented again on the test, without any alteration. Van Someren et al. (1994) recommend giving participants practice problems that resemble the task they will actually perform. Yet, there was a specific reason for presenting this personality item twice ("Have a soft heart."). The item's generality and lack of direct reference to most jobs implied that it would not be transparent in terms of its job or social desirability. Also, people could and did view it as being both a benefit and liability in the workplace. These characteristics allowed me to assign a "correct" answer that would be accepted by most participants without serious question. Regardless of the response given during the think-aloud practice, respondents were told that the research has found that the best workers in any job tend to select 'moderately inaccurate' as their answer. To reduce any suspicion that might be aroused if the "correct" answer was given for just this item, a correct answer was also provided for the first practice item as well ("What is 32 x 7?"), immediately preceding this item.

Measures

International Personality Item Pool (IPIP). Goldberg's (1999) IPIP measure of the Big Five personality dimensions was used. Items were available on the Internet (at www.ipip.org) in a 50- and 100-item version, with the latter version including all items of the former. As the 50-item version has become increasingly popular in research, it was administered in Part II (referred to as the posttest), and the remaining 50 items of the 100-item version were administered on the Internet in Part I (pretest). Although some of the items within each dimension look very similar, no item was administered in both the pre- and posttest, as a way of reducing potential practice or order effects.

Later, I address whether score changes between the two tests represent the magnitude of faking that individuals performed. For the moment, however, it is sufficient to say that the test developers do not claim that the two sets of 50 IPIP items are parallel (L. Goldberg & J. A. Johnson, personal communication, April 11, 2006). Based on data from participants in all conditions, the Big Five scale internal consistency reliabilities for the pre- and post- tests were generally acceptable, ranging from .77 and .84 except for the Agreeableness pretest (α = .72), Openness pretest (α = .69), and Agreeableness pretest (α = .68) scales. More detailed information is presented in Table 4 below.

Biographical data (Biodata). A measure of general background and life history for selecting college applicants (Oswald, Schmitt, Kim, Ramsay, & Gillespie, 2004) was adapted for use in this job applicant simulation. The questions were rephrased to assess experiences in college rather than high school, but one item was missed during this process (i.e., asking about volunteer experience in high school). Ten items each were taken from 4 of the 12 subscales of Oswald et al.'s work: Citizenship, Perseverance, Leadership, and Artistic (and cultural) appreciation.

Although it would have been interesting to compare deception processes for different types of scales measuring the same construct (i.e., if the biodata measured personality), a biodata test of the Big Five was not accessible at the time of the study. The perseverance scale was selected because its content overlapped with the information given to participants, regarding the characteristics desired by the organization. The leadership scale was selected to provide some intermediate level of contrast because this quality was not mentioned as being important to the organization, but was directly connected to the organizations' stated need for teamwork. The artistic scale was chosen

Table 4

Internal Consistency Reliability Estimates For Study Variables

Test	k	α	Honest (*n*=33)	Fake Good (*n*=33)	Fake Real (*n*=31)	Ideal App. (*n*=30)
Part I - Online						
Conscientiousness (Pre)	10	.84	.83	.87	.84	.82
Extraversion (Pre)	10	.83	.92	.85	.79	.79
Agreeableness (Pre)	10	.72	.73	.72	.65	.79
Emotional Stability (Pre)	10	.82	.78	.83	.84	.75
Openness (Pre)	10	.69	.64	.75	.63	.52
Deception Efficacy	4	.57				
Deception Skill	8	.85				
Part II - In session						
Conscientiousness (Post)	10	.80	.72	.66	.69	.54
Extraversion (Post)	10	.77	.90	.55	.44	.60
Agreeableness (Post)	10	.68	.71	.51	.66	.77
Emotional Stability (Post)	10	.78	.78	.66	.47	.67
Openness - (Post)	10	.80	.84	.73	.71	.68
Situational Judgment (SJ)	5	.39	.52	.01	.08	.02
Biodata: Perseverance	10	.80	.78	.62	.56	.57
Biodata: Artistic & Cultural	10	.89	.88	.86	.86	.89
Biodata: Citizenship	10	.82	.63	.74	.63	.78
Biodata: Leadership	10	.87	.83	.86	.77	.73
Impression Management (IM)	19	.85	.80	.86	.73	.67
IM - Extreme Scoring	19	.85	.73	.82	.84	.72
Bogus Items	3	.75	.23	.45	.66	.85
Ideal Applicant Knowl.	6	.62				
Percept. of Lying Norms	10	.74				
Attitudes Toward Lying	5	.85				

Note. N = 192 for Part I; N = 127 for Part II. k = number of items in the scale, α = internal consistency based on total sample

to provide a large degree of contrast, in that it was never mentioned as being important to the organization and was rarely viewed as job-relevant, but still measured characteristics that are generally regarded as socially desirable. Together, these scales provide breadth in this investigation because they allowed participants a number of opportunities to select and implement different kinds of faking strategies. Appendix B contains sample items for this proprietary measure. Alpha reliability coefficients for the subscales were good for the total sample, with all above .81 (see Table 4 for specific reliability information).

Situational judgment (SJ) items. Five situational judgment items were randomly selected from a test developed by Oswald et al. (2004). These items were designed to measure college experiences relevant to the development of students across a broad set of academic, extracurricular, and personal domains. This test was chosen because 1) it was accessible and 2) students would be able to understand the dilemmas better than if they were presented with job-specific situations. For each item, a hypothetical dilemma is described along with a set of 5 to 7 possible responses to the situation. The test respondent is then asked to choose both what he/she would be "most" and "least likely to do." Based on empirical key of expert responses, a respondent receives +1 point for choosing the experts' choice, -1 point for choosing the experts' choice for the opposite question (e.g., respondent chooses the experts' best choice as their "least likely" answer), and no point for all other cases. The score for each item is a sum of the scores to the two questions, ranging from -2 to +2. Because the alpha coefficient for the 5 items was only .39, I did not interpret these items as a scale, but did conduct some exploratory analyses using the sum of the item scores because it was simpler than interpreting each item separately. Appendix B contains sample items for this proprietary measure.

Bogus (Biodata) items. Three items were presented amongst the biodata to detect blatant lies (Appendix B). These items asked about fictional events and "bogus" skills that no participant could endorse if they were being honest *and* accurate. Items were adapted from Anderson et al. (1984) and Ramsay, Schmitt, Oswald, Kim, and Gillespie (in press). The words in each item all had a genuine meaning by themselves (i.e., words were not gibberish), but did not make any sense together as a broader statement. A scale score was calculated as the mean of these items, and their internal consistency reliability was .75.

The bogus items were scattered throughout the test, but their placement was not random. They were presented in the latter two-thirds of the biodata to increase the likelihood of obtaining usable verbal protocols, since not every participant was expected to be comfortable thinking-aloud at the start. Also, earlier placement of the bogus items could have created unnecessary confusion about what the biodata questions were generally asking.

Impression management. The impression management scale of Paulhus' (1991) Balanced Inventory of Desirable Responding (BIDR Version 6 – Form 40A) was administered. Due to their potentially intrusive nature, two of the 19 items were not used: "I have sometimes doubted my ability as a lover" and "I never read sexy books." Items from the BIDR's self-deception scale were also administered, but are not discussed because their meaning is beyond the intended scope of this study, as explained in the introduction.

There is some uncertainty regarding the best way to score the BIDR. Because Paulhus (2002) aims to measure people who give "overly positive self-descriptions," he

specifies that items should be scored dichotomously (1994). As the Mother Teresa catch implies, this decision should reduce the number of false positives produced by the scale. However, Stöber, Dette, and Musch (2002) showed that continuously scored items had better psychometric properties. I examined both types of scoring and found that they produced nearly identical results, both having alpha coefficients of .85. Scale scores produced by these two methods were correlated .94. This strong similarity may be the result of the strong faking manipulations used. Also, one should note that the scale scores for both scoring approaches rest on a continuum; Paulhus only suggest scoring items dichotomously rather than the overall scale score (i.e., as a cutoff score to divide fakers from honest responders). Consistent with the theoretical focus of this study, I used continuously scored items since false positives posed only a marginal threat to internal validity.

Deception training. After the verbal protocol had been collected and the test battery was completed, participants provided an open-ended response about how they would train another person to "fake, lie, or exaggerate answers to the questions" to get a high score (see Appendix B). This item was used to examine the kinds of deceptive strategies that actually occur to people. Because participants were not given time to prepare an elaborate deception strategy, as they might in a real situation, the question was not limited to actions performed during the experiment. Participants were told to list both performed behaviors and new insights that did not occur to them during testing. In this way, the question was pertinent to the experience of all participants, even those in the honest condition. Framing the item as a question about training another faker also enhances the external validity of these data because it is reasonable to assume that fakers will ask others about a test's content or about advice for getting a high score.

Mock interview questions. Three open-ended items resembling interview questions were included in the test for a supplementary investigation of deception beyond multiple-choice questions. Verbal protocols were also collected for these items. The results for these data, however, are not presented because they are beyond the focus of this study. The items were administered last in the test battery and should have a negligible impact on the study results, though insights about faking these questions were mentioned in response to the deception training item.

Cognitive and verbal ability. College admissions test scores were obtained from the university's admissions office, with the participants' permission. ACT composite and section scores were obtained for 124 participants; the composite scores ranged from 15 to 31. Thirty-one composite and section scores were obtained for the SAT (a less popular test for this institution and its geographical region), and composite scores ranged from 810 to 1390. Although this sample was slightly range restricted on these tests as compared to the national population of test-takers, the score range was similar to that of the institution's undergraduate population, on both tests. To create a single index of general cognitive ability, all SAT scores were converted to equivalent ACT scores, using a chart provided by the test designer (SAT program information – SAT/ACT comparison, n.d.). The converted composite scores were then averaged with any ACT composite scores taken by the same participants. The resulting 131 ACT values were then standardized using the 2004 national norm for the ACT (ACT, n. d.) to present the results in a simpler, standardized format.

I then computed an index specifically of general verbal ability because people with greater verbal ability tend to produce lengthier verbal protocols (Ericsson & Simon,

1993). The index was comprised of relevant sections scores from the SAT and ACT. Though one should not presume that these tests are strictly equivalent, Dorans (1999) found theoretical similarities in content and a high degree of concordance in scores between the reading and English sections of the ACT, as a set, and the verbal section of the SAT. Accordingly, I summed the reading and English ACT section scores, and averaged these with the converted values corresponding to SAT verbal scores (using Dorans' Table 14), when people had taken the SAT in place of or in addition to the ACT. The index of verbal ability was then standardized.

Perception of subjective norms about lying. Ten Likert-type items adapted from (Christie & Geis, 1970) and McFarland (2000) were used to measure people's perceptions about how lying is viewed in general (see Appendix B). These responses also indicate views of ethical behavior in some sense. The purpose of measuring this was to determine whether certain people would be unmotivated to engage in the test faking task at the outset. However, participants who stated feeling uncomfortable with lying before the task started were told to view the task as a challenge rather than as a moral dilemma. Beyond this purpose, these items were not of direct interest in this study, because participants in the manipulated conditions were given direct motivation to fake their answers. The alpha reliability estimate of the averaged item scores was .74.

Attitudes toward lying. Five items from McFarland (2000) were administered to measure whether people had a general positive or negative view of faking and lying. Participants were given the item stem: "I think that lying on a selection test is..." Then they rated five 5-point scales representing different qualities (see Appendix B). This measure was examined for exploratory purposes to see if scores and if the type of

strategy used was linked to whether people enjoyed lying on the test. Items were averaged into a scale score, which showed an internal consistency of .85.

Deception capability. An experimental, 15-item self-report measure of deception capability (Appendix B) was administered. Because theory has not specified what deceivers actually do, questions could not be aimed at how well specific behaviors are performed. Instead, the ability to produce deception "well" was based on past experience in performing and succeeding at deception, as a proxy for skill. This measure is expected to show some conceptual overlap with self-monitoring self-presentation scales. Such scales evaluate how well a person recognizes his or her own tendencies to actively construct a social persona (Paulhus, 2003). What distinguished the capability measure is that it is intended to assess *only* qualities and experiences directly related to the use of deception, though this measure is influenced to some extent by respondents' propensity to engage in deception. This measure had been examined in a previous pilot study and refined such that items showing no variance among college students had already been removed.

Two subscales were formed based on prior exploratory factor analyses. Eleven of the capability items refer to general deception skills and success at performing deception. Four self-efficacy items referred specifically to perceptions about how well one could perform deception on a test (since efficacy is defined as a task-specific concept). Based on the 192 participants who completed measure in Part I of the study, internal consistency estimates were $\alpha = .88$ for skill and $\alpha = .57$ for efficacy. An exploratory factor analysis with varimax rotation produced a three-factor solution because one item ("It would be difficult for me to fake a personality test.") had small but approximately

equal cross-loadings on all factors and another item loaded cleanly onto a factor by itself ("I usually get caught when trying to deceive someone."). A two-factor solution was supported after removing these items, but the alpha reliability of either subscale showed virtually no change. Because both items had substantive meanings consistent with their subscales, they were retained here.

Manipulation checks. Three pairs of items on a 5-point multiple-choice scale served as checks of the study manipulations and aided in data cleaning procedures (see Appendix B). The pairs asked about thinking aloud, pretending to be a job applicant, and faking/being honest, respectively. In each pair, the first item specifically asked about ease of performing the given instructions and the second item asked about the extent to which one's test performance was influenced by these factors. Three decision rules were used to identify participants who should be excluded from the analyses based on responses to these items: anyone 1) marking '5' for both items in a pair, 2) marking '5' for the 2nd item in a pair, and '4' for the first, and 3) anyone marking a '5' or '4' for all six of the items. (Two exceptions were made for the first rule for two participants in the honest condition who marked '5' for the last pair of questions because their VPs clearly indicated that they still provided honest responses, with no mention of deception.)

Knowledge of "ideal" others. Five items (Appendix B) were created to assess people's general level of knowledge of high-performing peers. Two items asked about affiliations with people who would be ideal job applicants, while the remaining items asked about people who had the same qualities as those desired by the hiring organization. These items were included as possible control variables to obtain a general sense of what people in the ideal applicant condition perceived to be ideal.

Demographic variables. General demographic information about age, year in school, ethnicity, sex, and English as a primary language was collected to help describe the sample.

Transcription

Recorded verbal protocol statements were transcribed into a spreadsheet through a combination of manual typing and computer dictation software. Transcription allows coders to evaluate the same datum, presumably enhancing reliability (van Someren et al., 1994). As can be seen in the results section, pound signs were used to denote portions that could not be understood or that were inaudible and ellipses to denote pauses (approximately 3-5 seconds). All irrelevant sounds were left out (e.g., "ah" or "um"). Punctuation was sometimes added to help readers understand what the statements sounded like when spoken, but not to impose a correct grammatical structure on the statements. However, basic grammatical and spelling mistakes were corrected in some cases to enhance legibility and to preserve participants' anonymity by substituting formal terms for slang typically used by people of particular ethnic backgrounds. When participants reread a question or response option aloud, it was transcribed as "*question*" or "*response option*" to reduce the amount of text and to help the coders understand that these words did not have unique literal meanings (every time they were read). This was not expected to impact coding decisions, since the content of parroted statements indicates nothing about deception processes, though notes of their occurrence can be informative.

Since participants in the control group were strictly instructed to respond "accurately and honestly," their VPs were of little use in this study other than to provide a

baseline level of scores for evaluating how much people faked in the other conditions, and for comparing test scores between conditions. Generally, the vast majority of honest VP statements involved either 1) attempts to retrieve honest answers from memory, 2) statements of selected answers, before writing them down, or 3) details supporting the content of the answers given. Also, research on the think-aloud technique has found that people make few verbal statements during pure retrieval processes, often skipping over thoughts (Ericsson & Simon, 1993). Therefore, honest VPs were only examined for the bogus items because they revealed how people generally reacted to these unclear items.

It also can be safely assumed that some participants in the faking conditions would be able to choose the "best" answer honestly, *at least* some of the time. Thus, one can expect some portion of the VPs in the faking conditions to be short as well. These statements were transcribed, but not always retained for coding (see next section) because they tended to provide information that was too ambiguous or sparse to be useful. In the end, 98.5% of the 7,925 comments/VPS were transcribed, with the rest either being inaudible or missing data. Two complete transcripts are provided in Appendix D. Each block of statements represents one verbal protocol describing a response to one question on the test battery. Though these participants tended to provide more insightful responses over the duration of the test, the individual VPs they produced are prototypical of ones analyzed in this study.

Coding Process

I used what could be termed a "minimally interpretative descriptive coding" process (Thompson, McCaughan, Cullum, Sheldon, & Raynor, 2004, p. 16), whereby five raters assigned dichotomous codes to each item response/VP depending on whether it could serve as evidence of a certain type of cognition, based on predefined categories. Generally, the analysis of qualitative, verbal protocols is resource intensive (Ericsson & Simon, 1993; Marques & McCall, 2005). This study was no exception. The complete set of 7,805 transcribed VPs needed to be pared down to a manageable number for coding. I describe in some detail the decisions made to accomplish this, and also explain the two-step process of selecting and defining categories to be coded for use in quantitative analyses. In the first step, categories were derived directly from the strategic theory of deception. In the second, categories were created from participant data.

Some transcriptions were not coded because they offered little to no information about the process of deception. One kind merely repeated the question or multiple-choice options. Another kind elaborated on personal feelings toward an event (e.g., "I really don't like going to concerts"). Affective statements can indicate that the characteristic under consideration was of some social value, or worth to the hiring organization, but they might only reflect a personal preference. Other VPs qualified the circumstances under which an item response would be accurate or selected (e.g., "I talk to people at parties 'very often,' as long as I'm with my friends."). These comments can be viewed as answers to hypothetical alternative questions, as opposed to the ones presented. (While this pattern of responding is noted as perhaps being a general way to "stretch" or "spin" the truth, it refers less to deception than it does to rationalizations that allow some honest information to fit the question.) Like honest VPs, some faked ones consisted only of

details supporting the content of a selected answer (e.g., just listing organizations when asked about number of volunteer hours worked). In the situations above, data were too ambiguous or vague to describe deception. Although only 44.6% of the VPs were retained for coding after eliminating these situations, the noncoded VPs were still included in later analyses since they would have been assigned "null" values, had they been coded. That is, these VPs were not coded for the sake of efficiency, not because they represented missing data. In this way, the decision not to code these statements should have a minimal impact on the study results.

In addition, extremely short VPs simply stating that an item or characteristic was "good"/"bad" or "important"/"unimportant," with no elaboration, were ignored. Though these comments are informative to a small degree and might reflect judgments about social or job desirability, they were too vague. Again, the same statements might only reflect the respondent's personal opinion (e.g., "That's important [to me]"). A similar argument holds for terse statements like "That *should* be true." These were also excluded from coding. (Fortunately, only a small subset of participants tended to give these short VPs.) By contrast, similar but marginally more informative statements were retained for coding (e.g., "That is not a good *quality* to have." or "*You shouldn't* do that." or "That would *look* really bad.") because their linguistic structure and diction slightly reduced the number of possible meanings they could have, at least when interpreted literally. Brief statements about qualities being "relevant" were also retained because it was more likely than not that relevance judgments were made in reference to the job applicant simulation, given the strong experimental context.

Unlike typical verbal protocol studies that break down transcribed statements into single thoughts or grammatical structures, VPs here were left as complete responses to each item because this study was focused on strategic processes that might not be discernible when broken down into discrete cognitions. This decision is consistent with the Kilpatrick (1968) approach to protocol analysis described and endorsed by Ericsson and Simon (1993, see pp. 298-299). Keeping item responses intact also facilitated coding because redundant thoughts regarding the same task did not need be coded multiple times, given my interest in the *types* of thoughts that people have as opposed to the order in which they occur.

Category development. Coding categories were first developed to cover key deception process components in Figure 4. Motivational goals to fake (and eventually get the job) and to avoid detection were coded, labeled "job goal" and "caught goal," respectively. However, neither image discrepancy nor the decision to deceive was coded simply because participants were instructed to answer either honestly or deceptively (based on their condition). General ability and deception capability in Figure 4 were measured quantitatively and did not need to be identified in the verbal protocols, but a category was created to measure any use of the test knowledge given to participants (about the qualities the test measures). The rest of the initial categories were created to capture specific instances of the different strategies.

The *reproduction strategy* was defined in a straightforward manner and measured fairly objectively. Moreover, it was expected to be used only with the aforementioned practice think-aloud question. (Later, it became evident that some people reproduced prior answers because they mistakenly believed that other items were repeated.)

Evidence of generalization was operationally defined as *socially desirable responding*, *job desirable responding*, and using some kind of test-taking rule ("rule use"). In the first two cases, respondents must first possess some knowledge about what is valued by society or the hiring organization and some sense of what the item is measuring for generalization to be possible. In the third case, respondents are just applying specific rules to determine which answer to select. For the most part, it can be assumed that these rules were acquired prior to the experiment, but in a few isolated cases, people created rules during in the session (i.e., interpolation) and applied them (i.e., generalization) on subsequent questions.

The E/I strategy is more complex and requires two pieces of evidence. Because it was intended to describe situations in which respondents could not just generalize existing knowledge, comments were coded when people were uncertain and attempted to determine what the question was measuring (i.e., extrapolation phase), labeled "dimension assessment." The interpolation portion of the strategy deals with the formulation of some response to be given, after the question's meaning has been extrapolated. Either the generation of a new response rule or the adoption of a role allowing an answer to be derived without test rules (see Table 3) was considered to be a sign of interpolation. Finally, evidence of the conditional randomization strategy would have been straightforward, but no comments showed any clear evidence of it.

This initial set of 8 categories was intended to allow for a weak test of the deception theory (i.e., that people use multiple strategies). As mentioned, this study also adopted an exploratory, theory-building/refinement approach. In this vein, I developed additional categories based on general observations of the data. I made a cursory

examination of the transcripts and identified certain types of cognitions that *seemed* to occur frequently enough to warrant coding. One category, labeled "memory," was added to capture instances when people selected answers based on their inability to recall relevant information (e.g., about their past achievements), perhaps indicating operation of the "availability heuristic" (MacLeod & Campbell, 1992). Another category, *leadership*, was created to measure instances when respondents assumed that leadership was a job desirable characteristic, despite it not being mentioned in the task instructions. Other VPs were coded when people considered the kinds of behaviors that are normative before selecting an answer, which occurred for impression management items in particular. Lastly, a category (labeled "uncertainty") was created to measure general confusion and an inability to understand the question asked, which was witnessed primarily with bogus items.

Next, I examined responses to the deception training item to identify new categories that would help describe the data more fully. To reiterate, participant-generated responses to the training item included both new insights and retrospective reports about deceptive actions performed during the experiment. To demonstrate that personal biases did not cause me to overlook new categories that would not fit with my proposed strategies (Glaser & Strauss, 1967), I present every independent idea collected in Appendix F for the reader to consider. In Appendix F, similar ideas were aggregated and organized within a loose structure, but I recognize that alternative groupings could also be valid. My intention was merely to identify new categories of interest given the focus on strategic behaviors. Using the list, I created six more categories: honesty, response patterns, concealment, excuses, demonstrating familiarity, and curvilinearity.

First, people recommended being honest, or close to honest, in a number of specific and general situations, even if faking the test overall. So, I coded VPs when participants explicitly stated the decision to give an honest answer instead of a deceptive one, or when they described the response given as being true while recognizing that it was not the "best" answer to select for getting the job. Clear instances of people selecting the best answer just because it also happened to be true were not coded since such behaviors are not consistent with the underlying meaning of this category.

Second, people recommended monitoring and controlling the general pattern of responses given throughout the test, versus focusing on responses to single items. Some of these might be viewed as "meta-strategic considerations" (ideas guiding the use of basic strategies) or concerns about strategy execution. However, most of the cognitions coded in this category were focused on maintaining consistency in answers to similar items or on compensating for general tendencies by purposefully creating inconsistency across items (e.g., choosing a low answer because too many extreme ones were chosen before).

Third, some people warned against admitting to doing or being certain things. Unlike general social or job desirable responding, thoughts in this category were focused on concealing the truth, not on selecting the best or most desirable answer. Thus, a person only focusing on concealment may be content to select a minimally acceptable (e.g., neutral) answer, as long as negative traits are not endorsed. For all intents and purposes, this category is an example of the defensive suppression strategy. Since suppression was not under direct investigation, concealment was a category coded only for purposes of describing the data more fully.

Perhaps for things that could not be completely concealed, training ideas also suggested fabricating specific excuses that could be given later to justify the answers given on the test. This category was coded, but the eventual identification of just a few comments rendered it relatively useless as a quantitative variable. So, this behavior is only reviewed qualitatively with selected VP comments.

Also rare but interesting, the category of demonstrating familiarity was created to capture a very specific strategy used when items (usually bogus items) were unclear and/or referred to "unusual" events/tasks/skills. Specifically, people recommended selecting a moderate to low answer as a way of demonstrating familiarity with the content of the item, but not expertise, based on the notion that novices would not be expected to perform a related task later (after the test) or to perform it well.

The last category, curvilinearity, was based on recommendations to avoid selecting extreme answers to avoid appearing "arrogant" or "obnoxious." Unlike acts related to "caught goals" (of avoiding detection), people actually believed that the best or most appropriate answer to certain items was a moderate one on the response scale. Thus, people were still selecting what they perceived to be the "best" answer, even though it was a moderate option on the response scale. While specific cognitions related to arrogance could have been coded, it seemed more useful to expand this logic to any situation in which respondents' thoughts directly implied that an item should be scored curvilinearly. (I say "directly implied" because no respondent actually used the word "curvilinear.")

In summary, the coding categories above were established to measure cognitions about different types of proposed deception strategies, other actions based on training

recommendations, and other typical decisions about test responses. Next, I describe the process of coding the data. Resource constraints and the sheer size of the data set required coding to occur in three phases (though the phases occurred nearly simultaneously in time). In the first phase, I coded all of the VPs, along every category. In the second phase, a pair of graduate students in psychology coded the VPs for 9 participants (3 per faking condition) who provided the greatest number of VPs that were not already excluded in the steps already described. Ratings were made for the categories of job goal and caught goal.

In the third phase, a medical student unfamiliar with psychology and a master's level graduate in psychology coded a larger subset of the data, along different categories (dimension assessment, JDR, SDR, and rule use). VPs for the 9 participants coded in phase 2 and for all 33 participants in the honest condition (for the three bogus items) were rated. In addition, twenty participants were added based on the following criteria. Three people were chosen from each faking condition because they generally had a large number of "codable" VPs. Two people providing a small number of codable VPs were then selected from each faking condition, in case the categories themselves were correlated with the tendency of people to provide codable VPs. Finally, 5 additional participants were randomly selected from each faking condition. This resulted in approximately 30% of the entire sample being coded in phase 2. For the latter two criteria, however, I did establish a cutoff (at 25% codable VPs) so that the participants selected for coding provided enough useful information to be worth coding at all.

Raters, excluding myself, were given minimal information about the background of the study and a set of written instructions that included category descriptions and

simple decision rules (Appendix G). For each category of cognitions, raters then assigned dichotomous codes to every VP because there is currently no basis for judging the quality (and therefore unique levels) of these deception-relevant cognitions. So, each VP was coded based on whether it did or did not represent an instance of cognitions in each rated category (somewhat like a behavioral checklist, but with cognitions). After evaluating the consistency between ratings, I summed the VPs across items, within persons, to indicate the frequency of cognitions in each category, to be used later in quantitative analyses.

One final note is that the coding of SJT VPs warranted a slightly modified procedure. On the whole, these VPs were more difficult to interpret because they consisted of two answers to the same item (i.e., "most" and "least" likely actions). They were lengthier and less ordered, with participants often shifting their attention back and forth between the two questions, item stem, and response options. Another factor affecting interpretability is that the response options do not fall on a clear continuum. Participants had greater difficulty comparing the options because each one could be associated with a unique set of advantages or disadvantages in the dilemma. In fact, it was not uncommon for people to evaluate the merit of each option independently, almost as if it was a separate item. To facilitate coding, the responses to the "most likely" and "least likely" SJT questions were collapsed into one unit, though they were transcribed separately.

After all raters coded their respective portions of the data, for their respective categories, it became clear that the job goal category was too broad, including virtually any mention of the job. For the 9 participants coded, more than half of the VPs qualified

as cognitions about the job goal. It also seems likely that participants were thinking about the job during other questions, but did not always bother to state such an obvious thing, especially if they had mentioned it on the previous question. Thus, the category was refined (labeled "job goal-revised") to measure VPs explicitly stating the job or performing well on the test as a motivational influence instead of as a reference to the general context (e.g., "I want to get the job" or "since this is for an interview, I'll put..." vs. "I might have to do that on the job", respectively). One other category was also created to narrow the meaning of the caught goal category, which included any concern about having deception detected. *Realism goals* were coded specifically for thoughts about selecting realistic answers, since this specific goal was assigned to people in the fake realistic condition. I then rated these new categories for all of the VPs.

Rater Consistency

It is important to establish that the coding process is reliable before drawing any conclusions about the category variables produced. Generally, raters should be able to assign the same category code (absence or presence) to each VP. Although a number of statistical indices of interrater reliability and agreement could be computed, the use of dichotomous ratings in this study means that most reliability indices will produce similar, but not necessarily identical, patterns of results as compared to agreement indices (see Cohen, 1960, and Schuster & Smith, 2005, for further explanation).

Nonetheless, it is important to recognize that an appropriate index in this study will describe consistency of the *rating process*, which includes the participating raters, instructions given, and operational category definitions used. An index should not necessarily describe how reliable ratings are across multiple items for each participant,

since there is no expectation that the categories will represent any kind of stable individual difference variable. Because prior theory and research do not warrant an expectation that participants will exhibit a general tendency to use one strategy over another, or to engage in a single type of process, measures of internal consistency might have little meaning.

Given these considerations, the proportion of agreement in judgments across the VPs for all participants was deemed the most appropriate measure. Posner, Sampson, Caplan, Ward, and Cheney (1990) provide an agreement statistic (P_o) for comparing two or more raters to a reference rater (Abedi, 1996; Cantor, 1996), or "standard" (Light, 1971). I averaged the percentage of agreement between myself and each of the other two raters, for the respective categories (in phases 2 and 3). This statistic then provides an indirect measure of the reliability of my ratings for all categories and VPs in phase 1.

Multi-rater kappa (Light, 1971; von Eye, 2006) was also computed, because it provides a holistic index of agreement (von Eye & von Eye, 2005) that is simpler to interpret when more than two raters are involved. Kappa is also useful because it accounts for the level of agreement that would be expected purely due to chance. Like the alpha coefficient for reliability, researchers typically evaluate kappa values according to a heuristic. Landis and Koch (1977) proposed that values of .4 or greater indicate moderate to good agreement, and values of approximately .7 or greater indicate excellent agreement (von Eye & von Eye, 2005).

Because kappa is biased downward when raters' marginal distributions are extremely asymmetric (Andrés & Marzo, 2004; Cohen, 1969), Brennan and Hays (1992) recommend interpreting consistency by considering both proportion of agreement and

kappa together, qualitatively. Table 5 presents these statistics, which are broken down by scale because it is reasonable to suspect that coding may be more difficult for certain kinds of VPs, given that the VPs did seem to differ for the different types of questions. As mentioned above, the VPs for SJT items were lengthier and not very systematic. For the personality questions, it was also apparent that VPs were generally shorter than those for the SJT or biodata, perhaps because the items themselves were terse and requiring participants to remember detailed information less often.

For the most part, the proportion of agreement between pairs of raters was high, well above the level expected by chance, but these figures are greatly influenced by the large numbers of null cases (i.e., VPs did not represent cognitions from a given category). Still, kappa values were also good for the categories of job goal, dimension, and JDR, and acceptable for most judgments of SDR and the caught goal. For SDR, rules, and caught goal, however, the base rate of these events was very low in the sample. Table 6 presents some descriptive statistics for these variables, regarding the maximum frequency of cognitions for any participant, and the average frequency across participants.

Rater consistency was not assessed for the other categories. This was partly due to resource limitations. It was also due to expectations that certain categories would have low base rates (i.e., reproduction strategy, response patterns over the test, demonstrations of familiarity with item content, and concealment) and that certain categories could be coded using fairly objective criteria. The categories of test knowledge, knowledge-leadership, realism goal, and memory were coded essentially based on the presence of keywords; uncertainty, on key phrases (e.g., "I don't understand this question"). While judgments about item curvilinearity, norm consideration, honesty, are somewhat more

Table 5

Consistency Indices: Multi-rater Kappa and Percentage of Agreement

			Dimension		JDR		SDR		Rules	
Question Type	Items	VPs	κ	%	κ	%	κ	%	κ	%
Phase 3										
Biodata	20	323	.53	.79	.70	.86	.48	.96	.34	.92
Personality	35	602	.45	.78	.73	.87	.40	.90	.20	.96
Situational Judg.	5	92	.47	.78	.71	.88	.33	.86	.00	1.00[b]
Impression Manag.	10	159	.48	.85	.77	.89	.50	.87	.28	.98
Bogus Items[a]	3	192	.00	.98	.62	.94	.20	.98	.14	.92
All scales	73	1368	.50	.82	.74	.88	.43	.92	.26	.95

			Job Goal		Caught Goal	
Phase 3			κ	%	κ	%
Biodata	20	157	.65	.86	.46	.92
Personality	35	283	.59	.80	.55	.98
Situational Judg.	5	38	.52	.78	-	1.00
Imp. Manag.	10	27	.71	.85	.23	.89
Bogus Items[a]	3	79	.85	.94	.42	.85
All scales	73	584	.64	.83	.44	.95

Note. Items = number of think-aloud items. VPs = number of verbal protocols coded by each rater. JDR = job desirable responding. SDR = socially desirable responding. Kappa could not be computed in one instance because no cognitions were identified in that category.

[a] Bogus items included responses from honest participants. Other statistics exclude them. [b] Denotes values rounded to the nearest hundredth.

subjective, a set of decision rules helped to delineate what kinds of comments would qualify.

Ultimately, a number of categories were endorsed too infrequently to be useful as quantitative variables in the analyses (see Table 6). The coding process then served only as a means of identifying possible examples of these cognitions. "Generating hypotheses requires evidence enough only to establish a suggestion – not an excessive piling up of evidence to establish a proof, and the consequent hindering of the generation of new hypotheses" (Glaser & Strauss, 1967, pp. 39-40). Instances of role taking, excuses, reproduction, and other rare cognitions, are discussed in qualitative analyses to inform future work about deceptive behaviors that are possible, but probably not normative. For the other categories, each VP was assigned a value based on the majority of the raters' codes, and then those values were summed within persons to create a frequency variable for the different categories of cognitions (as mentioned above). Although it would be ideal to require all three raters to reach consensus before assigning variable values, this criterion was too stringent.

Finally, it should be noted that my aim was only to assess the stability of the raters' judgments, with respect to each other. Consistency was not used to imply construct validity for the ratings. Within the framework of generalizability theory, the indices of interrater agreement used here do not account for rater-specific factors (Li & Lautenschlager, 1997). As a result, whether or not another researcher would rate the same VPs in a similar fashion, with or without the instructions, pertains to the broader issue of generalizability. Such issues are dealt with in an alternative manner by making the coding decisions in this study as transparent as possible and by presenting the raw,

Table 6

Descriptive Statistics for Verbal Protocol Coding Categories

Category variable	Max.	Mean[a]	SD	Time of Development
Job goal[b]	60	45.44	14.37	Derived from theory
Job goal - revised	5	.71	1.20	Category revision
Caught goal	10	.77	1.72	Derived from theory
Realism goal	4	.45	.90	Category revision
Norm consideration	7	1.04	1.45	Cursory examination of VPs
Test knowledge	45	4.52	8.93	Derived from theory
Knowledge - leadership	3	.25	.63	Cursory examination of VPs
JDR	66	14.45	17.14	Derived from theory
SDR	18	3.15	3.53	Derived from theory
Dimension assessment	44	8.12	10.89	Derived from theory
Rule use	7	.64	1.26	Derived from theory
Demonstrate familiarity	2	.04	.25	Deception training responses
Reproduction strategy	1	.19	.39	Derived from theory
Concealment	4	.24	.66	Deception training responses
Response patterns	3	.21	.57	Deception training responses
Uncertainty[c]	5	2.05	1.11	Cursory examination of VPs
Honesty	8	.75	1.36	Deception training responses
Memory	2	.09	.32	Cursory examination of VPs
Curvilinear answer	19	1.58	3.37	Deception training responses
Caught + realism	11	1.12	2.08	During analyses
Knowledge + leadership	45	4.69	9.06	During analyses

Note. $n = 91$ (participants in faking conditions) for all categories except uncertainty. Categories were measured as the frequency of specific cognitions per person. Max. = maximum frequency for a single person. Time of development = point in coding process when category was developed. Categories for role taking and excuses are not shown due to their extremely low frequency.

[a] Mean = average number of *total* cognitions per person, out of 73 possible verbal protocol opportunities.

[b] For job goal, $n = 9$.

[c] Since uncertainty occurred primarily with the bogus items, statistics are based on the total sample of 124.

qualitative data in conjunction with numerical results, so that readers may judge for themselves whether the results were presented in a credible manner. The potential limitations of this study with regard to the generalizability of its findings are covered in the discussion section.

RESULTS

Qualitative and quantitative findings are presented in the order that they pertain to the stated research questions, but I first offer some general observations about the verbal protocols collected and the overall patterns of responding exhibited by the participants. Following these observations, I present the results of the coding process to show that various goals were held and multiple strategies were used by the fakers. The main purpose of these results is to demonstrate that individual VPs supporting each coded category could be found in an unbiased manner. Some categories were related to the research questions more directly than others, and some categories could be aggregated because they had similar meanings in the context of this study. Subsequent analyses provide a more integrated view of the process of deception by addressing the interrelationships among the coded categories (i.e., deception and response process components), as well as the categories' relationships with deception capability and message outcomes (i.e., scale scores). Finally, I examine differences in deception-related cognitions across the different scales. Interpretations about the deception process as a whole, however, are reserved for the discussion section. Also, the temporal sequencing of each component is not examined since this exploratory study was focused on identifying and measuring the hypothesized deception components.

Before proceeding, the reader is forewarned that it is most appropriate to view the VPs and coded category variables as records of thoughts rather than behaviors for at least two reasons. First, people sometimes made basic errors such that the answer they intended to choose was not the one they wrote on their answer sheet. The protocols revealed some instances when participants remedied these mistakes, but other mistakes

surely went unnoticed. Second, coded category variables only measure whether certain cognitions occurred, not whether they solely determined a final answer. At times, people changed their response strategy or considered multiple response strategies before selecting one to be executed. Hence, the correspondence between coded cognitions and written test answers is imperfect. This would be the case even if the VPs could be coded in a perfectly reliable manner.

For the sake of brevity, transcribed VPs are sometimes presented as relevant fragments. The reader is assured that key information was not left out, especially when a VP described the changing of one's mind about a test answer or use of multiple response processes. For clarity, I present each VP with a reference to the participant's experimental condition and study ID, and to the scale of the item soliciting the comment (e.g., FR, 22, Agreeableness). Many of the VPs can be interpreted with minimal contextual information, but I present actual item content when a scale reference is too vague. To provide the reader with a richer picture of the data collected, I present a variety of VPs, but only when they are still representative of the data; few statements are shown more than once.

General Observations

Every participant had numerous opportunities to think aloud, a total of 73 when responses to the most and least SJT questions are viewed together as single units. Although the two participants in Appendix D provided codable VPs on a more consistent basis than was typical, their individual VPs resembled those produced by others in the sample. As already explained, non-codable VPs failed to describe deception, but are not missing data; nearly everyone thought-aloud for every question requiring it. For people in

the faking conditions (i.e., FG, FR, and IA), the mean of codable VPs was 34.1 (SD = 20.4). As for the extreme cases, three people provided fewer than 8 codable VPs, and three people provided greater than 68 VPs.

Generally, VPs within persons did not appear to vary much, particularly with respect to length. Except for a few occasions when participants provided brief answers because they referred to an earlier item or found one answer to be obviously best, large differences across scales and fatigue effects over the course of the test were not observed. Some participants certainly provided more information than others, but everyone appeared to be fairly consistent in the length of verbalizations they produced across items. (For reasons explained in the discussion section, exact word counts for each VP were not considered to be particularly meaningful measures and, thus, not used to make comparisons across scales.)

As expected, not all participants provided extensive statements of their thoughts. Yet, there is some evidence that the think-aloud procedure was implemented effectively. Most verbalized thoughts were about the testing task, as would be hoped for, but some participants also stated task-irrelevant thoughts such as being tired, being curious about what the experimenter (sitting behind) was doing, or thinking about plans for the day (e.g., parties). These show that spontaneously occurring thoughts were stated, per the experimental instructions. In addition, most people did not need to be reminded to keep thinking aloud, especially after the first 10-15 items. People who did require continual prodding were generally unable to perform the task as a whole and often failed the manipulation checks. Thus, the verbal protocols analyzed describe test taking processes with minimal interruptions, between the planned pauses.

Also, some participants performed in a role-playing type of manner supportive of the E/I strategy, but provided VPs that could not be analyzed. These participants treated the test battery like a face-to-face job interview. As such, they stated answers aloud with confidence and provided detailed answers, but censored any thoughts about the deception being performed. A few of these participants were identified through post-experimental questioning when this type of behavior was observed and then excluded from the analyses. Still, it is reasonable to assume that other participants may have acted in a similar manner through all or part of the test, especially since one person explicitly recommended this strategy for deception training: "try to seriously place yourself in the shoes of the applicant and imagine you are actually at the interview" (IA, 116). (Interestingly, this person did not appear to use this strategy much, or at all, and produced a fair number of interesting protocols, some of which are presented below.)

Multiple Goals

Job goal-revised. This section shows the types of comments that were coded as job goal cognitions (as defined and described in Appendix F). By itself, the job goal category does not carry much meaning in this study, since all participants were instructed to act like job applicants, even in the honest condition. However, it is a foundational component in the deception theory that logically drives all subsequent behaviors. By measuring it, analyses might be able to account for certain patterns of behavior, like how consistently fakers inflate their scores across test items. Although complex interactive relationships were not examined, this category also has the potential to explain instances when test fakers choose unrealistic or very extreme answers, even if they demonstrate an

awareness that their deception may be detected (i.e., they are extremely motivated to fake).

As explained above, the original job goal category indicated that most VPs included a direct or indirect reference to getting the job or performing well on the test. For the revised category focusing on motivation, 144 cognitions produced by 82 participants (i.e., 66.1% of the sample) were identified. It was also apparent that these cognitions tended to appear more often at the start and end of the experiment. The reasons for this pattern are not clear because it could be a function of the scale types presented or just of fatigue and memory loss prompting people to remind themselves of their goals. Here is a representative sampling of the VPs in this category:

> (IA, 91, Perseverance): "A great extent. Trying to get the job"
>
> (FR, 159, Perseverance): "I don't usually fail that much. I think... I want to get a job. ...That'll be rarely"
>
> (FG, 19, Conscientiousness): "In real life, I'm not really prepared. Because I'm trying to look good, I would say moderately accurate"
>
> (FR, 159, SJ-most likely): "Well... let me think. Usually, I'll probably skip it. Doesn't matter. But... if I want to get a job here, what would I do? *option A* No, that would be the least likely. *option B* that's a pretty good one. ... *response options* ... well, it might be B or C. Which one would be #? I would probably say C"

Caught goal. This category encompassed cognitions demonstrating some awareness of the possibility that an answer might be detected as fake. These codes were assigned more often to VPs pertaining to the bogus or impression management items, but could be found throughout the test. Some basic examples are shown below.

> (FG, 3, Citizenship): "This is something that I probably would do it. I don't... I don't wanna put every answer exactly what they want here because they'll think I'm lying. So I guess I'll just put somewhat likely, 'cause that's kind of the truth and... so I don't sugarcoat it...."

(FR, 79, Perseverance): "I... would say, probably... I don't know. 'Rarely' because it's not like it never happens to somebody. That wouldn't really be believable, and obviously I don't want to sound it. So I'd say rarely"

(FG, Leadership): "I'm going to say '4 or 5' because '5 or more' just kind of sounds like maybe you're faking"

(FR, 149, Bogus): "Wow. Matrix solvency files. Matrix -- put in order. Solvency files, publication standards. Okay... wow. What is this? ... If I put 'never,' I'll look inadequate. 'Five times or more,' seem like BS [i.e., bullshit]. *response options* I guess we'll go with 'twice'"

(IA, 81, Impression): "I'm going to go actually with F on that one because I don't want to seem like I'm trying to cover up everything I do... and lie"

(FG, 59, Impression): "See, I don't know. I don't... I feel like they'll know if I lie, but I don't know if they will. I'll just put D, somewhat true, because everyone speeds, I feel like"

Interestingly, the VPs in this category offered a range of reasons regarding how and when a person would be caught. Outlandish or unrealistic answers, trick questions, lie detector tests, post-interview tests of skills, and general observation on-the-job were mentioned as potential problems to be avoided. The only information actually given to participants (in the fake realistic condition only) was that a computer would analyze the general pattern test responses. Here are some exemplary comments of the different types of concerns observed:

(IA, 116, SJ-most likely): "Oh wow. So this is cut and dry. ... They're gonna be looking for someone that gives the... best lying answer. ... *option F* that sounds alright. And it's not the most, you know... made up, phony answer."

(IA, 158, Impression): "I usually don't like to talk about people # like that. I'm sure I have been guilty of that before, but it's not something that... I want other people to know that... I mean, not that I didn't #, but not something that you want to exemplify when you're in an interview, but I don't want to completely lie and sound inhuman, but..."

(FR, 129, Impression): "They don't, you don't... *question*. They're gonna know it's true... but you don't want it all the way true. So, we'll say D"

(FG, 80, Impression): "I think that everybody's done that. So maybe employers aren't looking for someone -- they might think that you are lying if you said that you hadn't. So I'll say that's somewhat true"

(IA, 116, Bogus): "well, let's see. I don't even know what that is. [...] Responsible, hard-working -- I don't think has to do with me. This question could also be kind of a trick or that it has nothing to do with my job. [...] So, I have to be honest and say never because, I mean, if that comes up again, I can't really prove that I know what I'm talking about."

(IA, 70, Conscientiousness): "it depends... depends on importance of the chore. If someone's going to be there to observe whether or not it's done right away. But for this, since I'm... be the 'ideal' video editor guy. We'll just go ahead and put very accurate. Hopefully there will be no lie detector test involved"

(FR, 118, Bogus): "Oh jeez. I'll stick with 'never' because I don't know what that is. And I don't want to say something I can't back up"

(FG, 2, Bogus): "I have no idea what that means. So... I have to go with... never, in case they ask me to do it"

(FG, 20, Bogus): "I don't even know what that is. So I'll go with never. Might as well keep it honest because you can't lie about something like that and then have them show it to me [to perform]"

(IA, 105, Bogus): "This program, I've never heard of it. So... I have no idea what this program is, man. So I have to just -- it might be important, but if they ask me how to work at it, so I'm just going to be screwed. So, I'm gonna say 'never.' This time I have to be honest."

(FG, 3, Artistic): "I guess I don't want to make myself out to be someone that's really into art and then when they meet me and try to have conversations with me about art, realize I don't know anything about it...."

(FR, 18, Extraversion): "I'll say moderately inaccurate... because the people at work will see who... I'm not the most outgoing person"

Another interesting situation occurred occasionally when people demonstrated an awareness of the potential to be detected, but decided that the threat was small.

(FG, 120, Bogus): "Ooh, since I'm... I've never actually used that program. ... But, I can probably say I've done it once, two or three times, and not know it that well and they'd probably, I'd probably get away with it. So, E"

Overall, these results show that people feared being caught faking. The examples above represent the bulk of data falling into this category and have central features in common. Yet, the reader is reminded that the kappa statistics were barely acceptable and suggest that some comments were difficult to code reliably. The other major finding is that different types of fears were mentioned. In a later section, I examine whether people showing these cognitions successfully lowered their test scores to avoid detection.

Realism. This category describes thoughts about providing realistic answers. Given the experimental context, realism goals could be viewed as a subset of caught goals, as instrumental for avoiding detection. I viewed the realism and caught goal categories as complementary and as having created similar types of motivation, but coded them separately due to methodological concerns about bias and reliability. Together, the categories represent a more exhaustive measure of the motivation to avoid detection, and they are aggregated in some later analyses. Another potential benefit of viewing this category separately is that realism might serve as one indication of the quality of faked answers that people create. Although not every deceptive message needs to be of high quality to produce a reward, it can be argued that the highest-quality messages, in theory, would appear real in every possible way, despite not being genuine. Examples of the 41 cognitions (in the total sample) coded as having realism goals are:

> (FG, 162, Citizenship): "I would say D again, not to sound like... to sound more realistic, but still sound like you care"
>
> (FG, 74, Impression): "that wouldn't be following the law but... it's not realistic to say not true so I'll say B very rare"
>
> (FR, 118, Emot. Stability): "well... they probably are gonna care about that. That affects just about everything. ... but we'll keep it semi-realistic."

(FR, 11, Leadership): "well, if I say more than 3... that just sounds way too good to be true so I'll say C, two."

(IA, 7, Leadership): "not too many because that would be really unnatural so D."

(IA, 36, Openness): "you wouldn't want to have any difficulty at all. So that would be 'very inaccurate,' but that doesn't really sound like a human. So we'll do 'moderately inaccurate'"

(FR, 152, Impression): "I'll probably answer C because... unless your Jesus... I highly doubt you haven't done that"

As an aside, not all estimates of realism were accurate. When asked about hours of volunteer work performed in high school, one person stated, "I think it would be impossible to do... more than 75 would be a lot. So... 31 to 75" (IA, 85, Citizenship). This person was evidently wrong since 8 respondents in the honest condition reported "more than 75." One participant even appeared to be overly influenced by the fake realistic instructions and disregarded a true maximal answer! The person's VP was given also in response to the question about hours volunteered:

(FR, 149, Citizenship): "Worked for a summer company, with my friend. It's more than 80 hours. Worked the whole summer. So... but I don't want to look unrealistic, even if it's the truth. So I'll go with D"

Unfortunately, most VPs offered few insights about how people determined what was realistic. For instance, this VP suggests that realism judgments occur outside of consciousness: "I guess for them [the employers], if I would want awards based on leadership, I wouldn't want to say more than three and not be believable but I... I'd go with... maybe three, because it's believable and to look good" (FR, 79, Leadership). Others, on the other hand, showed signs of at least pondering the issue:

(FR, 79, Leadership): "I guess I'd definitely want to be realistic here. So we'll say... leadership positions... How many leadership positions is someone usually offered? Maybe 2."

(FR, 39, Impression): "No, that's definitely not something... Now, do kids normally steal?... Be realistic. Do kids normally steal? # that's definitely something that I don't think... they want to hear though, even if it is normal. So, we'll say not true"

(FR, 152, Leadership): "Well, the instructions say... that they're looking for people that are responsible, cooperative, and helpful in a team. So... I think two or three might be kind of unrealistic because -- I mean, how many things do really volunteer for or are you involved in, in a year? So I think 'one' would be a good answer [...]"

As this last comment implies, normative responses may be valuable for judging realism. This notion is captured specifically in the next category of norm consideration.

A few people thought about how their overall pattern of responses might appear. Though the first two examples below may be unremarkable because they were from the fake realistic condition, the third occurred spontaneously.

(FR, 79, Agreeableness): "I would say moderately accurate because... I'm putting 'very accurate' for a lot of things and maybe it's going to start to sound... not believable"

(FR, 39, Citizenship): "I definitely have to be realistic about this because not too many high scores. I don't do volunteer all the time, but I am applying for a counselor. So... we'll say # to 30"

(FG, 155, Impression): "I say B... well... I don't know. I'm just thinking now like... what the person would think if you just say "I never do this", "I never do this." You would sound perfect. ..."

Another observation is that the realism codes were assigned more often for responses to two biodata items (about hours of volunteer work performed in high school and about leadership positions offered to the person) and one impression management item ("I have said something bad about a friend behind his/her back."). These items had 5, 6, and 7 VPs showing realism, respectively, whereas all other items (but one) had between 0 and 2 instances.

Situational Influences

Norm consideration. This category measured people who consider what kinds and levels of behaviors/events/characteristics are typical when answering items. As hinted at above, the consideration of norms often occurred in conjunction with caught and realism goals, but can be used by honest responders who want to be accurate and conscientious. Within the theoretical framework (Figure 4), this category best represents general "knowledge" that influences the strategies used, rather than direct motivation to deceive. Here are sample VPs for this category:

> (FR, 39, Citizenship): "we'll say once 'cause nobody really ever thinks about that"
>
> (FG, 87, Emot. Stability): "No. That's definitely another thing where I feel that's very inaccurate. I don't see myself as... feeling blue at all. Well... I mean, at times everybody has their moments but I definitely not often feel blue. So I would say very inaccurate"
>
> (FR, 42, Citizenship): "in high school I did a decent amount of volunteer work. ... Not many kids who wanted to become a gym teacher probably did a lot -- [it's] something that makes sense, I guess. So... '11 to 30' seems standard, in the middle"
>
> (IA, 26, Citizenship): "I think everyone in this state recycles. So, as often as possible. And I like to put my papers in that little [recycling] bin in the classroom. So I'm going to go with always"
>
> (FR, 84, Agreeableness): "I probably. I mean I know everybody probably has. So, I'm gonna say... neither inaccurate nor accurate because I know I have, but I don't really try to. So..."

Of the 95 VPs coded as showing norm considerations, 61% occurred with impression management questions. Sample VPs are:

> (IA, 16, Impression): "I think everybody does, a little bit. So I'll say F"
>
> (FG, 3, Impression): "well, if I have to. I think everyone lies at some point, but I don't know if that's good or bad in the workplace. ..."

(FG, 132, Impression): "Well, haven't we all? So we'll say somewhat true"

(FR, 152, Impression): "I'd probably say D for that one because... no one tells the truth 100% of the time. And if you're like 'No. That's completely false.' Then, they know you're... full of it."

Though norms were often used to assess realism and moderate answers downwards when necessary, some used them in more basic ways to achieve their primary goal of scoring highly to get the job.

(IA,29, Citizenship): "You want to have more than others, I would think. Show that you're dedicated, when you're not getting paid"

(FG, 155, Perseverance): "No... I mean, you still have to solve the problem, even if you're confused about it. So, but you don't want to spend too much time on it and waste time. So... I'd say...E, long period of time if you don't want to just be average"

As IM questions rely on fakers to select statistically deviant answers, the results here have obvious implications for the efficacy of the BIDR and similar IM scales, which are addressed in the discussion section.

Of course, an evaluation of normative behaviors is sometimes necessary just to answer certain questions. One think-aloud item had a scale ranging from "less than most people" to "more than most people." Since norm consideration is not necessarily related to deception, VPs for that item were not presented above (in this subsection). Also, the general instructions for the IPIP included this statement: "Describe yourself in relation to other people of the same sex and roughly the same age as you." Consequently, not all of the VPs coded for norm consideration described the deception on these items, since honest answers also needed to take normative behaviors into account. There is no simple way to distinguish when norm considerations were used to produce honest or deceptive responses just by looking at the VPs. So, these potential errors in coding were left in the

data because removing them would require creating a more complex category definition. Still, there did not seem to be too many of these kinds of errors among the 17 personality responses identified in this category. One example of the possible error is:

> (FG, 87, Conscientiousness): "That's definitely something I... I'm always neat and organized. So compared to others I would have to say that is very accurate, that I am always prepared"

In a more likely example of the error, a participant just expressed general satisfaction with this normative scale:

> (FR, 152): "I think this whole section is probably... really applicable if you're doing the sort of thing to be hired as an elementary schoolteacher, because you're relating yourself to others that's the same... like the same sex and roughly the same age as you, which could be your coworkers."

Test knowledge and leadership. One category indexed how often people showed an awareness or usage of the knowledge given to them by the experimenter, regarding job desirable and neutral qualities that were measured by the test. In part, test knowledge was developed as a more focused way of measuring the original job goal concept, in addition to job goal-revised. Test knowledge also overlaps with the job desirability category, but is more specific. Coding of this category was based primarily on keywords, but was done manually versus having a computer blindly count exact words without considering the context. Since this category is probably very reliable and straightforward enough that it warrants little explanation, I present just a few comments exemplifying more subtle demonstrations of the test knowledge.

> (FG, 131, Leadership): "... I'm looking at the [reminder] sheet: responsible hard-working, cooperative, helpful in a team. I'm going to go with once... but that again sounds bad to an employer. They'd probably want more leadership experience. So, I'm gonna go with C, twice"

> (FR, 129, Extraversion): "Thing [i.e., reminder sheet] says it's not important or unimportant. Probably gonna want... in the middle one"

(IA, 16, Extraversion): "I would say moderately accurate. It would be good to have an outgoing personality – again it's not the most important thing"

(IA, 98, SJ-most likely): "I am thinking D is the most responsible sounding to do. So, I'll put that."

(FR, 129, SJ-least likely): "You're gonna want to turn it in on time. So... *option A* No. *option B* *option C, D, E* Yeah, that one sounds good. *option F*... don't want you to be irresponsible. Probably *option C*"

Another category served an identical function, but was based on assumptions that leadership was valued by the organization. So, any mention of leadership as being important was counted. Though these codes were also based primarily on keywords, manual coding allowed certain kinds of statements to be ignored, such as when people were rereading parts of the question that contained the word "leadership." However, there are still potential errors in coding that resulted when people stated the word leadership, but did not imply that it was a job desirable characteristic.

As would be expected, most of the 23 cognitions identified were for biodata and SJT items about leadership experiences. Here are sample comments for this category:

(FG, 138, Leadership): "I guess...well, I have received awards, but not necessarily for leadership. But I guess you COULD say leadership was one of the qualities that they wanted. So I guess we'll go with like 'one'"

(IA, 69, Leadership): "well there's always recruitment for fraternity. You want to show you can... get people to join along with you, as a job applicants. Employers want people to show that they're leaders"

(IA, 36, Leadership): "that would show that leadership and responsibility. So I'm going to go with the most on that one"

Deception Strategies and Tactics

Job desirable responding (JDR). Codes for JDR were given whenever a participant attempted to discern whether the content of an item or the general

characteristic it reflected was important in the job or desired by the employer. As such, it included decisions based on test knowledge. It also included participants' judgments that a quality was neutral, undesirable, or irrelevant to the job, since judgments of irrelevance imply that a quality could not be desired by the employer. Examples are as follows:

> (FG, 120, Perseverance): "They probably want someone who is willing to work... a long time, like even if... I'm troubleshooting. And, I do like troubleshooting myself. So I'm going to say 'a long period of time'"
>
> (FR, 168, Perseverance): "that's good for hardworking. so... very long period of time"
>
> (FG, 80, Agreeableness): "well, to be a nurse, your whole life is other people's problems. So, I'll say very inaccurate"
>
> (IA, 69, Artistic): "they want this apparently. Just because you go to art exhibits, that doesn't mean you're good for the job"
>
> (IA, 148, Citizenship): "again, I don't know how much that would really have anything to do with being a counselor and being hired, unless they were looking for someone who's environmentally safe, but... I will go with... C, twice"
>
> (FR, 150, Impression): "[...] If you're like, 'Oh no. That's completely, it completely doesn't bias my thinking.' Obviously that's, that's a lie. So... I'm going for an answer that I think they would think is the correct answer. I would say C because it's not completely true but it's also not true.'

The agreement statistics already discussed show that this category was coded reliably. It is hoped that these prototypical comments show how this category can represent a meaningful variable when examined in later analyses.

One other observation is interesting. A few people took the test battery in a methodical, almost mechanical, manner. Each question was evaluated according to the qualities listed on the instruction reminder sheet (Appendix B) and then endorsed highly or lowly. Generally, cognitions across persons were not analyzed, but I present a

representative sample of the VPs for one participant (FR, 168) to demonstrate this finding.

> Question 10: "I'm looking at the reminder of instructions. A large extent"
> Q11: "that's good for hard-working. so... very long period of time"
> Q13: "look at the instruction sheet again. Probably... rarely"
> Q19: "that's # cooperative.... Twice"
> Q31: "that could be a characteristic of being helpful. So... more than three"
> Q41: "recycling shows that you're responsible. So... often"
> Q55: "that would probably bring the team morale down. So very inaccurate"
> Q57: "they don't want outgoing and extraverted. So, neither inaccurate nor accurate"
> Q93 (SJT): "*option A* I don't know if that's... probably... not. Ok. So, hardworking, responsible, cooperative, helpful... *option A*"
> Q126 (Impression): "that could hurt a team. So... C"

What is also interesting is that this method of responding persisted in the same way for all types of questions as shown above, where questions 1-45 were for biodata and 46-92 were for personality.

Socially desirable responding (SDR). SDR was coded much less frequently than JDR. In fact, SDR codes were not allowed to be given when the job was referenced. SDR cognitions were identified, but it is quite possible that at least a portion of these are closely overlapping with or even representative of JDR (e.g., IA, 29: "this would be a good quality to have. So E, five times or more" or FR, 168: "very inaccurate. Not good characteristic"). Below are sample VPs for the different types of questions.

> (FG, 162, Citizenship): "just because, to be a good person, I would say 'extremely likely.' I think this solution's generally a good characteristic"
>
> (FR, 79, Extraversion): "I would say moderately accurate because I don't want them to think that I always want to be the center of attention, because that's not a good quality. So..."
>
> (FR, 11, Citizenship): "I don't generally pick up trash when I see it unless it's right by the trash can. So I would say that... I would say somewhat likely, but someone might like to hear that I'm a better person than that. So, I'll say... very likely"

(IA, 29, Citizenship): "in accounting, you don't really deal with poor people. So... but it's nice to be generous sometimes, so maybe C"

(FG, 93, Agreeableness): "No. I kind of treat others the way I want to be treated. ... so, insulting's kind of an asshole thing. Sorry for my French"

(FG, 128, SJ-least likely): "bring a crossword puzzle because that kind of seems rude to the professor, to be working on puzzles while he's trying to teach other students and teach #. So... guess E"

(IA, 55, Impression): "not really. Stealing isn't... a good thing..."

(FR, 12, Impression): "Yeah, a couple of times in high school I did that...but I don't want to look bad. So I'll say somewhat not true"

Finally, here is an example of a comment that was coded as showing cognitions that were both attempts to assess JDR and SDR, with the respective portions underlined:

(FG, 20, Impression): "That's not really true. Usually I'm not too big on the whole revenge factor and it's probably what they're looking for too. Grudges aren't exactly something that's admirable in people so... not true, A. I'll go if not true, A."

Dimension assessment. This category was defined rather broadly and used to represent part of the E/I strategy, but the kappa statistics imply that it was clear enough. Some comments explicitly indicated that the person believed the item was measuring some broader characteristic, beyond just the item content. Sometimes these took the form of a question (e.g., "What is this item getting at?").

(IA, 158, Artistic): "It doesn't seem to be quite as related but they might be getting at your person qualities, to get a better idea at what kind of the person you are. So..."

(FG, 3, Citizenship): "Well I want to say... I don't know what this would have to do with me getting the job... it says maybe what kind of conscience I have and maybe what kind of interests I have, but it says nothing about how capable I am or would be of getting this job..."

(FG, 80, Artistic): "I've only attended one – that was for school – but five or more sounds like I'm more cultured"

(FR, 149, Impression: "I sometimes drive faster than the speed limit"): "I don't know if that really would play a part on me getting hired or not. ... But, I don't know. It can make you seem like a risk taker and I don't know if that would necessarily be a good thing, for a team. [...]"

People who focused on using their test knowledge were almost automatically classified in this category as well. Here are examples of these kinds of VPs:

(FG, 19, Leadership): "One of the qualities is to work, be helpful in a team and that's a pretty helpful thing to do, so I would say five times or more"

(FR, 168, Conscientiousness): "shows that you're responsible. So... very accurate, moderately accurate"

(IA, 148, SJ-most likely): "*option C* Well, especially if they're looking for someone hardworking. I don't even have to go on. That's probably the best answer for most likely to do."

As in the examples given to the raters in their coding instructions (Appendix G), many VPs not showing dimension assessment responded to the items based on their literal content. I present just a couple examples of contrasting pairs, where the first comment is more content specific.

Q: Take time out for others (Agreeableness)
(FR, 44): "I like to spend time with friends. I think it's very important. And it's nice to have other people do that for you too

(FG, 125): "Well they probably wanna know how well you'll be to – I don't know – put yourself aside. I guess helpful in the team is what they're looking for."

Q: I always declare everything at customs (Impression)
(FG, 93): "Like customs in an airport? Oh yeah. Of course I tell them what I have, even though I've never been outside of the country. But sure. So true"

(IA, 29): "You want a good honest person, working for you. Employers... would like to hear that you do."

Rule use. This category was determined to be of little use as a quantitative variable since there were few instances of it recorded and raters did not agree on these

instances. Part of the reason for this is that specific rules were not specified, making it difficult to identify the different kinds of rules participants used, as the examples below suggest. Still, these VPs provide some initial confirmation of the deception theory by showing that previously acquired knowledge can be generalized to select answers to new items.

Some of the "rules" were very simple and specified the general form of the answers that should be given (i.e., high/more, low/less, extreme, moderate, a particular answer choice):

> (IA, 36, Emot. Stability): "I don't think you'd want to be sad or anything like that. So I'm going to go with... well, moderately... I'll do very inaccurate. Look for extremes, to be better – it sounds a little impressive"
>
> (FR, 39, Conscientiousness): "That would be a... kind of being responsible. Eh, pretty much all of those. So that would be a good thing. So I wanna do the most extreme, very accurate"
>
> (IA, 36, Citizenship): "It shows caring, but I don't know how. Responsible or any of those # care for. So, I'm just going to go with – it'd show being a good person if you did it more often"
>
> (FG, 3, Leadership): "Well I guess I don't know. They didn't say anything about being leaders and they keep asking about leadership or being in charge... and I keep putting not the best answer but the second to best answer I would say C"
>
> (FR, 42, Extraversion): "I am the life of the party, but I don't want to tell them. So... when in doubt, go with 'C'"
>
> (FR, 152, Emot. Stability): "[...] so I'm going to say 'very inaccurate.' I tend to stay away from, well, I want to stay away from the extremes on the answers... but sometimes it's not possible"
>
> (FR, 118, Leadership): "Well, so far I've been saying 'one' on these type of questions. So I'll stick with one"
>
> (FR, 100, SJ-most likely): "There's a lot of options right there. ... Yeah, well, F and B look very similar, except one of them is 3 times as long. *option C* could be... #, depending on the situation, but the situation is VAGUE. So, what are you

> going to do? You're probably going to just *option F* A very complete answer, lengthy. I like it."

Or, the rule specified answer choices to eliminate from consideration. For example,

> (FR, 39, Extraversion): "I don't see how that really matters. So we'll say... I don't know if I want to say 'neither in accurate nor accurate' because that's probably not a good answer to put in a... survey. So we'll say... moderately inaccurate"

Another rule was slightly more general and indirect, describing how to create a substantive answer to the question:

> (FG, 138, Leadership): "Well, let's see. My examples: stats, stats... what else have I done? That FCC thing... and what else? ... I don't know. I think that's it. Ok, I guess we'll multiply that by 2. We'll go with 6. Ok, E"

Other rules resembled the reproduction strategy, but referred to items that were not identical. The examples below show how participants perceived certain items to be similar enough to previous ones that they should receive similar answers. As such, these response processes could arguably be called generalization instead of reproduction. Two examples are:

> (IA, 55, Leadership): "Probably about the same as the last question. They kind of ask the same thing. 3 or 4 times"

> (FG, 20, Artistic): "Another art question. What is the deal with the art questions? I put 'two' for the last one. I'll put 'two' for this one"

On a final note, some people had rules for responding based on the presence of extreme words in the item. For example:

> (FR, 150, "Have excellent ideas"): "Well, I think the word 'excellent' is subjective. I mean, obviously... you would think most of your ideas are good, but you don't want to be really... self-centered. So... I would say 'neither accurate nor inaccurate' because the word 'excellent' is kind of, kind of a strong word"

The use of this kind of rule was not coded in this category, however, because VPs did not clearly indicate that it was used to create a deceptive response; it is just as applicable to honest responding.

Demonstrate familiarity. This category was coded because it captured an interesting rule based on rather sophisticated logic, the purposeful selection of low answers to indicate familiarity with the content of an item rather than expertise. The rule can be applied in almost any situation, only requiring people to have a basic understanding of the response scale, not the question. Unfortunately, it occurred too infrequently to be informative in quantitative analyses; it is only explored in this section. The idea was explicitly recommended for deceptive training by one participant (IA, 7):

> "You may or may not want to fake know about certain analysis techniques or other things that could come back to haunt you. Don't pretend to be an expert, but familiar enough with things you've never even heard of."

Yet, three different people were coded as having thought about this rule/tactic. Their comments are presented below, with the latter two being stronger examples. Perhaps the most important thing to note is that no one in the fake realistic condition used it, demonstrating that even people told to be ideal or obtain the highest scores will moderate their answers for fear of being caught.

> (FG, 128, Leadership): "let's see. I guess... even if I wasn't offered any, I would... put I have. Just to show that I can be, or I am qualified to be leader. Two or three"

> (FG, 3, Bogus): "Again I have no idea even what that is... but it must be somewhat important. ... Oh man. ... If I say that I used it... and I haven't... I guess I could just figure out what it is, whenever I got the job. ... Lie about it and always sound that bad? ... I guess I would just indicate that I had used it before, period. Then I would say B, once."

> (IA, 148, Bogus): "Well, again, I'm not personally too familiar with that. So... if I was not knowing exactly what that is, you'd probably want to make it seem like

you have somewhat of experience, just in case that it has to be used at the job. So... I'll go with... once"

Excuses. Another interesting rule observed was the selection of an answer, for which an excuse could be made and given to the employer later, to avoid the consequences of having lied. Only two participants demonstrated this. As the examples do not really form a homogenous concept, they are presented only to show how people's deceptive answers can be based on sophisticated reasons.

> (IA, 70, Conscientiousness): "Yeah, I'm just going to go ahead and say that's moderately accurate, that I'm always prepared. That way... if something unexpected comes up, I can say, 'Well, I'm not always prepared, but I'm... moderately accurately always prepared.' And they can look back to this very application and say, 'Oh, oh. That checks out.'"

> (IA, 70, Bogus): "Well, I have no idea what that means. So, I'm just going to put... once. That way, if I make any mistakes on it, I'll say that I'm rusty"

> (IA, 70, Impression): "It's called... it's called... you know... you know... going with the flow of traffic as to not cause an accident. That's what it's called, and yes I do that. If the boss asks me about that, that's what I will say. But I don't speed all the time. ... I only speed when I'm late"

> (FG, 20, Bogus): "Another question that I do not know to what they're really asking. We'll go with once. That way, if they ask me, I can say I forgot"

> (FG, 20, Impression): "Well, ok. We'll go with... it's pretty much, it's true. I've done it before when I was younger, and you can always say that it's something that's in your past and that's something that's happened. We'll again, we'll go with... E. Sometimes honesty makes things look better anyway

Reproduction. By operationally defining instances of the reproduction strategy as a very specific behavior, comments exemplifying it were usually easy to identify. Seventeen comments were coded as evidence of this strategy, with 14 in reference to the repeated (agreeableness) item, given in practice and the test battery. Again, the 2nd to lowest response ("moderately inaccurate") was stated as being the best answer supported

by research. Prototypical comments coded as the reproduction strategy for this item are as follows:

> (FG, 120): "Well, you just recently told me that 2 is best worker. So, I'll just put that"
>
> (FG, 80): "We had that question earlier. So, if I say that that's 'moderately inaccurate,' I will sound more successful"
>
> (FG, 19): "I was told that '2' is the best, for any job. So I'm going to go with two"

For other items, comments in this category were stated like these examples:

> (FR, 78, Artistic): "Same question, same answer. Five or more"
>
> (FR, 159, Emot. Stability): "That's the same as the other one, and I put inaccurate for that one. So I'll put inaccurate for this one"
>
> (FG, 20, Openness): "Well this is the opposite of one of the other questions I just got. So we'll go with inaccurate. Guess # the opposite, same question."

These last two examples show that people treated some pairs of items as identical, even though they were not.

Interestingly, just 10 of the people coded as thinking about this strategy actually selected the "best answer" choice, with everyone being in either the fake good or ideal applicant condition. Also, these people represented exactly half of the fakers in the study who selected '2' as their answer. Although it was already established that thoughts in the VPs would not always correspond directly to people's final answer choices, this finding is noteworthy because some people rejected the "best" answer after recalling it, as demonstrated here:

> (FG, 155): "Well... if I want to do well, and most people that say about '2,' say they have a soft heart. I don't really...or... a harder heart or something. I still don't really agree with that, so I the teacher I'd say... probably like... I'd say four. So... E, very accurate."

Two other people coded for thinking about reproducing the best answer also selected answers in the opposite direction of the best one.

Taken together, these findings directly imply at least two things. First, the majority (10 of 14), but not all, of the people using the reproduction strategy selected the "best" answer to the repeated item, but an equal number of fakers selected the answer without showing evidence of this cognition. Second, most people in the experiment seem to have simply forgotten the best answer given to them, supporting memory as one aspect of deception capability (at least for this strategy).

To better understand the situation, I considered the possibility that fakers formed a more general rule (e.g., put a low number for this question) instead of encoding the exact best answer into memory. If this was common, fakers should produce lower scores than honest responders on average. The results support this possible scenario. The average scores of participants in the fake good and ideal applicant conditions for the repeated item were statistically lower ($p < .05$) than the average score in the honest condition. Still, it is not possible to rule out an alternative explanation, which is that "having a soft heart" would generally be seen as a liability in the workplace and be denied by any faker.

Concealment. Some cognitions and training item responses seemed to focus on not revealing certain pieces of information to the organization, rather than on selecting the best response. The concealment category is an attempt to identify these comments for the sake of completeness in describing the data, but this and other forms of suppression are not foci in this study. Below are samples of the 22 instances when it was coded.

(FG, 59, Impression): "No. Definitely don't want to tell them I do that. Not true"

(IA, 158, Emot. Stability): "Yeah, I guess I do tend to get stressed out usually, but I don't necessarily want them to know that"

(FG, 20, Conscientiousness): "That's exactly accurate [about forgetting to put things back in their place]... but we are gonna go with C because I don't think they want to know that about me right now"

(FR, 152, Impression): "Well, that's not really true anyway, but if you're applying for a job... in a school, that's probably not something you want to admit that you do, even if you do do it. So... you probably want to say not true to that one"

A few comments also addressed the need to conceal more generally, as shown here:

(FG, 58, Citizenship): "E, more than 3 times because... it shows that I would be kind and caring. ... It's very hard to describe your own personality, I'm finding because it still wants to come out when you answer the questions."

(IA, 148, Conscientiousness): "Well, I mean... counselors really kind of have to be more flexible than anything. I mean, I guess you could 'like order' and not really have it affect you, but... I don't know. I'm having difficulty answering these the way I really want to. It's kind of hard and... with some of these I'm definitely, I would be totally lying about..."

Not much more needs to be explained about this category in the context of this study, but establishing that suppression processes are involved and are, apparently, quite difficult to maintain for some fakers has direct implications for work on models of suppression and self-regulation in relation to deception (e.g., Vohs et al., 2005).

Research Question 1: Types of Strategies

Regarding RQ1 and the category-based results just presented, the verbal protocols provided strong evidence for the existence and use of the theorized reproduction and generalization strategies, but not the E/I or conditional randomization strategies. Unlike past test faking and coaching studies, the data reflect the use of different types of tactics that can still be grouped as a form of generalization, at least conceptually. Later analyses provide correlations between these tactics. The experiment provided people with the opportunity to use the reproduction strategy without encouraging it, and showed that

some people did use it. No evidence for conditional randomization was found, and possible reasons for this are included in the Discussion section. There was little evidence of the E/I strategies within individual protocols (i.e., responses to single items). However, some strong support for E/I could be found by looking across VPs within the same individual and by considering the admission of some participants that they *acted* like ideal job interviewees. Thus, I cover some additional results pertaining to the E/I strategy.

Original intentions to examine the E/I strategy as a combination of dimension assessment and either rule use or role taking had to be abandoned for practical and theoretical reasons. Role taking could not be coded reliably and occurred too infrequently. No VPs provided clear evidence of it, though some ambiguous ones reflected thinking about what the employer would consider to be a good test answer. As for rules, it became apparent that the *creation* of rules was difficult to distinguish from the mere usage of them because VPs were not interpreted across items within a single person to prevent biased interpretations, based on standard verbal protocol analysis techniques. There was definite evidence of rule usage, but it is not clear whether these rules were created prior to or during the test, especially since think-alouds were not required on all test questions. That is, mental processes were sampled rather than traced completely, for reasons already stated.

Despite these problems, other data provide reasonable evidence that E/I, as it was conceptualized, occurred on a broader level. Below are one participant's (FG, 3) VPs for all five biodata questions about artistic and cultural appreciation, for which no test knowledge was given. These thoughts show a clear progression of extrapolating meaning

from the broader set of questions (see underlined portions) to select new answers (an idea formulated and stated in Table 3, *prior* to conducting the experiment).

> Q9: "I don't... I have no idea why they would ask me this. I guess I'll just be honest about it and say B one to three times"
>
> Q15: "I really don't know why they would ask me these questions. I don't know. Maybe wanna know like personality-wise. Again, I guess I would just answer honestly. 'Cause I don't know why they would or wouldn't hire me because of what I answer on this question. So I guess I would just answer honestly and say more than most people"
>
> Q27: "I would... this is like a good question they asked about art and culture. Maybe it's more important than I thought it was. I would say a lot I guess because they keep asking me questions about it. So that means it is important to appreciate art, to value art and culture, although I don't know why, but obviously it means something so I would say B, a lot"
>
> Q34: "Why do they keep asking me these questions? I don't know what this would have to do with me getting a job. I guess I would say one... in the past year? Yeah, I would say one"
>
> Q36: "I guess I don't want to make myself out to be someone that's really into art and then when they meet me and try to have conversations with me about art, realize I don't know anything about it. So obviously that's somewhat important since they keep asking me questions about it. So I guess I'm somewhere in the middle... I would say see between three and five times"

Other evidence of role-taking was found rather frequently in response to the training items, even by people whose VPs did not demonstrate it. Some responses to this item were:

> (FR) Imagine or embody the role you are trying to portray.
>
> (IA) Try to imagine yourself as what or who you are trying to be/convey. Think what a perfect answer would be for the situation and then try to personalize it with a few details.
>
> (IA) Think outside of yourself. Imagine yourself as the most qualified and responsible person you know and answer the questions as this person.
>
> (FR) Imagine as though you were the one who was doing the hiring and ask yourself what you would want in an employee

One person did admit to using role-taking, but also explained that the need to think-aloud disrupted his ability to imagine being in a role:

> (FR) "In this experiment, I pretended to be applying for a teacher position. In order for my answers to be looked at in the best possible light, I answered the questions in a responsible, organized, but still honest way. Although I tried to fake most of these aspects, many of my actual characteristics were brought about because I had to talk out loud. If I just had to answer the questions by filling in the scantron, I believe it would have much easier for me to come across as a responsible and organized person who is qualified for the teaching position. It was difficult to imagine myself answering these questions as a teacher because I had to talk out loud, but for the most part, portraying myself as responsible and organized, but still honest was a good way to approach the experiment."

Regarding possible alternatives or additions to the strategies proposed in the deception theory, the large size of the dataset prevented any simple analysis to be performed that would directly address this issue. Based on efforts made to identify new coding categories with a general examination of the VPs and participants' training recommendations, new tactics were identified. However, they could either be placed into the existing theoretical framework (e.g., familiarity as a rule) or fell outside of the scope of this study (e.g., "excuses" designed to fool a target after the deceptive test response is detected).

Research Question 2: Strategy Prevalence

The second research question concerns prevalence rates for the strategies. Table 6 provides some relevant data. JDR thoughts occurred more frequently than other tactics. On average, fakers mentioned JDR in 20% of their VPs; the participant with the most JDR thoughts mentioned them for 90% of the think-aloud items. SDR thoughts occurred just 4% of the time on average, but in 25% of the VPs for the participant with the maximum. (Both variables were very negatively skewed.) Incidentally, the maximal case

for JDR was not the same person as for SDR, but other analyses showed that these dimensions correlated .45 in the sample. While rule use and creation could not be disentangled, any thoughts of applying general rules were explicitly stated much less so than JDR and SDR. However, it may be the case that extremely simple rules such as "select a high score whenever this word is encountered" may have been processed automatically and not stated aloud as often as it actually occurred.

As discussed, just 19% of the fakers thought about reproducing an answer, but not all of these instances pertained to the one item that was actually repeated. Furthermore, only 11% of the fakers demonstrating such cognitions also wrote the "best" answer. These results should not be taken as conclusive indicators of prevalence, however, given that participants were only given one opportunity to copy test answers. Since no single piece of evidence for the E/I strategy was convincing, it was difficult to estimate the prevalence rate in this study. Needless to say, this implies that E/I occurred much less frequently than the other strategies, unless one were to make the untenable assumption that most people excluded from the analyses at the outset were those who engaged in the acting/role-taking form of E/I and censored their thoughts about deception.

Table 7 shows the breakdown of the mean frequencies of each strategy variable by condition. Omnibus tests for ANOVAs performed on each of the strategy categories (i.e., JDR, SDR, rule use, and reproduction) revealed no significant differences between the faking conditions, except for realism, $F(2, 86) = 4.22$, $p < .05$. The pattern of results clearly shows that people in the fake realistic condition were more concerned about selecting realistic types and/or levels of answers. Of course, this "finding" is really best interpreted as a kind of manipulation check. In an exploratory sense, one other effect was

noticeable, but not supported statistically: reproduction thoughts were rarely found in the fake realistic condition.

Table 7, as well as Table 6, also reveals that there was a great deal of variance in the frequency of cognitions exhibited by participants. This finding is due in part to the fact that participants did not always think aloud every thought and/or verbalize information that was codable (i.e., descriptive of the response process). Additional explanations for variance in the use of different strategies are provided in the following section regarding peripheral category variables, including fakers' decisions to respond honestly despite faking the test overall and any inability to fake when test questions were not understood (primarily with the bogus items). Although this study was exploratory, the large degrees of variance observed imply that more refined methods of measuring deceptive (and honest) test taking cognitions and better classification rules of those cognitions would have been beneficial to this study.

Peripheral Category Variables: Situational Constraints and Potential Confounds

The following categories can be grouped in the sense that they focus on peripheral processes of or related to deception, with demonstrating familiarity perhaps being an exception. The *response patterns* category has to do with monitoring and adapting the use of strategies throughout the task. It seems to be a factor affecting strategy execution, though it was sometimes applied to determine single test answers. (This category is admittedly fuzzy, but it does seem to be distinguishable from the strategy components of deception.) *Uncertainty* describes situations when particular deception strategies are likely to occur, but is not deception in and of itself. *Honesty* describes the process that occurs when people prevent themselves from entering into deception after an assessment

Table 7

Mean Frequencies of Strategy-Relevant Categories By Faking Condition

	Fake good		Fake Realistic		Ideal Applicant	
Category	Mean	SD	Mean	SD	Mean	SD
Job goal	.65	.92	.75	1.56	.77	1.17
Caught goal	.74	1.84	1.00	1.96	.43	.82
Realism	.26	.58	.86	1.30	.30	.60
JDR	14.58	15.78	12.75	15.56	16.53	20.37
SDR	3.68	3.47	2.11	2.67	3.77	4.19
Rule use	.61	1.02	.79	1.45	.50	1.33
Reproduction	.23	.43	.07	.26	.27	.45

of image discrepancy or choose to end a deception process that was already initiated. *Memory* explains why fakers would select a less-than-ideal answer, but for reasons related to the truth. *Curvilinearity* refers mostly to characteristics of an item, regarding whether the "best" or most desirable response was not an extreme option on the response scale.

Response patterns (consistency and compensation). Nineteen items were coded as demonstrating a concern about how a *set* of responses would look. Most VPs focused on providing answers that would be consistent with previous ones. Note that the following examples do not provide enough information to demonstrate the use of the reproduction strategy. Given that none of these comments pertain to the one repeated test item, it is safe to assume that they should be coded in this category. Examples are provided here:

> (FR, 149, Leadership): "To be consistent, I'll go with 'three or four times,' like before"
>
> (FG, 162, Impression): "You're going to want to say 'very true.' so your answers sounds consistent"
>
> (IA, 70, Artistic): "Well, I'm having trouble with this question because it seems awfully repetitive with question number 7 and thusly I'm wondering why I'm being forced to answer essentially the same question twice. But [...] let's go with one to three times. I think that will be in accordance with my other answer."
>
> (FG, 3, Openness): "Well, I already answered a question about abstract ideas and... What did I say? I don't want to contradict myself. I said that [...]"

One person seemed to express frustration with the need to be consistent because previously selected answers to seemed too low later:

> (FR, 101, Artistic): "I don't know. We'll say... well, gotta be consistent, because I said zero on everything else. Damn. ..."

This finding, however, may not reflect typical testing behaviors if was caused by the fact that these participants were not allowed to revise old answers. Also, occasionally VPs

indicated that motivation for consistency was due to a fear of being caught faking. For instance:

> (FR, 118, Perseverance): "Well, they're definitely trying to catch me lying here. So, we're going to go back. ... I think we answered all of those with the second-best option. # And...yeah, one. So, we're going to go with the second-best option again."
>
> (IA, 70, Emot. Stability): "Since I said I feel blue sometimes on the other one. This is obviously trying to... # my ruse. It's a clever attempt to trick me -- make my answers appear to be inaccurate, conflicting..."

A smaller portion of the comments in this category could be viewed as reflecting the same type of thinking, but with a focus on the opposite end of the spectrum, compensation for previous answers. Here are some examples:

> (FR, 129, Artistic): "I'm going say once or twice. Since I put too many high ones"
>
> (FR, 152, 3rd Bogus item): "I don't even...what? I don't even know what that means. ... So... but 'according to publication standards' -- I feel like it's something that might be important. So I don't want to say 'never.' So, maybe 'once'? I'm not really sure, but I think I'm going to say once because on the last thing I really didn't know, I said 'never'... and I don't think I want to answer 'never' twice. So, I'm gonna answer 'once'"
>
> (FR, 79, Emot. Stability): "I'll say... moderately accurate to... switch it up so it doesn't sound like I'm... going too far"
>
> (FG, 20, Impression): "we'll go very true on that -- just so they know... I think sometimes most people drive over the speed limit and it's a flaw that people have, and if I point out some flaws here, maybe they won't look for more later."

Similar to this last comment, the one below shows that the respondent was considering how answers should generally appear. This cognition does seem to be related to rule usage, but clearly resides at a broader level of generalization; it seems to be a rule for applying more specific rules (e.g., deciding to select an extreme, once a specific rule indicated which end of the response continuum to aim for), or a kind of "meta-strategy."

(FG, 120, Conscientiousness): "Well, this is a personality test. So... they usually like... when you have strong opinions on certain things. So... and I am prepared most of the time. So, very accurate"

Uncertainty. As expected, most (92.1%) codes of uncertainty were assigned to VPs for the three bogus items. On average, participants in all conditions (including honest) stated being unable to understand the question 2.05 times, and only 12 participants (belonging to all conditions) never exhibited this thinking. Sample VPs with the bogus items are:

(H, 1): "I have no idea what that is so probably never"

(IA, 163): "I'm thinking about what this means. *question* I'm going to go in the middle. Well, I don't know what it asking me. So, again that's safe"

(FR, 3): "Well, I don't understand that question. I'll say never"

As for other types of questions, 6 people were confused by the impression management question: "I always declare everything at customs." Here are some examples.

(IA, 51): "Oh, I don't really understand that one. So... D"

(FR, 53): "Not really sure. ... *question* ... not really sure what that's asking... maybe somewhat true I guess."

(IA, 81): "I have no idea even what that means. So, I'm going to say somewhat true"

One person was confused by an extraversion item (i.e., "Keep in the background."):

(IQ, 72): "not really sure what that means. So, I don't know if that's good or bad. So I'll go with "neither inaccurate nor accurate" because I don't understand the question"

Although I did not attempt to conduct process tracing in this study, it was obvious and worth stating that uncertainty was almost always the first cognition mentioned in the VPs. This makes intuitive sense since a person cannot engage in a more direct faking strategy without understanding what is being asked. Furthermore, if the question is so

convoluted that even the true answer is not obvious, it would be theoretically impossible for people to lie. What is interesting about this category is that some comments provide evidence of conditional randomization. For the bogus items, some people gave responses that demonstrated uncertainty, an indication that the true answer should be never, and then a decision to put some other answer.

> (FG, 120): "I can't say I've ever done isometric analysis. I really have no idea. So, twice"
>
> (IA, 36): "I have no idea what that is. So I'm going to... well, truthful would be never, but since I don't know what it is, I'm just going to pick the middle number"
>
> (FG, 58): "I know what a matrix is. Don't know what solvency means. *question* that's probably an important thing to know for A job, I'm guessing. I don't want to say too many in case this is made up. ... I don't know whether to lie on this or -- of course, I'm lying"

The other clear implication that falls from these data is that generally confusing items will operate just like bogus items for would be-fakers. A complementary proposition is that participants with an abnormally small vocabulary or knowledge base may fake in erratic ways (as compared to honest or faking responders) when encountering standard personality items, if they fail to understand the question.

Honesty. Though thoughts of responding honestly really lie beyond the context of deception, they were coded to help explain instances when fakers intentionally selected less-than-ideal scores. Forty participants produced 70 VPs demonstrating this decision to select an honest answer instead of a deceptive one, as shown here:

> (FR, 78, Citizenship): "I really want this job and, again, you have to be honest. NJHS we were required to have... something like 70. So D"
>
> (IA, 136, Citizenship): "Always. Not always, often. I gotta be honest on that"

(FG, 3, Citizenship): "This is something that I probably would do it. I don't... I don't wanna put every answer exactly what they want here because they'll think I'm lying. so I guess I'll just put somewhat likely, cause that's kind of the truth and... So I don't sugarcoat it...."

(IA, 158, Leadership): "Well, you can't lie on that, but you do want to make it sound good. If you can think of anything that would qualify... but don't lie"

Other comments provide additional insights as to why people chose to answer honestly, despite being instructed to fake the test overall. The examples below were based on items perceived to be lacking job relevance:

(FG, 74, Artistic): "That doesn't really have to do with.. Speech Pathology so I'll be honest and say less than average. D."

(IA, 158, Artistic): "This one again doesn't seem necessarily directly related to what they are looking for this particular job, or in this particular field, or what kind of person I am. So I guess just be honest"

In other instances, it seems that honest answers were good enough (i.e., satisficing), such that participants did not bother to expend further effort enhancing their scores. Indirectly, this implies an evaluation of the discrepancy between one's actual and ideal states. When asked about hours of volunteer work performed, one person wrote: (FG, 120, Citizenship) "I think I... I think I got between 31 and 75. That works fine." The last phrase indicates a level of contentment with the truth, and it directly implies a decision not to distort the apparently true answer. In another example of discrepancy evaluation, the participant suddenly realized that the honest answer was different from the ideal answer, but still tried to maintain some degree of honesty:

(IA, 70, Emot. Stability): "Yeah, well who doesn't at my age for Christ's sake? All this stuff coming around #. Guess I'd say moderately accurate... job and stuff. #. No, that's honest. What do they want me to say?... They don't want someone who's stressed out easily. They want someone they can dump an enormous workload on -- have a big smile on their face. So, I'll say moderately accurate. That way it won't be a complete bald-faced lie"

Other people just seemed to prefer giving true responses at times. Two examples are provided:

> (IA, 139, Citizenship) "I consider this a lot. I think the environment's important. For this job, it doesn't seem to apply that much. So... but I still, the environment's important to me. So I'll pick 'five times or more.'"
>
> (FG, 155, Leadership): "Well, I just would probably say how many times I actually was. So it would probably be once or twice. So I'd say twice"

Also, a number of people made explicit concerns about being caught telling a lie. Note that such VPs would also be coded in the caught goal category, but are narrower because they specify the kind of action that needs to be taken – honest responding.

> (FR, 118, Conscientiousness): "Well, it's definitely not true. (I certainly feel bad about this whole... lying crap. And I'm gonna be in like, ... and then, I'm gonna be worried if they talk to me, you know... if they catch me. I don't want to worry about stuff like that.) so, I'll say that i'm never prepared. so let's, let's go with really not accurate or inaccurate. ... not a bald-faced lie"
>
> (FG, 94, Leadership): "and I wouldn't overexaggerate because if they try to look into it, they might see. So, I would probably be honest with that one. B, one"
>
> (FG, 2, Openness): "...Well I guess it's important to have a rich vocabulary, not necessarily use it but to have it to understand what other people are saying if they have better vocabulary that you. Well, I would say well I really don't want to lie about this. If you get the job and you don't have a really good vocabulary..."
>
> (FR, 152, Bogus): "I don't know what that is.... So... since I don't know what that is... I mean, I guess it's better to say never BECAUSE I don't know what it is than -- and just take the fall for one question -- than to say 'more than five times' and have that be something to do with like... I don't know... if it's something like the NRA ... I don't know. I think for this one, 'never' is probably a good answer because I don't know what it is. And I wouldn't say middle of the road because if it's something bad, then... then I don't want to say that I did it"

Memory. Just nine memory effects were found in the VPs. These were instances where people selected a low answer because they could not remember details that would support a higher, more desirable answer. This category could have served as a control variable for explaining why some people produce lower scores, whether responding

honestly or deceptively. However, it occurred too infrequently to be useful and, consequently, is not discussed further.

Curvinlinearity. Like memory, this category reflects honest responding rather than deceptive processes. However, the frequent observation of such thoughts in the VPs suggests that it would serve as an interesting control variable to assess when people lowered their scores, whether motivated deceptively or honestly, apart from explicit concerns about avoiding detection or being honest based on personal preferences.

> (IA, 36, Conscientiousness): "You gotta be with the good, the positive side. 'Accurate's' high. I'm going to go with 'moderate' because anyone who likes a lot of order probably wouldn't be that cooperative or responsible"
>
> (FR, 42, Artistic): "same thing as the rest of the stuff with art. I went to a lot so it seems like I'm interested. So... let's see. 'Five or more' – seems like I should be an art teacher. So, 'three or four'"
>
> (FG, 165, Agreeableness): "Well, that sometimes means that you're too easy... but it also means that... cooperative and you're helpful. No, I'm going to go with moderately accurate. #. It goes both ways, and that you are easy-going, but you're not too easy."
>
> (FR, 152, Agreeableness): I think I'm just going to say 'neither in accurate nor accurate' because... I mean, obviously you need to sympathize with others, but... you also don't want to take... sides, like if two people that you're with get in an argument, because that could be bad. That's kind of a tricky one. So I'm... I'm going to say C"
>
> (IA, 143, Impression): "Well, that goes back to the whole dangerous question that we had. ... I think they would like to prefer that you don't drive, but sometimes there are emergencies or you're late to work – and you don't want to be late to work either. So, I'm going to put... C, somewhat not true. Yeah."

The reasons for believing that the best answer is a moderate one varied quite a bit. Sometimes apparent reasons were specific to item content, as in the following examples:

> (FR, 152, "Take time out for others"): "I think moderately accurate is good because it's always good to take time out for others. I think it makes you feel good. I think it makes you learn more about yourself as a person. But at the same time, if you spend all of your energy helping out a colleague or one student in

particular, or anything like that, then you tend to put things aside that you need to get done for yourself. And you can't lose sight of that."

(FR, 101, "Don't mind being the center of attention"): "that's moderately accurate. Sometimes it's bad. Especially, if you're getting arrested. That's the center and that's bad"

(IA, 148, "When I was young, I sometimes stole things"): "Well, you don't really want to be known as someone that steals and is kind of delinquent behavior, and especially as a counselor. But, I don't know. That could also help... could help someone that you know, to show that like you went through... things that you could also relate to and help somebody with. ... So, I guess I can lean more towards be 'not true' side... so, I'll go with C. No, I'll go with B. just to make it seem like he might have done it once and then, you know, sort of relate to the kid by being like 'Yeah. I did it once, but I...'"

(FR, 150, "Use difficult words"): "Well... I would say moderately inaccurate because it's not like you want to, you're not completely illiterate. But there is no need... to show off. It's not like – you're an elementary school teacher – you need to... like impress other people that are trying to write like dissertations or anything. You're not a Harvard professor. You're with kids that say 'poo' all day. I mean, they're like... they're like little. So... I would probably say moderately inaccurate"

Sometimes participants believed that their answers would reflect arrogance or some other undesirable quality, even if that quality was not explicitly measured by the item. For example, for impression management item: "I sometimes try to get even rather than forgive and forget," a participant (IA, 36) stated, "I think... that it can't be a good thing when you're work as a group # cooperative, but you don't want to look like you don't have a backbone. So... I'm gonna go with C." Another examples is:

(FG, 58, "Have excellent ideas"): "Going to say moderately accurate because I don't want to sound... to think I'm boasting."

Comments based solely on concerns about being caught faking were not coded based on the decision rules (Appendix G), but cognitions of curvilinearity grounded in realism were coded, as shown below:

(IA, 143, Impression): "I think that's, yeah, I mean... I think even people who really try not to do that, still do to some extent. I mean, it depends on what you think [it is that] they're talking on your back. So it'd be unrealistic to say that you never do. But it's better that if, you don't. *question*. So, you're not going to put completely not true because that's not, unlikely. I'll put D"

(FR, 3, Impression): "I don't think they really want to see that on an application for a teaching job. ... but I'm not gonna say "not true," because I don't think anyone can say that honestly. So, we'll say B"

To conclude this section on the coded categories, at least some prototypical verbal protocols could be found in support of each category coded. It was found that people hold different types of goals, sometimes leading them to select maximal answers and sometimes moderate or even minimal ones. The reasons underlying these decisions also varied, but a portion of them was based on the avoidance of detection, even for people who were not given any kind of warning. If one accepts the categories as relatively accurate measures of their respective cognitions in the raw data, the relationships between different concepts can be investigated.

Correlates of Deception

For some analyses, aggregated categories were created because two categories seemed to impact deception similarly and represent aspects of a more general concept. Such was the case for caught goal and realism, and for test knowledge and leadership. Caught goal and realism ($r = .53, p < .01$) represented a certain type of motivation leading to lower item scores. Similarly, test knowledge and leadership ($r = .27, p < .05$) allowed people to understand the question clearly enough to apply generalization rules, even though leadership was not explicitly mentioned. The number of items overlapping for these category pairs was 9 and 7, respectively. To stay consistent with the dichotomous measurement of the other categories, a VP received a code of '1' if either

subcategory was mentioned and a '0' otherwise. The descriptive statistics for these aggregate categories are also included in Table 6 above.

Table 8 provides correlational data for the participants in the faking conditions (n = 89). Data were not analyzed within condition here because, although participants were given different goals, they were allowed to respond to the test in whatever way they wanted. Also included in the correlation table are the other measured scales. Note that correlations with the Big Five dimensions are based on participants pre-test scores, when the test was taken under honest conditions.

Research Question 3: Relationships Between Strategies and Goals

Results in Table 8 also provide answers to the third research question. Thoughts of reproducing the very best answer were statistically significantly related to job goal (r = .28), but not to goals about being caught or being realistic. Correlations between job desirable responding and job goal, caught goal, and realism, were substantial and statistically significant (.35, .38, & .39, respectively). On the other hand, socially desirable responding only produced a statistically significant correlation (.23) with realism goals. However, this correlation was not statistically different from those between SDR and job goals (.13) or caught goals (.17). Though occurring less frequently than

Table 8

Correlation Table for Coded Categories and Measured Scales of Covariates

Variables	1	2	3	4	5	6	7	8	9	10
Coded Categories										
1. Job goal-revised										
2. Caught goal	.17									
3. Realism	**.29**	**.53**								
4. Norms	.14	**.24**	.27							
5. Test knowledge	.05	.16	.12	-.01						
6. Leadership	.17	-.01	**.25**	.09	.27					
7. JDR	**.35**	**.38**	**.39**	**.21**	**.66**	**.44**				
8. SDR	.13	.17	**.23**	.15	**.29**	**.31**	**.45**			
9. Dimension assess.	.14	**.31**	**.32**	.09	**.85**	**.53**	**.85**	**.47**		
10. Rule use	**.37**	**.37**	**.41**	**.28**	**.28**	**.39**	**.40**	**.30**	**.41**	
11. Dem. familiarity	.11	**.36**	.15	-.01	.13	**.21**	**.34**	-.04	**.30**	.16
12. Reproduction	**.28**	.08	.04	.02	**.28**	-.02	**.25**	**.23**	**.22**	**.30**
13. Conceal	.11	.18	.17	.18	-.01	.02	**.24**	.12	.16	.09
14. Response patterns	**.36**	**.56**	**.40**	.10	.06	.02	.19	-.05	.11	**.49**
15. Uncertainty	**.27**	**.30**	**.34**	.19	.12	.17	**.35**	.14	**.28**	.19
16. Honesty	**.35**	**.52**	**.23**	**.27**	**.21**	-.01	**.30**	.20	**.29**	**.40**
17. Curvilinear	.17	**.46**	**.46**	**.44**	.16	**.28**	**.58**	.40	**.45**	**.47**
18. Memory	.06	-.02	-.07	**.21**	.10	.16	.13	.00	.13	-.06
19. Caught + Realism	**.26**	**.95**	**.75**	**.29**	.18	.12	**.47**	.20	**.38**	**.47**
20. Knowl. + Leader.	.06	.16	.13	-.01	**1.00**	**.32**	**.68**	**.30**	**.87**	**.29**
Measured Scales										
21. Conscientiousness	-.13	-.15	-.10	.02	.00	.15	-.09	-.06	.02	**-.22**
22. Extraversion	.11	**-.28**	.08	-.04	**.22**	.09	.14	**.25**	.17	.06
23. Agreeableness	-.09	.01	.05	-.01	**.22**	.06	.08	-.03	.14	.00
24. Emot. Stability	.11	-.03	.02	**.22**	.17	.18	.12	.11	.15	**.23**
25. Openness	.05	.03	-.03	-.06	**.26**	.00	.16	.07	.14	.01
26. Cognitive ability	.14	**.21**	.03	-.01	.00	.03	.00	.02	.03	**.24**
27. Decept. efficacy	.05	-.05	**-.24**	-.08	-.15	-.02	-.03	.08	-.11	-.11
28. Decept. Skill	.17	.00	-.14	.16	-.12	**-.27**	-.07	.16	-.16	.05
29. Test experience	.07	-.02	-.05	-.10	-.02	-.04	-.07	-.11	-.06	-.05
30. Knowl. of ideal others	-.13	-.15	**-.26**	.01	.02	.04	-.16	.01	-.07	-.15
31. Norms about lying	.17	-.10	-.05	.05	-.16	-.13	-.08	-.01	-.16	.02
32. Attitud. toward lying	**.21**	-.07	-.12	.05	-.17	-.12	-.04	.07	-.17	-.04

Note. *r*'s of .21 or greater are significant, $p < .05$ (two-tailed). *r*'s .27 or greater are significant, $p < .01$.. Significant correlations are bolded.

Table 8 (*cont.*)

Variables	11	12	13	14	15	16	17	18	19	20
12. Reproduction	.14									
13. Conceal	.07	.05								
14. Response patterns	.02	.15	.07							
15. Uncertainty	.06	.02	.17	.07						
16. Honesty	**.23**	.16	**.21**	**.27**	.18					
17. Curvilinear	**.24**	.06	**.51**	.19	**.31**	**.28**				
18. Memory	-.05	-.14	.01	-.10	**.21**	.13	.03			
19. Caught + Realism	**.37**	.10	.18	**.57**	**.36**	**.51**	**.52**	-.04		
20. Knowl. + Leader.	.14	**.27**	-.01	.06	.12	.20	.17	.11	.18	
Measured scales										
21. Conscientousness	-.07	-.05	-.07	**-.23**	-.04	-.06	-.13	.20	-.19	.00
22. Extraversion	-.14	**.28**	.03	-.09	.10	.01	-.07	-.17	-.17	**.22**
23. Agreeableness	-.01	.14	.02	.02	-.04	.08	-.04	-.01	.02	**.22**
24. Emot. Stability	.09	-.01	**-.24**	.06	-.02	.08	-.03	.00	.01	.18
25. Openness	.01	**.32**	.01	.03	-.07	.20	.05	-.08	.02	**.26**
26. Cognitive ability	.05	**.30**	-.05	**.23**	-.02	.16	.06	-.07	.20	.00
27. Decept. efficacy	.12	.18	-.06	-.12	-.12	-.06	-.08	-.19	-.11	-.15
28. Decept. Skill	-.15	**.24**	-.09	.06	-.03	.14	.02	-.12	-.04	-.13
29. Test experience	-.05	-.04	.14	-.07	.13	.00	-.05	-.08	-.03	-.02
30. Knowl. of ideal others	-.14	-.15	.06	**-.29**	-.09	.04	-.07	.04	**-.21**	.02
31. Norms about lying	-.20	.00	.13	-.04	.02	.09	.00	-.04	-.09	-.17
32. Attitud. toward lying	-.16	-.17	.08	-.06	.04	**.21**	.02	.08	-.10	-.18

Note. *r*'s of .21 or greater are significant, $p < .05$ (two-tailed). *r*'s .27 or greater are significant, $p < .01$. Significant correlations are bolded.

Table 8 (*cont.*)

Variables	21	22	23	24	25	26	27	28	29	30	31
22. Extraversion	.12										
23. Agreeableness	**.32**	**.37**									
24. Emot. Stability	.04	**.30**	.10								
25. Openness	**.21**	**.37**	**.25**	.08							
26. Cognitive ability	-.09	.05	.01	.12	**.26**						
27. Decept. efficacy	-.05	-.06	-.15	-.02	.12	.09					
28. Decept. Skill	**-.21**	.15	-.17	**.22**	**.28**	**.21**	**.33**				
29. Test experience	-.06	.03	.12	-.11	-.05	-.02	.07	.00			
30. Knowl. of ideal others	**.28**	.03	**.22**	.11	**.21**	-.13	.10	.08	.09		
31. Norms about lying	-.19	.00	**-.31**	-.14	-.05	.02	.12	**.35**	.11	-.19	
32. Attitud. toward lying	**-.22**	.03	**-.25**	-.03	-.02	-.17	.16	**.37**	.06	-.14	**.65**

Note. *r*'s of .21 or greater are significant, $p < .05$ (two-tailed). *r*'s .27 or greater are significant, $p < .01$. Significant correlations are bolded.

JDR and SDR, general rule use also produced relatively strong correlations with job (.37), realism (.37), and caught goals (.41). The specific rule of demonstrating familiarity, despite its low frequency, had a strong relationship with caught goal (r = .36), as might be expected. Incidentally, concealment produced small, but statistically nonsignificant correlations with the caught (.18) and realism (.17) goals, but explicit concealment cognitions were also stated infrequently.

Due to the sample sizes in each condition and low frequencies of various cognitions, it did not seem appropriate to conduct quantitative analyses of interactive effects. However, the various sets of condition-specific faking instructions provided goals to participants as well. Table 6 above also reveals that people held goals with relative frequencies that were consistent with the intended manipulations. All participants were focused on job goals, but participants in the fake realistic condition were focused much more on caught and realism goals. Interestingly, the average caught goal frequency was also high in the fake good condition, despite people being told to get as high a score as possible.

In any case, it is sensible to explore differences in correlational patterns between these manipulation conditions, while recognizing that the statistical power to detect most effects would be low (Aguinis, Beaty, Boik, & Pearce, 2005; Cohen 1988). For each condition, correlations between goals and strategies are shown in Table 9, excluding demonstrating familiarity due to its extremely low frequencies. For job goals, correlations with rule use and reproduction were clearly higher in the fake realistic condition, but only the former value was statistically greater than in the other conditions (r = .75 was greater than .17 and -.01, with $p < .01$, two-tailed). As for caught goal, it

Table 9

Correlations Between Goals and Strategies, By Condition

	Fake Good (n = 31)		
	Job Goal	Caught	Realism
JDR	.20	.56**	.34
SDR	.09	.15	.29
Rule use	.17	.35	.52**
Reproduction	.21	.12	.16
	Fake Realistic (n = 28)		
	Job Goal	Caught	Realism
JDR	.44*	.37	.64*
SDR	.14	.30	.52**
Rule use	.75**	.30	.40
Reproduction	.50**	-.07	.03
	Ideal Applicant (n = 30)		
	Job Goal	Caught	Realism
JDR	.40*	.37*	.36
SDR	.19	.28	.24
Rule use	-.01	.68**	.41*
Reproduction	.25	.43*	.21

* $p < .05$. ** $p < .01$

produced stronger positive correlations with rule use and reproduction in the ideal applicant condition, but neither correlation was statistically significantly greater than the other two. Comparisons of realism relationships across conditions did not reveal any striking differences, especially in light of the finding that caught goals and realism were strongly related.

Before addressing the next research question, I describe other results found in Table 8. To avoid redundancy, it is noted that all correlations equal to or greater than .21 are statistically significant at the $p < .05$ level. Ignoring correlations between the aggregated categories, the highest correlations of .85 in the table affirm my earlier qualitative observations that any mention of test knowledge (based primarily on keywords) was "almost automatically classified" as dimension assessment (as rated by three people) and that JDR "included decisions based on test knowledge." Since, JDR and test knowledge were correlated .66, these three variables seemed to capture the same event statistically, though conceptual distinctions can surely be made. Both dimension assessment and JDR were generally correlated with all other coded variables.

An interesting finding is that job goal ($r = .35$) was correlated with honest responding, despite all participants being in the faking condition. So, honesty seems to have been viewed as a partial means of eventually becoming hired. Yet, honest responding was not related to uncertainty ($r = .18$), which occurred primarily with the nonsensical bogus items. Despite not knowing whether the bogus items were referring to content that was desirable or undesirable, people still attempted to use JDR ($r = .35$ between uncertainty and JDR). Honesty was also related to a fear of being caught ($r = .52$).

In general, caught goal had similar relationships with the other categories as compared to job goal, despite the goals not covarying much in frequency (r = .17). One noticeable exception is that curvilinear assessments of items (when not mentioning a caught goal) were positively related to concerns about being caught (.46). This suggests that curvilinearity ratings may have been given when people were implicitly relying on caught goals. Caught goals were also associated more strongly with response patterns throughout the test (i.e., concerns about consistency and compensation). With personality, extraversion was negatively related to caught goals (r = -.28) and positively related to cognitive ability (r = .21). For the aggregate measure of caught goals plus realism, relationships were moderate to strong with most of the strategy variables except reproduction and concealment.

Test knowledge (plus leadership) and norm assessments were two ways for respondents to use information they already possessed to answer questions. Test knowledge was generally related to the deception strategy variables, and also to the personality dimensions (with correlations above .21 for extraversion, agreeableness, and openness). Conversely, norm assessments had only small, statistically significant correlations with JDR and rule use. Norm assessment was also related to curvilinear judgments, perhaps indicating that people use norms to determine what levels of a characteristic are "healthy."

As a final observation, the Big Five dimensions did not seem to be consistently or strongly related to the deception variables. If the coded variables are accurate and meaningful, this finding does not support existing propositions that faking is a form of

personality, when faking is viewed as a set of cognitions and behaviors instead of just motivation.

Research Question 4: Relationship Between Strategies and Deception Capability

For RQ4, an experimental measure of deception capability was created and analyzed. The efficacy component reflects people's views of their ability to deceive others specifically through tests, while the skill component is based mostly on successful attempts to deceive others in the past. However, this measure has not been validated, in large part because there is no standard for judging the *quality* of test deception. (Criteria used in lie detection studies are insufficient given the existence of the truth bias and research showing that no detection method is accurate at present.) Therefore, I first present some general results concerning the capability measure itself, which was collected online.

The reliability estimate of the efficacy subscale (α = .57) did not reach an acceptable level. Consequently, the correlational results shown in Table 8 are not discussed. As for deception skill (mean = 2.92, SD = .75), it was weakly related to cognitive ability, as researchers had proposed, but only weakly so (r = .21). Skill was negatively related to conscientiousness (r = -.21) and agreeableness (r = -.17), but positively related to openness (r = .28) and marginally to extraversion (r = .15). Perceptions of norms about lying and personal attitudes toward lying both produced the strongest correlations with deception skill (rs of .35 and .37, respectively), perhaps suggesting that practice increases skill more than a set of innate abilities.

In relation to the deception-relevant cognitions coded, skill only had weak statistically significant relationships with two variables. It was negatively correlated (-

.27) with assumptions that leadership was a job desirable characteristic and positively correlated (.24) with the reproduction strategy. The latter finding may be due to the influence of cognitive ability simply enabling people to remember the "best answer"; cognitive ability and the reproduction strategy were correlated .30. If cognitive ability is treated as a better or complementary measure of deception skill, the results show that it was related to thoughts about being caught faking (r = .21), rule use (r = .24), and concerns about consistent and compensatory response patterns (r = .23).

It is also important to consider the results for deception skill between the faking conditions since the act of spontaneously considering various goals and behaviors without being prompted may be a sign of skill (e.g., thinking about faking detection despite not being in the fake realistic condition). A one-way ANOVA indicated that mean values for deception skill did not differ across conditions, $F(2, 91) = .035$, *ns*. Similar results were produced for cognitive ability, $F(2, 83) = .60$, *ns*. Table 10 presents correlations for deception skill and (general cognitive) ability with relevant coded categories, by condition.

The condition sample sizes were too small to demonstrate that differences in correlations were statistically significant, but some general patterns emerged. Skill appeared to be related to the reproduction strategy only in the fake good condition, whereas ability seemed to show no differences between conditions. Also interesting is that skill had small negative relationships with caught and realism goals only, for conditions where participants were not told to be concerned about being detected. One direct interpretation of this pattern is that people self-reporting greater deception skill may not worry much about being caught. Another finding is that ability was strongly,

Table 10

Correlations Between Deception Skill, Cognitive Ability, and Coded Variables By Condition

	Skill			Ability		
	FG	FR	IA	FG	FR	IA
Job goal-revised	.38*	.21	-.04	-.12	.36	.12
Caught + Realism	-.09	.14	-.29	.16	.16	.26
Norms	.20	.11	.25	.12	-.16	.01
Knowl. + Leader.	-.26	.00	.13	-.11	-.11	.16
JDR	-.02	.04	-.20	-.08	-.10	.13
SDR	.09	.00	.27	-.24	.11	.24
Dimension assess.	-.25	-.10	-.14	-.05	-.12	.23
Rule use	-.08	.17	.03	.26	.04	.40*
Reproduction	.46**	.18	.06	.36	.31	.38*
Conceal	.06	-.24	-.37*	.16	-.42*	-.10
Response patterns	.05	.13	-.03	.36	.25	-.06
Uncertainty	-.06	.16	-.21	.06	.01	-.07
Honesty	.11	.42	-.08	.12	.34	.08
Curvilinear	.10	-.06	.01	.11	-.21	.34

Note. Sizes of the subsamples used to make these calculations ranged from 25 to 31. FG = fake good. FR = fake realistic. IA = ideal applicant.
* $p < .05$, ** $p < .01$

negatively related to concealment (r = -.42), but only in the fake realistic condition. To the extent that this condition allowed and promoted honest responding, this finding is consistent with propositions made by motivational theories (i.e., good applicants have less need to fake and hide the truth).

Out of curiosity, I examined whether the two fakers using the excuse tactic also reported having high skill levels for deception. For deception skill, participant #70 was in the 98th percentile and #20 was in the 87th percentile of the total sample completing the survey in Part I (N=192). In addition, both of these participants rated themselves extremely low on deception efficacy on tests; participant 70 was in the 2nd percentile and participant 20 was in the 9th percentile. Given that excuses were used as a "safety net" for fooling targets in-person after they detect poorly performed test deception, it seems reasonable that these participants would view themselves as having the general ability to fool people, but low efficacy for test deception.

I also examined relationships between skill and scores on the impression management and bogus item scales. Success at deceiving others on tests is partly determined by one's ability to avoid being caught. So, these relationships should be negative, except perhaps for the excuse-making behaviors observed. The correlation with impression management (-.11) was in the expected direction, but not statistically significant. The correlation with bogus items (.40, $p < .01$) was significant, but in the opposite direction.

Finally, I note that, even if deception capability was not deemed to be a meaningful concept, the verbal protocols revealed that self-perceptions of the ability to

lie still influenced test responses. Below are examples of VPs given in response to an impression management question (Q22: "I sometimes tell lies if I have to."):

> (FR, 60): "No. I'm a horrible liar and can't lie, even if I have to."
>
> (H, 63): "Usually white lies, so I don't have to hurt someone's feelings. But, in terms of important things, I am a horrible liar."
>
> (IA, 26): "I don't like to lie. I'm not really good at lying. ... But if I absolutely have to... I can."
>
> (H, 121): "I don't like to lie. I'm probably the worst liar in the world. So, even if I think I have to... I don't usually."

Research Question 5: Effectiveness of Strategies

This section addresses how various types of deception were related to the test scores people achieved. I first address the research question directly by correlating the strategic deception variables with scores on the "content" scales of the test battery (i.e., biodata, personality, and sum of situational judgment items). Note that scale scores were based on all items, not just the think-aloud items. The results (Table 11) indicate that the strategic deception variables, analyzed separately, generally produced small, nonsignificant correlations with the various scale scores. JDR, the most frequently measured type of cognition, was an example of this. The one statistically significant finding for rule use was that it was negatively related to Agreeableness scores (r = -.28, p < .01). SDR and reproduction, however, produced two moderate, positive correlations with the same biodata scales: perseverance and leadership, of about the same magnitude (rs all about .3, p < .01), despite SDR and reproduction not being strongly related to each other (r = .23, p < .05). Reproduction was also positively related to Emotional Stability scores (r = .25, p < .05). (All other coded category variables not included in Table 11 had statistically nonsignificant relationships with the scale scores.)

Table 11

Correlations For Faked Test Scores With Strategies and Other Categories

	Consc.	Extra.	Agree.	Emot.	Open.	Persev.	Art.	Citizen.	Lead.	Sit. Jud.
JDR	.14	.03	-.02	.18	.11	.20	.00	.02	.15	.02
SDR	.14	.06	-.10	.18	.09	.35**	-.01	.18	.31**	-.07
Rule use	-.05	.00	-.28**	.13	-.06	-.07	-.05	-.10	-.02	-.04
Reproduction	.07	.07	-.13	.25*	.02	.31**	.06	.03	.28**	-.03
Conceal	.04	.08	.01	-.07	.06	.14	.00	.04	.12	.02

Note. $n = 89$. The first five columns for personality dimensions (Conscientiousness, Extraversion, Agreeableness, Emotional Stability, and Openness) pertain to the posttest scores. The remaining columns are for the four biodata scales (Perseverance, Artistic/Cultural Appreciation, Citizenship, and Leadership) and summed situational judgment items.

* $p < .05$, ** $p < .01$

According to Figure 3, the analyses just described were aimed at linking observed deception (necessary) to observed scores at Layer 4. As McFarland (2000) and others have pointed out, it is important to examine the degree of faking that occurred by examining score changes within persons (from Layer 2 to Layer 4). In this scenario, score changes would indicate the degree to which the target's perception of image discrepancy was reduced, via test scores.

In this study, differences between honest and faked scores could only be computed for the personality items, which were taken honestly in Part I of the study. However, it is not safe to assume the corresponding pre- and post- test scales for personality are parallel or equivalent, since different items were used. So, even if the two versions of the personality test are both valid measures of the same constructs, differences in scaling (due to item content) could alter the conceptual meaning of difference scores. While difference scores cannot be computed in a meaningful manner, it is still possible to assess the effects of strategy use on test performance due to faking. To do this, I regressed faked posttest scores onto the strategy variables, after controlling for honest pre-test scores. Additional reasons for conducting such an analysis stem from other statistical problems associated with the use of difference scores (see Cronbach & Furby, 1970; Edwards & Cooper, 1990).

The results in Table 11 suggest that it is only worthwhile to conduct regressions on the Agreeableness and Emotional Stability dimension. Therefore, I conducted a separate hierarchical regression to predict each of the two posttest scales, entering the corresponding pretest score in step 1, and the strategy variables (minus concealment since it was not a variable of primary interest) in step 2. Results were similar for both

regressions; both steps explained incremental variance in both regressions. Consistent with Table 11, *t*-tests showed that rule use was the one statistically significant ($p < .05$) predictor of Agreeableness posttest scores, and that reproduction was the one significant predictor of Emotional Stability posttest scores ($p < .05$). These regression results show that the two strategies were not only related to people's final message outcomes, but to the degree to which they altered their appearance via test scores.

As with previous analyses, the statistical power to detect even moderate differences between correlations was low, but I made a qualitative examination of the correlations between strategies and scale scores, across the faking conditions. Generally, it seems that most of the non-null findings in Table 11 occurred only in the fake good condition. That is, correlations for the fake good condition were similar to those in Table 11, but larger in magnitude. The one statistical change was that the negative relationship between rule use and the Citizenship (biodata) scale was significant ($r = -.40$, $p < .05$); the positive relationship between SDR and Citizenship almost reached significance as well ($r = .36$, $p = .054$).

Research Question 6: Strategy Use and Question Type

Although the previous section presents results relevant to RQ6, concerning whether strategies are more effective on certain scales, I now discuss the propensity of people to use different strategies with different types of scales. I also examine the other types of cognitions witnessed to form a description of people's reactions not only to the content measures, but also to the detection measures (bogus and impression management items).

The results in Table 12 are broken down by scale and show the number of cognitions in each of the coded categories divided by the number of items within a scale. The number of items within each scale must be considered before comparing results between scales, since scales with a greater number of items provide people with more opportunities to exhibit various cognitions. In general, it appears that JDR occurred slightly less often with the biodata and bogus items. The same pattern seems to be true for SDR, except that it was used quite frequently for the situational judgment items. Rule use, on the other hand, seemed to be more frequent for biodata items. Finally, the reproduction strategy was found primarily for the agreeableness scale, as already discussed.

Regarding the other coded variables, people rarely stated their job goals during the personality items, but not much more often during the biodata. In contrast, job goals were noticeably more frequent on SJ items. Caught goals were mentioned with relatively the same frequency during the content measures, and more often with the faking measures, especially the three bogus items. One exception to this is that the leadership items also seemed to evoke caught goals. For concerns about realism, the most noticeable difference across measures is that realism was not mentioned during the bogus items.

Norm assessments were made considerably more often for the impression management items, and somewhat more often on biodata questions about artistic appreciation. Related to JDR, test knowledge was thought of quite frequently overall, but more so with the conscientiousness, extraversion, agreeableness, perseverance, citizenship, and leadership scales. These results require little interpretation given that all

participants were told that the hypothetical hiring organization valued corresponding traits. Given that emotional stability and intelligence (linked to openness) are both very desirable traits in general, these findings seem to be complementary to the SDR results since emotional stability and openness had slightly higher instances of SDR than the other dimensions, but lower instances of test knowledge. That leadership was mentioned most during the leadership items is only logical.

Dimension assessments occurred frequently, but somewhat less so with emotional stability and the bogus items; the latter finding was to be expected. Also expected were thoughts of uncertainty and confusion for the bogus items. A more interesting finding is that decisions to respond honestly instead of deceptively (as instructed) occurred more often for biodata than personality items. This finding was not predicted, but it was implied in the justification for using biodata items as a more objective and discrete measure of experiences than personality. However, honesty decisions were most frequent with the bogus items. Finally, curvilinear judgments had low, but non-negligible, frequencies on personality and biodata items.

Next, I address people's approach to the situational judgment items. For the most part, VPs indicated that respondents engaged in basic reasoning processes to determine the best solution for each dilemma presented. This typically involved attempting to understand the question and then evaluating each response option based on whether it would resolve the dilemma. Then people either impulsively selected the first acceptable option as their final answer or they narrowed down the acceptable choices based on various criteria until they were left with a single answer (to each of the item's most/least questions). Here are some examples:

Table 12

Number of Cognitions Per Item, For Each Coded Variable and Scale Type

	Personality						Biodata					
	Consc.	Extra.	Agree.	Emot.	Open.	Persev.	Art.	Citiz.	Lead.	SJ	Bogus	IM
Strategies												
JDR	19.6	23.3	19.9	17.6	23.6	15.6	17.2	14.8	15.2	19.6	11.7	17.2
SDR	3.4	2.3	3.1	4.6	5.9	1.2	1.0	3.0	1.4	9.6	2.7	6.3
Rule use	.4	.4	1.1	.7	.3	.4	2.4	1.8	1.4	.0	1.7	.2
Reproduction	.0	.0	2.0	.1	.1	.0	.2	.0	.0	.0	.0	.0
Conceal	.3	.1	.0	.6	.0	.2	.0	.4	.0	.2	.0	1.1
Other Categories												
Job goal-revised	1.3	.0	.4	.0	.0	1.8	.2	.6	.4	3.8	2.0	1.3
Caught goal	1.1	.1	.3	.4	.4	.8	.4	.8	2.0	.2	4.7	1.8
Realism	.6	.0	.1	.1	.3	.2	.4	1.4	2.4	.0	.0	1.1
Norms	.4	.1	.3	1.0	.6	.4	2.2	1.4	.0	.0	.0	5.8
Test knowledge	7.9	9.7	6.0	3.3	3.1	7.6	1.8	5.0	7.2	9.6	1.3	4.1
Leadership	.0	.6	.0	.0	.0	.2	.0	.2	2.8	.6	.0	.0
Dimension assess.	11.1	16.9	11.9	5.7	9.1	9.6	7.4	7.8	10.6	17.0	1.7	8.9
Resp. patterns	.1	.0	.0	.3	.4	.2	1.0	.0	.6	.0	.3	.3
Uncertainty	.1	.1	.6	.0	.0	.6	.0	.4	.2	.0	57.3	.8
Honesty	.1	.4	.0	.1	.4	.2	1.4	1.8	1.8	1.6	4.3	1.4
Curvilinear	2.1	3.4	2.7	.9	1.9	2.4	2.8	1.8	2.6	.0	.0	1.9

(IA, 16, SJ-most): "Well. *question* Let's see. A is a good idea. B is good. C you shouldn't do without their permission. D is not good. *option E*. I don't know. I don't think E. A, B, or F, I guess... would be the best answer. It's kind of a combination of the two."

In this way, VPs of fakers closely resembled those of honest people because people identified certain answers as being more objectively "correct" than others. That is, some answers seemed to solve the presented problem better than others. Since participants did not know that an empirical scoring key would be used, it is possible that some of the best looking options were actually scored less than the maximum.

In some comments, it is also apparent that the format of situational judgment questions can reduce transparency, and presumably, increase difficulty for fakers. Here are some excerpts from the VPs:

FG, P3, Q93: "Well, probably.... There's just like so many right answers. This is hard. They probably want someone who could find the cause of the problem and fix it, but... oh man. Or, do they want someone who would take care of the person that was being mistreated right away? I guess... Oh man. I guess I would say... I would say maybe inform the authorities, but I don't know if they want someone who's going to always run to the boss with the problem or someone who would just fix it themselves. I would say B."

IA, P7, Q97: "Well I don't know. Either A or B. I can't decide which is more interesting. Staying awake.... We'll say B because A's too much work."

FR, P11, Q93: "Probably A *option A*. I would be least likely to... not really liking these options. I would be least likely to inform.... Nope, this is hard. I think actually that I would most likely to *option E*"

FR, P53, Q93 & 94: "... I could see myself probably doing any of these. ... It's hard... hard to decide. I probably would do a mix of all these different things."

FR, P152, Q97: "All the other options [besides E] are pretty decent options. That's really hard because..."

Still, some people believed the answer to be obvious (even when they were not actually correct in their assumptions):

(IA, 29, SJ-least): "obviously, *option E*. Businesses want you to be occupied on the task at hand, not something else"

(IA, 7, SJ-least): "so definitely skipping [the class] because skipping would not be... sound good"

(FG, 20, SJ-least): "least likely to do with the crosswords. Of course, it looks bad. So I wouldn't do it "

From a test designer's point of view, it is also noteworthy that participants sometimes viewed response options as being nested within others. At other times, they ignored subtle differences between response options and viewed them as equivalent, although the overlap in content was not complete. Both phenomena imply that respondents may answer in ways that the test designer would not expect, when response options are constructed to be independent. See the sample comments below:

(FG, 74): "*option A* because that kind of includes all of the other answers."

(IA, P85): "Oh wait... can I do all of these things? Well, I would do like... I'd do *option A* which would be... all of the things, help the person that was mistreated and inform the appropriate people and.... Yeah, all of them."

(IA, P8): "I think the most... It's a lot of things on here that I would do, but if I had to pick one I would probably... *option F*. I think just F because it covers more things than the other ones do."

As for the impression management items, key findings were already covered. Here, I note a few other results. First, people sometimes complained in their VPs that the items were (for lack of better terms) presumptive or were soliciting answers that were conditional upon parts of a question being true. For example, the question: "I sometimes tell lies if I have to" lead to these statements:

(H, 32): "Why would you have to? ... I guess... I guess I, everybody kind of does, sometimes, maybe."

(FR, 41): "No. I do not lie. I got that that cannot... stand liars. It really irritates me. And I don't think you ever have to lie about something."

(IA, 140): "Well... I don't think in this job you would have any... you wouldn't, you really shouldn't lie about the important things, and honesty would be very important. So, I can't really think of why you would have to tell a lie. ... So, I would say not true"

In a second, related issue, a few items were just not applicable to some participants in this sample. For instance, the question: "I always declare everything at customs" assumes that people traveled outside of the country and had items that should be declared.

(FG, 19): "I've never had to take anything that [I had to] like to declare so I would put very true"

(FG, 49): "I've never had to go through customs, but I would figure if I did, I would, because I don't feel like getting caught"

(FG, 132): "I've never been through customs. So, we'll just say true."

Furthermore, a considerable proportion of the sample had only been through customs once or twice, allowing them to select a high answer while still being honest.

Third, the data show that some people did make normative assessments before responding to the IM items. Though it did not reach statistical significance, a negative correlation (-.16) between norms and impression management scores implies that people were occasionally able to avoid being detected by these items. However, the VPs also revealed that people were often unsure about how much to adjust their initial answer after assessing norms. The majority of comments indicated that people adjusted their answer by just one or two points on the 7-point response scale, causing them to still be detected by the impression management scale a fair portion of the time.

Lastly, I discuss the bogus items. Again, most VPs began with some indication (measured by the "uncertainty" category) that the respondent did not understand the

question. The decisions made after this admission of uncertainty varied quite a bit for fakers, but usually resulted in correct ("none" or "never") answers for people in the honest condition. Some fakers also disengaged from the deception process with these items, which could have resulted from a fear of being caught or from the inability to perform JDR or generalize some other rule.

Other people seemed to guess or assume that the item reflected something positive/negative and put extremely high/low answers. Others applied general test taking rules such as "when in doubt, go with 'C'" (i.e., the moderate answer). Just a few people demonstrated a pattern of thinking consistent with the rationale underlying the use of bogus items, as captured in the following comment:

> (FG, 58, Bogus): "not really sure what that is, but it sounds like something you should know, to have a job. So... I don't want to say 5 times or more, in case that's NOT something you want, that's not positive. So I'll say, D, 3 or 4 times"

Some participants first attempted to discern the meaning of a question by examining its component words. Due to the fictitious nature of the items, this strategy was necessarily ineffective. Yet, the VPs show how people's vocabulary level and general reading comprehension affected their interpretation of these items. To be more specific, the first question contained a meaningless abbreviation (i.e., AJMR), but made sense otherwise. The latter two questions made references to "isometric analysis" and "solvency files," respectively. A number of respondents were able to identify the general meaning of "isometric" (i.e., having equal measurement), or part of it, despite not knowing its definition and despite the contextual words providing no clues.

> (FR, 149): "Isometric. Okay, so 'metric' is probably having something to do with measuring. 'Iso,' isotope - that's something to do with protons. I don't remember."

> (FG, 122): "I would say... twice? I believe that isometric analysis is some type of... math term, and I've taken a lots of math classes. So I'm sure I've done that, at least a few times. So, I would say twice."

One participant even knew of the word as a term from kinesiology, which is synonymous with the concept of static tension.

> (IA, 13): "Isometric analysis. I'm not exactly sure what that is. I know it's from isometric as far as... in the subject of kinesiology, but... I've never heard that terminology before.... 'Isometric is'... movement without changing limbs I believe"

As for the term "solvency," it is apparently part of the language of formal (speech) debates. However, people often assumed its meaning was rooted in the word "to solve." The last general observation was that few people ever stated a suspicion that the bogus items were designed to catch lies, including people in the fake realistic condition who were warned of a different detection method.

Supplemental Results for Differences in Scale Scores Between Conditions

Because they have been the focus of much past research, I present basic results regarding differences in test scores between the manipulation conditions. These results (in Table 13) are not directly pertinent to this study's research questions because they primarily link motivation and test outcomes while largely ignoring the process variables. However, people in the fake realistic condition were given information about the detection mechanism (i.e., it searches for a broader pattern of unrealistic answers), in addition to motivational goals. To a lesser degree, ideal applicants were informed to produce test results that resemble those produced by "an actual person." These manipulations were intended to be strong, but they did not force participants to act in a particular manner. Instead, the instructions gave participants a certain degree of discretion, while encouraging them to *consider* faking in non-maximal ways.

When evaluating the statistics in Table 13, the reader should take into account the results in Table 4 showing that the scales seemed to have differential reliabilities (i.e., for conscientiousness, extraversion, and emotional stability). However, I only discuss the results as they are contained in Table 13. The results show that mean pretest scores did not differ across conditions. A corresponding MANOVA predicting all pretest personality scores produced a nonsignificant result for condition. In contrast, posttest personality scores did differ across the conditions, as should be expected if the manipulations worked at all. Consistent with past research and intuition, the fake good and ideal applicant manipulations generally lead to higher personality scores than the fake realistic condition, but the fake realistic condition scores were still higher than the honest condition scores. The same pattern was found for biodata and situational judgment. Cohen's effect size values, *d,* show that these results are consistent with, but more descriptive of, meta-analytic effect size estimates of faking in Viswesvaran and Ones (1999). The *d*-values also indicate that the five situational judgment questions were not more difficult to fake, relative to the other measures.

For the bogus items, the faking conditions produced higher scores. The mean score for the fake good condition did not differ greatly from the mean of the fake realistic condition, but both of those were considerably lower than ideal applicant scores. Thus, the "ideal applicants" would have been more likely to be caught. Differences between the faking conditions were less prominent for the impression management scale; the results do not indicate that faking realistic by these particular participants would lead to low detection rates.

Table 13

Comparisons of Scale Scores Between Conditions

	Honest		Fake good			Fake Realistic			Ideal Applicant		
Scale	Mean	SD	Mean	SD	*d*	Mean	SD	*d*	Mean	SD	*d*
Part I - Pretest											
Conscientiousness	3.41	.70	3.42	.72		3.48	.66		3.56	.63	
Extraversion	3.42	.78	3.38	.67		3.51	.59		3.53	.60	
Agreeableness	4.00	.54	4.03	.48		4.16	.41		3.97	.55	
Emotional Stability	2.95	.66	2.96	.69		3.25	.72		3.22	.58	
Openness	3.57	.45	3.78	.52		3.67	.45		3.73	.42	
Part II Measures											
Conscientiousness	3.64	.63	4.53	.42	1.66	4.26	.48	1.10	4.63	.35	1.92
Extraversion	3.50	.82	3.97	.43	.72	3.70	.42	.30	4.04	.44	.81
Agreeableness	4.09	.53	4.37	.37	.61	4.23	.47	.28	4.31	.54	.41
Emotional Stability	3.33	.69	4.18	.46	1.45	4.03	.41	1.22	4.24	.50	1.50
Openness	3.63	.68	4.24	.48	1.04	3.85	.53	.36	4.27	.48	1.08
Perseverance	3.75	.45	4.47	.38	1.73	4.12	.30	.96	4.49	.30	1.92
Artistic Appreciation	2.82	.84	3.60	.76	.97	3.25	.74	.54	3.62	.81	.97
Citizenship	3.00	.56	4.09	.54	1.98	3.79	.45	1.55	4.01	.57	1.79
Leadership	2.75	.77	3.93	.75	1.55	3.38	.55	.94	4.08	.52	2.01
Situational Judgment	.82	.58	1.27	.30	.97	1.11	.48	.54	1.18	.37	.73
Bogus Items	1.29	.48	2.39	.98	1.43	2.24	.86	1.38	2.93	1.35	1.65
Impression Management	4.27	.80	5.48	.86	1.46	5.21	.69	1.26	5.59	.68	1.77

Note. Sample sizes for the conditions are: 33 for honest and fake good, 31 for fake realistic, and 30 for ideal applicant. *d* = standardized mean difference from honest condition in Part II, with positive values meaning scores were higher in the given faking condition.

DISCUSSION

To summarize the aims of this research, I attempted to build a theoretical framework that adequately describes the motivational, cognitive, and behavioral components of the process of deception. Research on deceptive techniques, test faking, and impression management tactics in interviews suggested that people perform a variety of actions when attempting to influence others, but rarely placed them into a causal framework or tied them to general psychological constructs. In an attempt to integrate existing theoretical assumptions and propositions and to describe past and future empirical findings, I proposed a theory focused on the selection and implementation of various deception strategies.

At present, a strictly confirmatory test of the overall sequence of various components in the theory (in Figure 3) would be premature, due to the present dearth of process-based research on deception and accurate measures of deception behaviors, message quality, and overall effectiveness. Therefore, I collected and analyzed direct self-reports of the process-relevant cognitions that people experience, prior to and during the creation of a lie or misrepresentative test answer. Analyses examined whether cognitions exemplifying the theorized deception strategies could be identified. Analyses also explored the relationship between certain types of cognitions with each other, situational influences, the potential correlates of cognitive ability and deception skill, and proximal message outcomes.

Summary of Findings

At the most general level of observation and interpretation, the total set of results showed that deception can be a complex process. In many cases, people not only

attempted to reduce the discrepancy perceived by a hypothetical target (i.e., hiring organization), but also acted to avoid detection despite not being manipulated to hold caught goals. The results also showed that the sample, as a whole, used a variety of strategies and tactics to perform deception. Furthermore, at least some participants demonstrated an awareness of certain complexities involved with creating an appropriate deceptive response (e.g., recognizing that the bogus items failed to indicate which end of the response scale should be endorsed or acknowledging the possibility that certain responses would be implausible or not normative).

Yet, the aggregate evidence also suggests that most participants are naïve in their understanding of psychological tests and the principles driving those tests. Despite being informed that certain characteristics were measured by the test items in general, no participant in any condition stated that items were purposefully grouped into scales. In many cases, participants treated each question independently and provided answers specific to the literal content of the item. Others even complained about the degree of redundancy on the test due to similar items being asked.

These observations have direct implications for how tests should be designed, with respect to concerns about faking. For instance, Paulhus (1991) wishes test respondents to answer items based only on their specific content, treating answers influenced by anything else as biased. Yet, this logic is inconsistent with general psychometric theory, in which constructs are of central interest and variance specific to individual items is treated as error. In a related matter, McFarland et al. (2002) found that faking effects were larger when items measuring the same construct were presented together instead of randomly among items measuring other constructs, and that tests

ordering items randomly had similar factor structures when answered honestly or deceptively. Verbal protocols in this study suggest that, although test takers may come to develop a deeper understanding of the test with additional knowledge or practice, they may be influenced more by similarities in the appearance, content, and face validity of items.

Specifically in regard to research question 1, at least some evidence for three of the offensive strategies was found in the verbal protocols and/or training recommendations, as well as for the defensive strategy of concealment. Selective presentation did not occur since no one verbalized a decision to skip a question and little data were found to be missing. Also, verbal protocols did not support directly the conditional randomization strategy, though it may have occurred with the bogus items, given the frequent admissions of uncertainty stated by participants who still "faked" their answer. However, the basic act of guessing (without first eliminating the truthful answer) could have produced the same result and remains a plausible alternative explanation.

Evidence for the reproduction strategy was clearest since the experimental design provided just one opportunity for its legitimate occurrence, but the data also showed that this strategy was applied to other questions, for which the correct answer was not already known. Also, people who recognized the opportunity to copy a test answer (i.e., cheat) did not always take it. Taken together, these findings imply that strategy choice and execution can be important determinants of deception outcomes, while being somewhat independent. Moreover, that the majority of people failed to acknowledge any knowledge of the "best" answer given to them earlier in the experiment strongly suggests that memory is a critical aspect of the ability to perform this deception strategy, as one might

expect. This finding is probably likely to have a large effect in situations when fakers/cheaters attempt to memorize a long list of test answers.

Verbal protocols most often indicated the occurrence of JDR processes. Many JDR judgments were made using test knowledge received during the experimental instructions. Some people applied this knowledge in a strict, problem-solving manner (see the Results subsection on JDR), demonstrating that JDR can be used directly as a rule, and therefore, as a form of generalization. Unfortunately, VPs of other participants did not reveal how judgments about an item's desirability were made. In a similar manner, the data showed that people engaged in SDR as well, though less frequently. However, the SDR category may not have been as accurate as JDR. SDR codes were only assigned for verbal protocols that did not mention the job explicitly (e.g., "that's a good quality"), but the participants stating those thoughts may have still involved desirability in the context of the hiring organization.

The data also showed that other kinds of rules were used (see Appendix F for a summary of the different kinds observed). Rules referred to different types of knowledge and situations (e.g., questions measuring certain dimensions). Some rules involved choosing the middle answer while others involved picking either the desirable or undesirable extreme. Although the heterogeneity found in this category is useful in an exploratory sense – suggesting tactics that can be examined in future work, it limited the types of statistical analyses that could be conducted using this category.

Evidence for the E/I strategies was found, but it was sparse and mostly indirect. The verbal protocols themselves did not reveal clear instances of role-taking, at least not ones reaching sufficient levels of agreement. In some sense, this is surprising since

people in the ideal applicant condition were told to look like an "actual person." No VP indicated an attempt to recreate the profile of someone the participant knew, which also precluded meaningful analyses with the "knowledge of ideal others" scale. Such a strategy would be interesting to explore, even if not occurring naturally, because it may be effective particularly if impossible or unlikely virtues are indicators of deception. It may also be useful for training people to use deception in certain situations (e.g., acting). In fact, the role-taking strategy was mentioned rather frequently by participants as a recommendation for deception training (in Appendix E), in the form of adopting the perspectives and/or enacting the roles of specific people (e.g., ranging from the employer to one's grandparents).

In all, the different types of responding found were consistent with previous empirical findings and theoretical propositions. As stated in the introduction, Kluger and Colella (1991) and Kroger and Turnbull (1975) described a similar range of behaviors and processes. The results are also interesting to view in light of Whyte's (1956) practical recommendations for "cheating" on personality tests. His recommendations were numerous, very specific in some cases, and of questionable effectiveness, but they captured the general variety of faking possibilities. They included aiming for scores between the 40th and 60th percentiles, demonstrating certain societal values (e.g., respecting one's parents, commitment to an employer), supporting the status quo, showing an indifference to "books or music," "stay[ing] in character," answering consistently across test questions, and matching personal characteristics to ones desired by the employer. Whyte even addressed cheating on situational judgment measures (of personality), but admitted that his own suggestions to "not split hairs" and to not think

too much were of limited value: "On this type of question, let me confess that I can be of very little help to the reader" (p. 454).

The theoretical framework proposed in this paper was not designed to supplant such prior work. It was created to organize existing propositions into a unified, theory-based description of deception. In terms of an empirical contribution, this study moves far beyond Robie et al.'s (2006) study of verbal protocols, which represents one of the only studies to examine the thoughts and behaviors of fakers directly, rather than inferring them based on experimental manipulations and test outcomes. From the data, it can be concluded that faking, and deception more broadly, needs to be examined more carefully. People are driven by different motivational factors and perform different kinds of behaviors, even if they tend to use certain strategies (i.e., generalization) more than others that exist in the realm of theoretical possibilities (e.g., E/I strategies). Apart from results concerning the relationships between strategies and outcomes, this study demonstrates that the form of "deception" that people plan and execute is variable. As such, researchers and practitioners who attempt to detect or measure it must have a sense of what forms of deception they expect to observe, either directly or through message outcomes.

In response to the second research question, the results provided some indication of the relative prevalence of each strategy to each other, but exact rates were difficult to calculate in most cases. For instance, people were only given one opportunity to apply the reproduction strategy appropriately. That few people actually used it provides some indication of what its natural prevalence would be, but this single finding is far from conclusive. Also, the design feature of the experiment (i.e., providing test knowledge,

simulating a job selection process instead of using tangible rewards) could have altered the prevalence rates as they would be in other experimental or applied situations.

Nonetheless, the data show that JDR and SDR were used often, as well as other types of rules. The evidence for E/I strategy use was more indirect, but seemed to be found less frequently. As mentioned, the reproduction strategy occurred only a portion of the time when possible. Real-world settings that provide more time for memorization and practice, as well as additional motivation, are likely to produce varying results, given the test cheating literature reviewed earlier.

As for research question 3, patterns of relationships were examined between people's job, caught, and realism goals and their strategic and tactical cognitions. Many specific correlations were found to be statistically significant, but the overall patterns observed support an intuitive expectation that the task-relevant goals in this job applicant simulation were related more strongly to JDR and rule use than to SDR. It is also logical to assume that the specific rule of demonstrating familiarity was used primarily when people were attempting to avoid being caught faking. Also consistent with intuition, reproduction was related to job goal, but not caught or realism goals. Finally, the meta-strategy of attending to one's general "response patterns" was strongly related to all goals. That correlations were found despite this category measuring opposing actions of consistent and compensatory responding provides some support that deceptive actions should be viewed as part of a broader, goal-directed process.

Overall, the results show that cognitions about deceptive judgments and behaviors were intentional and tied directly to concerns about faking and good test performance. It was important to establish these linkages since I proposed that strategies could explain

how motivation affects test outcomes (cf. Baron & Kenny, 1986), ultimately representing a way to integrate the vast research on motivational factors affecting faking and deception with the large body of work on deception detection.

The results for research question 4 were difficult to interpret because they did not conform to any pattern that might be expected with simple theories. The deception skill measure was general enough that it certainly had face validity (i.e., asking directly how good/poor people were at deceiving others). However, it showed few relationships with any of the coded verbal protocol categories and was not related (negatively) to scores on the faking (i.e., bogus and impression management scales). Overall, the results did not lead to any straightforward conclusions.

There are a few reasons that could be proposed to explain this. First, deception capability may not exist. Second, the measure of deception skill may lack validity. Third, the concept, as it was conceptualized, may exist, but only have a minimal influence on people's ability to avoid detection and to obtain rewards through deception. For instance, the existence of the truth bias implies that many people will experience and self-report much success at deceiving, but only because their targets fail to suspect or recognize the possibility of deception. Fourth, people may be using so many different kinds of deceptive approaches and behaviors that success is not easily explained by a global assessment of skill. Perhaps skillful deceivers simply use more strategies and tactics than other people. Though any of these reasons is possible, it is worth mentioning that the first seems highly unlikely if one views deceptive communication as a type of problem-solving process. Whenever a deceiver attempts to discern through rational means the kind of image/message/representation that the target wants to see before providing the reward,

general problem solving models (e.g., Newell & Simon, 1972) can be applied rather directly to describe deception. Thus, at least some forms of deception should be related to general problem solving abilities. Here, skill was correlated with cognitive ability, but weakly enough that statistical significance was barely reached.

Despite these results, it was apparent in some verbal protocols that people viewed their deception skills differently. A few people believed they could "get away with" stating faked answers, even if unrealistic, while others believed there was little they could do to lie convincingly. Other results already discussed implied that there exists variance in how well people choose strategies to match the situation (i.e., test question) and in how well they execute their intended strategy (e.g., mistakenly writing an answer different from one intended). Some of the training recommendations explicitly addressed the latter issue by telling people to read the questions carefully and to remember their job goals.

Research question 5 concerned the effectiveness of various strategies. The results provided little support that any of the strategies allowed people to improve their scores systematically. In general, the frequent JDR cognitions found did not seem to be related to higher faked test scores. SDR produced moderate positive correlations, but only with two biodata scales. Rule use was negatively related with the agreeableness scale and unrelated to others, but it was already pointed out that this category was fairly heterogeneous.

As expected, rule use was negatively related to the one scale using a repeated item, agreeableness, but the correlation was not statistically significant (r = -.13). (The correlation should be negative because the "best answer" was low.) Oddly, the reproduction strategy was significantly related to perseverance and leadership scores. The

reasons for this finding are unclear, since these scales did not provide any opportunities for people actually to copy with the best answers. It is possible, though not probable, that the results indicate some kind of memory effect such that people who remembered the "best answer" for the think-aloud practice item were able to fake the emotional stability, perseverance, and leadership questions better, for some reason. Alternatively, it is possible that personality is a partial determinant of thoughts of about the reproduction strategy, if the scores on these three scales actually reflect true scores to some extent. (Note that this latter explanation could also be used to explain any of the findings in Table 11.)

In the end, the strategies did not help to explain the scores that people actually produced. It may be the case that people used particular strategies too infrequently for any of them to have a large impact on an overall scale score, especially since the verbal protocols were collected for only a portion of the items in each scale. It is also possible that the deception skill targeted in research question 4 was not measured properly, but did have an impact on scores, such that poor skill lead to poor strategy execution, thereby attenuating the effectiveness of the strategies. Or, relationships between the strategies and scores might have been attenuated by the lack of variance in faked scores. For all of the test battery scales, including impression management and bogus items, score distributions were skewed more negatively than in the honest condition, with most scales showing relatively normal distributions in the honest condition. Finally, it is possible that people's scale scores were determined by multiple strategies used for different items, with some strategies increasing the scale score and others lowering it. If strategies do compete against each other in terms of their influence on the scale score, these findings suggest

that research focused solely on differences between scale scores (or other message outcomes) are failing to account for important moderators.

The last research question, RQ6, concerned differences in the scales faked. The results supported simple intuitive notions that JDR is performed more often when test knowledge is available, and SDR is considered more often when test knowledge is unavailable. Also logical, the reproduction strategy was used most often with the repeated item, while concealment was used most for the impression management scale which was not designed to measure primary, "content" constructs. Yet, the main implication of these results is that, given that there are different ways to fake a test, people may attempt to and/or need to match specific strategies and tactics to situational demands and constraints, an idea proposed in some of the communications literature, but rarely tested. Again, whether fakers will actually benefit from selecting matched strategies is unclear in these data.

Regarding the general set of findings, some tactics (i.e., demonstrating familiarity and excuses) seemed to be based on relatively sophisticated logic that has not been addressed in prior literature. Also, a number of factors were found to affect faked scores in a manner that might produce results similar to those of honest people, though these factors are not the result of faking itself. For instance, many judgments of curvilinearity were made, implying that fakers will aim for lower response options while believing that those lead to the highest scale scores. Unless scales are keyed in such a manner (e.g., empirically), these strategies actually end up hurting the fakers.

Also, fakers sometimes chose to go with their honest answer. This likely explains some of the discrepancies found in the magnitude of faking between past studies, both

experimental and applied. Whether enticed with a reward or given explicit instructions to fake a test, people may not always attempt to fake on each item. When fakers respond to some of the questions honestly (i.e., when the honest answer is not also the best answer) and are still able to inflate their test scores, it must be the case that fakers are putting extremely large answers for the items that are faked. In these situations, it seems that there is a greater potential to detect faked answers that are unrealistic. Differences between the frequency of faking and the magnitude of faking on specific items have not been examined in past research.

In reference to the lie scale (L) of the MMPI that is generally regarded as a less "subtle" detector of only naïve fakers, Paulhus (1991, p. 32) stated, "It is obvious to a sophisticated test-taker that, even if one is trying to appear desirable, it is unrealistic to deny such ubiquitous attributes." Yet, the verbal protocols collected in this study clearly indicate that the same criticism can be applied to his BIDR scale, for both impression management and self-deception (for which data were not presented here). At times fakers explicitly stated concerns that their answers might appear unrealistic or unbelievable, even when not warned in the fake realistic condition. However, the results also indicate that a portion of people recognizing the possibility of being caught still selected scores extreme enough to trigger the extreme BIDR scoring system (i.e., the two most extreme response options). In some cases, people offered justifiable reasons for moderating their extreme scores by just a small amount (e.g., "I have to be realistic but I still don't want to admit that"). At other times, this type of response might be the result of anchoring and adjustment biases, a type of natural processing error. Still, it may be an indicator of poor faking skill or poor strategy execution. Disentangling these different reasons could

provide some insights about the kinds of people who actually end up being detected, given that not all fakers are detected.

Limitations

This study was conducted primarily as an initial, exploratory investigation of direct information about components of the deception process. As such, it is important to discuss a number of design features that limited the types of conclusions that could be drawn from the results. At the same time, few equivocal conclusions were made, particularly ones going beyond just establishing the "existence" of concepts.

One potential limitation relates to the ideal applicant condition. It was assumed that people are not ideal applicants and that their overall pattern of behaviors would reflect faking (even if truthful for some questions). Since deception is defined as intentional, it would be more accurate to state that faking was assumed to occur only when the participants did not *perceive* themselves to be ideal applicants. Upon completion of the experiment, one participant cheerfully stated that she found the task easy because she already viewed herself as an ideal applicant. As a result, she was excluded from the analyses because her verbal protocols would reflect the processes of an honest person. Since no manipulation check was given to assess this scenario, it is possible that other people in the study acted in this manner and added noise to the results. However, since responses that sounded honest were often not coded since they failed to reveal any information about deception, the "ideal" participants would only affect some of the analyses.

Disclosure. Another potential study limitation that could impact results is self-censoring of thoughts by participants. Most participants seemed to be comfortable

disclosing at least minimally embarrassing information, but they were not asked very intrusive questions. Early in the experiment, people were asked the third think-aloud practice item: "Name three activities that you have done at a party." While alcohol consumption is by no means the most embarrassing thing for a student to admit having done and may even be something to boast about in college, 66% of the students named drinking alcohol as one of the activities, and 8% listed some form of drinking as two or three activities. Besides, many people hesitated or giggled before giving their answer, which can be taken as a reasonable sign of some level of embarrassment or shame, even if small. Three people stated using illegal drugs and two males reported sexual (not just romantic) activities.

During the test, one participant even offered information that could be seriously incriminating.

> (H, 151, Impression): "I try not to exceed the speed limit too much because... I don't really have insurance. So, I don't want to get caught."

As well, only four people did not provide consent to publish their verbal protocols in an anonymous form. Therefore, it does not seem likely that participants consciously censored thoughts when thinking aloud during the test. Given that participants in other types of research have answered questions of a more personal nature, I concluded that potential self-censoring effects probably had a minimal impact on the think-aloud procedures.

Think-alouds and verbal protocol. Ericsson and Simon (1993) discussed general advantages and limitations of think-alouds and verbal protocol analysis. I only address a few issues as they relate to this study's design and findings. All VPs were transcribed to help prevent rater biases that might occur if speech was coded (e.g., including pauses &

intonation). However, this large set of data went through an elaborate process of reduction, especially for analyses of interrater agreement. The choices made were not atypical of ones found in published studies, but they undoubtedly affected the results to some extent. The coded variables would probably be more accurate and generalizable if more raters (or different types of raters) had been used, if more data had been coded, or if training and refinements to the instructions had been made during multiple cycles of coding.

Yet, the biggest threat is that personal biases of myself and the raters, especially biases held in common, would lead me to overlook alternative strategies that do not fit with the theoretical framework of deception I proposed. As stated earlier, I have attempted to present the design decisions and results in a transparent fashion to allow the reader some discretion in the interpretation of the results (e.g., Appendices E and F contain additional information about data collected). I have also argued that the initial set of theorized strategies can describe the full data set. The strategies can be ordered based on how much situational information is available to the deceiver. It also appears that they can be roughly structured based on the degree of rational thinking performed (e.g., problem solving to satisficing to guessing). However, alternative groupings of deceptive components could yield other interesting taxonomies. For instance, the excuse-making tactic identified in this study suggests that deceptive behaviors could be organized meaningfully based on duration or the level of interaction required between the source and target.

Another concern is that VP data may be confounded with general cognitive or verbal abilities and affect the study results through those variables. In the same way that

the results of paper and pencil tests would be confounded with literacy, the VP results are obviously confounded with any ability that allows participants to state their thoughts aloud, accurately and clearly. Past research has shown there are individual differences in the length of verbal protocols that people produce (Ericsson, 2002).

Protocol length itself is of no interest in this study, but the percentage of usable protocols provided by different types of people is important to consider. (To the extent that shorter protocols provide unusable data, length is an indirect concern.) Although participants who were wholly unable to provide coherent verbal protocols were not analyzed, only small nonsignficant correlations were found between the percentage of usable protocols produced by participants and the potential confounds of general cognitive ability, general verbal ability, age, and gender. (These results were based on the total sample and on various subsamples.) In the unlikely event that all participants were "expert" deceivers and processed thoughts automatically, I also examined the correlation with deception skill; it was not statistically significant. These results are consistent with Ericsson and Simon's (1993, p.89) general conclusion that "few differences [in the completeness of think aloud data] that were encountered in published reports were not of major importance, and where differences appeared, they were very likely attributable to differences in the experimental conditions."

Also, readers may wonder if the length of verbal reports for honest responses differs from faked ones, in a manner similar to the research on response latencies during deception (e.g., Walzcyk et al., 2005). Though it may be tempting to conduct such an analysis, comparisons of verbal protocol length in this situation (e.g., in Robie et al., 2006) are likely to be wholly inappropriate because verbal reports are less likely to

produce thoughts about memory retrieval (Ericsson, 2002), which is arguably the primary cognition performed by honest participants when selecting a test response. So, honest verbal reports are expected to be shorter for confounding reasons.

The other major limitation of using the think-aloud procedure is that it may disrupt natural thought processes related to deception. This issue could not be examined since there do not appear to be alternative, equivalent ways of obtaining the same information, and preexisting measures of cognitive faking do not exist. Still, the people in the honest condition served as a control group for some of the outcome analyses presented. Since people in all conditions were told to think-aloud and imagine being job applicants, the effect sizes between honest and faked scores shown in Table 13 can be attributed solely to the faking manipulations.

Unintended consequences of experimental design. A couple of design choices may have created unintended effects that altered people's response patterns. One choice was to allow participants to select the job for which they would imagine applying. Most jobs did not conflict with the test knowledge given to participants, though extraversion may be valued more in some jobs than the instructions state. However, there were occasions when people responded to individual items by using job specific knowledge that was incompatible with the test knowledge or general assumptions about qualities that are socially desirable. Here is just one VP example for a person applying to be an FBI agent:

> (FG, 130, "I sometimes try to get even rather than forgive and forget."): "That is very true. ... Let's see. Does it have anything to do with hard work? No. ... I don't know. Forgive and forget... Is it always nice to forgive and forget? I mean, the FBI is a forceful place. I don't think they forgive a lot of people. So... yeah, I'm just going to stick with F, I guess. I mean, it really isn't a bad thing. There's nothing wrong with being mad with somebody."

Another design choice influencing test answers was that participants were told to imagine applying for a job right after finishing college. This instruction was intended to make people's chosen jobs somewhat comparable, with all being at a high enough level of complexity that a large test battery would not seem out of place in the selection process (e.g., as opposed to a fast food clerk position). Apart from the fact that people normally differ in their years of college before graduation, this decision particularly affected participants' answers to questions about experiences and achievements that occurred during the previous two or three years (i.e., prior to the hypothetical selection process). For these questions, younger students had to imagine being a graduating student before faking their answer. In contrast, older students could base their faked answers on their actual past.

From an experimental standpoint, this is one more factor that was not controlled. For the most part, this scenario seems to have had some effect, but not a large one since only a portion of the biodata items referenced experiences within a discrete time period, during the recent past. This was not an issue for the personality factors that are expected to be stable over long time spans. On the other hand, this design choice could be viewed as beneficial for increasing the generalizability of study results. Younger job applicants who have little experience will be the ones with the largest biodata discrepancies between their ideal and current states. Thus, they may also be the ones who are most likely to create deceptive responses in the selection process. Nonetheless, this design choice created unintentional effects and was not controlled for or measured directly.

Another design choice was related to the repeated personality item ("Have a soft heart") providing an opportunity to use the reproduction strategy. Providing this

opportunity could have created at least two unintended effects. First, telling participants that '2' was the best answer to the item may have influenced them to give low answers on all items measuring something similar. The item was chosen for repetition because it seemed to be less transparent than other items, but four participants stated being influenced by it on subsequent questions, regardless of whether or not those questions were actually from the same dimension (Agreeableness). One of these four referenced the "soft heart" question 3 times, while the others mentioned it just once. Therefore, it did not seem to be a pervasive problem, though it does alter the results slightly. Here are some examples of the VPs for items other than the repeated one:

> (FG, 59, Agreeableness): "I think that goes along with being helpful in a team and being cooperative. So, but I just said I don't have a soft heart... so, I guess I can put C again"
>
> (FG, 138, Openness): "very, wait. That soft heart thing wasn't good, then maybe this abstract idea understanding is bad too. ... C"

The second potential problem was not noticed in the VPs, but was still certainly possible. Because the "best answer" to the soft heart question was not an extreme, telling participants this information may have encouraged people to make curvilinear assessments of items. Thus, the instructions were confounded with some of the results for the curvilinearity variable. Still, many more participants made curvilinear judgments than mentioned the reproduction strategy. Also, Table 7 shows that the correlation between reproduction thoughts and curvilinearity was negligible (r = .06). Moreover, participants often provided very specific reasons for making curvilinear judgments and avoiding an extreme answer, whereas no explanation was given for the practice problem's answer, other than that it was supported by research. Therefore, it is reasonable to interpret the

curvilinear variable as a legitimate measure of problem solving processes, rather than as solely a consequence of this study's design.

Credibility and generalizability. Ultimately, the credibility of conclusions reached in this study will depend on whether a broad consensus could be achieved among other researchers, beyond the manageable number of raters in this study. As such, every attempt has been made to describe all design and analysis decisions fully so that replications and expansions of this type of work can be conducted and the stability of these findings can be evaluated, in similar or more general deception situations. The data collected by the training item provides at least one form of corroboration and confirmation that the strategies specified theoretically could account for most of the types of actions that participants thought of performing. As there is currently no means of identifying expert deceivers, a traditional approach could not be used whereby verbal protocols are evaluated by a group of experts. As an exploratory investigation, however, this study provides a number of findings that open the door for much future work. The present formulation of the strategic theory has been built using existing theory and research, but is open to revision and elaboration, as every scientific theory should be.

Future Directions

One primary recommendation for future work is to develop alternative measures of all faking and deception components (i.e., motivation, cognitions, behaviors, and outcomes). Based on the recommendations of my dissertation committee, I used the think-aloud procedure to obtain direct self-reports. This allowed me to collect unique data that are invaluable for exploratory purposes, especially since few others have studied deception in this way. However, future research might benefit from using alternative

measures that allow faking or lying to be studied when participants are under a reduced cognitive load. The cognitive demands in this study were quite high since people had to pretend like they were applying for a job, fake their answers (in three conditions), and say their thoughts aloud. As well, new measures of faking should not only assess the frequency of cognitions and behaviors, but also provide some indication of their magnitude and quality.

More generally, future examinations of the theoretical framework proposed should eventually be focused on testing confirmatory models of the causal relationships between deception components. This study lays some of the required foundation for that work by defining relevant components of the deception process theoretically and offering initial empirical support for them, but does not trace the process of deception over time.

One general insight that arose from the data is that few researchers have examined faking while accounting for general misperceptions that people may have about how a test actually works. It is unlikely that many job applicants will have learned about basic psychometric theory and test development principles, and the VPs in this study suggest that people will react to test items in different ways. For instance, people encountering items similar in appearance or content became confused, irritated, or suspicious of an attempt to detect their deception (i.e., through inconsistent responding). Each of these reactions might lead a test faker to react in a different way. It seems that existing research on general test reactions could be incorporated into models of faking to address such issues. It may also be possible to draw analogies between the structure of test questions and face-to-face communications (e.g., job or criminal interrogation interview), where general norms of social interaction have a greater influence.

In the end, this study was not designed to measure E/I strategies very well. That the strategy involves two phases is itself an indication that E/I might need to be studied over some length of time. The *task* in this study was adequate in that sense, but the *analysis* of individual items did not allow for patterns over time to be extracted. In future work, E/I might be better studied as a form of acting or role-playing, as long as there are rewards contingent upon a realistic (not theatrical or dramatized) performance. It would be interesting to examine whether E/I is more effective in certain dynamic interactions, when the deceiver is able to completely immerse him or herself in another persona.

I proposed that some degree of suppression occurs with every offensive act of deception. As such, it was not studied here. However, acts of deception that only require suppression may provide interesting results. General work on suppression and new studies tying it to impression management (e.g., Vohs et al., 2005) suggest that it can be countered with resource depletion techniques. It would also be interesting to examine whether there are individual differences in self-control and suppression, which might better represent the concept of deception capability.

Also related to capability, it may be beneficial to identify "expert" deceivers, if they exist. It seems simple enough to identify certain groups of people that are likely to repeatedly attempt deception (e.g., DePaulo & Pfeifer, 1986), usually because of situational presses. However, the motivation to deceive is not necessarily synonymous with the ability to perform it well. Unfortunately, a paradox exists such that people who tend to be identified as deceivers through any means other than self-reports are the ones who failed to avoid being detected in the first place.

The conditional randomization strategy was also largely ignored in this study based on the results, though it was not expected to occur often in real situations. Rather, it was created to make the strategic deception theory exhaustive (able to describe all possible situations). Yet, there may be a few situations when this strategy is used more often. If people take a test battery multiple times and only receive general feedback about their poor performances, they may develop the notion that putting their true answer will lead to low scores without gaining knowledge about how to achieve a high score. In this way, this strategy may be more beneficial for trial-and-error learning over repeated task trials when general feedback is available.

Other work could be focused on removing a confound in this study concerning the comparisons of results between measures. The measures used here were different not only in their format (i.e., biodata versus situational judgment), but also in terms of the constructs they measured. Future investigations of cognitive and behavioral differences across scales must determine whether those differences are caused by changes only in format, only in constructs, or both format and constructs.

Another topic of interest is how people react to measures of faking and lies. Most of the past research has examined the issue of detection from the perspective of the target. Little has been explained about how deceivers react to bogus items and impression management (assuming they recognize these measures as being different from content ones of personality, for example). Research is also lacking on the factors that lead people to avoid detection. In addition, the VPs indicated that a few participants in this study were confused by regular biodata and personality questions because they had a limited vocabulary. Thus, even regular questions might evoke reactions similar to the bogus

items, implying that test takers may not be surprised when seeing bogus items despite their use of very obscure terms – an advantage for designers of faking measures.

On the contrary, one person in this study tricked himself into believing that a personality item was attempting to detect faking because he incorrectly applied test knowledge about job desirable characteristics to an item of a different dimension. (Note that this participant was in the ideal applicant condition and was never warned about being detected.)

> (IA, 70, Agreeableness – "Feel others emotions"): "Yeah, but... I don't have to be very outgoing or extroverted... for this job. So, I think they're trying to trick me. So I'm going to put moderately accurate."

In a related matter, the study of warning statements is of interest from the perspective of test designers who wish to thwart faking attempts. Yet, this study has shown that people will consider being caught and associated penalties even without being prompted. Thus, any study of faking or other type of deception should at least consider the possibility that effect of instructions to "fake good" or lie will be moderated by people who automatically develop goals related to avoiding detection.

Finally, an obvious direction for future research is to apply this theoretical framework in new domains, beyond just test faking. Given the varied types of literature reviewed, it is not necessary to list all the situations that could be studied. However, it is interesting to note that deception in the workplace has been rarely studied, despite the popularity of test faking research aimed at lies communicated prior to employment. Already mentioned, new research might also investigate ways of training deception for socially valued ends.

When considering the direct practical implications that arise from this study, it is useful to consider recommendations from the literature, which primarily apply to the detection of deception. Many of these were noted in the introduction and are currently being studied by other researchers. While these methods may be generally better than chance in their detection rates, research has not supported any method as being accurate enough to be used in practice with a high degree of efficacy. This study provides some insights as to why and how false positives occur, which cannot just be attributed to errors of the detection method; people can outwit the impression management and bogus scales at least some of the time. Based on the verbal protocols, deceivers are sometimes aware of the very logic that is used to catch them, including the need to be consistent in answers over time, to avoid unrealistic answers, and to perform tasks later when information is being verified.

The BIDR studied here exhibited a number of limitations. First, there were a number of "Mother Teresa" cases based on answers of honest respondents and their explanations in VPs. For instance, some reported "not true" or rarely true to an impression management item about speeding when driving because they lacked the money to pay for a traffic ticket, were they to speed and be caught. Others had already gotten tickets in the past and seem to have developed an honest aversion to speeding. Second, some items were not applicable in this sample because people were too young and inexperienced for negative things to be true, particularly when asked about their experiences with "customs" when traveling and about regularity in voting behaviors.

Although the data for the self-deception scale were not reported, one person's thoughts revealed an obvious conceptual flaw in the item: "I am a completely rational

person." It begs the question: "Can irrational people make accurate self-assessments about their rationality?" If the answer to this question is "No," then irrational people cannot be expected to give an accurate answer, even when answering honestly.

Other participants complained in their VPs about items not fitting with the response scale in a literal sense. For instance, the inclusion of the word "always" in the item "I always declare everything at customs" means that the item is either wholly true or wholly untrue. The middle answers should not be chosen if the item is to make literal sense.

Taken together, these flaws lead me to recommend severe modifications to the BIDR and similar SDR scales. The first is that items not making conceptual sense should be discarded, even if they can discriminate groups based on statistical deviance. Items should also be consistent with the response scale used, such that virtually all items need to imply or state that the content occurs "sometimes." To reduce the number of false positives, extreme answers according to any scoring key must reflect nearly impossible events or trait levels, unless a large number of items are used to compensate for people with a few ideal answers. Also, scale scores should be normative and administered to age appropriate populations, so that everyone has had a sufficient number of opportunities to demonstrate a "virtue" before being tested for "unlikely" virtues.

The curvilinearity category results provide another practical implication for general testing. They showed that fakers sometimes determined the best answer to be a moderate one. For questions with answers on a subjective scale (e.g., rate how well you _____), it is reasonable to suspect that honest people also think along the same lines. Such behaviors represent a bias toward moderate answers that will increase errors in

measurement if the items are scored linearly. Thus, it may be useful to explain the scoring key/system to test takers in certain situations, assuming that test transparency and faking are not affected by this knowledge.

Given that deceivers are active agents in the deception process and not merely behaving automatically, it may be possible to outwit fakers and trick them into endorsing undesirable or neutral items that an honest person would not. A dissertation by Bernal (1999) provides an example of this type of approach, but the basic logic can be applied in other crafty ways. In a sense, this approach is similar to the use of bogus items, but is intended to entice fakers into endorsing a certain type of item or characteristic as opposed to just providing them with an opportunity to do so. However, some of the obvious drawbacks of this approach might be complaints of unfairness if people feel they were manipulated unreasonably, honest misinterpretations of the question by people (as occurs with the bogus items), and questions about the test designers' ethics (based on logic similar to the concept of "entrapment" in the legal domain).

However, I would like to conclude this section by offering some potentially positive uses of deception. Emotional expression and social skills training may involve the use of deception. As an example, people in certain occupations who must perform "emotion work" (Zapf, 2002) may benefit by using certain deception strategies and not others. Some have also suggested that deception is beneficial for people who must protect others' privacy and confidentiality when questioned (e.g., therapists). For these purposes, understanding the relevant components of deception and monitoring them may help people identify where in the process they can make improvements.

Overall Conclusion

In this study, a number of key aspects of deception were identified with verbal protocol data. Though the thoughts considered and the behaviors performed by deceivers varied quite a bit, they could be grouped according to their underlying psychological processes. All instances of deception witnessed here appear to be consistent with the general definitions of deception offered by myself and others (refer to the Introduction). When viewed in a unified theoretical framework, generalization strategies seemed to occur more frequently than other behaviors, but evidence was found for multiple strategies, offensive and defensive. The offensive strategies and tactics were linked to people's motivational goals, but did not produce clear relationships with scale scores representing the deceptive message outcomes.

Nonetheless, the general conclusion implied by this study is that deception *is* a process. In fact, it seems to be a fairly complicated one that involves multiple goals, multiple strategies, and considerations of various factors situational factors that eventually lead to the production of a message outcome. As such, prior studies focusing on just motivation or detection may be important in their own right, but appear to be inadequate for understanding deception and faking. The closing quotation, in reference to a bogus item, represents the unique kinds of insights into deception that this study is able to provide:

> (IA, 70): "The thing about 'isometric analysis' is I don't know what it means. So, once again, I'm just going to put 'once.' That way they won't be expecting me to solve stuff by isometric analysis a lot and therefore, it won't matter I don't know what it means. ... That's called strategic thinking my friends"

Appendix A
Experimental Protocol

[*Seat the person so that you are behind them. They can turn around in their chair during the instructions. Place the recorder in the specified location – in front, but out of direct view*]

Hello. My name is _____. Thanks for participating in Part II of this research study. To ensure that our experiments are consistent, I will read instructions from this sheet. However, feel free to ask me questions at any point.

This session will last about 105 minutes. During the session, you will answer a set of questionnaires about your personal characteristics, background, and judgment. I will also ask you to say thoughts aloud, as I will explain in a moment.

Now, I will go over the consent form with you and answer any questions.
[*Pass out consent and data release form*]

All of your work will be recorded on paper and a digital voice recorder, but it will be kept completely confidential. Also, you can leave this experiment at any time and still receive credit for your participation. Please read over the consent form and sign it.

[**IF they do not want to be recorded, they **cannot participate** in part II and cannot receive credit for this portion of the study.* *]

On the following page is a release form that allows us to obtain your college admissions test scores. You do not need to sign this to participate in the study and may leave it blank, but it would be very helpful for us to have to analyze your data properly.

[**Explanation given if asked for one: Past research has shown that verbal skills are related to the amount people talk. So we want to control for this using your ACT scores.*"]

Before proceeding, I want to emphasize again that everything you do in this study, including what you say and write, will be kept strictly confidential. Although you will be asked to make verbal statements to help us understand how people answer questionnaires, know that we are NOT interested in revealing anything secretive, like unconscious or hidden thoughts.

[***Start audio recorder***]

In this experiment, I am interested in what you think about when answer certain types of questions. So, I am going to ask you to THINK ALOUD as you work on the question given. What I mean by "think aloud" is that I want you to tell me EVERYTHING you are thinking from the time you first see the question, until you choose a final answer.

Appendix A
(continued)

Please talk aloud **CONSTANTLY** from the time you see each problem, until you decide on your final answer.

[*Pause*]

I don't want you to try to plan out what you say or explain to me what you are saying. Just act as if you are alone in the room, speaking to yourself. It is extremely important that you keep talking, even if your thoughts do not seem important. If you are silent for a period of time, I will remind you to keep talking. Do you understand what you are supposed to do?"

[*Check to see they understand.*]

Good, now we will begin with some practice problems. Remember to talk aloud and say whatever you are thinking until you decide on a final answer. You won't write anything down for now. So just let me know when you have a "final answer."

1. What is 32 x 7? [Don't let them write!]
[*After they answer: "Just to let you know, the correct answer is 224. For now, we weren't concerned if you got the right answer; this is just to give you practice talking aloud. Also, there will be no math problems later."*]

2. On a scale from 1 to 5, with 5 being Very Accurate and 1 being Very Inaccurate, how would you rate this statement: You have a soft heart.
[*Although there is no "correct way to be"... just for your information... research suggests that people who answer 2, or Somewhat Inaccurate, tend to be the most successful workers, regardless of their specific jobs."*]

3. Name 3 activities that you have done at a party.

4. How many faucets (or sinks) are their in your parents' or guardians' home?
[*"I will keep track of the number for you."*]

5. Again on a scale from 1 to 5, with 5 being Very Accurate, rate this statement: I enjoy reading books about history.

Ok. Now we will begin the study. First, imagine that you are applying for a job that you would start right after finishing college. So, think of a professional job you would want that requires a college degree. Please write down the name of the job, or a brief phrase describing it, on the scantron sheet (in the box where is says your name, course, etc.)."

Appendix A
(continued)

[*If the job they want doesn't involve any contact with coworkers of any kind, ask them to choose another one. If it requires higher education, ask them to downgrade it to something like an assistant, or to choose a different job.*]

Good. Throughout this study, imagine that you are really trying to get this job. Next, I will give you a series of questions that assess different qualifications and characteristics. Imagine that an organization offering the job you want is using these questionnaires to hire someone for the job.

[*Pause*]

Additional instructions based on condition:
Honest: Remember to act as if you are applying for a job that you really want. At the same time, it is important for purposes of this experiment that you answer the questions accurately and honestly. The responses that you give should reflect your actual characteristics and behavior.

Fake Good: Remember to act as if you are applying for a job that you really want. Answer the questions by trying to produce responses that would result in the best scores, so that you would be hired for the job. Do not worry about what your real answers to the questions would be. Just try to fake the best answers.

Fake Realistic: Remember to act as if you are applying for a job that you really want. Answer the questions by trying to produce responses that would result in the best scores, so that you would be hired for the job. Do not worry about what your real answers to the questions would be. However, if your pattern of responses appears to be too good, abnormally high, or unrealistic, your faked responses may be detected by a special formula that will be applied to your scores with a computer. If you are caught faking by this formula, you will lose your chance to get the job.

Ideal Applicant: Remember to act as if you are applying for a job that you really want. Answer the questions as if you are trying to look like as good an applicant as possible for the job you want. Do not worry about what your real answers to the questions would be. Instead, choose answers that would make you look like an actual person who would definitely get the job. So try to look like an ideal applicant.

While you do not know which exact questions are the most important and relevant for the specific job you want, you do know that the organization is looking for job applicants who are hardworking and responsible. You also know that being outgoing and extraverted is not particularly important for the job, although being cooperative and being helpful in a team is important. Also, know that the questionnaires do measure other applicant characteristics that may or may not be important for getting the job.

Appendix A
(continued)

Now, I know I have just given you a lot of instructions. So here is a reminder sheet to help you remember these instructions throughout the experiment. You can keep that throughout the experiment.

[*Give instruction reminder sheet.*]

For some of these questions, you will be asked to THINK ALOUD just as you practiced. You should THINK ALOUD whenever you see a picture of a dialog box next to the question. It looks just like this.

[*Show dialog box on actual survey.*]

Whenever you see this, please start talk aloud CONSTANTLY until you write in your final answer. If you don't see the dialog box, you can just take this like a normal test and be silent. ... Finally, once you write in your answer, please don't go backwards in the test and change old answers. Are there any questions?

[*IF NOT the honest condition, say this portion*: "Before we start, I'll give you a minute to think about how you might fake the best answers / fake realistic answers / look like this ideal applicant."]

Okay, you may now begin the survey. Remember to pretend you are graduating from college and applying to be a _____ [*say their chosen job*]. Talk aloud whenever you see the dialog box next to an item.

You don't have to read the question out loud, but some people find it easier to talk aloud if they do. If you are silent, I will gently remind you to keep talking.

[*Give a prompt whenever they don't talk aloud, or when they're quiet for more than 5 seconds. DO NOT let them change old answers.*]

[*After the first stop sign...*]
Okay. Whenever you see a stop sign, I'm going to give you a short break and a chance to have some water if you like. I'm also going to repeat the instructions. It may seem a little repetitive after a while, but please pay attention since this is part of the procedure.

*At EVERY "**STOP SIGN**" on the questionnaire, **repeat these instructions**:*

1. When you see the dialog box, think aloud and talk constantly, until you make a final answer.
2. Act as if you are a applying to be a _____.
3. You can look at the reminder sheet whenever you like.
4. *Condition* instructions... [*see below*]

READ THESE WORD FOR WORD

Honest: Answer these questions accurately and honestly. Your responses should reflect your actual characteristics and behavior.

Fake Good: Produce responses that would result in the best scores. Do not worry about your real answers. Just try to fake the best answers.

Fake Realistic: Produce responses that would result in the best scores. Do not worry about your real answers, but make sure your faked responses are not too unrealistic.

Ideal Applicant: Look like as good a job applicant as possible. Do not worry about your real answers. Just answer like an actual person who would definitely get the job.

[*At the "**interview questions**," read the instructions for those questions in addition to the general instructions.*]

[*At the "**training item**," **stop recorder** and emphasize that the recording portion of the study is over. They do not need to talk aloud anymore and should no longer pretend to be a job applicant. They should answer honestly as themselves.*
Also, emphasize that they should be as specific as possible when answering this item.]

[*After training item, tell them they can just **fill out rest of survey** on their own.*]

[*Collect tests when done*]

Before we finish, I would like to request your permission to use the recorded verbal statements that you gave, when reporting the findings of this study. No personally identifying information will be used; anything that may identify you will be edited out. You do not have to give us your permission, but it would help other researchers understand this study and its findings more clearly. Here is the official form in more detail. Please read over and sign it if you consent.

[*Give data request form*]

That's the end of the experiment. Thank you very much for your time. Let me give you this debriefing form that explains the purpose of everything you did in this experiment. Basically, we are just trying to understand how people decide to answer these questions, either honestly or deceptively. So only some people in this study are told to fake their answers.

Appendix A
(continued)

I'll give you a moment to look it over and then answer any questions that you might have. If you have questions later, you can contact the primary researcher for this study... information is included on the debriefing sheet.

I do want to ask you kindly to not discuss the specific task you did with anyone who might participate in this study later. We don't want to bias anyone because not everyone in this study will be doing the same thing. So it is very important that you don't tell future participants that they will be answering the test falsely or honestly. However, you can tell them about how you had to say your thoughts out loud while answering survey questions.

Thanks!

Note to experimenter: After the participant leaves, make sure all written materials collected have the person's name on them!

*Extra note: If the participant asks you about the bogus items, tell them to act as if they are taking the test alone in the room (as stated in the initial instructions) and to just answer the questions as best as they can.

Appendix B

Measures Used in Part I and II of the Study

The measures are shown in the order in which they were administered.

Part I online survey:
Scale: **International Personality Item Pool**
Response scale: 5 points: Very Inaccurate, Moderately Inaccurate, Neither Inaccurate nor Accurate, Moderately Accurate, Very Accurate

1. Grumble about things.
2. Continue until everything is perfect.
3. Know how to comfort others.
4. Find it difficult to get down to work.
5. Love to help others.
6. Neglect my duties.
7. Am not easily bothered by things.
8. Am hard to get to know.
9. Do things in a half-way manner.
10. Show my gratitude.
11. Know how to captivate people.
12. Love order and regularity.
13. Think of others first.
14. Like to tidy up.
15. Wait for others to lead the way.
16. Carry the conversation to a higher level.
17. Rarely get irritated.
18. Am skilled in handling social situations.
19. Take charge.
20. Take offense easily.
21. Find it difficult to approach others.
22. Love to read challenging material.
23. Get overwhelmed by emotions.
24. Catch on to things quickly.
25. Am easily disturbed.
26. Love children.
27. Love to think up new ways of doing things.
28. Am a very private person.
29. Will not probe deeply into a subject.
30. Seldom get mad.
31. Have difficulty imagining things.
32. Can handle a lot of information.
33. Avoid difficult reading material.
34. Do things according to a plan.
35. Feel at ease with people.
36. Am on good terms with nearly everyone.

Appendix B
(continued)

37. Bottle up my feelings.
38. Am indifferent to the feelings of others.
39. Leave a mess in my room.
40. Feel threatened easily.
41. Often feel uncomfortable around others.
42. Have a good word for everyone.
43. Try to avoid complex people.
44. Inquire about others' well-being.
45. Get caught up in my problems.
46. Make plans and stick to them.
47. Make friends easily.
48. Am good at many things.
49. Waste my time.
50. Panic easily.

Scale: **Deception capability**
Response scale: 5 points, from Strongly Disagree to Strongly Agree

Rate how much you agree that the following statements accurately describe you. Please answer honestly, knowing that your responses are completely confidential and only examined for purposes of the study.

51. Anyone can easily put fake answers to get a better score on the personality test I took.
52. It would be easy for me to choose the right answers on a personality test, so as to make myself appear better than I actually am.
53. I usually get caught when trying to deceive someone.
54. It is easy for me to fool people.
55. It is easy for me to reproduce fake emotions.
56. It would be difficult for me to fake a personality test.
57. The "best" answers were obvious for the personality test items.
58. I can easily put on a false personality when I want others to think of me differently.
59. I often get away with little white lies.
60. I have a difficult time role playing and pretending to be something I'm not.
61. It would be easy for me to lie on a personality test.
62. I can easily get people to think I'm better at something than I really am.
63. I have a difficult time deceiving others.
64. I am a good liar.
65. I am a good actor.

Note. Items 1, 2, 6, and 7 belong to the Deception Efficacy subscale. The remaining items belong to the Deception Skill subscale. Items 3, 6, 10, and 13 should be reverse coded.

Appendix B
(*continued*)

Scale: **Demographics**
Response scale: Variable by question
Please answer the questions about general demographic information.

66. How old are you?
    18
    19
    20
    21
    22
    23
    24
    Over 24

67. What is your gender?
    Female
    Male

68. What is your year in school?
    Freshmen
    Sophomore
    Junior
    Senior
    5th year or more

69. Which of the following best characterizes your citizenship?
    U.S. Citizen
    Non-citizen – Canadian
    Non-citizen – other

70. Is English your primary language?
    No
    Yes

Appendix B
(continued)

71. What ethnicity do you consider yourself to be?
    Mexican American
    Puerto Rican
    Other Hispanic
    Native American or Alaskan native
    Native Hawaiian or Pacific Islander
    Asian
    Black / African-American
    White / Caucasian (not of Hispanic origin)
    Multi-racial
    Other

72. What is your college major(s)?

Part II test battery:
As shown in Appendix B and the sample questions below, participants were shown a picture of a dialog box when they were to think-aloud during an item.

Items 1-42 were biodata & bogus items. Biodata items cannot be shown for proprietary reasons.

Scale: **Biodata (sample items).** (The second item shown was not used.)
Response scale: 4 or 5 points, Variable by item

*For Leadership subscale:

How many times in the past year have you tried to get someone to join an activity in which you were involved or leading?
    a. never
    b. once
    c. twice
    d. three or four times
    e. five times or more

*For Citizenship subscale:
    How often have you signed a petition for something you believe in?
    a. very often
    b. often
    c. sometimes
    d. seldom
    e. never

Appendix B
(continued)

Think-alouds were required for biodata items: 3, 4, 5, 8, 9, 10, 11, 13, 15, 17, 19, 20, 23, 25, 27, 28, 31, 33, 34, 36, 38, 40, and 41

Scale: **Bogus Items**
Response scale: 5 points: Never, Once, Twice, Three or four times, Five times or more

17. How often in the past year have you programmed in AJMR?
25. In the past six months, how often have you resolved disputes by isometric analysis?
40. In the past six months, how often have you matrixed solvency files according to publication standards?

Scale: **International Personality Item Pool**
Response scale: 5 points: Very Inaccurate, Moderately Inaccurate, Neither Inaccurate nor Accurate, Moderately Accurate, Very Accurate

43. Get angry easily.

44. Am full of ideas.

45. Spend time reflecting on things.

46. Am always prepared.

47. Am relaxed most of the time.

48. Keep in the background.

49. Make people feel at ease.

50. Have a soft heart.

51. Get stressed out easily.

52. Have little to say.

53. Often forget to put things back in their proper place.

54. Don't talk a lot.

55. Often feel blue.

56. Have difficulty understanding abstract ideas.

Appendix B
(*continued*)

57.  Don't mind being the center of attention.

58. Worry about things.

59.  Avoid doing my duties.

60.  Use difficult words.

61.  Am the life of the party.

62.  Am not interested in other people's problems.

63. Am not really interested in others.

64.  Like order.

65.  Change my mood a lot.

66.  Start conversations.

67.  Am not interested in abstract ideas.

68. Feel little concern for others.

69.  Get chores done right away.

70.  Am quick to understand things.

71.  Have frequent mood swings.

72. Make a mess of things.

73.  Feel others' emotions.

74.  Pay attention to details.

75.  Have a vivid imagination.

76.  Get upset easily.

77. Do not have a good imagination.

Appendix B
(continued)

78.  Talk to a lot of different people at parties.

79.  Insult people.

80.  Take time out for others.

81.  Have excellent ideas.

82. Don't like to draw attention to myself.

83.  Sympathize with others' feelings.

84. Am interested in people.

85.  Am quiet around strangers.

86.  Have a rich vocabulary.

87. Leave my belongings around.

88.  Feel comfortable around people.

89. Am exacting in my work.

90.  Get irritated easily.

91.  Seldom feel blue.

92.  Follow a schedule.

Items 93-102 were for the most and least responses to the situational judgment items, which cannot be shown for proprietary reasons. All required think-alouds.

Scale: **Situational judgment (sample items).** (Neither item below was used.)
Response scale: 4 or 5 points, Variable by item

Appendix B
(continued)

An important class project you have been working on with a group of other students is not developing as it should because of petty differences and the need of some members to satisfy their own agenda. How would you proceed?

a. Try to solve the group problems before starting on the work.
b. Work hard by yourself to make sure the project is finished, taking on others' share of the work if necessary.
c. Talk to the professor and get suggestions about solving the problem. If that doesn't work, try to switch groups or have an independent project.
d. Schedule a number of meetings, forcing the group to interact.
e. Take charge and delegate tasks to each person. Make them responsible for their part of the project.
f. Talk to the group and demand that they start working together.

What are you most likely to do?
What are you least likely to do?

There is a concert coming up that you think will be fantastic. No one you know is interested in going with you. What would you do?

a. Go by yourself and find someone else at the concert that went alone.
b. Try to find someone else to go with you, but if you cannot then you would not go.
c. Ask your best friend to go even if you knew that he/she wasn't as excited as you were.
d. Get two tickets and offer a free ticket to anyone you know that might want to go.

What are you most likely to do?
What are you least likely to do?

Scale: **Balanced Inventory of Socially Desirable Responding (Version 6, Form 40A)**
Response scale: 7 points covering: Not True, Somewhat True, Very True
Self-deception included items 103-121. Impression management included items 122-140

Instructions: Use the scale to indicate how true each statement is. Remember to respond as if you are applying for a job.

103. My first impressions of people usually turn out to be right.

104. It would be hard for me to break any of my bad habits.

Appendix B
(*continued*)

105. I don't care to know what other people really think of me.

106. I have not always been honest with myself.

107. I always know why I like things.

108. When my emotions are aroused, it biases my thinking.

109. Once I've made up my mind, other people can seldom change my opinion.

110. I am not a safe driver when I exceed the speed limit.

111. I am fully in control of my own fate.

112. It's hard for me to shut off a disturbing thought.

113. I never regret my decisions.

114. I sometimes lose out on things because I can't make up my mind soon enough.

115. The reason I vote is because my vote can make a difference.

116. My parents were not always fair when they punished me.

117. I am a completely rational person.

118. I rarely appreciate criticism.

119. I am very confident of my judgments

120. It's all right with me if some people happen to dislike me.

121. I don't always know the reasons why I do the things I do.

122. I sometimes tell lies if I have to.

123. I never cover up my mistakes.

124. There have been occasions when I have taken advantage of someone.

125. I never swear.

126. I sometimes try to get even rather than forgive and forget.

127. I always obey laws, even if I'm unlikely to get caught.

128. I have said something bad about a friend behind his/her back.

Appendix B
(continued)

129. When I hear people talking privately, I avoid listening.

130. I have received too much change from a salesperson without telling him or her.

131. I always declare everything at customs.

132. When I was young I sometimes stole things.

133. I have never dropped litter on the street.

134. I sometimes drive faster than the speed limit.

135. I have done things that I don't tell other people about.

136. I never take things that don't belong to me.

137. I have taken sick-leave from work or school even though I wasn't really sick.

138. I have never damaged a library book or store merchandise without reporting it.

139. I have some pretty awful habits.

140. I don't gossip about other people's business.

Scale: **Deception training item**
Response scale: Open-ended

Instructions: You no longer need to imagine that you are applying for a job or think aloud. Please answer the REST of the questions HONESTLY, as they relate to yourself or your experiences. Do NOT fake these questions.

Training Item: Please write a quick set of instructions or rules for teaching another person how to fake, lie, or exaggerate answers to questions like the ones you were presented with here. You should include things you actually did that you thought were effective. You can also mention things you did not do today, but would do if given another chance. Please try to be specific in the rules or tips that you would provide.

(You do not need to write in full sentences or paragraphs.)

Appendix B
(*continued*)

Scale: **Knowledge of ideal others**
Response scale: 6 points: Never, Once, Twice, Three to five times, Six or more, *I don't know*

Circle the response that best describes your previous experience with tests like the ones you took today.

141. I have taken multiple-choice tests about my background experiences and/or prior achievements.
142. I have taken a formal personality test.
143. I have answered questions about my judgment in hypothetical situations (like the questions I answered earlier).
144. I have seen or answered a "lie scale" before.

Scale: **Knowledge of ideal others**
Response scale: 5 points, from Strongly Disagree to Strongly Agree

145. I know many people (personally) who would be ideal applicants for the job I pretended to apply for on the previous tasks.
146. Most of my close friends at MSU are hardworking and responsible.
147. Most of my close friends at MSU are outgoing and extraverted.
148. Most of my close friends at MSU are cooperative and helpful when working in groups.
149. Among people I know well, hardly anyone who would be considered an "ideal job applicant."
150. A lot of the people I spent time with during the last 2 years would be considered lazy.

Scale: **Perception of subjective norms about lying**
Response scale: 5 points, from Strongly Disagree to Strongly Agree

151. People should never lie.
152. I think it is sometimes necessary and ethically acceptable to lie to other people.
153. The best way to handle people is to tell them what they want to hear even if it is not true.
154. In life, honesty is the best policy.
155. There is no excuse for lying to someone else.
156. Generally, it is better to be honest than to be dishonest.
157. If I lied on a real selection test, I would feel disappointed in myself later.
158. People who are important to me think it is NOT OK to lie on a selection test.

Appendix B
(continued)

159. Most people who are important to me will look down on me if I lie on a selection test.
160. My parents would approve of me lying on a selection test.

Scale: **Attitudes toward lying**
Response scale: 5 points, Variable by question

Consider the statement below and select a response that best describes how you feel.
"I think that lying on a selection test is...

161.

A	B	C	D	E
Bad				Good

162.

A	B	C	D	E
Unpleasant				Pleasant

163.

A	B	C	D	E
Foolish				Wise

164.

A	B	C	D	E
Useless				Useful

165.

A	B	C	D	E
Unattractive				Attractive

Scale: **Manipulation check items**
Response scale: Variable by question

In this experiment, you answered a number of questions while pretending to apply for a job (questions before the "job interview" questions). Please answer the following statements about those questions.

166. How difficult was it for you to talk out loud?
A. Very Easy (Not Difficult at All)
B. Easy (A Little Difficult)
C. Not Easy or Difficult

D. Difficult
E. Very Difficult

167. How often did you think of things that you did not say out loud?
A. Never
B. Rarely
C. Sometimes
D. Often
E. Always

168. How difficult was it for you to remember to act like a job applicant when answering the questions?
A. Very Easy (Not Difficult at All)
B. Easy (A Little Difficult)
C. Not Easy or Difficult
D. Difficult
E. Very Difficult

169. While answering the questions, how often did you forget to act like you were applying for a job?
A. Never
B. Rarely
C. Sometimes
D. Often
E. Always

170. In this experiment, you were told to answer the questions either honestly or to get a high score by faking your answers or acting like an ideal applicant. How difficult was it for you to remember this instruction while answering the questions?
A. Very Easy (Not Difficult at All)
B. Easy (A Little Difficult)
C. Not Easy or Difficult
D. Difficult
E. Very Difficult

171. While answering the questions, how often did you forget this instruction?
A. Never
B. Rarely
C. Sometimes
D. Often
E. Always

Appendix C
Reminder Sheet of Instructions Given to Participants

Reminder of Instructions

The hiring organization is looking for **job applicants** who are:

- Hardworking
- Responsible
- Cooperative
- Helpful in a team

BUT being **outgoing and extroverted** is **neither** important nor unimportant.

*Other characteristics measured by the questions may or may not be important for getting the job.

Whenever you see the dialog box  , **immediately** start saying your **thoughts out loud** until you provide a final, written answer to the question.

Do NOT go backwards and change your answer to a previous question.

Appendix D

Transcripts of Verbal Protocols for Two Participants

Participant 152, Fake Realistic condition

3. I think it's interesting that... they're asking us about "picking up litter, that you come across and carry it until you find a trash can." I'm not really sure what that has to do with the job, but... ironically, that's something that I always do. And I think it's weird that people don't throw out their trash, and litter actually REALLY bothers me. So... in real life, I would probably say "extremely likely," but... for being an elementary schoolteacher, I'm not sure if that answer would matter or not. So, I would probably say "very likely," just because... I don't want them to think I'm too OCD or anything, but if I was going to be an elementary school teacher, they... probably know that I need to pick up after my students

4. this one, I think, is more applicable... and... I think... for the job of an elementary schoolteacher. *question* from a teacher standpoint, it's not always about success but it's about the trying and the process that you use to get there. So, I would say... just "important," but not "very important"... because obviously not every child is the same. And success is kind of subjective in a lot of cases

5. this one, I don't know... why it would be terribly important, because... because it says being outgoing and extroverted doesn't actually matter? ... But I think those things matter, especially if you're going to be a teacher.... So, I would probably answer... one

8. this one I think is is a good way of asking this question, of how you work with others or how you... get others to follow you as the leader. I think that's important for almost any job you're applying for. But you might sound really pompous if you answer "five times or more." If you to say "never," that's, probably wouldn't... that maybe you're not as good with team stuff or that you're not good with leading. so, I think the middle of the road is... is kind of good, maybe. Maybe saying "twice" ...because you don't want to seem pushy

9. I think this one's interesting, that they're asking about the arts. Because that seems like, especially in the business world, that's something that no one cares about losing. It seems to be a pretty dying thing, which personally bothers me because I'm a music therapy major. But... if I was answering honestly, I would say "more than 10 times" because that's true. But...for this purpose, again, I would probably go middle-of-the-road, to show that I'm supportive but, not like, overly enthusiastic. So I'd probably say C again

10. in applying to any job, I think that they want to know how you interact with other people because... I can't really think of very many jobs where you would... be by yourself all the time. So... I would probably say... I would probably say D, a large extent

Appendix D
(continued)

11. at first, that question kind of confused me... but I think as a teacher, that's something that you have to... do, because... it's not like you just have one problem when you're teaching something, like a subject or an idea, you have 25 problems because each kid is completely different, and you have to reach them in a different way. So... you would need it longer than average, I think. So, I would say "a long period of time"

13. this one, I think, they want to know how driven you are. I don't think the overly driven matters much, if you're a teacher. Like, I don't think they want you to be overly driven because it's not like... I don't know, I just, it's not like you're working for Donald Trump or something. You're working for a principal. So... and "driven" is so subjective, for teaching. So... I would probably say... I would say "rarely"

15. realistically, I'd say WAY more than most people. But, being a teacher, that's... something that's... probably important because unless you're working in, like, Hartford Connecticut with a bunch of WASPy kids that go to a private school, you're probably going to get people that aren't white. So... I would definitely, for this... say B

17. I don't know what that is.... So... since I don't know what that is... I mean, I guess it's better to say never BECAUSE I don't know what it is than -- and just take the fall for one question -- than to say "more than five times" and have that be something to do with like... I don't know... if it's something like the NRA ... I don't know. I think for this one, "never" is probably a good answer because I don't know what it is, and I wouldn't say middle of the road because if it's something bad, then... then I don't want to say that I did it

19. when applying for a teacher, I think this is important because... you... they probably want to hire someone that's used to being able... to be in the position of dealing with lots of people -- not necessarily dealing, but...Just from my personal experience, being a leader isn't always just about telling people what to do, but it's about delegating. But, also knowing what you need to do. and being a teacher, you have to work with not only other faculty members and your students, but their parents. So I think having your peers elect you is something that's good, even though it says being "outgoing and extroverted" is neither [un/important] ... it always looks good to have others your age elect you for stuff. So... I think... but "in the past year" -- that makes it hard because -- I mean, how many clubs are you really in, in one year? Some people are in one. Some people are in 10. So... I think... I think, even "once" is a good answer because that's not... it's not too showy offy, but it's also pretty feasible. Because even if you're in three clubs, the odds of getting voted to an official or representative position in all of those might be a little far-fetched

20. well, I remember when I was in elementary school, they went through a really big push about... about the environment, and they did a whole, like, reduce-reuse-recycle

thing. And I actually AM really environment-conscientious. But for being... a teacher in elementary school, I think that would be at least sort of important because you're trying to mold these young minds, and this is kind of their first experience of getting exposure to the fact that... being conscious about the environment is a good thing. So, I would probably go middle of the road on this too. So, C

23. I don't think a project is really all that worthwhile if it doesn't have something sort of difficult to it. So... and if you're going to be a teacher in elementary school... and assigning projects... I mean, really your kids are your projects and that's going to be difficult. So, I would probably say often

25. well, I don't really know what isometric analysis is. ... I don't know. I don't really know what that would have to do with... elementary school teaching. So I'm going to say never

27. for being an elementary schoolteacher, I would probably say "a great deal," because when I was in elementary school, I know that, even in Spanish, they had us sing songs and do things. And, there's just been so much research going on how good music is for students. And, even in like toddlers, getting them to express themselves through music and art, and it's good for their motor skills. so especially for an elementary schoolteacher, I would say "a great deal"

28. for this I would definitely say "five times are more" because that's what being a teacher's about. Especially... I mean, that's what lesson plans are. So, I think that would be extremely applicable

31. I mean, I think it's always good to lead by example, if you're a teacher. And I know... in at least all of the schools that I've been through, they've, at least once a month, had some sort of either fundraiser for the school or some sort of philanthropy project. And I was on student council, so we did that like once a month. So, I would definitely say more than three. I think that's just... good for the economy anyway

33. well, the instructions say... that they're looking for people that are responsible, cooperative, and helpful in a team. So... I think two or three might be kind of unrealistic because -- I mean, how many things do really volunteer for or are you involved in, in a year? so I think "one" would be a good answer... because being offered things is still flattering, whether it's 1 or 10. It's better than none

34. well, depending on what you're teaching in elementary school, this might be important. But I guess it's always good to go and check out museums and exhibitions because you take your kids on field trips based on what you're studying. So I think... I don't know how I'm debating between "two" and "three or four" because "two" kind of

doesn't seem like enough. And "four" kind of just seems like a lot, because shouldn't you be checking out other things like science museums and other things for field trips? so, I want to say "three or four," but I think the best answer for this would be "two." So i'm gonna say 2

36. well, I think that's important. I think this is a great question, actually. I mean, some people might think this is not very applicable to things, but... I think it's great and important, as a teacher or as a leader in anything, to make sure that you stay up-to-date on anything, and that you're still being taught because you learn by people you observe and... and by taking classes and furthering your own knowledge. So, I think that's important. I think "three and five times" might be a stretch because if you're teaching, I mean, unless you're a very new teacher, and... you're working towards your Masters. I'm not really sure. I think, for this one... *question*. I think "once or twice" would be a good answer because... I don't know. That just seems like the best answer for this question

38. well... if I was answering honestly, I would say more than 75 because... our school required that to graduate. You did 45 hours of community service. But I was in student council and you had to have 10 hours a month. So, I would answer more than 75. Plus, that just seems reasonable anyways. If you go through four years of high school, 75 hours is like... it's like four days. So, hopefully you did that much. I would say E. ... it shows you're willing to give yourself to others

39. I don't even...WHAT? I don't even know what that means. ... So... but "according to publication standards" -- I feel like it's something that might be important. So I don't want to say "never." So, maybe "once"? I'm not really sure, but I think I'm going to say once because on the last thing I really didn't know, I said "never"... and I don't think I want to answer "never" twice. So, i'm gonna answer "once"

41. always. I guess that matters on... where you live, but... recycling's just good practice anyway

46. I think this whole section is probably... really applicable if you're doing the sort of thing to be hired as an elementary schoolteacher, because you're relating yourself to others that's the same... like the same sex and roughly the same age as you, which could be your coworkers. so *question*. i would say... i would say very accurate because you don't, you don't want them thinking that you're going to go in, to be in charge of 25 kids, and not be prepared. That would... that would be retarded. So, I would say very accurate

48. well... even if this is... in relation to other people, I know it says being outgoing and extroverted is neither important nor on important, but being... helpful in a team is important. So... you wouldn't want to be completely in the background because then you wouldn't be contributing anything to your colleagues or... to whatever job you are applying for. So... I would probably say moderately inaccurate

Appendix D
(continued)

49. I think that's important, no matter what, because... if people are really tense... they tend not to perform as well and they tend to shut down mentally and physically. So... I would say "very accurate." That's something, especially if you're a teacher -- whether you're with your colleagues or parents or your superiors -- that's... that's important. You don't want people to feel uncomfortable around you because you need to be approachable

50. yeah, I think that... that's something that they would want. That is very accurate. ... I think that's something that's hard... to put on a scale of something that goes from "very inaccurate" to "very accurate" because... I don't know. I just don't see having a soft heart has been described in... in those kinds of words. That seems... I think that's hard to answer and that's why I don't really like these types of questions... because I think everyone gets confused and they're forced to just... choose A good answer

51. for this... a little bit of stress is good because it pushes you... but you don't want to get stressed out easily when working with your colleagues. So... you would probably say... very accurate just because the word "easily" is there. If it just said "get stressed out," probably say "moderately inaccurate," but you don't want to give them any reason to think that you could stress easy, because that's probably not good. so i'm going to say "very inaccurate." I tend to stay away from, well, I want to stay away from the extremes on the the answers... but sometimes it's not possible

53. well, if you're a teacher, that's just annoying for everyone because then if you're sick and one of your colleagues has to take over your classroom or a substitute has to take over your classroom, or even for kids, for finding stuff that's... I don't know. That's just not very good. So... I would say... again, I would say very inaccurate because the word "often"... is that I question. And having a word like "often," would force me to answer A

55. again, I would have to say A because of the word "often." And hopefully, you're in this job because you like it and that wouldn't make you feel blue

56. that would... I don't know. I # use "moderately inaccurate" or "very accurate" because... either way, I think it needs to be answered at as "inaccurate statement." ... Because... abstract ideas... i'm sure that's a lot of, there's a lot of that, in elementary school... just experimenting with different ways of teaching, at all different grade levels. And to keep it interesting... and not lose your touch, I think you have to constantly reinvent the way that you teach certain things. And sometimes I think that can get the abstract. So, I would say... I don't know. I want to say "moderately inaccurate," but I don't think that's strong enough. So, I'm going to put A

57. I mean, I guess...obviously... but... But these questions are relating to how you deal with others that are about your same age and same sex. So, I wouldn't say "very accurate"because I think that's kind of a really bold statement. So I would say "moderately

accurate" because it still, it's still true, but it's not like completely the absolute "I have to be the center of attention"

59. I would definitely say very accurate... when applying for, especially for applying for a job. The whole point is to get a job done. So, if anyone gave you duties, you'd have to do them.

60. well... I would say moderately inaccurate because it's not like you want to, you're not completely illiterate. But there is no need... to show off. It's not like -- you're an elementary school teacher -- you need to... like impress other people that are trying to write like dissertations or anything. You're not a Harvard professor. You're with kids that say "poo" all day. I mean, they're like... they're like little. So... I would probably say moderately inaccurate

61. I don't really know... I think that's hard because... because of the word "party." ... I know it's a figure of speech, but... I mean, you want to be... someone that people are drawn to and that's interesting. You definitely don't want to be a wallflower... but you also don't want to be... like, the class clown or the people that like -- at the party, or like everyone, or like at the board meeting, everyone like listens to you the whole time, but then when you turn your back, they're like "Man. I wish that person would just shut up." You want to be cooperative... and the instructions say that it's looking for a job applicant who is cooperative. ... So... this one, I think I would actually answer C, "neither inaccurate nor accurate" because... truthfully, I really don't know what would be a good answer, and that seems kind of a good, like, cop-out answer. So I'm going to answer C

62. you would want to answer "very inaccurate" because... you want to be helpful in a team... and cooperative. So, if you're not interested in other people's problems,... then.... I think THAT question alone would be able to not get you the job.

64. I would say moderately accurate because... in the elementary school, everything is going to be a little bit chaotic, but you need to have things in order and be able to maintain control.

65. I don't think that's helpful for your colleagues, or your students -- if you are like bipolar, or if you can have really big mood swings. So,... I would say... I would... I don't know. I kind of want to answer "moderately inaccurate," but... I think "very accurate" would be a better answer, because they probably want someone who can stay even tempered

66. That might be good... but it could be bad. I mean, you don't always want to be the instigator. ... But... if there's, if there's like dead silence or... if no one's answering and the people are just like "Buehler?... Buehler?", then you kind of... I think that would be a

good thing. ... I kind of want to say moderately accurate, but I also want to say C, "neither inaccurate nor accurate," because... I think sometimes it's good and sometimes it's bad. So... and it doesn't say "Start conversations A LOT" or "SELDOM start conversations." It just says "Start conversations." ... So, I think I'm going to say "neither inaccurate nor accurate" just because of the way the question's worded

67. For this, I would say very inaccurate because... at staff meetings and stuff, all these different teachers are going to be bringing all these different ideas to the table. And... you know, they always say two heads is better than one. ... And, certainly 15 heads is better than two. So... and I mean, ideas, I guess, can be too abstract, but I guess you never know unless you try. So... I would probably "say very... inaccurate" because... they don't want someone who's disinterested in things that other people have to say because you want to be... cooperative

69. I don't... I mean, obviously you don't want to put things on the back burner, because they pile up, and that's not good. But... being an elementary schoolteacher, I don't know... how important that is. So, I'm going to answer C because I'm kind of indifferent about that, and I think, again, that that's a good middle-of-the-road. It's not saying that I DON'T get them done right away, but it's also not saying that like... I'm so... compulsive that I always get things done... like lickety split, I guess.

70. I'm going to say... I, I kind of want. Originally, I wanted to say moderately accurate... because I didn't want to be to like "Oh, I'm really, really smart and I'm quick on the uptake." But... things in an elementary school kind of move at a fast pace. And... if your colleagues or boss... or your peers or... are trying to introduce ideas or talk to you about something that might be important... So, I'm going to say very accurate.

71. I would say "very inaccurate" because, again, you want someone that's even tempered. And having frequent mood swings isn't good for any party involved

73. I'm going to say "neither accurate nor inaccurate" because... obviously, you don't want to be completely stone-hearted. But if you, if you take things too much to heart, that can get in the way of a lot of stuff. So, I think a middle-of-the-road answer is good for that because you're not pinning yourself one way or the other

74. I think that's, that should be. I think they would want someone that would say "very accurate" for that. ... I'm not really sure why, but I just... I just feel like that would be the make answer for that one

75. I think... being an elementary schoolteacher... that would be good, but these questions are relating to people that are your age and the same sex. Either way, I think that would be very good. ... I don't know. Vivid's weird because it's... like, if you live

too much in imagination, I think it could be that bad because there's definitely a lot to be said for reality, obviously. And you need to know things that are going on, like current events, and understand the realness of stuff. But if you're an elementary schoolteacher, you're gonna have kids that are at that stage in their life where a vivid imagination is something that's still encouraged. So, I would answer... "very accurate" for this one

76. *skipped*

78. I think that's something... I think I'd say moderately accurate for that one because you don't just want to... only talk to the people in your department, or only talk to the people in the grade that you're teaching, because then you won't know how things are working as a whole. And that's the same thing as saying like "Well, I live in Michigan. What do I care who gets elected (you know) in France?" but.... You have to look at the whole. And I think... I think that's important

79. very inaccurate. That's a no-brainer

80. I think moderately accurate is good because it's always good to take time out for others. I think it makes you feel good. I think it makes you learn more about yourself as a person. But at the same time, if you spend all of your energy helping out a colleague or one student in particular, or anything like that, then you tend to put things aside that you need to get done for yourself. And you can't lose sight of that.

81. well, I think the word "excellent" is subjective. I mean, obviously... you would think most of your ideas are good, but you don't want to be really... self-centered. So... I would say "neither accurate nor inaccurate" because the word excellent is kind of, kind of a strong word

83. I think I'm just going to say "neither in accurate nor accurate" because... I mean, obviously you need to sympathize with others, but... you also don't want to take... sides, like if two people that you're with get in an argument, because that could be bad. That's kind of a tricky one. So I'm... I'm going to say C

85. well... I think that that would be... I think moderately inaccurate would be good because... like if you get someone new to the faculty, you want to be able to go up and introduce yourself and, you know, be like "Hey. Yeah, I'll show you around." and just the way things go. But you also want to take the time to learn about them and hear what they have to say. So... or like when you get a bunch of new kids in your classroom, obviously they're all strangers, and their parents are all strangers, but if you're quiet, that's not good. So I think moderately inaccurate is a good answer

86. I think neither inaccurate nor accurate is good for that because, again, you're working in an elementary school. You're not... trying to impress anybody

Appendix D
(continued)

88. you should probably say very accurate for that. You're going to be around people all day long

90. I think it's the word "easily" that makes me say very accurate. ... because obviously everyone gets irritated

91. I think moderately accurate is good for that... because everyone gets down a little bit

92. I think that's important because the whole school day is structured around a schedule... So, I'm going to say very accurate

93 (SJ item). I would probably say... *option B* because... depending on how they've been mistreated, it can get really sticky. It can be like a really sticky situation. And, i mean, you have to think: Is really worth losing your job over, if it came to that? But then again... I don't think it's right to do nothing. ... And if that makes you very angry. I don't know. I just think you should act on your good conscience and... I think helping and... Well... Actually, I think F's a better answer because it says *option F* because that way you're not overstepping like... What if they don't want to do anything? Like personally, I think if you're mistreated, something should be done. You should never, especially in a professional setting, you should not be put in a way that's like... I don't know, like if it was like sexual assault or racism or something like that. I definitely think something should be done. But also you need to know if it's not your place to do it. But, showing that you're willing to help, is good... because then you just put that out there and if they need help, they know, either then or in the future, that you're someone that they can go to. I think... I think if I was applying for a job, they would like that answer. So, I would be most likely to answer F

94. probably option D*... because then you're just assuming that that person was in the wrong, without even talking to them first. And that's just... that's backwards from the entire American judicial system. That's just weird

95. Well... *option B* would... would be a good idea because then you're... you're not just like looking at it, being like "Oh man. That's a shame" and not doing anything about it. ... But... then you're also not fixing, you're not doing like a permanent fix of the problem. People will still do it. So, I think *option D* is good, because floor meetings are good because everyone's there and... and then maybe you could tell them why it's not very courteous to do that. So I would say D

96. *option C* that's just not nice. I mean... that's just, I don't know. I think, I just think that's rude. So... I would say C

Appendix D
(continued)

97. All the other options [besides E] are pretty decent options. That's really hard because... If it's dull and boring like what if it's because the teacher's talking monotone? ... *option A* I don't think that would be the best answer because that's not making it more interesting. That's just something to help you stay awake. ... But I also don't think studying during the lecture, even if it's for that course, is good because that you're not paying attention to the teacher, and no teacher wants to look up and see that no one is listening to them. So, I think B would be the best answer, which is *option B* because then... and that's just an easier way to study it I think.

98. to skip the class

99. Well, being cooperative is important and helpful in a team is important, but being responsible is also important. And... like if everyone else in the class can do it... there's not really a good excuse for your group not doing it. But it's also hard because when you're working with other people, you're working with other schedules, you're working... there are other classes, there are other... you know, extracurricular activities, and their personal lives. And it's really hard to get... to plan time where everyone can be there... Well, if I was applying. Since this is for applying for the job... Well, usually you follow a chain of command, but I like F because it says *option F*. that's admitting that you don't know everything and... there's nothing wrong with asking for help. But usually, you want to... deal with it within your group and then, you know, like work up the chain. ... But, I mean, if you've been working on it, and it's just not happening, I think that would definitely be efficient way, of fixing it. So I'm going to say F

100. I guess just like *option A*. That's just kind of a... "eh" kind of answer. So, I'd say that that's, I'm least likely to do that. ... because then that's just like being like "Well, we'll try, but were really going to submit it late anyway." That's like really not doing anything about it

101. You would go. ... I would say B would be the best answer. *option B* I mean, everyone's kind of got that white elephant that's like... you know, it's just there. You gotta do it. That's part of life, that's a part of any job you'll ever do and... yeah. ... But I kind of like D. *option D* I don't know if that's being TOO bold though. But, then you wouldn't only, you'd not only be helping yourself, but you'd be helping probably a lot of other students too. ... That's kind of a proactive answer. I think I like, I think the committee would like that one. So, I'm going to answer that.

102. bring a crossword puzzle to class. ... that's just... the committee would not... think that's a good answer

Self-deception comments omitted

Appendix D
(continued)

122. I'd probably say D for that one because... no one tells the truth 100% of the time. And if you're like "No. That's completely false." Then, they know you're... full of it.

126. Well, that's not really true anyway, but if you're applying for a job... in a school, that's probably not something you want to admit that you do, even if you do do it. So... you probably want to say not true to that one

128. I'll probably answer C because... unless your Jesus... I highly doubt you haven't done that

129. I'd put that on more true side because if... some things are just meant to be kept in private, but sometimes you can't help it.

131. Yeah. ... I don't really know what that has to do... I guess, if you're doing this from a business perspective, that would be kind of important. But from a teacher perspective, I mean, I guess that says that you are a law-abiding citizen

132. You probably want to put that. You probably don't want to say that's completely untrue because: How do you remember everything from when you were young? ... Then, you're learning morals when you're young. So, I don't think you necessarily do it maliciously. So I think B would be a good answer

134. Well... yeah, you should probably say like somewhat true... because most people drive faster than the speed limit. ... But... you don't want to say you do it all the time. But it's not like you're going to be driving kids on a field trip because the bus drivers do that. So... yeah

137. Well, even if you have, you PROBABLY don't want to say that because... they probably don't want to hear that. So... you probably want to go towards the "not true." You can probably get away with answering B

138. Probably on the very true side would be, probably like an F. That'd probably be good

140. You probably want to put that more on the true side because gossip really isn't very good, but... I mean, everybody does it. It's kind of hard not to

Appendix D
(continued)

Participant 148, Ideal Applicant condition

3. well, I mean... I'm not really much of a picker-upper of litter. So, but, if I'm trying to get the job I would want. So, say somewhat likely

4. well, if you're looking for a job, you definitely want to say that you're pretty successful and succeed often. So, I would probably say -- oh, "how" important is it? So, that would be it would probably be extremely important for a job.

5. well, since more on, based on the team, it wouldn't be that like important. So, I'd probably go with like... a mid-range answer

8. well, again, I don't know. Well, I'd have to say that would kind of be relatively important. So, I would probably go with D, because if you didn't do that, then... you put like "five times or more," they might want to call you out on that one

9. what does that have to do with being hard-working? ... I guess I'll just go with B then

10. well. I would imagine that'd probably be somewhat important, in coming into a job, but... I don't know. That would probably be, probably more "to a large extent" because you kind of want to seem like a go-getter if you're a hard worker. So, I probably go to the D.

11. well, if you're hard working, then you'd probably want to spend a lot of time... on the question. And even me personally, I'm pretty sure that I -- like when I lose something or something like that, or even if I get like lost or confused in something -- I definitely stick with it. so I'd probably go for "long period of time"

13. well, that would look bad. So, I'd HAVE to go with "rarely"

15. again, I don't know how much that would go for counseling. But you don't want to say you don't value art or culture because then that would make you seem less diversified. So, I'd probably go with... "about average" to more than average. I'll go with more than average. I guess appreciating art and culture can't really hurt that bad

17. well, I don't know what that is, and I don't know what that would relate to counseling. I guess I could say that I did it once to have some kind of experience, but... I don't know exactly what that is. So....

19. again, I don't know how much that would probably be important to being a counselor, but it always looks good on my application to show that you're involved and that people look to you for being a leader. So, I'd probably go with twice, a little bit above average or around average

Appendix D
(continued)

20. again, I don't know how much that would really have anything to do with being a counselor and being hired, unless they were looking for someone who's environmentally safe, but... I will go with... C, twice

23. well, you'd want to look hard-working and that would mean that you definitely want to go with A, very often

25. well, being a counselor then... I'm sure if you resolved disputes, you probably used many different types of analysis. So... I'd probably go with D, three or four times

27. again, I don't know how much that would have to do with being hired as a counselor, but you want to seem like you're cultured. So, probably go with "somewhat" to "a lot," which -- I've basically been answering C the whole time. So, C again

28. well... if you set the time and schedule, it is again IT leadership position. So... I'd probably go with... D, three or four times

31. being a counselor, you probably wanna... be a somewhat giving, which shows that you care about like the people that you'd be counseling because if you're a school counselor, you're gonna have to deal with people who aren't necessarily the richest people or... people that are well off. So... you probably wanna go with D, three

33. well, you'd still kind of want to look like a leader, but you wouldn't want to go with the most, the largest answer, because they're still looking for someone who's relating more to the team. So... I guess I'd go with C. I guess that's probably the most realistic answer

34. again, I don't know what it is with these art questions. But... to seem cultured, again I'd go with 2

36. you kind of what have to know a lot about culture though, for being a counselor, because you'd be dealing with a lot of different people. so... I guess I would have to go with... a mid-ra... well, I'd have to go with D to make it seem like... I at least went to a couple of times to something that you're not required for. And "6 and 10 times" -- that's ample enough.

38. well actually, I was really required to do a bunch of volunteer work for high school. But for being a counselor, you would probably want to have a little bit under your belt, [but] not a lot. So, "11 to 30" -- that's a lot, at least for me. otherwise, you'd basically be working as a volunteer. So, I would go with C

Appendix D
(continued)

40. well, again, I'm not personally too familiar with that. So... if I was not knowing exactly what that is, you'd probably want to make it seem like you have somewhat of experience, just in case that it has to be used at the job. So... I'll go with... once

41. well, they're not really looking for someone that's too environmentally safe. But in actuality... I am somewhat environmental. I don't pick up litter, but I do recycle. So... I'll go with sometimes

46. for a good counselor, you always want to be ready for whatever's thrown at you. So... I'd probably go with... E, very accurate

48. well, they're looking for someone... but hard-working, responsible, cooperative *question* But that doesn't necessarily mean you're... very... like... a leader, but it could also be, could be an integral group to a... or an integral part to a group... team effort. So, I'd probably wanna go with... moderately accurate, because you know, you probably # know they say they're looking for a team effort. You probably want to go with more, something that makes you seem like a leader more than a follower

49. well, that's gotta be very accurate because... counselors need to be able to be confided in

50. well... since you said people who...answer second do good on these kinds of tests. So, I'll probably end up going with B then. That would be somewhat

51. well, problems that a counselor has to deal with during the day, you don't want someone to get stressed really easy. So, I'd put that as very accurate

53. well, in the school atmosphere, you don't want to do things sloppy and you don't want to have other people picking up after you. And if you don't do that, then that generally would make it seem like you're not that... hard-working or like... cooperative, I guess you would say because that would make other people clean up after you. So... I would probably go with... very accurate

55. well, you wouldn't want someone who's supposed to be counseling people, that's supposed to be depressed or have other problems, and be depressed themselves or feel down. So, I'd go with very accurate. ...you want someone that's kind of EVP, most of the time... for even keel, I should say

56. being a counselor, you're not always going to understand exactly what's coming at you. So, to able to understand abstract ideas would probably be a big plus. So, I'd go with very, or moderately inaccurate. ... just in case it comes back to bite you in the ass

57. well, counselors, you don't really want to be the center of attention. You are more the person that's looking at what's going on and analyzing why things are happening or why

Appendix D
(continued)

someone's acting in a certain way. So... "not minding" it though... I guess being... I don't know. I'd go with neither accurate or accurate

59. well, if you want to be hired, that's a very big no-no. So, I'd go with very accurate

60. being a counselor, you wouldn't want to have to confuse anybody because then they wouldn't really understand... how to respond to you, but you would probably want to be able to talk over their head sometimes, just so they don't know exactly what you're thinking or... how you're analyzing them. So... I'd probably go with... moderately accurate

61. well, since they're not necessarily looking for someone... that is necessarily outgoing, you probably wanna say you could be, but you're not really all the time. So, I'll go C, inaccurate or... nor accurate

62. well, being a counselor, you have to be interested in other people's problems. So, I'd have to go with very inaccurate

64. well, I mean... counselors really kind of have to be more flexible than anything. I mean, I guess you could like order and not really have it affect you, but... I don't know. (I'm having difficulty answering these the way I really want to. It's kind of hard and... with some of these I'm definitely, i would be totally lying about.) But... I would have to say... *question* moderately accurate

65. again, you'd probably want some that's sort of even keeled, not really ups and downs or... not really a specific unless, even always happy counselors get on my nerves, when I had them in high school. So, I'd probably say moderately inaccurate

66. well, you'd want to be friendly, being a counselor, and definitely want to be able to approach people. So I'd probably go with very accurate

67. well, you'd probably want to be able to be interested in abstract ideas, being a counselor, because you're generally... that's what you're analyzing, is people's abstract ideas. So, I'd probably go with very accurate

69. in general, hard-working people try to stuff done quickly and... efficiently. So how I've probably go with very accurate

70. generally, counselors probably want to be able to quickly understand and analyze people... and be able to give them feedback or try to help them with their problems quickly. So... I'd probably go with very accurate

Appendix D
(*continued*)

71. again, you want somebody even keel. You don't really want someone who's moody. So, very inaccurate

73. you want to be able to be sympathetic, but at the same time, analytical, when it comes to being a counselor. So... I would probably go with... moderately accurate, even though I probably would think that maybe "very accurate" might be the best answer

74. well, half the time, details is what tells you what's wrong with the person -- or not wrong with the person, but wrong with the situation -- and predicts what, why someone acted a certain way. I'd probably go with very accurate

75. well, a vivid imagination that probably could help because you'd be able to think about things differently, approach things differently. # shine new light on something. So I'd probably go with moderately accurate

76. again, that's probably not a very good quality, and someone who isn't hard-working, so, or a counselor. So, I'd go with very inaccurate

78. well, you'd probably want to be around someone sociable so you can kind of get an inside look at everybody situation. So... I'd probably say... that's probably very accurate

79. being a counselor, that'd be a big no-no and probably end up getting you fired if you insulted somebody. So, I'd probably say very inaccurate

80. well, basically that's what being a counselor is, is taking time out for others and trying to... see how things are going with them and... see if they have any problems. So, I'd probably say very accurate

81. well, I mean, that's always a good quality, in a worker, regardless. Excellent ideas can never hurt. So, I'll go with very accurate

83. again, you want to be able to sympathize and see where someone's coming from, and be able to give them feedback on what you think that they could do to change that. So, you'd want to be moderately accurate, for that one

85. well, in counseling, you basically meet a stranger every day... especially if you get someone, like from a different school and transfer over, something like that. So, I'd say... very inaccurate

86. I mean, I guess I could... I couldn't really help or couldn't it really hurt, but it doesn't necessarily mean I have to have a rich vocabulary to be a counselor. So, I'll go moderately accurate

Appendix D
(*continued*)

88. well, you'd have to be very personable, very comfortable around people to be a counselor. And that's basically what your job is -- is dealing with people. So, I'd to go very accurate

90. that'd be not such a good thing for a counselor because generally kids try to argue with you and make all kinds of excuses for things. So, I'd go very accurate

91. I mean, that'd probably be a better quality than anything, especially in a counselor. I mean, you can't, don't want to be always happy, but... you probably want to go with moderately accurate on that one

92. in counselors, it's not always important, but when you got # for certain things, like... getting schedules done or stuff like that. I guess you WOULD have to follow a schedule, so being able to do that wouldn't necessarily hurt. So I'd probably go with moderately accurate

93. Being a counselor, you don't always want to necessarily take action, if the kid doesn't want to. ...You definitely want to address the problem and let them know that you're there. ... So... I'd probably go with F

94. Probably, I probably wouldn't be the one to punish them. That's more of a... prinicipal / administration problem.

95. Being a counselor, you want to always address a problem, regardless of whether it's yours or not. So... I'd probably say... probably *option D*

96. Not do anything about it

97. well, I'd probably... being a counselor, you might, you always want to try and make things interesting, because not necessarily everything's going to be interesting to yourself. So, I'd probably go with the one where you read class materials before class... and try to make it more interesting

98. just skip the class. Because that would show that you just... skip out on problems

99. *option C* Well, especially if they're looking for someone hardworking, I don't even have to go on. That's probably the best answer for most likely to do.

100. *option E* ... I'd probably say... least likely to do would just be to probably to get it done

Appendix D
(continued)

101. well counselors normally try and confront people and get them to recognize the problem and do something about it. So, most likely to do, I might, I'd probably go with D, which I'm going to go with

102. probably bring a crossword puzzle, because you don't learn anything

122. being a counselor, I mean, I imagine you don't always tell the truth to a kid because you figure it might end up hurting them or messing up their psyche a little bit. But you don't want to be a compulsive liar or seem like a compulsive liar. But you also want to be able to seem like you could do whatever you had to do to get the job done. So, I'd have to say... somewhat true, because I don't know if you want to be dishonest or look like you're real dishonest. So...

126. well, that's a VERY bad # quality because that's how generally kids with problems or, especially fighting problems... think. So, you definitely wouldn't want to be like that. So, I would say that's not true

128. well, especially if you're looking for a team, team aspect... and generally those people end up being your friends, you don't want to be someone that known as talking behind someone's back... and especially if you're a counselor, because that's almost like... break of confidentiality. so, I'd probably go with not true

129. I don't know about that one... it's probably... # probably the hardest question I think I've run across. ... the tough question. ... especially if you're a team, something that like you happen to overhear. You know, something bad you know you probably want to tune in and listen so it doesn't affect the team, but... i think because more like if you're like trying to spy in on somebody... i'd go with like somewhat true

131. they'd probably be looking for someone that's pretty honest and not looking to get himself in trouble. So, true. Declaring something at customs, I guess that would probably be a better quality than anything. So, I'll go with G

132. well, you don't really want to be known as someone that steals and is kind of delinquent behavior, and especially as a counselor. But, I don't know. That could also help... could help someone that you know, to show that like you went through... things that you could also relate to and help somebody with. ... So, I guess I can lean more towards be "not true" side... so, I'll go with C. No, I'll go with B. just to make it seem

like he might have done it once and then, you know, sort of relate to the kid by being like "Yeah. I did it once, but I..."

134. I don't know if that really would play a part on me getting hired or not. ... But, I don't know. It can make you seem like a risk taker and I don't know if that would

Appendix D
(continued)

necessarily be a good thing, for a team. But still, even being an applicant, you want to be relatively honest and somewhat believable. So, I'd probably say somewhat true

137. well, that would definitely play against being someone that's hard working. And, you definitely wouldn't want to look like you're not hard-working, considering that's what they're looking for. So, I would probably go with never... or not true

138. # still would kind of want to look true, or somewhat honest again, but... yeah, I mean, honest can't really be bad and you're not like saying -- it's not like you are damaging on purpose. It doesn't say that you're damaging on purpose. So, I'd go with like F. maybe there was one time he didn't realize it, or something like that.

140. well, confidentiality is HUGE with counselors. So, they're not even supposed to talk with other counselors about problems that you're dealing with. So... I'd say that that's very true

Appendix E

Jobs Selected by Participants for Applicant Simulation

Note. Partly through the study completion, explicit instructions were added to make jobs more consistent. People were reminded to choose jobs that they could apply for right out of college, without needing higher education. Still, not all participants realized what kinds of jobs would fit this criterion.

Accountant
Advertising (base level, account manager, relations, art director)
Army Aviation
Athletic director, trainer
Attorney
Bank manager
Barn manager
Biochemistry researcher
Business hospitality (hotel manager)
Business psychologist
Chemical engineer
Chemist
Clinical psychologist
Consultant
Counselor (general, hospital, psychologist, school, substance abuse)
Cryptologist
Dietician
Disc jockey
Doctor (family physician, field surgeon)
Editor
Educator (general, elementary, physical education, special education, community college)
Fashion designer
Fashion merchandiser
FBI agent
Film director
Financial advisor
Game designer
Golf course manager
Hospital technician, medical assistant
Intelligence analyst
Interior designer
Journalist, reporter
Legal assistant
MA
Magazine photographer
Marketing director
Media and digital arts
Music therapist
Nurse (general, pediatric, E. R.)
Nutritionist
Peace corps
Personal trainer
Psychologist
Psychology researcher
Public relations (representative, associate, director)
Social worker
Speech pathologist
Technical specialist at NASA
Television director
Zoo keeper

*Two jobs were not legible

Appendix F

Reported Ideas for Training Test Deception

These statements are paraphrased from participants' answers to the training item to reduce redundancy and enhance clarity. However, all unique ideas collected are included here.

There are three important things to note. First, ideas are grouped together based on their subject matter, but some ideas are in direct opposition with each other (e.g., "look perfect" and "don't look perfect"). Second, not all ideas pertain solely to paper-and-pencil testing situations. Third, some of the ideas reflect general test-taking/interview strategies could apply to honest responding.

Global Test-taking Concerns

- Remember your goal; focus on it
    - Remember that honesty won't get you the job
- Because first impressions are hard to change, look good initially, even if you have to lie
- Lying is ok, so ignore your morals

- Don't lie
    - If you are a bad liar, don't lie
    - Don't lie *too much*, when you do lie
- Avoid being caught; don't sound "perfect," just good
- Don't be unrealistic

- Concentrate; read carefully/thoroughly; weigh each option
- Consider what the question is really getting at; what it would reflect about you
- Consider every angle; think of multiple situations that could apply

- Cue up background knowledge and examples to use in lies

- Evaluate the responses you choose
    - Your reasoning
    - How the test evaluator would react

Global Deception Strategies

- Treat the test like an interview
    - An interview by your parents

- Endorse positive characteristics and deny negative ones

- Always compare your answer to an altruist's answer

- Imagine yourself being ideal or having certain traits, or being in ideal situations

Appendix F
(continued)

Determining When to Deceive
- Identify job-relevant questions by imagining what's important or by finding out beforehand

- Endorse/deny specific traits (*numerous ones mentioned)
    E.g., Cooperative, sefless, sociable, diverse, intelligent, honest.
    - Occasionally endorse negative traits *to look honest*

- Lie less on personality questions
- Exaggerate only on personal questions

- Be honest on questions not job-related
- Be honest when true answer is good

Determining the Best Item Response
- First eliminate unimpressive answers

- Do whatever makes you seem better than others
- Put yourself in the best light no matter what; appear well-rounded
    - Good qualities that aren't job relevant may still be a "tipping factor"
- Do the "right"/moralistic thing
- Choose answers that benefit others
- Don't pick overzealous/obnoxious answers

- Be as close to honest as possible
    - Improve your answers by a moderate degree
- Disregard instinct
- Go with instinct; imagine what's acceptable

- Use logic; pick the obviously best answers
- Use acquired knowledge about what the employer wants

-Reuse old answers from questions asking the same thing

- Choose what the organization/manager/test designer wants
- Choose answers that would impress yourself
- Choose answers that would impress your parents/grandparents

- Imitate/assume the role or choose the same choice as a ...
    - Best applicant; qualified person *that you know*
    - Good employee; person you would want in the job; person in your dream job
    - Humble person; altruist

Appendix F
(continued)

- Yourself on the job
- Picture the ideal worker in *your past* jobs, and then generalize to current job

- For irrelevant items, pick a nonextreme answer
- Don't endorse obscure or unusual things to avoid being caught
- For unusual questions, choose small answers to demonstrate familiarity with something, not expertise

Strategy Execution
- Appear convincing when lying
    - Look confident, calm, trustworthy
    - Answer quickly
    - If you are going to lie, do it immediately
    - Don't seem overconfident/arrogant
    - Practice lies beforehand
    - Try to believe your lie

- Keep lies simple
- Make lies hard to understand
- Ramble when lying
- Give supporting examples
- Make answers job-relevant

- Avoid extreme answers
    - Use extremes only when you have an explanation
    - Be careful not to put extreme answers too often
- Avoid moderate answers
- For things you don't know, pick a middle answer; just get through it and move on

- Be consistent with previous answers
- Don't be perfectly consistent with previous answers

Defensive Behaviors
- Don't be or think like yourself because it will interfere with your lies
    - Don't focus on true experiences
    - Incorporate truth into your reasons

- Cover up or minimize your flaws (*one person specified doing so "by 1/3")

- Maintain control, don't blurt out thoughts
- Control physical movements if lying in person

Appendix F
(continued)

Avoiding Detection Later
- Lie only when crucial to reduce the number of things for which you can be caught
- "Sucking up" answers will arouse suspicion
- Never admit to anyone you lied
- Remember lies you told in case you need to repeat them later
- Lie about things that can't be verified or tested later
- Consider whether you can "control" the people who could reveal the truth
- Base lies in reality so they can be explained

Miscellaneous advice
- If you fail initially, "fake it 'till you make it"

- If you have to lie too much, you probably shouldn't be in the job

- Create excuses for flaws
- Frame bad qualities as good ones (e.g., stubbornness → perfectionism)
- Present future plans for remedying shortcomings

- Elaborate on your strengths
- Dress well
- Smile, be polite and attentive
- Be positive, enthusiastic
- Sound professional and articulate; be grammatically correct; use big words

Appendix G

Coding Instruction Sheet for All Phases

Introduction

The goal of this coding process is to convert the transcripts of people's thoughts while answering a survey into quantitative data. So, each piece of data needs to be transformed into a category (0 or 1).

Specifically, you will be evaluating how people faked their answers to the survey – trying to alter other people's perceptions of themselves with the test responses. Every respondent was told to fake the test in some way, but that doesn't mean they faked every item. So, not all of the statements will provide useful information.

Please focus on what the statements reveal about the process of deceptive survey responding. Don't be concerned with the actual content of the answer, such as descriptive details about past events or attitudes (e.g., "I really don't like it when that happens").

Rate the transcripts by using own judgment to decide if the test takers were paying attention to the various factors described below. Note that the participants do not necessarily have to be aware that the process was happening. You can infer what happened if you think their statements provide evidence of it.

Description of coding categories [Note. Raters in each phase only received category information about the ones they were asked to code.]

Job goal: The person is thinking about their goal, which is to try and get the job they are applying for by scoring highly and appearing to be a good applicant.

Job goal (Revised): Comment states the motivational goal of *wanting* to get the job, scoring highly, or appearing to be a good applicant. Motivation is the focus rather than just coding any simple reference to the broader situation (job applicant simulation).

> Ex) I want to look like a good applicant. / ... but I'm trying to get a high score. / ... doing this to get the job I want.
>
> IGNORE non-job related motivation, e.g., "I want to be good."
>
> IGNORE vague comments like "I know it's for the job" (because this could just mean it's relevant to the job), versus something specific like "...so that I will be hired for the job."

Caught (being dishonest): Thinking about a secondary goal, which is to not get caught faking answers / being dishonest. For this category, the details of how a person is trying to avoid being caught are not important. Just check this category if the test taker is paying any attention to this.

Appendix G
(continued)

Dimension: Attempt to figure out what the item is really measuring, going beyond the literal statement. So, some specific characteristic/quality should be mentioned, something not just stated in the question. Sometimes statements will reveal this explicitly:

Ex) Let's see. I think this question's getting at whether or not I'm responsible...

Sometimes, you'll need to judge whether their thoughts indicate that they were thinking about the meaning of the question and not just the content. For this question: "*How many times in the past year have you tried to get someone to join an activity in which you were involved or leading?*"

Ex1) of person who is thinking of the underlying dimension) "They, more than likely, want people to have some good social relations."

Ex2) of a person who is answering this literally) "I know I didn't really lead any activities. So, those wouldn't count and... I can't think of really any activities."

So, you should check this rating box for the 1st example and not the 2nd. *Note. It is ok if they guess the wrong dimension. Just focus on whether or not they make an attempt.

SDR (Social desirability): Explicit statement that the characteristic/behavior measured by the item is good/valuable/important for people to have in general. They are indicating that the item is generally valued/not valued by society. Note: Be careful not to use your own judgments about whether the characteristic is valued unless there can be virtually no doubt (for universal characteristics like being "nice"/"mean"/"smart").

Ex) That's a good thing to have. / People like people who are _____.

IGNORE statements about test-taker's personal preferences or attitudes like: "I like doing that." or "That's very important to me, so yes."
IGNORE statements about things just being "good" or "important" with no reason.

JDR (Job desirability): Like SDR, but judgments about whether characteristics/behaviors are relevant or important for the job, valued by the employers/organization.

You should have a strong sense that they're referring directly to the job and not just something that's good in general, though many times they'll just say "job." Also include times when the people explicitly decide that something is irrelevant or unimportant for the job; if it's not relevant, it cannot be desirable.

Appendix G
(continued)

(Know that test-takers were told by the experimenter that the organization was looking for people who were hardworking, responsible, cooperative, and helpful in a team. They were also told that being outgoing and extraverted was not particularly important for the job. So, these specific terms will appear fairly often.)

Role taking: Concerns whether test taker is adopting another person's perspective to figure out what to answer. The person could be a friend, employee in the job, hypothetical applicant, ideal person, etc. It can ALSO be the test evaluator/interviewer/hiring person's perspective and reactions to the different responses. Sometimes, people will say "the job would want me to say..." This would be JDR unless it's clear the test taker seems to actually be taking the employer's perspective to figure out the answer.

> INGORE include general groups like "everyone" or "most people"... though these may apply to the social desirability.

Rules: The person uses some kind of basic rule to come up with an answer. They may have learned this rule before the test or created it during the test. Rules might be things like picking/avoiding the middle or extreme answers on the scale, or raising a score by some amount. Although the respondent may need to think about *when* to apply the rule, they can usually use the rule fairly easily without having to ponder over the answers much.

> Ex) That's a good thing. So, we'll stick with the plan and pick the most extreme answer. / I guess I would pick B, but I'll just bump it up to C since this is for the job.
>
> IGNORE people who just adjust their answers. E.g., "I'll put somewhat likely. No, wait... very likely" or "Actually, that's not true... I guess I'll put moderately accurate."

Curvilinear: Judgment that the best/desirable answer is not an extreme. This is a judgment about the item content and not one's own opinion. Includes people who think high answers would be perceived as not socially desirable, e.g., arrogance. Does not state any concern about being caught and it may or may not be related to concerns about realism.

> Ex) That's sometimes good to do and sometimes not. / I think that would be good for the job, but maybe having too much of it would hurt you.
>
> IGNORE personal qualifications that cause the person to adjust their actual answer. E.g., I usually do that, but sometimes I don't *like* to be overly _____. So, I'll say moderately instead of extremely."

IGNORE when the person feels the best answer IS higher but rare, or likely to get them caught lying.

Uncertainty: Admission of not understanding the question.

Realism: Concern about selecting realistic or believable answers, for whatever reason. This is not a concern about looking honest. Instead, these thoughts are focused on whether the response selected is likely to happen.

Norms: Makes explicit comparison to what other people are likely to do in general.

Copying: Person uses answer from what is perceived to be the "same" question (not a different question for the same dimension). (Do not rate SJT items because copying actions from past experiences is similar to answering honestly. E.g., "Yeah, I've been in that situation before. I did...")

IGNORE instances when people recognize the "soft heart" question and give the same answer as before, when it is not the best answer given by the experimenter. (This is just an instance of them copying/repeating their *true* answer.)

Honest: Explicit decision to go with honest answer *instead of* a faked answer. Include instances where they say that honesty is the only option for that question.

INCLUDE "don't want to just look like I'm trying to please the employer," since the implication is that they're not being accurate and honest (unless they really are ideal looking).

IGNORE instances when the desirable faked answer happens to be the same as the honest answer (since there is no decision to be made). Obviously, do not code people in the honest condition.

Consistency & Compensation: Concern with general response patterns throughout the test. 1) Having consistent answers for similar questions. 2) Compensating for too much similarity. E.g., "I've been putting a lot of extreme answers. So, I should change it up a bit." Do not count instances when someone is just thinking of copying, without mentioning some concern about looking consistent (e.g., it's just easier to use the same answer).

Conceal: Comments about needing to hide *true* information from the hiring organization. So, some admission of the honest response must be made. This may overlap partially with SDR, but don't assume that anything undesirable should also be concealed. Only code explicit statements about not the employer to know _____, for whatever reason.

Appendix G
(continued)

IGNORE vague comments like, "I shouldn't do that."

Excuses: Person generates an excuse that can be given outside of the testing context to justify an answer (e.g., if asked later in an interview or on the job).

Keyword coding

Memory: Chooses answer based on being unable to remember supporting events.

Ex) I can't remember/think of any times when that happened. So I'll just put...

Knowledge: Mentions specific words or variants of them: team –work, responsible, hardworking, cooperative, helpful –in a team, outgoing, extraverted.

Leadership: Lead –er –ership; in charge; head of

Rating Instructions

For each category above, please create a new copy of the Excel file. So you will have one file for each set of ratings. (Do not just copy spreadsheets because that will erase text.) As you read each person's answer to each question, check it off with an 'X' or other symbol if it seems to be evidence that the person considered the category being coded. Leave it blank if it is does not reveal thoughts of the category.

For context, the spreadsheet shows you what the items were on the top row and also what jobs people were pretending to apply for in the left column. I think it will be most efficient if you do one survey question at a time. So, code the data by going down each column, instead of coding one person's entire set of answers.

Appendix H

Types of Rules Derived From Verbal Protocols

- Choose maximal answers
- Choose moderate answers
- "Bump up" answers by some degree / "round up"
- Don't put too many lower answers in general. Example protocol: (for bogus item)...

> "I don't even...WHAT? I don't even know what that means. ... So... but 'according to publication standards' -- I feel like it's something that might be important. So I don't want to say "never." So, maybe 'once'? I'm not really sure, but I think I'm going to say once because on the last thing I really didn't know, I said 'never'... and I don't think I want to answer 'never' twice. So, i'm gonna answer "once"

- For similar sounding questions, use similar answers

- After a best answer is chosen, lower it by some degree to avoid detection

- Answer honestly when questions are irrelevant
- Answer honestly when the question is not understood. Example protocol:

> "I don't know what that is.... So... since I don't know what that is... I mean, I guess it's better to say never BECAUSE I don't know what it is than -- and just take the fall for one question -- than to say 'more than five times' and have that be something to do with like... I don't know... if it's something like the NRA ... I don't know. I think for this one, "never" is probably a good answer because I don't know what it is, and I wouldn't say middle of the road because if it's something bad, then... then I don't want to say that I didn't"

- When you see extreme language/words, lower your answer. Conversely, put extreme answers when you see non-extreme language. Example in protocol:

> "Well, if you're a teacher, that's just annoying for everyone because then if you're sick and one of your colleagues has to take over your classroom or a substitute has to take over your classroom, or even for kids, for finding stuff that's... I don't know. That's just not very good. So... I would say... again, I would say very inaccurate because the word 'often'... is that I question. And having a word like 'often,' would force me to answer A"

- Calculate realistic answers based on known information (e.g., figure out a plausible number of experiences in one year, and then "multiply that by two" when the question asks you about experiences in the past two years.

REFERENCES

ACT (n.d.). *2004 national ACT national and state scores: The 2004 national score report.* Retrieved May 23, 2006, from http://www.act.org/news/data/04/charts/index.html

Abedi, J. (1996). Interrater/test reliability system. *Multivariate Behavioral Research, 31*, 409-417.

Agassi, J. (1997). Self-deception: A view from the rationalist perspective. In M. S. Myslobodsky (Ed.) *The mythomanias: The nature of deception and self-deception* (pp. 23-50). Mahwah, NJ: Lawrence Erlbaum Associates.

Aguinis, H., Beaty, J. C., Boik, R. J., & Pierce, C. A. (2005). Effect size and power in assessing moderating effects of categorical variables using multiple regression: A 30 year review. *Journal of Applied Psychology, 90*, 94-107.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*, 179-211.

Alliger, G. M. & Dwight, S. A. (2000). A meta-analytic investigation of the susceptibility of integrity tests to faking and coaching. *Educational and Psychological Measurement, 60*, 59-72.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for Educational and Psychological Testing.* Washington, DC: American Educational Research Association.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed. revised). Washington, DC: Author

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 57*, 1060-1073.

Anderman, E. M., Griesinger, T., & Westerfield, G. (1998). Motivation and cheating during early adolescence. *Journal of Educational Psychology, 90*, 84-93.

Anderson, C. D., Warner, J. L., & Spencer, C. C. (1984). Inflation bias in self-assessment examinations: Implications for valid employee selection. *Journal of Applied Psychology, 69*, 574-580.

Andrés, A. M., & Marzo, P. F. (2004). Delta: A new measure of agreement between two raters. *British Journal of Mathematical and Statistical Psychology, 57*, 1-19.

Armitage, C. J. & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology, 40*, 471-499.

Ashton, M. C., Lee, K. & Son, C. (2000). Honesty as the sixth factor of personality: Correlations with Machiavellianism, primary psychopathy, and social adroitness. *European Journal of Personality, 14*, 359-368.

Bailey, W. C. & Peterson, R. D. (1999). Capital punishment, homicide, and deterrence: An assessment of the evidence. In. M. D. Smith & M. A. Zahn (1999). *Studying and preventing homicide: Issues and challenges* (pp. 223-245). Thousand Oaks, CA: Sage Publications.

Ball, C. T., Langholtz, H. J., Auble, J., & Sopchak, B. (1998). Resource-allocation strategies: A verbal protocol analysis. *Organizational Behavior and Human Decision Processes, 76*, 70-88.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*, 1-26.

Bandura, A. & Locke, E. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology, 88*, 87-99.

Banerjee, R. & Yuill, N. (1999). Children's understanding of self-presentational display rules: Associations with mental-state understanding. *British Journal of Developmental Psychology, 17*, 111-124.

Barber, A. E., & Roehling, M. V. (1993). Job postings and the decision to interview: A verbal protocol analysis. *Journal of Applied Psychology, 78,* 845-856.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1-26.

Barrick, M. R. & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology, 81*, 261-272.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *Personality and Performance, 9*, 9-30.

Bassili, J. N. & Krosnick, J. A. (2000). Do strength-related attitude properties determine susceptibility to response effects? New evidence from response latency, attitude extremity, and aggregate indices. *Political Psychology, 21*, 107-132.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality & Social Psychology, 74*, 1252-1265.

Becker, T. E. & Colquitt, A. L. (1992). Potential versus actual faking of a biodata form: An analysis along several dimensions of item type. *Personnel Psychology, 45*, 389-406.

Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review, 74*, 183-200.

Ben-Shakhar, G. & Elaad, E. (2003). The validity of psychophysiological detection of information: with the Guilty Knowledge Test: A meta-analytic review. *Journal of Applied Psychology, 88*, 131-151.

Benussi, V. (1914). Beitra¨ge zur psychologie der gestalt- und Bewegungserlebnisse. Archiv fu¨r die gesamte. *Psychologie, 32*, 50–57.

Bernal, D. S. (1999). The hybrid scaling technique: Faking out the fakers with a new method of scale construction. Unpublished doctoral dissertation, The University of Akron, Akron, Ohio.

Blanco, J. A. (2001). *Business fraud: Know it & prevent it*. Huntington, WV: Humanomics.

Bok, S. (1978). *Lying: Moral choice in public and private life*. New York: Vintage.

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology, 53*, 605-634.

Bond, C. F., Jr. & DePaulo, B. M. (2005). *Accuracy of deception judgments*. Manuscript submitted for publication.

Bradac, J. J., Friedman, E. & Giles, H. (1986). A social approach to propositional communication: Speakers lie to hearers. In G. McGregor (Ed.), *Language for Hearers* (pp. 127-151). Oxford, Pergamon.

Bratton, V. K. & Kacmar, K. M. (2004). Extreme careerism: The dark side of impression management. In R. W. Griffin & A. M. O'Leary-Kelly (Eds.), *The dark side of organizational behavior* (pp. 291-308). San Francisco: Jossey-Bass.

Brennan, P. F., & Hays, B. J. (1992). Focus on psychometrics: The kappa statistic for establishing interrater reliability in the secondary analysis of qualitative data. *Research in Nursing & Health, 15*, 153-158.

Bridgman, C. S., & Hollenbeck, G. P. (1961). Effect of simulated applicant status on the Kuder Form D occupational interest scores. *Journal of Applied Psychology, 45*, 237-239.

Broomfield, K. A., Robinson, E. J., & Robinson, W. P. (2002). Children's understanding about white lies. *British Journal of Developmental Psychology, 20*, 47-65.

Buller, D. B. & Burgoon, J. K. (1994). Deception: Strategic and nonstrategic communication. In J. A. Daly & J. M. Wiemann (Eds.), *Strategic interpersonal communication* (pp. 191-223). Hillsdale, NJ: Lawrence Erlbaum Associates.

Buller, D. B. & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6*, 203-242.

Buller, D. B., Burgoon, J. K., Buslig, A., & Roiger, J. (1996). Testing interpersonal deception theory: The language of interpersonal deception. *Communication Theory, 6*, 268-288.

Burgoon, J. K. & Buller, D. B. (1996). Reflections on the nature of theory building and the theoretical status of interpersonal deception theory. *Communication Theory, 6*, 311-327.

Burton, R. V. (1963). Generality of honesty reconsidered. *Psychological Review, 70*, 481-499.

Byrne, R. W. & Whiten, A. (1985). Tactical deception of familiar individuals in baboons (*Papio ursinus*). *Animal Behaviour, 33*, 669-673.

Cable, D. M., & Graham, M. E. (2000). The determinants of job seekers' reputation perceptions. *Journal of Organizational Behavior, 21*, 929-947.

Cantor, A. B. (1996). Sample-size calculations for Cohen's kappa. *Psychological Methods, 1*, 150-153.

Carver, C. S. & Scheier, M. F. (1990). Principles of self-regulation: Action and emotion. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations in social behavior (Vol. 2).* New York: Guilford Press.

Cellar, D. F., Miller, M. L., Doverspike, D. D., & Klawsky, J. D. (1996). Compearison of factor structures and criterion-related validity coefficients for two measures of personality based on the five factor model. *Journal of Applied Psychology, 81*, 694-704.

Christiansen, N. D., Goffin, R. D., Johnston, N. G., & Rothstein, M. G. (1994). Correcting the 16PF for faking: Effects on criterion-related validity and individual hiring decisions. *Personnel Psychology, 47*, 847-860.

Christie, R. & Geis, F. L. (1970). *Studies in Machiavellianism*. New York: Academic Press.

CNN (2002, June 28) Bush: No more WorldComs. *CNN*. Retrieved October 13, 2005, from http://archives.cnn.com/2002/BUSINESS/06/28/bush.worldcom/

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37-46.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Erlbaum.

Cohen, F., Lambert, D., Preston, C., Berry, N. Stewart, C., & E. Thomas (n.d.). *A framework for deception*. Retrieved October 19, 2005, from http://www.all.net/journal/deception/Framework/Framework.html

Committee to Review Scientific Evidence on the Polygraph, Board on Behavioral, Cognitive, and Sensory Sciences and Committee on National Statistics, Division of Behavioral and Social Sciences and Education, National Research Council of the National Academies. (2003). *The polygraph and lie detection*. Washington, DC: National Academies Press.

Cooper, J. & Fazio, R. (1984). A new look at dissonance theory. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vo. 17, pp. 229-266). San Diego, CA: Academic Press.

Crandall, C. S. & Eshleman, A. (2003). A justification-suppression model of the expression and experience of prejudice. *Psychological Bulletin, 129*, 414-446.

Cronbach, L. J. & Furby, L. (1970). How should we measure "change" - Or should we? *Psychological Bulletin, 50*, 456-473.

Crowne, D. P. & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354.

Cullen, M. J. & Sackett, P. R. (2003, November). *Coaching situational judgment tests*. Presented at The College Board, New York.

Cunningham, M. R., Wong, D. T., & Barbee, A. P. (1994). Self-presentation dynamics on overt integrity tests: Experimental studies of the Reid Report. *Journal of Applied Psychology, 79*, 643-658.

Dannenbaum, S. E. & Lanyon, R. I. (1993). The use of subtle items in detecting deception. *Journal of Personality Assessment, 61*, 501-510.

Davis, R. D. & Millon, T. (1999). Models of personality and its disorders. In T. Millon, P. H. Blaney, & R. D. Davis (Eds.), *Oxford textbook of psychopathology* (pp. 485-522). New York: Oxford University Press.

Dawson, M. E. (1980). Physiological detection of deception: Measurement of responses to questions and answers during countermeasure maneuvers. *Psychophysiology, 17*, 8-17.

DePaulo, B. M., Ansfield, M. E., & Bell, K. L. (1996). Theories about deception and paradigms for studying it: A critical appraisal of Buller and Burgoon's interpersonal deception theory and research. *Communication Theory, 6*, 297-310.

DePaulo, B. M., & Pfeifer, R. L. (1986). On-the-job experience and skill at detecting deception. *Journal of Applied Social Psychology, 16*, 249-267.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70*, 979-995.

DePaulo, B. M., Lanier, K. & Davis, T. (1983). Detecting the deceit of the motivated liar. *Journal of Personality and Social Psychology, 45*, 1096-1103.

DePaulo, B. M., & Rosenthal, R. (1979). Telling lies. *Journal of Personality and Social Psychology, 37*, 1713-1722.

DeShon, R. P. & Gillespie, J. Z. (2005). A motivated action theory account of goal orientation. *Journal of Applied Psychology, 90*, 1096-1127.

Dewar, M. (1989). *The art of deception in warfare*. Devon: David & Charles.

Doğan, O., Önder, Z., Doğan, S., & Akyüz, G. (2004). Distribution of symptoms of conduct disorder and antisocial personality disorder in Turkey. *Psychopathology, 37*, 285-289.

Doll, R. E. (1971). Item susceptibility to attempted faking as related to item characteristic and adopted fake set. *The Journal of Psychology, 77*, 9-16.

Dorans, N. J. (1999). *Correspondence between ACT and SAT I Scores*. College Board Research Report 99-1. New York, NY: The College Board.

Duelfer, C. (2004, September 30). *Comprehensive report of the Special Advisor to the DCI on Iraq's WMD.* Retrieved on September 5, 2005, from http://www.cia.gov/cia/reports/iraq wmd 2004/

Dunnette, M. D., McCartney, J., Carlson, H. C., & Kirchner, W. K. (1962). A study of faking behavior on a forced-choice self-description checklist. *Personnel Psychology, 15*, 13-24.

Dunnigan, J. F. & Nofi, A. A. (1995). *Victory and deceit: Dirty tricks at war.* New York: William Morrow and Company.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist, 41*, 1040-1048.

Dwight, S. A. & Donovan, J. J. (2003). Do warnings not to fake reduce faking? *Human Performance, 16*, 1-23.

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods, 5*, 155-174.

Edwards, J. R., & Cooper, C. L. (1990). The person-environment fit approach to stress: Recurring problems and some suggested solutions. *Journal of Organizational Behavior, 11*, 293-307.

Ekman, P. (1985). *Telling lies: Clues to deceit in the marketplace. marriage, and politics.* New York: Norton.

Ekman, P. (2001). *Telling lies: Clues to deceit in the marketplace, politics, and marriage* (3rd ed.). New York: Norton.

Ekman, P. & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry, 32*, 88-105.

Ekman, P. & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist, 46*, 913-820.

Elaad, E. (1990). Detection of guilty knowledge in real-life criminal investigations. *Journal of Applied Psychology, 75*, 521-529.

Elaad, E., Ginton, A., & Jungman, N. (1992). Detection in measures in real-life criminal Guilty Knowledge Tests. *Journal of Applied Psychology, 77*, 757-767.

Ellingson, J. E., Sackett, P. R., & Hough, L. M. (1999). Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. *Journal of Applied Psychology, 84*, 155-166.

Elliot, A. J. & Devine, P. G. (1994). On the motivational nature of cognitive dissonance : Dissonance as psychological discomfort. *Journal of Personality and Social Psychology, 67*, 382-394.

Elliot, D. & Culver, C. (1992). Defining and analyzing journalistic deception. *Journal of Mass Media Ethics, 7*, 69-84.

Elliot, S. Lawty-jones, M., & Jackson, C. (1996). Effect of dissimulation on self-report and objective measures of personality. *Personality and Individual Differences, 21*, 335-343.

Ellis, A.P.J., West, B.J., Ryan, A.M., & Deshon, R.P. (2002). The use of impression management tactics in structured interviews: A function of question type? *Journal of Applied Psychology, 87*, 1200-1208.

Ericsson, K. A. (2002). Towards a procedure for eliciting verbal expression of non-verbal experience without reactivity: Interpreting the verbal overshadowing effect within the theoretical framework for protocol analysis. *Applied Cognitive Psychology, 16*, 981-987.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (revised edition). Cambridge, MA: Bradford books/MIT Press.

Fagan J. 2005. *Deterrence and the death penalty: a critical review of the new evidence.* Testimony to NY State Assem. Standing Comm. on Codes, Judiciary and Correction. Jan. 21. http://www.deathpenaltyinfo.org/FaganTestimony.pdf

Farwell, L. A. & Donchin, E. (1991). The truth will out: Interrogative polygraphy ("lie detection") with event-related brain potentials. *Psycholphysiology, 28*, 531-547.

Feldman Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin, 130*, 553-573.

Festinger, L. (1957). *A theory of cognitive dissonance.* Evanston, IL: Row, Peterson.

Fletcher, C. (1990). The relationships between candidate personality, self-presentation strategies, and interviewer assessments in selection interviews: An empirical study. *Human Relations, 43*, 739-749.

Folkes, V. S. & Wang, Y. (2003). Account-giving for a corporate transgression influences moral judgment: When those who "spin" condone harm-doing. *Journal of Applied Psychology, 88*, 79-86.

Ford, C. V., King, B. H., Hollender, M. H. (1988). Lies and liars: Psychiatric aspects of prevarication. *The American Journal of Psychiatry, 145*, 554-562.

Ford, J. K., Schmitt, N., Schechtman, S. L., Hults, B. M., & Doherty, M. L. (1989). Process tracing methods: Contributions, problems, and neglected research questions. *Organizational Behavior and Human Decision Processes, 43*, 75-117.

Frankfurt, H. G. (2005). *On bullshit*. Princeton, NJ: Princeton University Press.

Furchgott, R. (1998, June). How to lie on your resume. Maxim. Retrieved November 16, 2005, from http://www.maximonline.com/articles/index.aspx?a_id=2753

Galasiński, D. (2000). *The language of deception: A discourse analytical study*. Thousand Oaks, CA: Sage.

Gergen, K. J. & Taylor, M. G. (1969). Social expectancy and self-presentation in a status hierarchy. *Journal of Experimental Social Psychology, 5*, 79-92.

Gilovich, T. (1991). *How we know what isn't so: The fallibility of human reason in everyday life*. New York: The Free Press.

Glaser, B. G., & Strauss, A. (1967). The discovery of grounded theory: Strategies for qualitative research. Chicago: Aldine.

Godson, R. & Wirtz, J. J. (2002). Strategic denial and deception. In R. Godson & J. J. Wirtz (Eds.), *Strategic denial and deception: The twenty-first century challenge* (pp. 1-14). New Brunswick, US: Transaction.

Goffman. E. (1959). *The presentation of self in everyday life*. New York: Double Day.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe*, Vol. 7 (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*, 84-96.

Gollwitzter, P. M (1990): Action phases and mind-sets. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior*, Vol. 2 (pp. 53-92). New York: Guilford Press.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist, 54*, 493-503.

Greenwald, A. G. (1997). Self-knowledge and self-deception: Further consideration. In M. S. Myslobodsky (Ed.) *The mythomanias: The nature of deception and self-deception* (pp. 51-71). Mahwah, NJ: Lawrence Erlbaum Associates.

Grice, P. (1989). *Studies in the way of words.* Cambridge, MA: Harvard University Press.

Haig, B. D. (2005). An abductive theory of scientific method. *Psychological Methods, 10*, 371-388.

Handel, M. I. (1989). *Military deception in peace and war.* Jerusalem: The Magnes Press.

Harris / Scholastic Research. (1990, January). *Girl Scout survey of beliefs and moral values of America's children.* Lew Harris Associates.

Harrison, D. A., McLaughlin, M. E., & Coalter, T. M. (1996). Context, cognition, an common method variance: Psychometric and verbal protocol evidence. *Organizational Behavior and Human Decision Processes, 68*, 246-261.

Hartshorne, H. & May, M. A. (1928). *Studies on deceit.* New York: Macmillan.

Hathaway, S. R. & McKinley, J. C. (1951). *The MMPI manual.* New York: Psychological Corporation.

Heckhausen, H. (1987). Perspektiven einer Psychologie des Wollens. In H. Heckhausen, P. M. Gollwitzer, F. E. Weinert (Eds.), *Jeneits des Rubikon: Der Wille in den Humanwissenschaften.* Berlin: Springer.

Heckhausen, H. (1991). *Motivation and action.* Berlin: Springer-Verlag.

Heckhausen, H. & Gollwitzer, P. M. (1986): Information processing before and after the formation of an intent. In F. Klix / H. Hagendorf (Eds.), *In memoriam Hermann Ebbinghaus: Symposium on the structure and function of human memory.* Amsterdam: Elsevier/North-Holland. pp. 1071-1082.

Henig, R.M. (2006, February 5): Looking for the lie. *New York Times Magazine*, 47-82.

Heuer, R. J. (2002). Elements of strategic denial and deception: Commentary. In R. Godson & J. J. Wirtz (Eds), *Strategic denial and deception : The twenty-first century challenge* (pp. 33-35). New Brunswick, USA: Transaction.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94*, 319-340.

Higgins, E. T., Roney, C. J. R., Crowe, E., & Hymes, C. (1994). Ideal versus ought predilections for approach and avoidance: Distinct self-regulatory systems. *Journal of Personality and Social Psychology, 66*, 276-286.

Higgins, E. T. & Tykocinski, O. (1992). Self-discrepancies and biographical memory: *Personality and Social Psychology Bulletin, 18*, 527-535.

Hilton, J. L., Fein, S., & Miller, D. T. (1993). Suspicion and dispositional influence. *Personality and Social Psychology Bulletin, 19*, 501-512.

Hopper, R. & Bell, R. A. (1984). Broadening the deception construct. *Quarterly Journal of Speech, 70*, 288-302.

Hough, L. M. (1998). Effects of intentional distortion in personality measurement and evaluation of suggested palliatives. *Human Performance, 2-3*, 209-244.

Hughes, J. F., Dunn, J. F., & Baxter, B. (1956). The validity of selection instruments under operating conditions. *Personnel Psychology, 9*, 321-324.

Hyman, R. (1989). The psychology of deception. *Annual Reviews of Psychology, 40*, 133-154.

Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2004). *Criminal Interrogation and Confessions* (4th Edition). Boston: Jones and Bartlett.

Isenberg, D. J. (1986). Thinking and managing: A verbal protocol analysis of managerial problem solving. *Academy of Management Journal, 29*, 775-788.

Jacobs, S., Dawson, E. J., & Brashers, D. (1996). Information manipulation theory: A replication and assessment. *Communication Monographs, 63*, 70-82.

Janis, I. L. & Gilmore, J. B. (1965). The influence of incentive conditions on the success of role playing in modifying attitudes. *Journal of Personality and Social Psychology, 1*, 17-27.

Jones, E. E. & Pittman, T. S. (1982). Toward a general theory of strategic self-presentation. In J. Suls (Ed.), *Psychological perspectives on the self* (Vol. 1, pp. 231-262). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kashy, D. A. & DePaulo, B. M. (1996). Who lies? *Journal of Personality and Social Psychology, 70*, 1037-1051.

Kilpatrick, J. (1968). Analyzing the solution of word problems in mathematics: An exploratory study (Doctoral dissertation, Stanford University, 1967). *Dissertation Abstracts International*, 28, 4380-A. (University Microfilms No. 68-6442).

Kim, S.C. (1985). Family therapy for Asian Americans: A strategic-structural framework. *Psychotherapy, 22*, 342-348.

Kluger, A. N. & Colella, A. (1993). Beyond the mean bias: The effect of warning against faking on biodata item variances. *Personnel Psychology, 46*, 763-780.

Kluger, A. N., Reilly, R. R., & Russell, C. J. (1991). Faking biodata tests: Are option-keyed instruments more resistant? *Journal of Applied Psychology, 76*, 889-896.

Kraut, R. (1980). Humans as lie detectors: Some second thoughts. *Journal of Communication, 30*, 209-216.

Kroger, R. O. & Turnbull, W. (1975). Invalidity of validity scales: The case of MMPI. Journal of *Consulting and Clinical Psychology, 43*, 48-55.

Kuusela, H. & Paul, P. (2000). A comparison of concurrent and retrospective verbal protocol analysis. *The American Journal of Psychology, 113*, 387-404.

LaFrenière, P. (1988). The ontogeny of tactical deception in humans. In R. W. Byrne & A. Whiten. (Eds.), *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans* (pp. 238-252). Oxford: Clarendon Press.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.

Langer, E., Blank, A., & Chanowitz, B. (1978). The mindlessness of ostensibly thoughtful action: The role of "placebic" information in interpersonal interaction. *Journal of Personality and Social Psychology, 36*, 635-642.

Langer, E. & Newman, H. M. (1979). The role of mindlessness in a typical social psychology experiment. *Personality and Social Psychology Bulletin, 5*, 295-298.

Larson, J. A. (1932). *Lying and its detection: A study of deception and deception tests.* Montclair, NJ: Patterson Smith.

Leary, M. R. & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin, 107*, 34-47.

Lee, S. T. (2004). Lying to tell the truth: Journalists and the social context of deception. *Mass Communication and Society, 7*, 97-120.

Levin, R. A. & Zickar, M. J. (2002). Investigating self-presentation, lies, and bullshit: Understanding faking and its effects on selection decisions using theory, field research, and simulation. In J. M. Brett & F. Drasgow (Eds.) *The psychology of work.* Mahwah, NJ: Lawrence Erlbaum Associates.

Levine, T. R., Anders, L. N., Banas, J., Baum, K. L., Endo, K., Hu, A. D. S., & Wong, N. C. H. (2000). Norms, expectations, and deception: A norm violation model of veracity judgments. *Communication Monographs, 67*, 123-137.

Levine, T. R., Park H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect." *Communication Monographs, 66*, 199-218.

Levitt, S. D. & Dubner, S. J. (2005). *Freakonomics: A rogue economist explores the hidden side of everything*. New York: Harper Collins.

Lewicki, R. (1983). Lying and deception. In M. Bazermann & R. Lewicki, (Eds.), *Negotiating in organizations* (pp. 68-90).

Li, M. F. & Lautenschlager, G. (1997). Generalizability theory applied to categorical data. *Educational and Psychological Measurement, 57*, 813-823.

Liberty, P. G., Jr., Lunneborg, C. E., & Atkinson, G. C. (1964). Perceptual defense, dissimulation, and response styles. *Journal of Consulting Psychology, 28*, 529-537.

Light, R. J. (1971). Measures of response agreement for qualitative data: Some generalizations and alternatives. *Psychological Bulletin, 76*, 365-377.

Linderholm, T. & van den Broek, P. (2002). The effects of reading purpose and working memory capacity on the processing of expository text. *Journal of Educational Psychology, 94*, 778-784.

Lykken, D. T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology, 43*, 385-388.

Lykken, D. T. (1960). The validity of the guilty knowledge test: The effects of faking. *Journal of Applied Psychology, 44*, 258-262.

Lykken, D. T. (1998). *A tremor in the blood: Uses and abuses of the lie detector*. New York: Plenum Press.

MacLeod, C. & Campbell, L. (1992). Memory accessibility and probability judgments: An experimental evaluation of the availability heuristic. *Journal of Personality and Social Psychology, 63*, 890-902.

MacLaren, V. V. (2001). A quantitative review of the guilty knowledge test. *Journal of Applied Psychology, 86*, 674-683.

Marques, J. F. & McCall, C. (2005). The application of interrater reliability as a solidification instrument in a phenomenological study. *The Qualitative Report,10*, 439-462.

Mattson, M., Ryan, D. J., Allen, M., & Miller, V. (2000). Considering organizations as a unique interpersonal context for deception detection: A meta-analytic review. *Communication Research Reports, 17*, 148-160.

McCabe, D. L. & Bowers, W. J. (1994). Academic dishonesty among males in colleges: A thirty year perspective. *Journal of College Student Development, 35*, 5-10.

McCornack, S. A. (1992). Information manipulation theory. *Communication Monographs, 59*, 1-16.

McCornack, S. A. (1997). The generation of deceptive messages: Laying the groundwork for a viable theory of interpersonal deception. In J. O. Greene (Ed.), *Message production: Advances in communication theory* (pp. 91-126). Hillsdale, NJ: Lawrence Erlbaum.

McCornack, S. A. & Levine, T. R. (1990). When lovers become leery: The relationship between suspicion and accuracy in detecting deception. *Communication Monographs, 57*, 219-230.

McCornack, S. A., Levine, T. R., Solowczuk, K. A., Torres, H. I., & Campbell, D. M. (1992). When the alteration of information is viewed as deception: An empirical test of information manipulation theory. *Communication Monographs, 59*, 17-29.

McCornack, S. A. & Parks, M. R. (1986). Deception detection and relationship development: The other side of trust. *Communication Yearbook, 9*, 377-389.

McCrae, R. R. & Costa, P. T. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology, 51*, 882-888.

McDaniel, M. A., Bruhn-Finnegan, E. B., Morgeson, F. P., Campion, M. A., & Braverman, E. P. (2001). Predicting job performance using situational judgment tests. *Journal of Applied Psychology, 86*,730-740.

McFarland, L. A. (2000). *Toward an integrated model of applicant faking*. Unpublished doctoral dissertation, Michigan State University, East Lansing.

McFarland, L. A. & Ryan, A. M. (2000). Variance in faking across noncognitive measures. *Journal of Applied Psychology, 85*, 812-821.

McFarland, L. A., Ryan, A. M., & Ellis, A. (2002). Item placement on a personality measure: Effects on faking behavior and test measurement properties. *Journal of Personality Assessment, 78*, 348-369.

McFarland, L. A., Ryan, A. M., & Kriska, S. D. (2003). Impression management use and effectiveness across assessment methods. *Journal of Management, 29*, 641-661.

McLean, B. & Elkind, P. (2003). *The smartest guys in the room: The amazing rise and scandalous fall of Enron.* New York: Penguin Group.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100*, 254-278.

Meehl, P. E., & Hathaway, S. R. (1946). The K factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. *Journal of Applied Psychology, 30*, 525-564.

Metts, S. (1989). An exploratory investigation of deception in close relationships. *Journal of Social and Personal Relationships, 6*, 159–179.

Miles, H. L. (1986). How can I tell a lie?: Apes, language, and the problem of deception. In R. W. Mitchell & N. S. Thompson (Eds.), *Deception: Perspectives on human and nonhuman deceit* (pp. 245-266). Albany, NY: State University of New York Press.

Miller, G. R. (1983). Telling it like it isn't and not telling it like it is: Some thoughts on deceptive communication. In J. I. Sisco (Ed.), *The Jensen lectures: Contemporary communication studies* (pp. 91-116). Tampa: University of South Florida.

Miller, G. R. & Stiff, J. B. (1993). *Deceptive communication.* Newbury Park, CA: Sage.

Mitchell, T.R. & Daniels, D. (2002). Motivation. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.) *Comprehensive Handbook of Psychology, Volume Twelve: Industrial and Organizational Psychology.* New York: John Wiley.

Muraven, M. & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin, 126*, 247-259.

Murphy, K. R. & Balzer, W. K. (1989). Rater errors and rating accuracy. *Journal of Applied Psychology, 74*, 619-624.

Murphy, K. R., & DeShon, R. (2000). Interrater correlations do not estimate the reliability of job performance ratings. *Personnel Psychology, 53*, 873-900.

Myslobodsky, M. S. (1997). Living behind a façade: Notes on the agenda. In M. S. Myslobodsky (Ed.) *The mythomanias: The nature of deception and self-deception* (pp. 1-22). Mahwah, NJ: Lawrence Erlbaum Associates.

Newberger, E. H. (2003, December). Why do students cheat? Retrieved May 20, 2005, from http://www.school-for-champiions.com/character/newberger_cheating2.htm

Newell, A. & Simon, H. (1972). *Human problem solving*. Englewood Cliffs, New Jersey: Prentice-Hall.

Nickels, B. J. (1994). The nature of biodata. In G. S. Stokes, M. D. Mumford, & W. A. Owens (Eds.), *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction (pp. 1-16)*. Palo Alto, CA: CPP Books.

Nisbett, R. E. & Ross, L. D. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice Hall.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259.

Ones, D. S. & Viswesvaran, C. (1998). Integrity testing in organizations. In R. W. Griffin, A. O'Leary-Kelly, & J. M. Collins (Eds.), *Dysfunctional behavior in organizations: Violent and deviant behavior. Monographs in organizational behavior and industrial relations* (Vol. 23, pp. 243-276). Stamford, CT: JAI Press.

Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: A red herring. *Journal of Applied Psychology, 81*, 660-679.

Oswald, F. L., Schmitt, N., Kim, B. H., Ramsay, L. J., & Gillespie, M. A. (2004). Developing a biodata measure and situational judgment inventory as predictors of college student performance. *Journal of Applied Psychology, 89*, 187-207.

Panone, R. D. (1984). Predicting test performance: A content valid approach to screening applicants. *Personnel Psychology, 37*, 507-514.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*, 838-852.

Paulhus, D. L. (1986). Effects of response styles on the report of psychological and somatic distress. *Journal of Consulting and Clinical Psychology, 54*, 309-313.

Paulhus, D.L. (1991). Measurement and control of response bias. In J.P. Robinson, P.R. Shaver, & L.S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp.17-59). New York: Academic Press.

Paulhus, D. L. (1994). *Balanced Inventory of Desirable Responding: Reference manual for BIDR Version 6.* Unpublished manuscript, Department of Psychology, University of British Columbia, Vancouver, BC.

Paulhus, D. L. (2002). Socially desirable responding. The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement (pp. 49-69).* Mahwah, NJ: Erlbaum.

Paulhus, D. L. (2003). In R. Fernandez-Ballesteros (Ed.), *Encyclopedia of psychological assessment* (pp. 858-861). Thousand Oaks, CA: Sage.

Pennebaker, J. W., Mehl, M. R., Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology, 54,* 547-577.

Pezdek, K., Morrow, A., Blandon-Gitlin, I., Goodman, G. S., Quas, J. A., Saywitz, K. J., Bidrose, S., Pipe, M., Rogers, M., & Brodie, L. (2004). Detecting deception in children: Event familiarity affects criterion-based content analysis ratings. *Journal of Applied Psychology, 89,* 119-126.

Posner, K. L., Sampson, P. D., Caplan, R. A., Ward, R. J., Cheney, F. W. (1990). Measuring interrater reliability among multiple raters: An example of methods for nominal data. *Statistics in Medicine, 9,* 1103-1115.

Potosky, D. & Bobko, P. (1997). Computer versus paper-and-pencil administration mode and response distortion in noncognitive selection tests. *Journal of Applied Psychology, 82,* 293-299.

Premack, D. (1988). 'Does the chimpanzee have a theory of mind' revisited. In R. Byrne & A. Whiten (Eds.), *Machiavellian intelligence. Social expertise and the evolution of intellect in monkeys, apes, and humans* (pp. 160-179). Oxford: Clarendon Press.

Pulakos, E. D., Arad, S., Donovan, M. A., & Plamondon, K. E. (2000). Adaptability in the workplace: Development of a taxonomy of adaptive performance. *Journal of Applied Psychology, 85,* 612-624.

Ramsay, L. J., Schmitt, N., Oswald, F. L., Kim, B. H., & Gillespie, M. A. (in press). The impact of situational context variables on responses to biodata and situational judgment inventory items. *Psychology Science.*

Robie, C., Brown, D. J., & Beaty, J. C. (2006). *Do people fake on personality inventories?: A verbal protocol analysis.* Paper presented at the 21st Annual Conference of the Society of Industrial and Organizational Psychology, Dallas, May.

Robinson, S. L., & Bennett, R. J. (1997). Workplace deviance: Its definitions, its manifestations, and its causes. *Research on Negotiations in Organizations, 6*, 3-27.

Rosenfeld, J. P., & Bessinger, G. T. (1990). Feedback-evoked P300 responses in lie detection. *Psychophysiology, 27* (Suppl. 4A), S60.

Rosenfeld, J. P., Cantwell, B., Nasman, V. T., Wojdac, V., Ivanov, S., & Mazzeiri, L. (1988). A modified event-related potential-based guilty-knowledge test. *International Journal of Neuroscience, 42*, 157–161.

Rosenzweig, S. (1933). The experimental situation as a psychological problem. *Psychological Review, 40*, 337-354.

Rosenzweig, S. (1934). A suggestion for making verbal personality tests more valid. *Psychological Review, 41*, 400-401.

Rynes, S. L., Colbert, A. E., & Brown, K. G. (2002). HR professionals' beliefs about effective human resource practices: Correspondence between research and practice. *Human Resource Management, 41*, 149-174.

Sackeim, H. A. & Gur, R. (1978). Self-deception, self-confrontation, and consciousness. In G. E. Schwartz & D. Shaprio (Eds.), *Consciousness and self-regulation: Advances in research* (Vol. 2, pp. 139-197). New York: Plenum Press.

Sackett, P. R., Burris, L. R., & Callaghan, C. (1989). Integrity testing for personnel selection: An update. *Personnel Psychology, 42*, 491-529.

Sackett, P. R., Burris, L. R., & Ryan, A. M. (1989). Coaching and practice effects in personnel selection. In C. L. Cooper & I. Robertson (Eds.), *International review of industrial and organizational psychology*. Chichester, NY: John Wiley & Sons.

SAT program information – SAT/ACT comparison. (n.d.). Retrieved May, 23, 2006, from http://www.collegeboard.com/sat/cbsenior/html/stat00f.html

Savage-Rumbaugh, E. S. & McDonald, K. (1988). Deception and social manipulation in symbol-using apes. In R. W. Byrne & A. Whiten (Eds.), *Machiavellian intelligence* (pp. 224-237). Oxford: Clarendon Press.

Schab, F. (1991). Schooling without learning: Thirty years of cheating in high school. *Adolescence, 26*, 839-847.

Schlenker, B. R. (1980). *Impression management: The self-concept, social identity, and interpersonal relations*. Monterey, CA: Brooks/Cole.

Schlosser, E. (2004). Fast food nation: The dark side of the all-American meal. *New York*: Harper Collins.

Schmit, M. J., & Ryan, A. M. (1993). The Big Five in personnel selection: Factor structure in applicant and nonapplicant populations. *Journal of Applied Psychology, 78*, 966-974.

Schmitt, N. & Chan, D. (in press). Situational judgment tests: Method or construct? In J. Weekley & R. E. Ployhart (Eds.), *Situational judgment tests*. Mahwah, NJ: Erlbaum.

Schmitt, N. & Kunce, C. (2002). The effects of required elaboration of answers to biodata questions. *Personnel Psychology, 55,* 569-587.

Schmitt, N. & Oswald, F. L. (2006). The impact of corrections for faking on the validity of noncognitive measures in selection settings. *Journal of Applied Psychology, 91*, 613-621.

Schrader, A. D., & Osburn, H. G. (1977). Biodata faking: Effects of induced subtlety and position specificity. *Personnel Psychology, 30*, 395-404.

Schuster, C. & Smith, D. A. (2005). Dispersion-weighted kappa: An integrative framework for metric and nominal scale agreement coefficients. *Psychometrika, 70*, 135-146.

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist, 54*, 93-105.

Selling, L. S. (1942). The psychiatric aspects of the pathological liar. *Nervous Child, 1*, 335-350.

Seymour, T. L., Seifert, C. M., Shafto, M. G., & Mosmann, A. L. (2000). Using response time measures to assess "guilty knowledge." *Journal of Applied Psychology, 85*, 30-37.

Shafir, E., & Tversky, A. (1995). Decision making. In E. E. Smith & D. N. Osherson (Eds.), *Thinking: An invitation to cognitive science* (2nd ed., Vol. 3, pp. 77-100). : Cambridge, MA: MIT Press.

Shulsky, A. (2002). Elements of strategic denial and deception. In R. Godson & J. J. Wirtz (Eds.), *Strategic denial and deception: The twenty-first century challenge* (pp. 15-32). New Brunswick, US: Transaction.

Sklansky, D. (1999). *The theory of poker*. Las Vegas: Creel Printing.

Smith, D. L. (2004). *Why we lie*. New York: St. Martin's Press.

Smith, D. B. & Ellingson, J. E. (2002). Substance versus style: A new look at social desirability in motivating contexts. *Journal of Applied Psychology, 87*, 211-219.

Snell, A. F., Sydell, E. J., Lueke, S. B. (1999). Towards a theory of applicant faking: Integrating studies of deception. *Human Resource Management Review, 9*, 219-242.

Society for Industrial and Organizational Psychology. (2003). *Principles for the Validation and Use of Personnel Selection Procedures*. Bowling Green, OH: SIOP.

Sonnentag, S. (1998). Expertise in professional software design: A process study. *Journal of Applied Psychology, 83*, 703-715.

Stark, S., Chernyshenko, O. S., Chan, K., Lee, W. C., & Drasgow, F. (2001). Effects of the testing situation on item responding: Cause for concern. *Journal of Applied Psychology, 86*, 943-953.

Stiff, J. B. (1996). Theoretical approaches to the study of communication: Comments on interpersonal deception theory. *Communication Theory, 6*, 289-296.

Stöber, J., Dette, D. E., & Musch, J. (2002). Comparing continuous and dichotomous scoring of the Balanced Inventory of Desirable Responding. *Journal of Personality Assessment, 78*, 370-389.

Stokes, G. S. & Toth, C. S. (1996). Background data for personnel selection. In R. S. Barrett (Ed.), Fair employment strategies in human resource management (pp. 171-179). Westport, CT: Quorom Books.

Sun, T. (1991). *The art of war*. Boston: Shambahla.

Sutton, S. (1998). Predicting and explaining intentions and behavior: How well are we doing? *Journal of Applied Psychology, 28*, 1317-1338.

Tangney, J. P. (1992). Situational determinants of shame and guilt in young adulthood. *Personality and Social Psychology Bulletin, 18*, 199-206.

Taylor, L., Gitttes, M., O'Neal, E. C., & Brown, S. (1994). The reluctance to expose dangerous lies. *Journal of Applied Social Psychology, 24*, 301-315.

Thompson, N. S. (1986). Deception and the concept of natural design. In R. W. Mitchell & N. S. Thompson (Eds.) *Deception: Perspectives on human and nonhuman deceit*. Albany, NY: SUNY Press.

Thompson, C., McCaughan, D., Cullum, N., Sheldon, T. A., & Raynor, P. (2004). Increasing the visibility of coding decisions in team-based qualitative research in nursing. *International Journal of Nursing Studies, 41*, 15-20.

Turner, R. E., Edgley, C., & Olmstead, G. (1975). Information control in conversations: Honesty is not always the best policy. *Kansas Journal of Sociology, 11*, 69-89.

Tversky, A. &Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131.

Van Natta (2006, March 27) Bush was set on path to war, British memo says. *The New York Times*. Retrieved July 17, 2006, from http://www.nytimes.com/2006/03/27/international/europe/27memo.html?ei=5088&en=be186887fe0c83a2&ex=1301115600&partner=&pagewanted=all

van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1994). *The think aloud method: A practical guide to modeling cognitive processes*. New York: Academic Press.

Vasek, M. E. (1986). Lying as a skill: The development of deception in children. In R. W. Mitchell & N. S. Thompson (Eds.), *Deception: Perspectives on human and nonhuman deceit* (pp. 271-292). Albany, NY: State University of New York Press.

Vasilopoulos, N. L., Cucina, J. M., & McElreath, J. M. (2005). Do warnings of response verification moderate the relationship between personality and cognitive ability? *Journal of Applied Psychology, 90*, 306-322.

Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement, 59*, 197-210.

Vohs, K. D., Baumeister, R. F., & Ciarocco, N. J. (2005). Self-regulation and self-presentation: Regulatory resource depletion impairs impression management and effortful self-presentation depletes regulatory resources. *Journal of Personality and Social Psychology, 88*, 632-657.

von Eye, A. (2006). An alternative to Cohen's κ. *European Psychologist, 11*, 12-24.

von Eye, A. & von Eye, M. (2005). Can one use Cohen's kappa to examine disagreement? *Methodology, 1*, 129-142.

Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and the implications for professional practice*. New York: John Wiley.

Vroom, V. H. (1964). *Work and motivation*. New York: Wiley.

Waid, W. M. & Orne, M. T. (1981). Cognitive, social, and personality processes in the physiological detection of deception. *Advances in Experimental Social Psychology, 14*, 61-106.

Waid, W. M., Orne, E. C., & Orne, M. T. (1981). Selective memory for social information, alertness, and physiological arousal in the detection of deception. *Journal of Applied Psychology, 66*, 224-232.

Walczyk, J. J., Roper, K. S., Seemann, E., & Humphrey, A. M. (2003). Cognitive mechanisms underlying lying to questions: Response time as a cue to deception. *Applied Cognitive Psychology, 17*, 755-774.

Walczyk, J. J., Schwartz, J. P., Clifton, R., Adams, B., Wei, M., Zha, P. (2005). Lying person-to-person about life events: A cognitive framework for lie detection. *Personnel Psychology, 58*, 141-170.

Wegner, D. M., Erber, R., & Zanakos, S. (1993). Ironic processes in the mental control of mood and mood-related thought. *Journal of Personality and Social Psychology, 65*, 1093-1104.

Whiten, A. & Byrne, R.W. (1988a). The manipulation of attention in primate tactical deception. In R. Byrne & A. Whiten (Eds.), *Machiavellian intelligence. Social expertise and the evolution of intellect in monkeys, apes, and humans* (pp. 211-223). Oxford: Clarendon Press.

Whiten, A. & Byrne, R.W. (1988a). The manipulation of attention in primate tactical deception. In R. Byrne & A. Whiten (Eds.), *Machiavellian intelligence. Social expertise and the evolution of intellect in monkeys, apes, and humans* (pp. 211-223). Oxford: Clarendon Press.

Whiten, A. & Byrne, R.W. (1988b). Tactical deception in primates. *Behavioural and Brain Sciences, 11*, 233–244.

Whyte, W. H., Jr. (1956). *The organization man*. Garden City, NY: Doubleday & Company.

Wiersma, D. (1933). On pathological lying. *Character and Personality, 2*, 48-61.

Williamson, J., Ranyard, R. & Cuthbert, L. (2000). A conversation-based process tracing method for use with naturalistic decisions: An evaluation study. *British Journal of Psychology, 91*, 203-221.

Wilson, S. R. (1997). Developing theories of persuasive message production: The next generation. In J. O. Greene (Ed.), *Message production: Advances in communication theory* (pp. 15-43). Hillsdale, NJ: Lawrence Erlbaum.

Xin, K. R. (2004). Asian American managers: An impression gap?: An investigation of impression management and supervisor-subordinate relationships. *The Journal of Applied Behavioral Science, 40*, 160-181.

Yeung, L. N. T., Levine, T. R., & Nishiyama, K. (1999). Information manipulation theory and perceptions of deception in Hong Kong. *Communication Reports, 12*, 1-11.

Zackay, D. & Bentwich, J. (1997). The tricks and traps of perceptual illusions. In M. S. Myslobodosky (Ed.) *The mythomanias: The nature of deception and self-deception* (pp. 73-103). Mahwah, NJ: Lawrence Erlbaum Associates.

Zapf, D. (2002). Emotion work and psychological well-being. A review of the literature and some conceptual considerations. *Human Resource Management Review, 12*, 237-268.

Zerbe, W. J. & Paulhus, D. L. (1987). Socially desirable responding in organizational behavior: A reconception. *Academy of Management Review, 12*, 250-264.

Zickar, M. J., Gibby, R. E., & Robie, C. (2004). Uncovering faking samples in applicant, incumbent, and experimental data sets: An application of mixed-model item response theory. *Organizational Research Methods, 7*, 168-190.

Zickar, M. J. & Robie, C. (1999). Modeling faking good on personality items: An item-level analysis. *Journal of Applied Psychology, 84*, 551-563.

Zuckerman, M., DePaulo, B., & Rosenthal, R. (1981). Verbal and non-verbal communication of deception. In L. Berkowitz (Ed.), *Advances in experimental psychology* (Vol. 14, pp. 1-59). New York: Academic Press.

Zuckerman, M. & Driver, R. E. (1985). Telling lies: Verbal and nonverbal correlates of deception. In A. W. Siegman & S. Fieldstein (Eds.), *Multichannel integrations of nonverbal behavior* (pp. 129-148).