This is to certify that the
dissertation entitled

WHEN PERFORMANCE FAILS: EXPERTISE, ATTENTION, AND PERFORMANCE UNDER PRESSURE

presented by

Sian Leah Beilock

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Kinesiology and Psychology

R. H. Carr Deborah L. Feltz
Major Professor's Signature

April 14, 2003
Date

WHEN PERFORMANCE FAILS:
EXPERTISE, ATTENTION, AND PERFORMANCE UNDER PRESSURE

By

Sian Leah Beilock

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Departments of Kinesiology and Psychology

2003

perf

perf

poo

perf

indi

pro

stru

199

in v

dev

pha

uti

do

(A

ABSTRACT

WHEN PERFORMANCE FAILS:
EXPERTISE, ATTENTION, AND PERFORMANCE UNDER PRESSURE

By

Sian Leah Beilock

This work explored the cognitive mechanisms underlying pressure-induced performance decrements. *Performance pressure* is defined as an anxious desire to perform at a high level (Hardy, Mullen, & Jones, 1996). *Choking*, or performing more poorly than expected given one's level of skill, tends to occur in situations fraught with performance pressure (Baumeister, 1984; Beilock & Carr, 2001; Lewis & Linder, 1997).

Self-focus or *explicit monitoring theories* of choking suggest that pressure-induced performance decrements result from the explicit monitoring and control of proceduralized knowledge that is best run off as an uninterrupted and unanalyzed structure (Baumeister, 1984; Beilock & Carr, 2001; Lewis & Linder, 1997; Masters, 1992). Conversely, *distraction theories* propose that pressure creates a dual-task situation in which skill execution and performance worries vie for the attentional capacity once devoted solely to primary task performance (Lewis & Linder, 1997; Wine, 1971).

To date, explicit monitoring theories have accounted quite well for the choking phenomenon (see Appendix A and B). However, the extant choking literature has solely utilized sensorimotor skills as a test bed. Well-learned, proceduralized sensorimotor skills do not possess the right task control structures to *choke* according to distraction theories (Allport, Antonis, & Reynolds, 1972). Furthermore, unpracticed sensorimotor skills,

althou
Wiere
storag
senso
all (E
via th
conv

prob
decr
con
mo
the
pro
fail
sol
pro
dif

ma
go
di
a

although based, in part, on explicitly accessible declarative knowledge (Beilock, Wierenga, and Carr, 2002), may not demand the type of processing and information storage that make a task susceptible to choking via distraction. Indeed, novice sensorimotor skills do not appear to be negatively impacted by performance pressure at all (Beilock & Carr, 2001). It remains an open possibility then, that choking may occur via the mechanisms proposed by distraction theories in certain tasks – for example, complex cognitive tasks not based on an automated or proceduralized skill representation.

Four experiments examined performance under pressure in the mathematical problem solving task of *modular arithmetic* (MA). Exp. 1 demonstrated that performance decrements in difficult, unpracticed MA problems occurred under high pressure conditions. Exp. 2 demonstrated that these pressure-induced failures only occurred for the most difficult and capacity demanding unpracticed equations. Exp. 3 further explored these performance failures both early and late in learning. Similar to Exp. 2, only difficult problems with large on-line working memory demands choked. Furthermore, these failures were limited to problems early in practice when capacity-demanding rule-based solution algorithms governed performance. In Exp. 4, participants performed MA problems once, twice, or 50 times each, followed by a high pressure test. Again, only difficult problems that had not been highly practiced showed performance decrements.

These findings support *distraction theories* of choking in the domain of mathematical problem solving. This outcome contrasts with sensorimotor skills, such as golf putting, in which the data have uniformly supported explicit monitoring rather than distraction theories (Beilock & Carr, 2001; Lewis & Linder, 1997). This contrast suggests a taxonomy of skills based on the nature and representation of their control structures.

To my brother, Mark Beilock, as a symbol of the fact that we can accomplish anything we work hard at…and that we should not let anxiety, worry, and the "what ifs" stand in our way.

not
Fel
me
res
sup

my
Ps
me
in
in
Th
tha

Ste
an
Wi
an

rea

ACKNOWLEDGMENTS

There are many people whom, without their love and support, this work would not have been possible. First of all, I would like to thank my advisors, Tom Carr and Deb Feltz. Both Tom and Deb have been instrumental in my academic development, pushing me to not only to work hard, but to have confidence in myself and my abilities as a researcher. More than that, however, they have also provided invaluable emotional support and guidance at times when my anxieties and worries were getting the best of me.

In addition to my advisors, countless others helped to keep me grounded during my tenure as a graduate student. My fellow graduate students in both Kinesiology and Psychology consistently remind me that there was a world outside my lab – and pushed me to get out and enjoy it every once in a while: Karin Allor Pfeiffer and Kevin Stefanek in Kinesiology and Monica Castelhano, Mareike Wieth, Laurie Carr, and Kate Arrington in Psychology. And outside of the academic arena, my friends Scott Kopchinski and Liz Thornton were always around to make sure that I didn't take myself, or anything else for that matter, too seriously.

My family offered unconditional support – my mother, Ellen Beilock, my father, Steve Beilock, my brother, Mark Beilock, my grandparents, George and Sylvia Elber, and Phyllis Beilock and Lou Baca, and my soon-to-be-husband Allen McConnell. Without your words of encouragement and your ability to make me feel better in almost any situation, I would have been lost.

I would also like to mention all the strong women in my life who, whether they realized it or not, were amazing role models and inspirations to me. Women such as my

advi
succ
who
the

my
wou
not
am
wor

advisor, Deb Feltz, and Janet Starkes and Rose Zacks consistently demonstrated that a successful academic life was obtainable for a woman. And my mother and grandmothers who, through their own lives and explicit teachings, made sure that I was always aware of the fact that I had the ability to navigate any hurdles standing between me and my goals.

Finally, I must mention Berkeley the dog. Anyone who knows me is familiar with my love/obsession for her. And at the same time, my friends and family also realize that I would have gone mad a long time ago without Berkeley's companionship. Berkeley is not only there to comfort me when I am stressed or sad, but she serves a purpose that I am not sure any human could – Berkeley ensures that I actually leave the lab and my work once in a while! For that, I owe her my sanity.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 I

pers
Har
resu
ofte
has
than
whe
& C
bask
golf
in in
unm

supe
dom
Brow
1993
1995
relat
situat
into t

1 INTRODUCTION

The desire to perform at an optimal skill level in situations with a high degree of personally felt importance is thought to create *performance pressure* (Baumeister, 1984, Hardy, Mullen, & Jones, 1996). Paradoxically, despite the fact that performance pressure results from aspirations to function at one's best, environments fraught with pressure are often where suboptimal skill execution is most visible. The term *choking under pressure* has been used to describe this phenomenon. Choking, defined as performing more poorly than expected given one's level of skill, is thought to occur across diverse task domains where incentives for optimal performance are at a maximum (Baumeister, 1984; Beilock & Carr, 2001; Lewis & Linder, 1997; Masters, 1992). We often refer to the "bricks" in basketball free throw shooting when the game winning shot is on the line, or the "yips" in golf putting when an easy 3 foot putt to win the tournament stops short, and to "cracking" in important test-taking situations where a course grade or college admission is at stake as unmistakable instances of such incentive or pressure-induced performance decrements.

Surprisingly, while research concerning the cognitive mechanisms governing superior task performance is abundant across both cognitive and sensorimotor skill domains (Allard & Starkes, 1991; Anderson, 1982; 1983; Anderson & Lebiere, 1998; Brown & Carr, 1989; Ericsson & Charness, 1994; Ericsson, Krampe, & Tesch-Romer, 1993; Fitts & Posner, 1967; Logan, 1988; Newell & Rosenbloom, 1981; Proctor & Dutta, 1995; Reimann & Chi, 1989; Rosenbaum, Carlson, Gilmore, 2001; Staszewski, 1988), relatively less attention has been devoted to suboptimal skill execution – especially in situations in which optimal task performance is not only desired, but expected. Insight into the mechanisms governing execution failure is important, as it will not only serve to fu

re

pr

of

1.

fo

th

ab

an

th

no

19

La

19

a d

ex

wh

att

19

me

the

further our understanding of the variables responsible for skill decrements, but those responsible for success as well. A careful cognitive analysis of the choking under pressure phenomenon may open a new kind of window to the organization and operation of the information processing mechanisms that underlie skilled performance.

1.1 Theories of Choking under Pressure

Why do skills fail in high pressure situations? Two main theories have been put forth as explanations for the choking phenomenon. *Self-focus* or *explicit monitoring* theories propose that performance pressure increases anxiety and self-consciousness about performing correctly, which in turn enhances the attention paid to skill processes and their step-by-step control. Attention to performance at this component level is thought to disrupt the proceduralized or automated processes of high level skills that normally run outside the scope of working memory during performance (Baumeister, 1984; Beilock & Carr, 2001; Butler & Baumeister, 1998; Kimble & Perlmuter, 1970; Langer & Imber, 1979; Lewis & Linder, 1997; Marchant & Wang, in press; Masters, 1992). D*istraction theories*, on the other hand, suggest that performance pressure creates a distracting environment that competes with the attention normally allocated to skill execution (Wine, 1971). In essence, pressure serves to create a dual-task environment in which controlling execution of the task at hand and performance worries vie for the attentional capacity once devoted solely to primary task performance (Lewis & Linder, 1997). Thus distraction and explicit monitoring theories offer contrasting accounts of the mechanisms responsible for performance decrements under pressure. While distraction theories suggest that pressure creates a distracting environment that draws attention away fro

pre

1.2

not

test

und

goa

by e

decr

conc

ordir

chok

envir

expo

activi

press

perfor

theori

consci

being

design

from primary skill execution, explicit monitoring theories suggest the opposite – that pressure prompts *too much* explicit attention to performance processes and procedures.

1.2 Support for Explicit Monitoring Theories of Choking

Training studies. To date, explicit monitoring theories have received the most noted support in accounting for choking under pressure. For example, in an attempt to test the two main theories of choking, Beilock and Carr (2001) examined skill execution under pressure in a golf putting task (for the full text of this paper, see Appendix A). The goal was to determine whether practice at dealing with the causal mechanisms proposed by each theory (i.e., explicit attention vs. distraction) would reduce performance decrements in high pressure environments.

Individuals were trained to a high putting skill level under a variety of learning conditions and then exposed to a pressure situation. The first training condition involved ordinary single-task practice, which provided a baseline measure of the occurrence of choking. The second training condition involved practice in a distracting, dual-task environment (while monitoring an auditory word list for a target word) designed to expose performers to being distracted from the primary task by execution-irrelevant activity in working memory – the specific cause of performance decrements under pressure according to distraction theories. The third training condition exposed performers to the particular aspects of high pressure situations that explicit monitoring theories of choking propose as the cause of performance decrements. In this "self-conscious" or "skill focus" training condition, participants learned the putting task while being videotaped for subsequent public analysis by experts. This manipulation was designed to expose performers to having attention called to themselves and their

perfo
Follo
by a p

single
envir
signif
imme
Howe
this g
imme
trainin
perfor
consci
of ove
mecha
decren

skill in
negativ
trained
awarene
analysis

performance in a way intended to induce explicit monitoring of skill execution. Following training, all groups were exposed to the same high pressure situation created by a performance-contingent monetary award.

Choking occurred for those individuals who were trained on the putting task in a single-task isolated environment, and also for individuals trained in a dual-task environment that simply created distraction. That is, both of these groups putted significantly less accurately in the high pressure situation than they had in an immediately preceding block of putts during which no pressure had been applied. However, choking did not occur for those trained in the self-conscious condition. Indeed, this group performed slightly better in the pressure situation than they had in the immediately preceding no-pressure trials. Beilock and Carr (2001) concluded that training under conditions that prompted attention to skill parameters served to adapt these performers to the type of attentional focus that often occurs under pressure. That is, self-consciousness training served to inoculate individuals against the negative consequences of over-attending to well-learned proceduralized performance processes – the precise mechanisms that explicit monitoring theories suggest are responsible for performance decrements in high pressure situations.

Previously, Lewis and Linder (1997) had also found that learning a golf putting skill in a self-awareness-heightened environment inoculates individuals against the negative effects of performance pressure at high levels of practice. Participants were trained on a golf putting task under either normal, single-task conditions or under a self-awareness condition (in which individuals putted while being videotaped for later analysis by golf professionals) and then exposed to a high pressure situation. Similar to

Beil
those
the n
intro
press
Mull
induc
perfo
expli

as the
conse
the fa
memo
& Po
expli
level
practi
evide
autom

exami
learne

Beilock and Carr (2001), Lewis and Linder demonstrated that pressure caused choking in those individuals who had not been adapted to self-awareness (i.e., participants trained in the normal, single-task situation). Furthermore, Lewis and Linder found that the introduction of a secondary task (counting backward from 100) while performing under pressure helped to alleviate these performance decrements (for confirmatory data, see Mullen & Hardy, 2000). Because the secondary task served to prevent the pressure-induced instantiation of maladaptive explicit attention to well-learned proceduralized performance processes, choking under pressure was assuaged – a finding consistent with explicit monitoring theories.

If explicit monitoring theories are correct, then high level sensorimotor skills such as the golf putting task described above should be more susceptible to the negative consequences of performance pressure than less practiced performances. This is due to the fact that the former, but not the latter, are thought to operate outside of working memory, largely devoid of step-by-step attentional control (Anderson, 1982, 1993; Fitts & Posner, 1967; Proctor & Dutta, 1995). To the extent that performance pressure prompts explicit attention to execution, those skills not normally attended in real time (e.g., high level sensorimotor skills) should be more harmed by pressure-induced control than less practiced skills. This finding would be consistent with the idea that a majority of the evidence for choking has been derived from well-learned sensorimotor tasks that automate via proceduralization with extended practice (Marchant & Wang, in press).

To test this prediction, Beilock and Carr (2001) conducted a second study examining performance under pressure at both low and high skill levels. Participants learned a golf putting skill to a high level and were exposed to a high pressure situation

both e
perfor
emerg
affect
pressu
explici

(1997)
inocul
of prac
mecha
directl
Recent
focus o
novice
comple
predict
mecha

skill-fo
conditi
partici
swing.

both early and late in practice. Early in practice, pressure to do well actually facilitated performance. At later stages of learning, performance decrements under pressure emerged. Thus, the proceduralized performances of experts appear to be negatively affected by performance pressure. However, novice skill execution is not harmed by pressure-induced attention to execution, as less skilled performances are already explicitly attended in real time.

Attentional focus studies. Both Beilock and Carr (2001) and Lewis and Linder (1997) have demonstrated that skill training that induces attention to performance may inoculate individuals against the negative effects of performance pressure at high levels of practice. While these types of training methods lend insight into the cognitive mechanisms driving skill failure in high stakes situations, it may also be possible to more directly assess the processes responsible for pressure-induced performance decrements. Recently, Beilock, Carr, MacMahon, and Starkes (2002) manipulated the attentional focus of experienced golfers while performing a putting task and the attentional focus of novice and experienced soccer players while performing a soccer dribbling task (for a complete report of this work, see Appendix B). The goal was to directly test the predictions of explicit monitoring and distraction theories regarding the causal mechanisms of choking in sensorimotor skills.

In Beilock et al.'s (2002) first study, experienced golfers took a series of putts in a skill-focused attention condition and a dual-task attention condition. The order of these conditions was counterbalanced across participants. In the skill-focused condition, participants were instructed to attend to a particular component of their golf putting swing. Specifically, individuals were instructed to monitor the swing of their club and at

the e
to a s
draw
moni
simu
Parti
speci
inten
mech

in the
puttin
condi
signif

(Ande
Carr, 2
proces
audito
deman
which
(e.g., lo
target o

the exact moment they finished the follow-through of their swing, bringing the club head to a stop, to say the word "stop" out loud. This skill-focused condition was designed to draw attention to a component process of performance, coinciding with explicit monitoring theories. The dual-task attention condition involved putting while simultaneously listening to a series of recorded tones being played from a tape recorder. Participants were instructed to monitor the tones carefully, and each time they heard a specified target tone, to say the word “tone” out loud. The dual-task condition was intended to distract attention away from skill execution, in line with the choking mechanisms proposed by distraction theories.

Results demonstrated that the experienced golfers performed significantly better in the dual-task condition in comparison to the skill-focused condition. Additionally, putting in the skill-focused condition was less accurate than a single-task practice condition used as a baseline measure. Performance in the dual-task condition did not significantly differ from this practice condition.

Because well-learned golf putting does not require constant on-line control (Anderson, 1982, 1983, 1993; Beilock, Bertenthal, McCoy, & Carr, in press; Beilock & Carr, 2001; Fitts & Posner, 1967; Proctor & Dutta, 1995), attention is available for the processing of secondary task information if necessary (such as monitoring a series of auditory tones). As a result, performance does not suffer with the addition of dual-task demands. It should be noted that this may not hold true for dual-task environments in which the tasks draw upon similar processes and hence create structural interference (e.g., looking at a golf ball while lining up a putt and visually monitoring a screen for a target object). When prompted to attend to a specific component of the golf swing,

howev
task co
chokin
attendi

and du
and im
Additi
differin

both qu
differen
support
attende
1967;
the task
operate
Keele &

that exp
differen
executi
these in

however, experienced performance degrades in comparison to both single-task and dual-task conditions. This pattern of results coincides with explicit monitoring theories of choking suggesting that well-learned performances may actually be compromised by attending to skill execution.

In a second study, Beilock et al. (2002) again explored the impact of skill-focused and dual-task attention conditions, but in a movement skill that uses different effectors and imposes different temporal demands than golf putting – soccer dribbling. Additionally, the effects of dual-task and skill-focused attention on performance at differing levels of soccer skill proficiency were also explored.

Skill acquisition is believed to progress through distinct phases characterized by both qualitative differences in the cognitive structures supporting performance and differences in performance itself. Early in learning, skill execution is thought to be supported by a set of unintegrated control structures that are held in working memory and attended to one-by-one in a step-by-step fashion (Anderson, 1983, 1993; Fitts & Posner, 1967; Proctor & Dutta, 1995). With practice, however, procedural knowledge specific to the task at hand develops. Procedural knowledge does not require constant control and operates largely outside of working memory (Anderson, 1993; Fitts & Posner, 1967; Keele & Summers, 1976; Kimble & Perlmuter, 1970; Langer & Imber, 1979).

Because novices must devote attentional capacity to task performance in ways that experts do not (Fitts, 1964; Fitts & Posner, 1967), novices and experts may be differentially affected by conditions that either draw attention away from, or toward, skill execution. Specifically, the capacity-demanding performance of novices may not afford these individuals the attentional resources necessary to devote to secondary task demands

if the si
executi
conditi
This is
perforn
perforn
executi
to com

were a
second
execut
focuse
recentl
perforn
experin
particip

distrac
skill in
were in
that m

if the situation requires it. However, the novice, who must attend to the steps of skill execution in order to succeed, might not be harmed or could perhaps be helped by conditions that focus attention more squarely on the skill and prevent it from wandering. This is in contrast to the impact of skill-focused and dual-task attention on experienced performance, as seen in the golf putting study described above. Here, the proceduralized performances of experts are not harmed by dual-task conditions. Yet, high level execution is negatively impacted by skill-focused conditions designed to prompt attention to component parts of performance not normally attended to in real time.

In Beilock et al.'s (2002) second study, novice and experienced soccer players were asked to dribble a soccer ball through a series of pylons while either performing a secondary auditory monitoring task (designed to distract attention away from skill execution, in line with distraction theories' proposed choking mechanisms) or a skill-focused task in which individuals were asked to monitor the side of the foot that most recently contacted the ball (designed to draw attention to a component process of performance, coinciding with explicit monitoring theories). As in the first golf putting experiment, the order of these attention conditions were counterbalanced across participants.

Similar to the results from the golf putting study described above, performing in a distracting, dual-task environment did not harm experienced soccer players' dribbling skill in comparison to the skill-focused condition. However, when the soccer players were instructed to attend to step-by-step performance (i.e., monitoring the side of the foot that most recently contacted the ball), their dribbling skill deteriorated in comparison to both the dual-task condition and a single-task practice condition used as a baseline measure.

Novices showed the opposite pattern. These less skilled individuals performed at a lower level in the dual-task condition, designed to distract attention away from performance, in comparison to the skill-focused manipulation, designed to draw attention toward the task at hand. Furthermore, novices performed at a higher level in the skill-focused condition than in the single-task practice condition.

Consistent with the evidence presented above in support of explicit monitoring theories of choking, step-by-step attention to skill processes and procedures does not appear to harm novice sensorimotor skill execution that is already explicitly attended in real time. However, this same type of attentional control disrupts or slows down well-learned sensorimotor skill performance thought to normally operate largely outside of working memory (Baumeister, 1984; Beilock & Carr, 2001; Lewis & Linder, 1997). The negative effects of enhanced attention to highly skilled performance can not only be seen in complex tasks such as golf putting and soccer dribbling, but in more basic skills we use everyday. For example, Wulf and colleagues have suggested that directing performers' attention to their movements through "internal focus" feedback on a dynamic balance task interferes with the automated control processes that usually control balance movements outside of conscious scrutiny (Wulf & Prinz, 2001).

It should be noted that novice soccer dribbling performance was harmed by dual-task demands. This result suggests that distraction theories might be a viable explanation for performance decrements under pressure in novel sensorimotor skill execution. However, as seen in Beilock and Carr's (2001) work presented above, novice motor skill execution does not appear to choke under pressure. It may be that explicit monitoring theories are the most appropriate explanation for performance decrements under pressure, or, alternatively, that novice sensorimotor skills do not demand the type of processing and information storage that makes a task susceptible to choking via distraction. This is an issue to which we turn to below.

1.3 Is the Issue Settled? Differences Due to Task Control Structure

The work reviewed above suggests that maladaptive explicit monitoring may be responsible for choking under pressure. Given the consistency of this evidence, it may seem unlikely that distraction theories could provide any additional insight into the choking phenomenon. However, the automated or proceduralized sensorimotor skills predominantly used in the extant choking research may not possess the right task control structures to be susceptible to pressure-induced performance decrements according to distraction theories. These proceduralized sensorimotor skills are thought to run outside of working memory and are largely robust to performance decrements as a result of distracting, dual-task situations (Allport, Antonis, & Reynolds, 1972; Beilock et al., 2002; Beilock, Wierenga, & Carr, 2002, in press; Keele & Summers, 1976; Kimble & Perlmuter, 1970; Langer & Imber, 1979; Leavitt, 1979; Smith & Chamberlin, 1992). Furthermore, even unpracticed sensorimotor skills, although based, in part, on explicitly accessible declarative knowledge (Beilock, Wierenga, and Carr, 2002) may not require the type of sequentially dependent interweaving of processing and information storage that make a task susceptible to choking via distraction. Thus, one reason why distraction theories of choking may not have received much support is because they have not been tested in the appropriate skill domains.

1.4 Tes

T
pressure
and info
anxiety
intrusive
Eysenck
Because
devoted
availabl
on proce
this decr
(Ashcraf
Holyoak,

R
limited c
math anx
task invo
Difficult
in an atte
math anxi
increased
anxiety, w

1.4 Test Anxiety Literature

There is a literature that we can look to, however, for clues concerning how a pressure situation might influence skills with sequentially dependent on-line processing and information storage demands susceptible to capacity limitations. Within the test anxiety literature researchers have suggested that anxiety manifests itself in the form of intrusive thoughts or worries about the situation and its outcome (Ashcraft & Kirk, 2001; Eysenck, 1979, 1992; Eysenck & Calvo, 1992; Eysenck & Keane, 1990; Wine, 1971). Because these thoughts are attended to, a portion of working memory capacity normally devoted to primary skill execution is consumed by such thoughts, and therefore not available for the processing of task-relevant cues. For tasks with interdependent demands on processing and storage that rely heavily on working memory for on-line execution, this decrease in capacity is thought to result in suboptimal performance outcomes (Ashcraft & Kirk, 2001; Darke, 1988; Leon, 1989; Sorg & Whitney, 1992; Tohill & Holyoak, 2000).

Recent research in the math problem solving literature has found support for this limited capacity theory. Ashcraft and Kirk (2001) examined the ability of low and high math anxious individuals to simultaneously perform a mental addition task and a memory task involving the maintenance of random letter strings in memory for later recall. Difficulty levels of both the primary math and secondary memory tasks were manipulated in an attempt to examine the effects of task difficulty on performance as a function of math anxiety. Results demonstrated that performance was lowest (mainly in the form of increased math task error rates) in instances in which individuals, regardless of math anxiety, were required to perform both a difficult math and difficult memory task simultaneously. Furthermore, in comparison to less anxious individuals, those participants high in math anxiety showed an exaggerated increase in performance errors under the difficult math and memory task condition. The authors concluded that performance deficits under demanding dual-task conditions were most pronounced in individuals high in math anxiety, as anxiety, similar to the instantiation of a secondary demanding task, drains the attentional capacity that might otherwise be available for primary skill performance.

Anxiety has also been shown to lead to performance decrements in capacity-demanding analogical reasoning tasks. Tohill and Holyoak (2000) explored the relationship between anxiety and the ability to make attributional and relational mappings between objects. Attributional mapping involves mappings based on the physical characteristics of individual objects (e.g., a woman in picture A might map to a woman in picture B). In contrast, relational mappings are based on relationships linking multiple objects together (e.g., a dog being held by a woman in picture A might map to a baby being held by a woman in picture B, as both the dog and the baby are being held by a woman). Because attributional mapping requires that only a single object characteristic be held in memory, while relational mapping requires that an individual maintain a number of object relations in memory, this latter form of mapping is thought to impose a heavier load on working memory (Halford, 1993; Tohill & Holyoak, 2000). If anxiety serves to limit working memory capacity through the instantiation of attention-demanding situational worries (Ashcraft & Kirk, 2001; Eysenck, 1979, 1992; Eysenck & Calvo, 1992; Eysenck & Keane, 1990; Wine, 1971), then high anxiety individuals may have difficulty performing memory-intensive relational mappings more so than attributional mappings, as both the relational mapping task and anxiety should compete for attentional capacity.

Tohill and Holyoak (2000) presented individuals with pairs of analogical pictures and asked them to perform mappings between a specific object in one picture to an object in a second picture. Highly anxious individuals were less likely to generate relational mappings in comparison to less anxious individuals. This pattern of results occurred even when individuals were given explicit instructions to use relational mapping techniques. Thus, similar to the effects of anxiety on the simultaneous performance of a difficult mental addition and memory task, in memory manipulation and maintenance intensive relational mapping procedures, anxiety also appears to restrict the working memory capacity required for successful skill execution.

Research stemming from the test anxiety literature lends support to the notion that performance decrements may result from anxiety-induced worries that decrease task-relevant processing resources. If, as proposed by distraction theories, pressure serves to create a distracting environment via worries about the situation and its consequences, then pressure may impose constraints on working-memory-intensive tasks similar to those of anxiety as outlined above. Under this view, skills that have become proceduralized with extended practice and do not rely heavily on explicit attentional processes for successful skill execution (e.g., the well-learned golf putting or soccer tasks mentioned above), as well as skills that do not possess the appropriate types of inter-dependent processing and information storage demands (e.g., the novice soccer task mentioned above), may not be harmed by performance pressure. The lack of support for

distra
under

1.5 (

task w
accord
automa
levels
analysi
accoun
into po

Modula
number
and b is
(mod 4
(i.e., 51
equatio
the equ

questio
the mo
applied

distraction theories in the choking literature then, may be a function of the types of skills under investigation.

1.5 Overview of the Current Research

The aim of the current work was to examine skill performance under pressure in a task with working memory requirements that are likely to make it susceptible to choking according to distraction theories – at least at low levels of learning prior to skill automatization. It may be that such a task would also be susceptible to choking at higher levels of practice, but due to explicit monitoring rather than distraction. This type of analysis will not only shed light on the ability of these theories to successfully predict and account for performance decrements in high pressure situations, but may also lend insight into possible differences in their domains of applicability.

I chose Gauss's (1801) *modular arithmetic* task (Bogomolny, 1996) as a test bed. Modular arithmetic is defined as a particular sequence of arithmetic operations. Two numbers a and b are said to be equal or congruent *modulo N* if the difference between a and b is exactly divisible by N. For example, to verify the modular equation "$51 \equiv 19 \pmod{4}$," the middle number of the equation is first subtracted from the far left number (i.e., 51 - 19), and then this answer is divided by the far right number of the original equation (i.e., $32 \div 4$). Because the result of this division step is a whole number (i.e., 8), the equation is true.

How does the modular arithmetic task change with practice? To answer this question we may need to draw on a different conceptualization of skill automaticity than the model of automaticity via proceduralization described in the introduction and often applied to sensorimotor skill acquisition (e.g., Fitts & Posner, 1967). According to

Logan's (1988) instance-based theory of automaticity – which was originally developed to account for the data from mental arithmetic-type tasks – a rule-based algorithm should be initially employed to solve unpracticed modular equations. In this unpracticed state, problem solutions are dependent on the explicit application of a capacity-demanding step-by-step process that must be maintained and controlled on-line by working memory during execution. With practice on particular problems, the reliance on this procedure decreases and past instances of problem solutions are retrieved directly or "automatically" from long term memory, whereas new problems continue to engage the algorithm.

Logan's model proposes an alterative view of automaticity to traditional theories of proceduralization and has been quite successful in accounting for changes in speed and accuracy of performance with practice on cognitive tasks such as alphabet arithmetic, lexical decision, and semantic categorization. Nonetheless, contrary to Logan's (1988) predictions, there is some evidence that unfamiliar problems still based on algorithmic computations do become more efficient with practice of the algorithm (Touron, Hoyer, & Cerella, 2001). Even practiced algorithms, however, may not be governed by the same type of control structures as, for example, proceduralized motor programs. An algorithmic solution procedure, regardless of each component's efficiency, is based on a hierarchical and sequentially dependent task representation in which initial steps must be held and acted on in working memory in order to generate subsequent processes and final solutions. Well-learned sensorimotor skills that operate largely outside of working memory (Fitts & Posner, 1967; Proctor & Dutta, 1995) are most likely not governed by the

wil

an

of

of

ind

esp

of d

and

mem

press

arithm

harme

harm u

not neg

on-line.

actually e

execution

In

group prior

contrast to E

differed in te

the same type of working-memory-dependent representation. This is an issue to which we will return later.

In Experiment 1, participants were assigned to either a low or high pressure group and performed a series of unpracticed modular arithmetic equations. From the standpoint of distraction theories, modular arithmetic is a good candidate for choking at low levels of practice when the attentional demands of problem solving are high and pressure-induced worries compromise task-relevant processing resources. This should be especially true for the problems utilized in Experiment 1. All of the equations consisted of double-digit numbers with a borrow operation (e.g., 43 = 28 (mod 5)). Large numbers and borrow operations are thought to increase task difficulty and on-line working memory demands (Ashcraft, 1992; Ashcraft & Kirk, 2001). Thus, if performance pressure impinges on the processing resources needed to successfully solve modular arithmetic equations, difficult problems of the type used in Experiment 1 are likely to be harmed. In contrast, explicit monitoring theories would suggest that pressure should not harm unpracticed modular equations – as pressure-induced attention to execution should not negatively impact information that is already explicitly attended to and maintained on-line. Indeed, as mentioned above, Beilock and Carr (2001) have found that pressure actually *enhances* the performance of novices in golf putting, despite harming the execution of more practiced individuals.

In Experiment 2, participants were again assigned to either a low or high pressure group prior to performing a series of unpracticed modular arithmetic equations. In contrast to Experiment 1, the modular arithmetic equations utilized in Experiment 2 differed in terms of the demands these problems placed on working memory. If, as

dis
res
und
dem
Con
prob
regar
press
atten

pressu
arith
levels
induce
depen

answe
predict
of prac
behavi
memor
perform
perform

distraction theories would propose, pressure-induced limitations on working memory are responsible for choking, then difficult equations with large capacity demands should fail under pressure. However, easier problems that do not impose such heavy attentional demands should be less susceptible to pressure-induced performance decrements. Conversely, if explicit monitoring theories are correct in the domain of mathematical problem solving, then the unpracticed modular equations utilized in Experiment 2, regardless of problem difficulty, should not fail under pressure. As mentioned above, pressure-induced attention to skill execution should not harm novel equations normally attended on-line.

Experiment 3 extended the first two experiments' exploration of choking under pressure to highly practiced equations. Distraction theories propose that modular arithmetic should be most susceptible to performance decrements under pressure at low levels of practice. Following extended practice on specific problems however, pressure-induced distraction should not harm execution as answer derivation is no longer dependent on the intermediate memorial maintenance of task information.

To the extent that the control structure of modular arithmetic changes to automatic answer retrieval with practice, explicit monitoring theories might make a different prediction – at least according to Masters, Polman, and Hammond (1993). At high levels of practice, pressure "may result in a return to an explicit, algorithmic-based control of behavior through disruption of automatic retrieval of skill-based information from memory" (Masters et al., 1993, p.664). Such a regression, if it occurred, would slow performance and increase the opportunity for error – which would create poorer performance. Furthermore, if rather than a shift to automatic answer retrieval as Logan

(1988

algori

press

induc

rehea

differ

Indivi

expos

repea

instan

probl

exper

is due

then r

those

retrie

Logan

perfor

oppos

have b

on alg

(1988) would propose, modular arithmetic automates via proceduralization of the algorithm, explicit monitoring theories would still predict performance decrements under pressure at high levels of practice. In this case, such failures might be due to pressure-induced attentional control that increases the time or error associated with maintaining, rehearsing, or acting on the well-learned algorithm.

Experiment 4 further examined susceptibility to choking under pressure following different amounts of exposure to specific problems within the modular arithmetic task. Individuals performed over 800 modular arithmetic practice problems and were then exposed to a high pressure environment. Specific problems were presented either once, repeated twice, or repeated fifty times each during practice. According to Logan's (1988) instance-based theory of automaticity, the task control structures of modular arithmetic problems should only change as a function of specific problem exposure, not necessarily experience at performing many different modular arithmetic problems. Thus, if choking is due to pressure-induced capacity limitations, as distraction theories would propose, then regardless of how many different problems individuals have been exposed to, only those equations that have been practiced enough to produce instance-based answer retrieval (a minimum of 36 to 72 exposures according to Klapp, Boches, Trabert, and Logan, 1991), should be inoculated against the detrimental capacity-limiting effects of performance pressure. In contrast, explicit monitoring theories would again make an opposite prediction. Namely, pressure-induced attention will harm those problems that have been repeatedly practiced to the level of automatic answer retrieval.

As mentioned above, there is some evidence that unfamiliar problems still based on algorithmic computations do become more efficient with practice of the algorithm

(Tour
impo
unde
comp
mem
perfo
probl
decre
the m

possib
fail un
monito
serves t
much lik
to some
memory
maladapti
problems
instantiatio
performance
the fact that

(Touron, Hoyer, & Cerella, 2001). However, even practiced algorithms should still impose attentional demands, and thus show susceptibility to performance decrements under pressure according to distraction theories. This is due to the fact that the component parts of any algorithmic solution procedure must still be held in working memory during on-line performance. Such a prediction can be tested by examining performance decrements under pressure as a function of problem difficulty. If novel problems based on practiced algorithms are susceptible to pressure-induced performance decrements via distraction, then the more difficult and capacity-demanding the problem, the more vulnerable the equation should be to capacity-limited failure.

If practice increases the proficiency of algorithmic computations, then it is also possible that unfamiliar problems based on highly practiced algorithmic procedures might fail under pressure as a result of the attentional control mechanisms proposed by explicit monitoring theories. That is, such problems might fail via pressure-induced attention that serves to disrupt or slow down well-learned and highly efficient performance processes – much like choking under pressure in well-learned sensorimotor skills. This should be true to some extent across all problems, regardless of equation difficulty level or working memory demands. While easier problems may be less susceptible to failure as a result of maladaptive attentional control than their more demanding counterparts, as these problems have fewer independent steps and thus less of an opportunity for the instantiation of pressure-induced control, there should still be at least some sign of performance decrements under pressure in these less demanding problems. This is due to the fact that all novel equations based on practiced algorithms require several steps to achiev

induce

pressu

proper

learnin

achieve a problem solution and thus should be at least somewhat susceptible to pressure-induced attention that disrupts proceduralized algorithmic processes.

I now turn to Experiment 1 in which I initially explored performance under pressure at low levels of practice in modular arithmetic – a skill that should have the right properties to be susceptible to pressure-induced performance decrements during initial learning according to distraction, but not explicit monitoring theories.

2 CH

2.1 E

high pr
equati
the sec
perfor
was in
pressu

equati
proble
might
consis
menti
difficu
is resp
most

mathe
skill e
I sho
execu

2 CHOKING UNDER PRESSURE IN MATHEMATICAL PROBLEM SOLVING

2.1 Experiment 1

In Experiment 1 individuals were randomly assigned to either a low pressure or high pressure group prior to performing three blocks of novel modular arithmetic equations. The first block of equations served as a pretest measure of performance and the second block served as a small amount of practice at the algorithm to stabilize performance. Immediately preceding the last block of equations, the low pressure group was informed that they would be performing another set of equations, while the high pressure group was given a scenario designed to create a high pressure environment.

From the standpoint of distraction theories, unpracticed modular arithmetic equations that are dependent on a working-memory-intensive rule-based algorithm for problem solutions should be susceptible to performance decrements under pressure. This might be especially true for the problems utilized in Experiment 1, as all of the equations consisted of double-digit numbers with a borrow operation (e.g., 43 = 28 (mod 5)). As mentioned above, large numbers and borrow operations are thought to increase task difficulty (Ashcraft, 1992; Ashcraft & Kirk, 2001). Thus, if pressure-induced distraction is responsible for choking, difficult equations with large capacity demands should be most likely to fail.

If, however, explicit monitoring theories are correct in the domain of mathematical problem solving, and performance pressure prompts explicit attention to skill execution processes, then the unpracticed modular equations utilized in Experiment 1 should not be harmed by increased pressure. Here, pressure-induced attention to execution should not impact information that is already explicitly attended on line.

2

v

c

a

g

a

fi

ra

de

wa

of a

a se

that

exam

possil

proble

standa

arithmet

2.1.1 Method

Participants. Participants were students enrolled at Michigan State University who were not math majors, had taken no more than 2 introductory college-level math courses, and had no previous exposure to the modular arithmetic (MA) task. Participants were randomly assigned to either a low pressure group (n=18) or a high pressure group (n=18) provided their accuracy in the pretest block was greater than 80%. A minimum accuracy criterion was implemented in order to assure that the low and high pressure groups in Experiment 1 consisted of individuals with similar levels of modular arithmetic ability. This allowed for the inference that any difference in MA performance as a function of pressure group could be attributed to our pressure manipulation rather than to random disparity in group ability.

Procedure. Participants filled out a consent form and a demographic sheet detailing previous math experience. They were informed that the purpose of the study was to examine how individuals learn a new math skill. Participants were set up in front of a monitor controlled by a standard laboratory computer and introduced to MA through a series of written instructions presented on the computer screen. They were informed that they would be judging the validity of MA equations and provided with several examples. Participants were instructed to judge the validity of the equations as quickly as possible without sacrificing accuracy and when they had derived an answer to the problem presented on the screen, to press the corresponding “T” or “F” keys on a standard keyboard set up in front of them.

The stimuli were digits, the word “mod” designed to denote the modular arithmetic statement, and a congruence sign (≡). Each trial began with a fixation point

expo
repla
"T" c
word
wheth
ms in

separa
digit b
acros
were f
each p
counte
that pe
equatio

perforn
simply
possibl
high pre
precedir
group w

exposed for 500 ms in the center of the screen. The fixation point was immediately replaced by an equation, which remained on the screen until the participant pressed the "T" or "F" key. When the participant responded, the equation was extinguished and the word "Correct" or "Incorrect" was displayed on the screen for 1,000 ms, indicating whether the problem had been solved correctly. The screen then went blank for a 1,000 ms inter-trial interval.

All participants performed three blocks of 24 modular arithmetic problems each, separated by a short break of approximately 1 minute. All equations required a double-digit borrow subtraction operation (e.g., $51 \equiv 19 \pmod 4$) and were presented only once across the entire experiment. Half of the problems within each block were true and half were false. Equations within each block were presented in a different random order to each participant. Additionally, the equations presented in the last two blocks were counterbalanced across participants. This counterbalancing was done in order to assure that performance in the last block of equations was independent of the particular equations to which individuals were exposed.

The first block of equations served as a pretest measure of modular arithmetic performance (pretest block) for both the low and high pressure groups. Individuals were simply informed to perform as best they could – solving the equations as quickly as possible without sacrificing accuracy. Similar instructions were given to both the low and high pressure groups prior to the second, practice block of equations. Immediately preceding the last block of equations (posttest block), individuals in the low pressure group were simply informed that they were going to be performing another set of

prob
creat

form
score
to the
that r
that t
$5, n
set of
partici
had i
impro
did n
Finall
situat
perfo

partic
both
comp
off th

problems while individuals in the high pressure group were given a scenario designed to create a high pressure situation.

The high pressure group participants were informed that the computer used a formula that equally takes into account reaction time and accuracy in computing an "MA score." Participants were told that if they could improve their MA score by 20% relative to the preceding practice trials, they would receive $5. Participants were also informed that receiving the monetary award was a "team effort." Specifically, individuals were told that they had been randomly paired with another participant, and in order to receive their $5, not only did the participant presently in the experiment have to improve in the next set of problems, but the individual they were paired with had to improve as well. Next, participants were informed that their partner had already completed the experiment, and had improved by the required amount. If the participant presently in the experiment improved by 20%, both participants would receive $5. However, if the present participant did not improve by the required amount, neither participant would receive the money. Finally, participants were told that their performance would be video taped during the test situation so that local math teachers and professors in the area could examine individuals' performances on this new type of math task.

The experimenter set up the video camera on a tripod directly to the left of participants approximately 0.61 m away. The field of view of the video camera included both the participant and the computer screen. Participants in the high pressure group completed the last block of modular arithmetic equations. The experimenter then turned off the video camera and faced it away from the participants.

awar
wheth
pressu
seen i
of indi
academ
terms

pressur
anxiety
Trait A
State-T
consisti
in time.
"I feel a
State F
anxiety
Holyoa

percept
point s
posttes

It is an empirical question whether different types of pressure (e.g., monetary awards, peer pressure, or social evaluation) have equivalent effects on performance, and whether all pressures exert their effects via similar processes. The goal of the current pressure manipulation was to impose a high level of pressure using sources commonly seen in everyday life. For example, in athletics, team success is based on the performance of individual athletes and this performance is often scrutinized by others. And in more academic arenas, college entrance exam performance has monetary consequences in terms of scholarships and future educational opportunities.

Following completion of the MA task, participants in both the low and high pressure groups filled out a number of questionnaires designed to assess their feelings of anxiety and performance pressure. Individuals first filled out the State Form of the State-Trait Anxiety Inventory (Spielberger, Gorsuch, & Lushene, 1970). The State Form of the State-Trait Anxiety Inventory (STAI) is a well-known measure of state anxiety, consisting of 20 questions designed to assess participants' feelings at a particular moment in time. Individuals are instructed to assign a value to questions such as, "I feel calm" and "I feel at ease" on a 4 point scale ranging from 1 (not at all) to 4 (very much so). The State Form of the STAI has been used in a number of studies investigating the impact of anxiety on complex task performance (e.g., in analogical reasoning ability, see Tohill & Holyoak, 2000).

Following the STAI, participants answered a number of questions related to their perceptions of performance in the posttest. Specifically, individuals were asked on a 7 point scale (a) how important they felt it was for them to perform at a high level in the posttest – ranging from 1 (not at all important to me) to 7 (extremely important to me),

(b) hc
rangir
and (c
poor)
award

2.1.2

absolu
althou

3.94. S
perforn

more *p*
the low

signific
compar
F(1,34)

importa
that it w
anxiety

(b) how much performance pressure they felt to perform at a high level in the posttest – ranging from 1 (very little performance pressure) to 7 (extreme performance pressure), and (c) how well they thought they performed in the posttest – ranging from 1 (extremely poor) to 7 (extremely well). Individuals were then fully debriefed and given the monetary award, regardless of their performance.

2.1.2 Results

Questionnaires: The high pressure group (M = 38.83, SE = 2.73) showed higher absolute levels of *state anxiety* than the low pressure group (M = 33.89, SE = 1.44), although this difference did not reach significance, F(1,34)=2.57, p<0.12.

The low pressure group (M = 4.33, SE = 0.40) and the high pressure group (M = 3.94, SE = 0.33) did not significantly differ in their *perceptions of the importance* of performing at a high level in the posttest, F<1.

Participants in the high pressure group (M = 4.33, SE = 0.29) felt significantly more *performance pressure* in the posttest equation block in comparison to participants in the low pressure group (M = 3.39, SE = 0.31), F(1,34)=4.85, p<0.04, MSE=1.66.

Finally, participants in the high pressure group (M = 3.94, SE = 0.34) had significantly worse *perceptions of their performance* in the posttest equation block in comparison to participants in the low pressure group (M = 5.22, SE = 0.26), F(1,34)=8.91, p<0.01, MSE=1.65.

Thus, while the low and high pressure groups did not differ in terms of how important they thought it was to perform at a high level in the posttest – both believing that it was moderately important to succeed – the high pressure group reported more state anxiety and significantly heightened perceptions of performance pressure in comparison

to the
perfor
count
wheth
its ob

equati
correc
perfor
requir
1979;
Furthe
block
of 12

ANOV
10^5, nc
F(1,34
block f
posttes
group (
545.40

to the low pressure group. Furthermore, the high pressure group thought that they performed at a lower level on the posttest in comparison to their low pressure counterparts. The reader can now turn to the actual performance data to determine whether participants' self-reports of performance pressure and its consequences parallel its objectively measured impact.

Reaction time and accuracy. Reaction times (RT) were computed for each equation and retained for only those equations answered correctly. Using RT's for only correct equations helps to guard against possible speed-accuracy trade-offs in performance, allowing for the interpretation of RT data as representative of the time required for successful modular arithmetic execution (Lachman, Lachman, & Butterfield, 1979; Pachella, 1974; Sternberg, 1969). Accuracy data was analyzed separately. Furthermore, RT's more than 3 SD below or above an individual's mean RT for each block of equations were considered outliers and removed. This resulted in the dismissal of 12 RT and corresponding accuracy scores from the entire data set.

A 2 (low pressure group, high pressure group) x 2 (pretest block, posttest block) ANOVA on RT indicated a main effect of block, $F(1,34)=92.89$, $p<0.01$, $MSE=14.74 \times 10^5$, no main effect of group, $F(1,34)=1.74$, ns., and no block x group interaction $F(1,34)=2.56$, ns. Reaction times significantly decreased from the pretest to the posttest block for both the low pressure group (pretest: $\underline{M} = 8582.07$ ms, $\underline{SE} = 660.48$ ms; posttest: $\underline{M} = 6282.25$ ms, $\underline{SE} = 430.13$ ms), $t(17)=6.23$, $p<0.01$, and the high pressure group (pretest: $\underline{M} = 10141.88$ ms, $\underline{SE} = 802.09$ ms; posttest: $\underline{M} = 6925.79$ ms, $\underline{SE} = 545.40$ ms), $t(17)=7.36$, $p<0.01$.

F(1,34

MSE=

signifi

93.03

accura

85.56

signifi

accura

high p

F(1,34

A similar ANOVA on accuracy revealed no main effect of block, $F<1$, or group, $F(1,34)=3.61$, ns., but a significant block x group interaction, $F(1,34)=6.28$, $p<0.02$, $MSE=0.01$. As can be seen in Figure 1, while the low pressure group's accuracy significantly increased from the pretest (M = 89.44%, SE = 1.14%) to the posttest (M 93.03%, SE = 1.35%), $F(1,17)=6.10$, $p<0.03$, $MSE=0.01$, the high pressure group's accuracy declined from the pretest (M = 90.00%, SE = 1.07%) to the posttest (M = 85.56%, SE = 2.74%), $F(1,17)=2.42$, $p<0.14$, $MSE=0.01$, although this decrease was not significant. Additionally, the low and high pressure groups did not differ in terms of MA accuracy in the pretest, $F<1$. This was not the case in the posttest however, in which the high pressure group had significantly worse accuracy than the low pressure group, $F(1,34)=5.99$, $p<0.02$, $MSE=0.01$.

F
in
er

2.

sh

the

unp

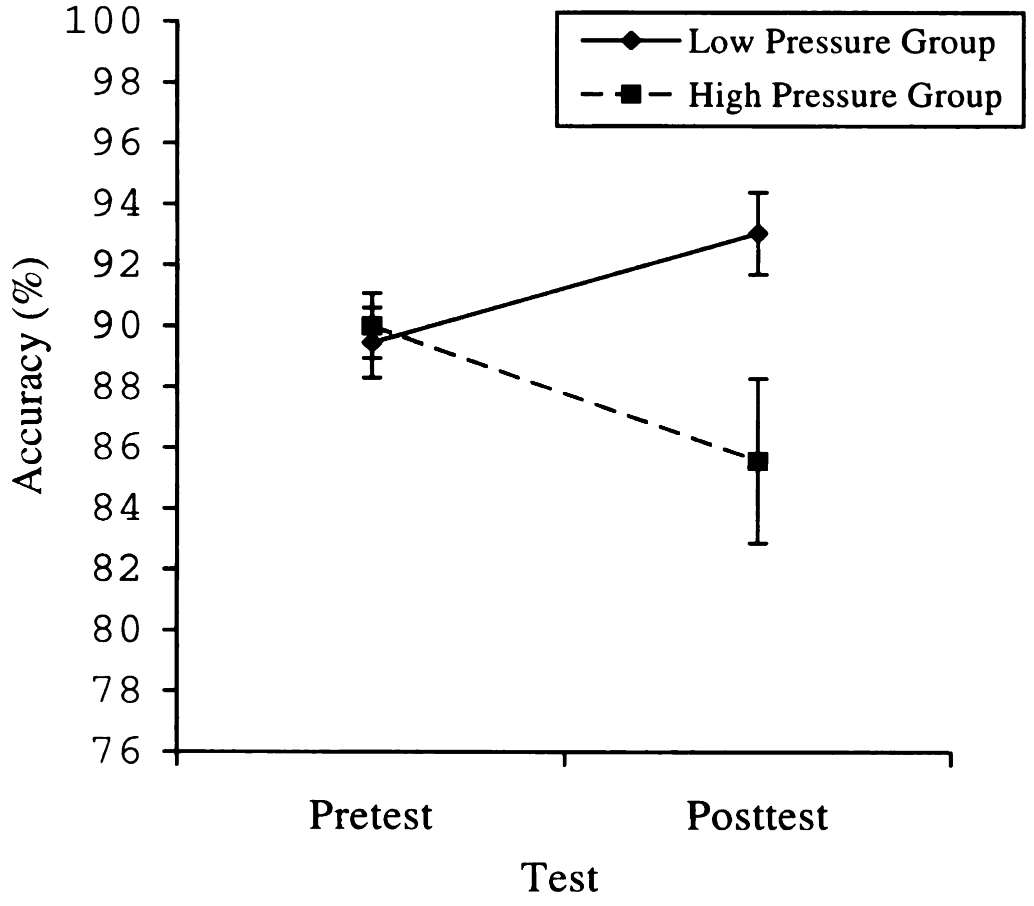

Figure 1. Mean accuracy (% correct) for the low pressure group and high pressure group in the pretest and posttest equation blocks in Experiment 1. Error bars represent standard errors.

2.1.3 Discussion

Experiment 1 was designed to examine performance under pressure in a task that should be susceptible to choking according to distraction, but not explicit monitoring theories. Individuals assigned to either a low pressure or high pressure group performed unpracticed modular arithmetic equations requiring both large number manipulations and a borrow

1992; Ash

Re

pressure g

accuracy f

Instead, p

following

Pa

the high p

higher lev

the low pr

it was equ

This findi

variations

D

arithmeti

during pro

distractio

in the hig

group on

pressure s

If

modular a

a borrow operation – two variables thought to increase problem difficulty (Ashcraft, 1992; Ashcraft & Kirk, 2001).

Reaction times decreased from the pretest to the posttest for both the low and high pressure groups. While this decrease in reaction time was accompanied by an increase in accuracy for the low pressure group, this was not the case for the high pressure group. Instead, participants in the high pressure group declined in modular arithmetic accuracy following the instantiation of a high pressure situation.

Participants' self-reports mirrored these performance differences. Individuals in the high pressure group felt significantly more performance pressure, had moderately higher levels of state anxiety, and thought that they performed significantly worse than the low pressure group. Furthermore, both the low and high pressure groups reported that it was equally important to perform at a high level on the posttest block of equations. This finding suggests that differences in motivation were not responsible for the variations in modular arithmetic performance reported above.

Distraction theories predict that choking is most likely to occur in novel modular arithmetic problems that require a rule-based algorithm to be held in working memory during problem solution – as these are the equations most susceptible to pressure-induced distractions. This pattern of data was clearly born out in Experiment 1 in that individuals in the high pressure group showed performance decrements relative to the low pressure group on unpracticed, difficult modular equations following the introduction of a high pressure situation.

If distraction is responsible for performance decrements under pressure in modular arithmetic by way of decreased capacity for task-relevant problem solving, then it foll

choke

alterin

equati

2.2 E

In

the ex

were s

Experi

proble

to crea

operati

attentic

and bor

Ashcra

increasi

Ac

pronour

memory

(e.g., pro

susceptil

intensive

it follows that equations with higher on-line attention demands should be more prone to choke under pressure than less demanding problems. Experiment 2 tested this notion by altering the working memory demands of modular arithmetic through the manipulation of equation difficulty.

2.2 Experiment 2

Individuals carried out the exact same experimental procedure as Experiment 1 with the exception that modular arithmetic problems with varying working memory demands were substituted for the difficult modular arithmetic problems solely utilized in Experiment 1. Working memory demands were manipulated through two different problem difficulty levels: Single-digit problems without a borrow operation were thought to create the least on-line capacity demands and double-digit problems with a borrow operation (the same type of equations used in Experiment 1) were assumed to be the most attention demanding. As mentioned above, problem size (single-digit vs. double-digit) and borrow operations were chosen as the means to establish these comparisons as Ashcraft (1992) has suggested that these are the two variables most associated with increasing math task problem difficulty.

According to distraction theories, performance decrements should be most pronounced in equations that possess the heaviest on-line executive control and working memory maintenance demands. Under this view, difficult modular arithmetic equations (e.g., problems that require both large numbers a borrow operation) should be more susceptible to performance decrements under pressure than easier, less working-memory-intensive equations (e.g., single-digit problems without a borrow operation). This pattern

of res
doma

2.3.1

who v
course
were r
(n=40

Experi
groups
pressur
arithme
Experin
in any w

modular
the MA
modular
borrow su
digit borr
with inter
operation (

of results would replicate Experiment 1 and further support distraction theories in the domain of mathematical problem solving.

2.3.1 Method

Participants. Participants were students enrolled at Michigan State University who were not math majors, had taken no more than 2 introductory college-level math courses, and had no previous exposure to the modular arithmetic (MA) task. Participants were randomly assigned to either a low pressure group (n=40) or a high pressure group (n=40).

Unlike Experiment 1, a minimum accuracy criterion was not implemented in Experiment 2. A larger sample size in Experiment 2 reduced the initial variability across groups. Thus, a minimum accuracy criterion was not necessary to assure that the low pressure and high pressure groups consisted of individuals with similar levels of modular arithmetic ability. However, implementing the same minimum accuracy criterion used in Experiment 1 would not have significantly altered the pattern of results in Experiment 2 in any way.

Procedure. The procedure was identical to Experiment 1 with the exception that modular arithmetic (MA) problems with varying levels of difficulty were substituted for the MA equations utilized in Experiment 1. Specifically, in each of the three blocks of 24 modular arithmetic problems in Experiment 2, eight equations required a single-digit no borrow subtraction operation (e.g., $7 \equiv 2 \pmod 5$) and eight equations required a double-digit borrow subtraction operation (e.g., $51 \equiv 19 \pmod 4$). An additional eight equations with intermediate attentional demands, requiring a double-digit no borrow subtraction operation (e.g., $19 \equiv 12 \pmod 7$), were also included in each of the equation blocks.

These

contr

probl

partici

Additi

counte

that pe

equatic

2.3.2

signific

1.21), F

pressur

percept

F(1,78)

least "m

F

more *pr*

pressure

F

significan

These intermediate level equations served as filler problems, intended to diminish the contrast between the easy single-digit problems and the difficult double-digit borrow problems. Half of the problems within each operation were true and half were false.

Equations within each block were presented in a different random order to each participant and across the entire experiment, each problem was presented only once. Additionally, as in Experiment 1, the equations presented in the last two blocks were counterbalanced across participants. This counterbalancing was done in order to assure that performance in the last block of equations was independent of the particular equations individuals were exposed to.

2.3.2 Results

Questionnaires: The high pressure group (M = 42.65, SE = 1.87) showed significantly higher levels of *state anxiety* than the low pressure group (M = 32.13, SE = 1.21), F(1,78)=22.35, $p<0.01$, MSE=99.15.

Participants in the low pressure group (M = 4.63, SE = 0.21) and the high pressure group (M = 5.03, SE = 0.19) did not significantly differ in terms of their *perceptions of the importance* of performing at a high level in the posttest equation block, F(1,78)=1.95, ns. As in Experiment 1, on average, both groups reported that it was at least "moderately important" to perform at a high level on these equations.

Participants in the high pressure group (M = 5.08, SE = 0.21) felt significantly more *pressure* to perform at a high level in the posttest than individuals in the low pressure group (M = 3.95, SE = 0.24), F(1,78)=12.44, $p<0.01$, MSE=2.03.

Finally, participants in the high pressure group (M = 4.03, SE = 0.20) had significantly worse *perceptions of their performance* in the posttest equation block in

compar
F(1,78)

While
import
that it
signifi
perfor
high p
postte
demor
perfor
increa

equat
RT's
equat
and c

digit
test, F
F(1,7
and n

comparison to participants in the low pressure group (M = 4.98, SE = 0.19), $F(1,78)=11.55$, $p<0.01$, MSE=1.56.

The questionnaire results described above replicate Experiment 1's findings. While the low and high pressure groups in Experiment 2 did not differ in terms of how important they thought it was to perform at a high level in the posttest – both believing that it was moderately important to succeed – the high pressure group reported significantly higher levels of state anxiety and significantly heightened perceptions of performance pressure in comparison to their low pressure counterparts. Additionally, the high pressure group thought that they performed at a significantly lower level on the posttest than the low pressure group. Similar to Experiment 1, the questionnaire results demonstrate that our manipulation was successful in increasing participants' feelings of performance pressure. Again, we can turn to the behavioral data to determine if these increased perceptions of pressure parallel actual modular arithmetic performance.

Reaction time and accuracy. Reaction times (RT) were computed for each equation and retained for only those equations answered correctly. As in Experiment 1, RT's more than 3 SD below or above an individual's mean RT for each block of equations were considered outliers and removed. This resulted in the dismissal of 18 RT and corresponding accuracy scores from the entire data set.

A 2(low pressure group, high pressure group) x 2(pretest, posttest) x 2(single-digit problems, double-digit borrow problems) ANOVA on RT revealed a main effect of test, $F(1,78)=48.77$, $p<0.01$, $MSE=15.03 \times 10^5$, a main effect of problem difficulty, $F(1,78)=615.54$, $p<0.01$, $MSE=40.68 \times 10^5$, no main effect of group, $F(1,78)=3.06$, ns., and no test x pressure group x problem difficulty interaction, $F<1$.

As can be seen in Table 1, reaction times significantly decreased from the pretest to the posttest equation blocks for the single-digit problems and the double-digit borrow problems for the low pressure group, $t(39)=5.05$, $p<0.01$, and $t(39)=2.60$, $p<0.02$, respectively, and the high pressure group, $t(39)=7.71$, $p<0.01$, and $t(39)=5.20$, $p<0.01$, respectively.

Low F

Sing

Dou

High I

Sing

Dou

Table
Low F
Doubl
Exper

	Pretest		Posttest	
	(M)	(SE)	(M)	(SE)
Low Pressure Group				
Single-Digit				
RT (ms)	2444.10	94.65	1982.75	91.22
Accuracy (%)	93.13	1.55	98.44	0.80
Double-Digit Borrow				
RT (ms)	8814.71	449.81	7816.94	404.85
Accuracy (%)	81.88	2.83	85.31	1.95
High Pressure Group				
Single-Digit				
RT (ms)	2530.52	109.14	1846.02	84.78
Accuracy (%)	94.69	1.33	98.44	0.66
Double-Digit Borrow				
RT (ms)	8117.57	423.13	6432.11	357.65
Accuracy (%)	80.00	2.32	74.06	2.66

Table 1. Mean Modular Arithmetic Reaction Time (ms) and Accuracy (% correct) for the Low Pressure Group and High Pressure Group for the Single-Digit Problems and the Double-Digit Borrow Problems in the Pretest and Posttest Equation Blocks in Experiment 2.

A
digit prob
significan
MSE=0.0

A
the single
MSE=0.0
respective
posttest) .
F<1, and
by a signi

A
in accurac
t(39)=2.9
difficult a
increased
increase v
accuracy
Additiona
double-di
however,
digit borr

A 2(low pressure group, high pressure group) x 2(pretest, posttest) x 2(single-digit problems, double-digit borrow problems) ANOVA on accuracy revealed a significant test x pressure group x problem difficulty interaction, $F(1,78)=4.01$, $p<0.05$, $MSE=0.01$.

A 2(low pressure group, high pressure group) x 2(pretest, posttest) ANOVA on the single-digit MA problems revealed a main effect of test, $F(1,78)=14.60$, $p<0.01$, $MSE=0.01$, no main effect of group, and no test x pressure group interaction, $F\text{'s}<1$ respectively. In contrast, a 2(low pressure group, high pressure group) x 2(pretest, posttest) ANOVA on the double-digit borrow problems revealed no main effect of test, $F<1$, and a main effect of group, $F(1,78)=4.88$, $p<0.05$, $MSE=0.04$, which was qualified by a significant pressure group x test interaction, $F(1,78)=6.65$, $p<0.02$, $MSE=0.01$.

As can be seen in Table 1 and Figure 2, the easier single-digit problems increased in accuracy from the pretest to posttest equation blocks for both the low pressure group, $t(39)=2.98$, $p<0.01$, and the high pressure group, $t(39)=2.40$, $p<0.03$. In terms of the more difficult and capacity-demanding double-digit borrow problems, the low pressure group increased in accuracy from the pretest to the posttest, $t(39)=1.17$, ns., although this increase was not significant. In contrast, the high pressure group significantly declined in accuracy from the pretest to the posttest equation blocks, $t(39)=2.77$, $p<0.01$. Additionally, the low and high pressure groups did not differ in terms of accuracy on the double-digit borrow problems in the pretest, $F<1$. This was not the case in the posttest however, in which the high pressure group was significantly less accurate on the double-digit borrow problems than the low pressure group, $F(1,78)=11.64$, $p<0.01$, $MSE=0.02$.

Figure 2
for the s
the prete
errors.

2.3.3 D

In

performed

problem c

state anxie

of the high

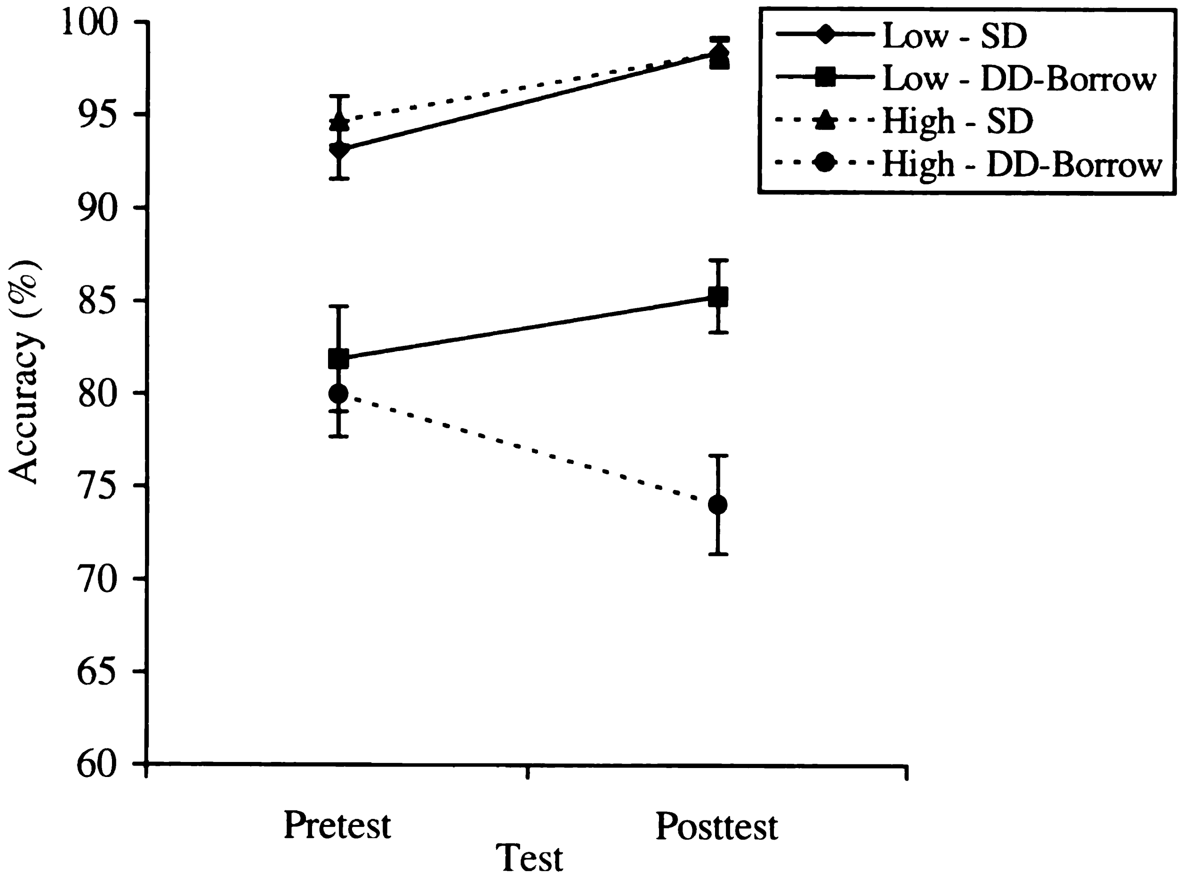

Figure 2. Mean accuracy (% correct) for the low pressure group and high pressure group for the single-digit problems (SD) and the double-digit borrow problems (DD-Borrow) in the pretest and posttest equation blocks in Experiment 2. Error bars represent standard errors.

2.3.3 Discussion

In Experiment 2, individuals assigned to either a low or high pressure group performed novel, unpracticed modular arithmetic equations that varied as a function of problem difficulty. Individuals in the high pressure group reported increased levels of state anxiety and heightened feelings of performance pressure following the instantiation of the high pressure scenario in comparison to participants in the low pressure group.

Additionally, individuals in the high pressure group performed at a significantly lower accuracy level on the MA equations after receiving the pressure scenario in comparison to both their pretest performance and the performance of their low pressure counterparts. Analysis of these performance failures as a function of problem difficulty revealed that only the most difficult equations, requiring both large number manipulations and a borrow operation, were performed at a lower accuracy level under pressure.

This pattern of results replicates and extends Experiment 1's support for distraction theories of choking in the domain of mathematical problem solving. Unpracticed modular arithmetic equations, whose solutions require the maintenance of intermediate problem steps and their products in working memory, choke under pressure. Furthermore, these performance failures are limited to those equations with the heaviest on-line maintenance demands (i.e., double-digit borrow problems). According to distraction theories, pressure-induced limitations in working memory capacity cause choking. Thus, novel equations with large on-line attentional demands should be precisely the type of equations for which performance failures under pressure are most likely to occur.

The first two experiments lend support to distraction theories as an explanation for the choking phenomenon in the domain of mathematical problem solving. However, it is still possible that performance decrements under pressure may occur at high levels of practice via the mechanisms proposed by explicit monitoring theories. Experiment 3 was designed to test this notion.

2.3 E

memo
extend
deman
proble
and d
deman
were t
pressu
consid

suscep
Here,
worki
decre
that in
higher
are bei
related

via the
based o

2.3 Experiment 3

In Experiment 3, individuals performed modular equations with varying working memory demands under low and high pressure conditions both prior to and following extended modular arithmetic practice. Similar to Experiment 2, working memory demands were manipulated through different problem difficulty levels: Single-digit problems without a borrow operation thought to create the least on-line capacity demands and double-digit problems with a borrow operation assumed to be the most attention demanding. Participants had no previous exposure to the specific problems on which they were tested prior to the first high pressure situation. By the time of the second high pressure test, however, participants had received 49 exposures to each problem under consideration.

As seen in Experiments 1 and 2, modular arithmetic problems should be susceptible to choking under pressure early in learning according to distraction theories. Here, pressure-induced reductions in the attentional capacity needed to carry out working-memory-intensive problem solving processes should result in performance decrements. Furthermore, these failures should be most pronounced in difficult problems that incur the highest working memory load (e.g., double-digit borrow problems). At higher levels of problem-specific practice, when answers to now well-practiced problems are being retrieved from memory rather than computed via the algorithm, such capacity-related failures should no longer occur.

However, it is still possible that choking might happen at high levels of practice via the mechanism proposed by explicit monitoring theories. Well-learned problems, based on the stimulus-driven retrieval of past problem instances from memory, may fail un

(M

hig

prob

to pr

retrie

still f

and hi

2.3.1

Univers

level ma

I

detailing

was to ex

of a mon

MA task

In

order to ea

operation (

subtraction

intermedia

procedure (

under pressure because pressure-induced attention disrupts automatic answer retrieval (Masters, Polman, & Hammond, 1993). If so, then choking should be observed for all highly practiced problems regardless of difficulty – at least to the extent that practiced problems are solved via automatic answer retrieval. Furthermore, even if practice serves to proceduralize the algorithm, rather than shift performance to automatic answer retrieval as Logan (1988) would propose, practiced modular arithmetic problems might still fail via pressure-induced attention that serves to disrupt or slow down well-learned and highly efficient algorithmic processes.

2.3.1 Method

Participants. Participants (N=20) were students enrolled at Michigan State University who were not math majors, had taken no more than 2 introductory college-level math courses, and had no previous exposure to modular arithmetic (MA).

Procedure. Participants filled out a consent form and a demographic sheet detailing previous math experience. They were informed that the purpose of the study was to examine how individuals learn a new math skill. Individuals were set up in front of a monitor controlled by a standard laboratory computer and introduced to the same MA task used in Experiments 1 and 2.

Individuals first performed 12 practice problems, presented in a different random order to each participant. Four of the equations required a single-digit subtraction operation (e.g., $7 \equiv 2 \pmod{5}$) and four equations required a double-digit borrow subtraction operation (e.g., $51 \equiv 19 \pmod{4}$). An additional four equations with intermediate attentional demands, requiring a double-digit no borrow subtraction procedure (e.g., $15 \equiv 10 \pmod{3}$), were also included. As in Experiment 2, these intermediate level equations served as filler problems, intended to diminish the contrast between the easy single-digit problems and the difficult double-digit borrow problems. Half of the equations within each operation were true, half were false, and each equation was presented only once.

Following the practice problems, individuals completed a 12 equation low pressure test (LP1) and a 12 equation high pressure test (HP1), separated by a short break of approximately 1 minute. The equations in LP1 and HP1 were presented in a different random order to each participant. Each equation appeared only once in either LP1 or HP1 and the equations in LP1 and HP1 were counterbalanced across participants. Within both LP1 and HP1 there were four equations with a single-digit subtraction operation and four equations with a double-digit borrow subtraction operation. Four equations with a double-digit no borrow subtraction operation, that served as intermediate difficulty filler problems, were also included. Half of the equations within each operation were true and half were false.

To the participant, LP1 appeared to be just another series of practice equations. Following LP1, participants were given a scenario designed to create a high pressure situation. The same high pressure scenario and video camera situation used in Experiments 1 and 2 was presented to participants in Experiment 3, with the exception that individuals in Experiment 3 were informed that they were about to enter the first of two test situations in the experiment and that they had to improve their MA performance by the required amount in both test situations in order to receive the monetary award for themselves and their partner.

Following HP1, individuals were informed that they would be performing a series of practice MA equations (MA training). Participants were presented with 12 new equations. Four of these equations required a single-digit subtraction operation and four equations required a double-digit borrow subtraction operation. An additional four equations with a double-digit no borrow subtraction operation were also included. Half of the equations within each operation were true and half were false. Each equation within this MA training session was repeated 48 times for a total of 576 trials, separated into 3 blocks of 192 equations each, with a short break of approximately 1 minute after each block. Within each block, each equation was repeated 16 times. Trials were presented in a different random order to each participant.

Participants then took part in the second 12 equation low pressure test (LP2) and the second 12 equation high pressure test (HP2). The equations within LP2 and HP2 were the same 12 equations that were presented 48 times each in the training session. Each participant received the equations within LP2 and HP2 in a different random order. There was no need to counterbalance these equations across the second low and high pressure tests, as the same 12 equations were used in both LP2 and HP2.

As in LP1, participants were not made aware of the LP2 test situation. They were not told that this block would be shorter, nor were they given any other cues. To the participant, LP2 appeared to be just another series of practice problems. The experimenter then informed participants that they were about to take part in the second test situation, repeated the high pressure scenario, and turned the video camera on. Participants completed HP2. Individuals were then fully debriefed and given the monetary award regardless of their performance. In total, participants were exposed to 50 presentations of each of the 12 training problems (48 exposures during the training session, 1 exposure in LP2, and 1 exposure in HP2).

The self-report measures of performance pressure and anxiety administered in Experiments 1 and 2 were not utilized in Experiment 3. Experiment 3 was a completely within participants design. While one might imagine that the questionnaires could be administered multiple times in order to create a within participant comparison (i.e., once prior to and once following the instantiation of each high pressure scenario), pilot testing revealed that asking for such off-line measures of performance pressure prior to the introduction of a high pressure situation significantly increased participants' skepticism regarding the validity of the pressure manipulation. Therefore, in order to present the strongest pressure manipulation possible, the questionnaires were excluded. As will be seen, however, behavioral evidence of choking made it clear that the manipulation was again effective.

2.3.2 Results

Reaction times (RT) were computed for each equation and retained for only those equations answered correctly. As in the previous experiments, RT's more than 3 SD below or above an individual's mean RT for each experimental condition were considered outliers and removed. A total of 3 RT and corresponding accuracy measures were dismissed from the entire data set.

A 2 (low pressure test, high pressure test) x 2 (before training, following training) ANOVA on RT indicated a main effect of time, $F(1,19)=61.17$, $p<0.01$, $MSE=28.89 \times 10^5$, no main effect of pressure, $F(1,19)=1.83$, ns., $MSE=24.41 \times 10^4$, and no time x pressure interaction, $F<1$. Reaction times were slower prior to MA training in LP1 ($\underline{M}$ =

4114.70 ms, SE = 415.27 ms) and HP1 (M = 4001.97, SE = 432.09), than following training in LP2 (M = 1178.44 ms, SE = 100.87 ms) and HP2 (M = 992.71 ms, SE = 63.04 ms).

A similar ANOVA on accuracy revealed a main effect of time, F(1,19)=11.22, p<0.01, MSE=0.01, no main effect of pressure, F(1,19)=2.68, ns., MSE=0.01, and a significant time x pressure interaction, F(1,19)=6.24, p<0.03, MSE=0.01. As can be seen in Figure 3, while MA accuracy significantly declined from LP1 (M = 92.92 %, SE = 2.28 %) to HP1 (M = 86.67 %, SE = 2.66 %), t(19)=2.26, p<0.04, accuracy improved somewhat from LP2 (M = 96.25 %, SE = 0.95 %) to HP2 (M = 97.92 %, SE = 1.03 %), t(19)=1.45, ns., although this improvement was not significant. Thus MA performance decrements under pressure occurred prior to extended problem training, but not following it. This pattern of data supports the predictions of distraction theories as an explanation for the choking phenomenon in that only novel MA problems, requiring the instantiation of a capacity-demanding rule-based solution algorithm, choked under pressure. After extended practice of the problems being tested, behavioral evidence of choking was no longer observed.

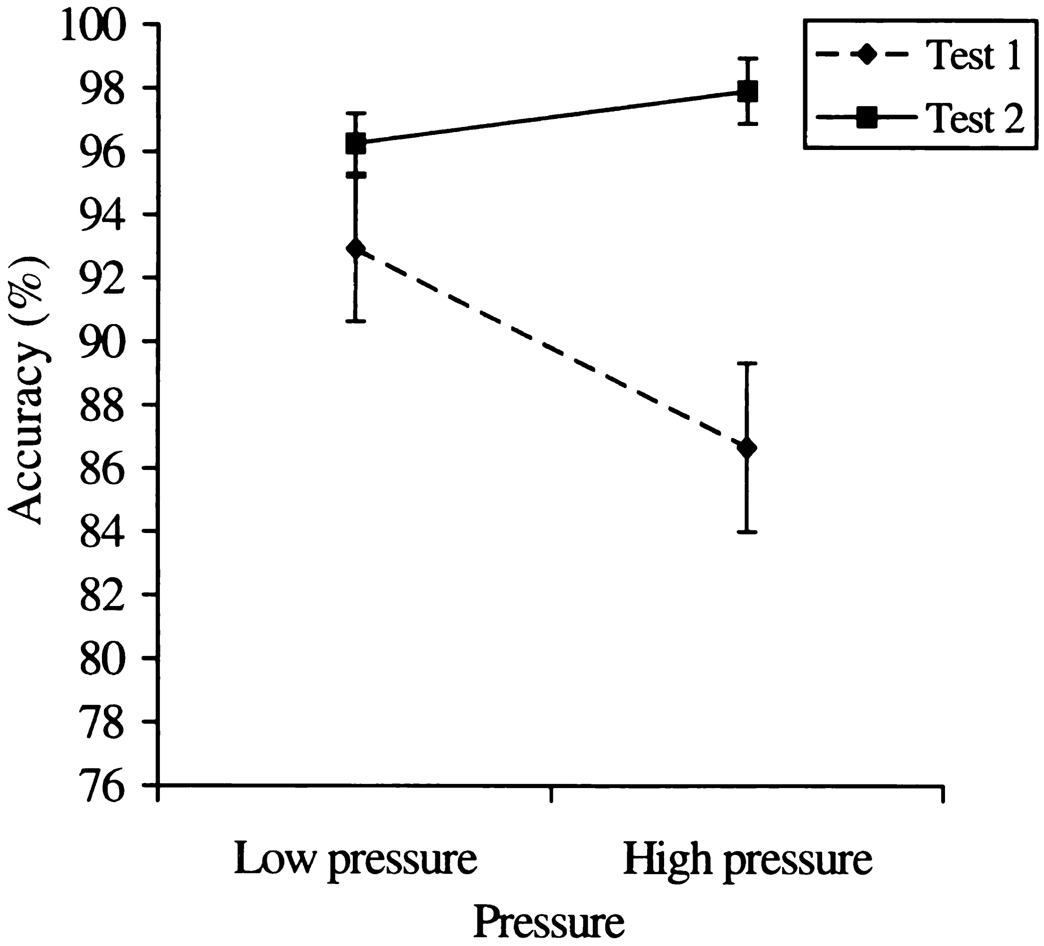

Figure 3. Mean accuracy (% correct) for the low and high pressure tests prior to modular arithmetic training (Test 1) and following modular arithmetic training (Test 2) in Experiment 3. Error bars represent standard errors.

If distraction theories are correct, then the pressure-induced performance decrements just observed for novel MA problems should be most pronounced in equations that possess the heaviest on-line executive control and working memory maintenance demands (i.e., double-digit problems that require a borrow operation). In order to explore this possibility, we compared RT and accuracy of the least demanding single-digit problems and most difficult double-digit borrow problems across the low pressure and high pressure tests administered prior to MA, where performance decrements under pressure were shown to have occurred.

Applied to RT, this 2 (single-digit, double-digit borrow) x 2 (LP1, HP1) ANOVA produced a main effect of problem difficulty, F(1,19)=27.64, $p<0.01$, MSE=36.87 x 10^7, no main effect of pressure, and no pressure x difficulty interaction, F's<1 respectively. Reaction times increased as a function of problem difficulty during both LP1 (single-digit: M = 2263.13 ms, SE = 183.33 ms; double-digit borrow: M = 6393.77 ms, SE = 884.09 ms) and HP1 (single-digit: M = 2262.02 ms, SE = 165.12 ms; double-digit borrow: M = 6718.47 ms, SE = 873.23 ms).

A similar ANOVA on accuracy produced significant main effects of problem difficulty, F(1,19)=14.46, $p<0.01$, MSE=0.03, and pressure, F(1,19)=6.17, $p<0.03$, MSE=0.02, and a significant problem difficulty x pressure interaction, F(1,19)=12.67, $p<0.01$, MSE=0.02. This interaction is shown in Figure 4. Paired sample t-tests performed as post hocs revealed that the single-digit problems did not significantly differ in accuracy from LP1 (M = 95.00 %, SE = 2.29 %) to HP1 (M = 96.25 %, SE = 2.74 %), t(19)=0.44, ns. In contrast, the accuracy of the double-digit borrow problems got significantly worse from LP1 (M = 91.25 %, SE = 3.28 %)to HP1 (M = 72.50 %, SE = 4.76 %), t(19)=3.30, $p<0.01$. Furthermore, there were no significant accuracy differences across problem difficulty levels in LP1, t(19)=1.00, ns. During HP1 however, an effect of problem difficulty occurred in which the double-digit borrow problems were significantly less accurate than the single-digit problems, t(19)=4.50, $p<0.01$. Thus, as can be seen in

Figure 4, the most difficult MA problems requiring a borrow operation showed performance decrements under pressure while the least difficult, single-digit problems did not. This finding parallels work in the math anxiety literature demonstrating that mental arithmetic problems possessing a carry operation are most susceptible to performance difficulties as a result of decreased working memory capacity (Ashcraft & Kirk, 2001).

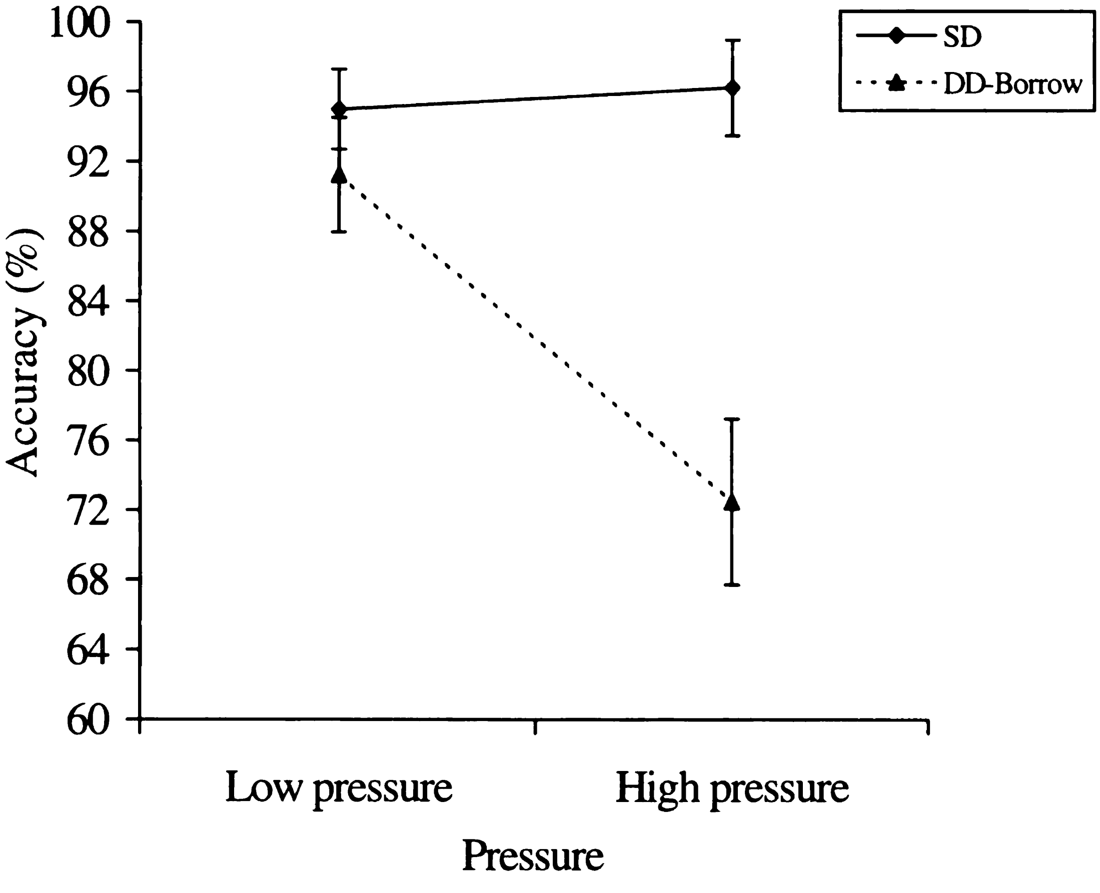

Figure 4. Mean accuracy (% correct) for the low and high pressure tests prior to modular arithmetic training for the single-digit problems (SD) and the double-digit borrow problems (DD-Borrow) in Experiment 3. Error bars represent standard errors.

Finally, in order to demonstrate that the MA equations were fully automated following extended training, and hence should have shown choking if explicit monitoring theories were applicable, I performed an analysis of reaction time and accuracy as a function of problem difficulty (i.e., single-digit, double-digit borrow). This kind of analysis has been used in other types of mental arithmetic tasks – for example, Logan's (1988) alphabet arithmetic – to diagnose the extent to which the control structures of performance have shifted from a working-memory-intensive counting algorithm that produces a significant effect of problem difficulty, to automatic memory retrieval which is independent of equation difficulty level (Klapp et al., 1991).

A significant interaction of problem difficulty by HP1 versus HP2 for RT was found, $F(1,19)=26.55$, $p<0.01$, $MSE=36.15 \times 10^5$. Prior to MA training, RT was significantly faster for the simplest single-digit problems (M = 2262.02 ms, SE = 165.12 ms) in comparison to the most difficult double-digit borrow problems (M = 6718.47 ms, SE = 873.23 ms), $t(19)=5.26$, $p<0.01$. In contrast, following MA training, there was no significant difference in RT between single-digit problems (M = 914.23 ms, SE = 55.61 ms) and double-digit borrow problems (M = 989.79 ms, SE = 67.59 ms), $t(19)=1.84$, ns.

A similar analysis of problem difficulty by HP1 versus HP2 on accuracy also revealed a significant interaction of problem difficulty by HP1 versus HP2, $F(1,19)=21.11$, $p<0.01$, $MSE=0.01$. Again, prior to MA training, accuracy was significantly higher for the simplest single-digit problems (M = 96.15% , SE = 2.74%) in comparison to the most difficult double-digit borrow problems (M = 72.50%, SE = 4.76%), $t(19)=4.50$, $p<0.01$. In contrast, following MA training, there was no significant difference in accuracy between the single-digit problems (M = 97.50%, SE = 1.72%) and the double-digit borrow problems (M = 98.75%, SE = 1.25%), t(19)=0.57, ns. Thus, following repeated exposure to MA problems, there appears to be a relative independence between MA performance (whether measured by reaction time or accuracy) and problem difficulty. This is a sign that performance following MA training was automated – based on the direct retrieval of answers from long term memory rather than working-memory-intensive algorithmic computation (Klapp et al., 1991).

2.3.3 Discussion

The results of Experiment 3 replicate the finding of the first two experiments that novel modular arithmetic equations, whose solutions require the maintenance of intermediate problem steps and their products in working memory, decline under pressure. Furthermore, as in Experiment 2, analysis of these performance decrements prior to training revealed that only the most difficult equations requiring both large number manipulations and a borrow operation failed under pressure. In contrast, well-learned modular arithmetic problems, thought to be supported by the one-step direct retrieval of past problem instances from memory, showed no signs of choking.

According to distraction theories, modular arithmetic should be most susceptible to pressure-induced performance decrements at low levels of practice when working memory demands are greatest and pressure-induced worries impinge on task-relevant processing resources. This prediction was clearly borne out. Not only was choking solely observed prior to modular arithmetic training, but similar to Experiment 2, performance decrements under pressure were limited to difficult problems that incurred the highest working memory load (i.e., double-digit borrow problems).

The first three experiments provide support for distraction theories and argue against explicit monitoring theories as an explanation for choking under pressure as observed in the working-memory-intensive task of modular arithmetic. The purpose of Experiment 4 was to further explore performance under pressure in this task by considering the role of general practice at the algorithm, through a comparison of infrequently practiced problems with heavily practiced equations under pressure, at similarly high overall levels of general algorithmic practice.

2.4 Experiment 4

Participants performed over 800 modular arithmetic problems over 3 days of practice prior to being exposed to a low and high pressure situation. Specific equations within this practice period were presented either once, repeated twice, or repeated 50 times each. As previously discussed, the task control structures of modular arithmetic problems should change most dramatically as a function of specific problem exposure, not necessarily experience at performing many different modular arithmetic problems (Logan, 1988). Thus, if choking is due to pressure-induced capacity limitations, as distraction theories would propose, then regardless of how many different problems individuals have been exposed to, only those equations that have been repeated enough to produce instance-based answer retrieval should be inoculated against the detrimental capacity-limiting effects of performance pressure.

Furthermore, even if algorithmic computations do become more efficient with practice (Touron, Hoyer, & Cerella, 2001), implementation of the algorithm should still impose attentional demands as novel problem information must be maintained and manipulated on-line in working memory during performance. This prediction can be tested by examining performance decrements under pressure as a function of problem difficulty level. If novel problems based on practiced algorithms are susceptible to pressure-induced performance decrements via distraction, then the more difficult and capacity-demanding the problem, the more vulnerable it should be to capacity-limited failure. This earmarks difficult problems repeated only a few times during practice, and still based on problem-solving algorithms, as candidates for choking under pressure according to distraction theories.

It should be noted that if practice increases the proficiency of algorithmic computations, it is also possible that performance failures under pressure (for novel problems based on practiced algorithms) could be explained by explicit monitoring theories of choking. Specifically, pressure-induced attention may serve to disrupt highly efficient, proceduralized algorithmic computations. If so, this type of skill failure should be evident, at least to some extent, across all problem difficulty levels, as practiced algorithms, regardless of working memory demands, should be harmed by the instantiation of explicit attentional control mechanisms that slow down or disrupt highly efficient computations. This difference between theories concerning whether choking should depend on the capacity demands of the problems being performed gives Experiment 4 some further leverage in distinguishing distraction from explicit monitoring as a source of performance decrements under pressure in modular arithmetic.

2.4.1 Method

Participants. Participants (N=22) were students enrolled at Michigan State University who were not math majors, had taken no more than 2 introductory college-level math courses, and had no previous exposure to the modular arithmetic (MA) task.

Procedure. Participants filled out a consent form and a demographic sheet detailing previous math experience and were informed that the purpose of the study was to examine how individuals learned a new math skill over several days of practice. Individuals performed the same MA task used in first three experiments over three days of practice. On Days 1 and 2, participants performed three blocks of 120 equations each, separated by short breaks of approximately 1 minute. On Day 3, individuals performed 90 equations, for a total of 810 practice equations over the 3 days of practice. Ten practice equations were repeated 50 times each (multiple repeats) over the 3 days of practice (22 presentations of the multiple repeat equations on Days 1 and 2; 6 presentations on Day 3), 100 equations were repeated once (once repeats) over the three practice sessions (80 of these occurred on Days 1 and 2; 10 occurred on Days 1 and 3; 10 occurred on Days 2 and 3), and 110 equations were presented only once (no repeats). Practice equations were presented in a different random order to each participant.

Following practice on Day 3, participants took part in a 30 equation low pressure test and a 30 equation high pressure test. The low pressure test consisted of the 10 equations that were repeated 50 times each during practice (multiple repeats), 10 equations that were repeated once during practice (once repeats), and 10 equations not previously presented during practice (no repeats). The high pressure test consisted of the 10 multiple repeats, 10 new once repeats, and 10 new no repeats. Equations within these tests were presented in a random order to each participant.

To the participant, the low pressure test appeared to be just another series of practice equations. Participants then completed the high pressure test, consisting of the same high pressure scenario and video camera situation utilized in the first three experiments. Participants were then fully debriefed and given the monetary award regardless of their performance.

2.4.2 Results

Reaction times (RT) were computed for each equation and retained for only those equations answered correctly. There were no RT and corresponding accuracy measure outliers in either the low or high pressure test for any participants utilizing the 3 SD outlier criterion established in the first three experiments. However, 5 equations in the low pressure test (2 once repeats; 3 no repeats) and 3 equations in the high pressure test (1 once repeat; 2 no repeats) were discarded from the subsequent analyses because the accuracy of these equations across participants was not significantly different from chance.

I began by comparing once repeat problems to no repeat problems in order to determine if a small amount of problem-specific exposure changes performance. A 2 (low pressure, high pressure) x 2 (once repeats, no repeats) ANOVA on accuracy produced no main effects of pressure, or problem repetition, and no pressure x repetition interaction, F's<1 respectively. Accuracy was relatively high and did not differ across once repeats and no repeats for either the low pressure test (once repeats: M = 87.50%, SE = 2.97%; no repeats: M = 87.01%, SE = 3.51%) or high pressure test (once repeats: M = 89.39%, SE = 1.20%; no repeats: M = 89.20%, SE = 2.88%).

A similar ANOVA on reaction time revealed main effects of pressure, F(1,21)=15.04, p<0.01, MSE=29.18 x 10^4, and problem repetition, F(1,21)=12.38, p<0.01, MSE=36.07 x 10^4, and no pressure x repetition interaction, F<1. As can be seen from Figure 5, RT significantly increased from the low pressure test to high pressure test for both once repeats and no repeats, t(21)=2.61, p<0.02, and t(21)=2.41, p<0.03 respectively. No repeats were significantly slower than once repeats during the low pressure test t(21)=2.94, p<0.01, but did not differ from once repeats during the high pressure test, t(21)=1.90, ns. This latter outcome makes it tempting to speculate that once repeats were more susceptible to pressure than no repeats, but as already reported, the pressure x repetition interaction produced an F<1.

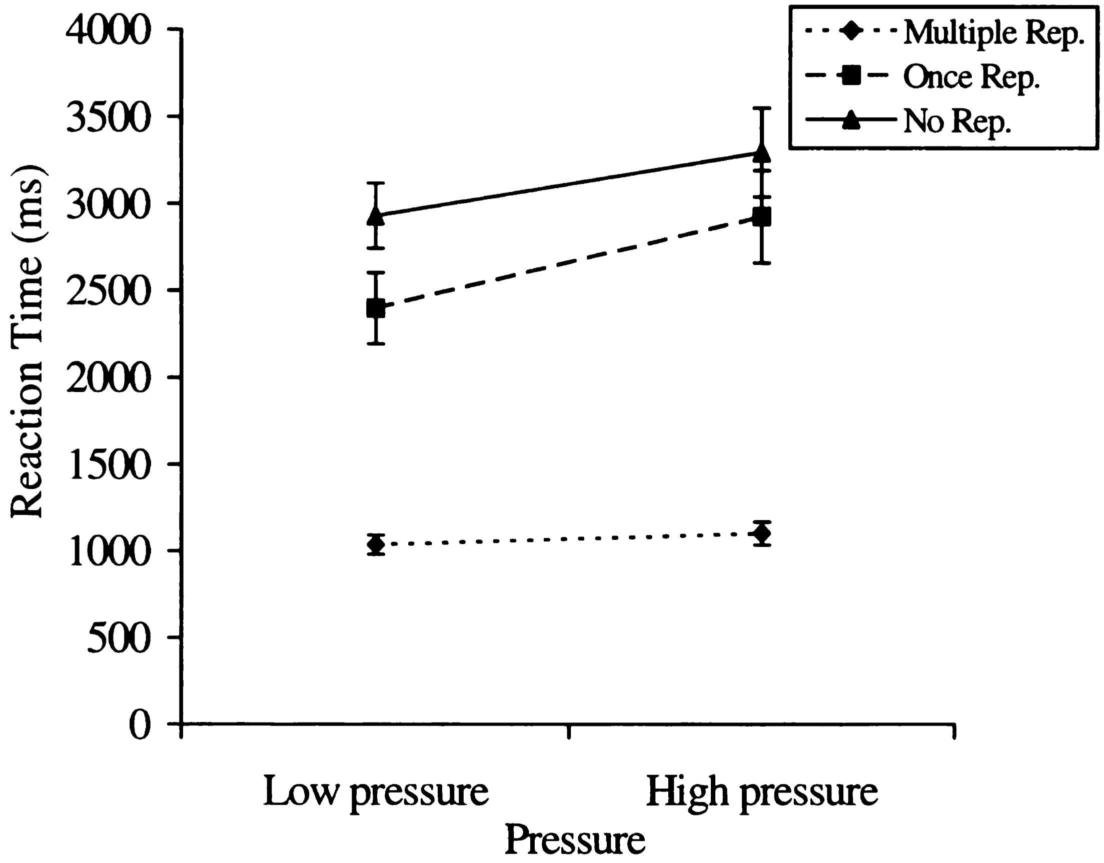

Figure 5. Mean reaction time (ms) for the low and high pressure tests for the multiple repeat problems (Multiple Rep.), the once repeat problems (Once Rep.), and no repeat problems (No Rep.) in Experiment 4. Error bars represent standard errors.

Thus, although the once repeats were somewhat faster than the no repeats, both showed similar patterns of susceptibility to performance decrements under pressure. This is precisely the pattern of results that distraction theories would predict given that neither the once repeats nor the no repeats had been practiced to the extent necessary to be dependent on instance-based answer retrieval in the high pressure situation. However, it is also possible that the once repeats and no repeats failed under pressure due to maladaptive explicit monitoring that served to slow down or disrupt highly practiced algorithmic execution. One way to discriminate between these two possibilities is to examine the performance of these relatively novel problems under pressure as a function of problem difficulty. If distraction theories are correct, then the more difficult and capacity-demanding the problem, the more vulnerable it should be to capacity-limited failure. In contrast, if explicit monitoring theories apply to the performance of the once repeat and no repeat problems under pressure, then failure should be evident to some extent across all problem difficulty levels. This is due to the fact that any practiced algorithm should be harmed by the pressure-induced instantiation of attentional control that serves to slow down or disrupt highly efficient computations.

In order to explore these possibilities, the no repeat problems were next used in an ANOVA comparing the performance of the least practiced problems with the most practiced problems as a function of performance pressure and problem difficulty. Comparing multiple repeat problems (that should be instance-based as result of extended practice) with no repeat problems (that have had the least opportunity to become reliant on automatic answer retrieval) is the strongest test of performance under pressure as a function of task practice. However, as can be seen from Figure 5, because the once repeat and no repeat problems behaved similarly under pressure, utilizing once repeat problems in the subsequent analyses would not have altered the pattern of results.

A 2 (multiple repeats, no repeats) x 2 (low pressure, high pressure) x 2 (no borrow problems, borrow problems) ANOVA on accuracy revealed a main effect of problem repetition, F(1,21)=5.24, p<0.04, MSE=0.02. No other interactions or main effects reached significance. As can be seen from Table 2, the multiple repeat equations were higher in accuracy than the no repeat problems across both the low and high pressure tests, regardless of problem difficulty.

Problen

NR – N

RT (n

Accur

NR – B

RT (n

Accu

MR – N

RT (n

Accu

MR – I

RT (

Accu

*Diffe
improv
Positiv
perfor

Table
High I
(MR)
for Ex

	Low Pressure		High Pressure		Difference*	
Problem	(M)	(SE)	(M)	(SE)	(M)	(SE)
NR – No Borrow						
RT (ms)	2311	157.43	2061	164.94	250	132.41
Accuracy (%)	90.91	2.84	92.05	3.03	-1.14	2.47
NR – Borrow						
RT (ms)	3274	255.11	3938	325.36	-664	171.95
Accuracy (%)	87.88	4.68	89.09	3.15	-1.21	4.61
MR – No Borrow						
RT (ms)	967	57.56	1036	56.28	-69	47.53
Accuracy (%)	94.77	2.17	96.82	1.63	-2.05	2.92
MR – Borrow						
RT (ms)	1075	74.65	1212	116.60	-137	107.50
Accuracy (%)	93.93	3.56	95.45	2.50	-1.52	2.66

*Difference = Low Pressure - High Pressure. *RT*: Positive represents performance improvement under pressure; negative represents performance decrement. *Accuracy*: Positive represents performance decrement under pressure; negative represents performance improvement.

Table 2. Mean Modular Arithmetic Reaction Time (RT) and Accuracy for the Low and High Pressure Tests for the No Repeat Problems (NR) and Multiple Repeat Problems (MR) with a Borrow Operation (Borrow) and without a Borrow Operation (No Borrow) for Experiment 4.

interact
pressure
no press
pressure
signific
Paired s
borrow
t(21)=1
from th

A similar ANOVA on RT produced a significant repetition x pressure x difficulty interaction, $F(1,21)=15.43$, $p<0.01$, $MSE=12.76 \times 10^4$ (Figure 6). A 2 (low pressure, high pressure) x 2 (no borrow, borrow) ANOVA within the multiple repeat problems revealed no pressure x problem difficulty interaction, $F<1$. In contrast, a 2 (low pressure, high pressure) x 2 (no borrow, borrow) ANOVA within the no repeat problems revealed a significant pressure x difficulty interaction, $F(1,21)=22.38$, $p<0.01$, $MSE=20.54 \times 10^4$. Paired sample t-tests performed as post-hocs demonstrated that RT for no repeat no borrow problems did not significantly differ from the low to high pressure test, $t(21)=1.89$, ns. In contrast, RT for no repeat, borrow problems got significantly slower from the low to high pressure test, $t(21)=3.86$, $p<0.01$.

Figure 6
repeat p
problem
borrow
(NR-bo

problem

problem

showed

problem

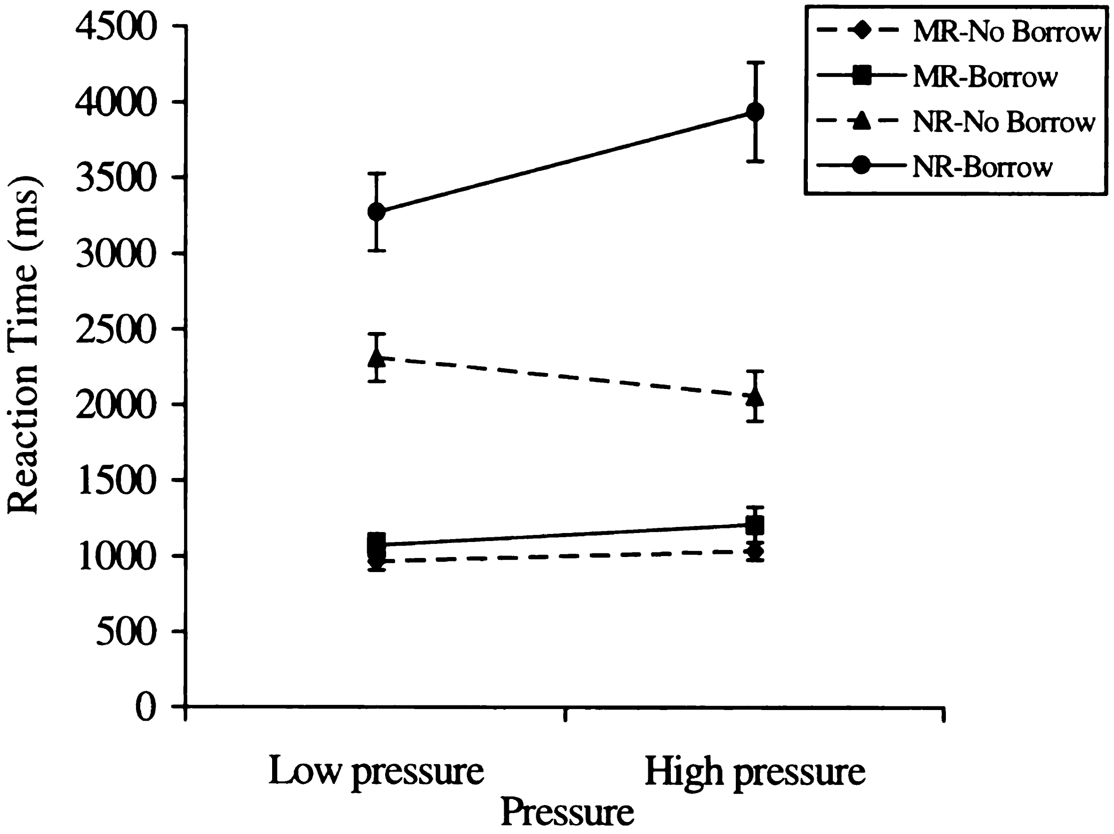

Figure 6. Mean reaction times (ms) for the low and high pressure tests for the multiple repeat problems without a borrow operation (MR-No Borrow) and the multiple repeat problems with a borrow operation (MR-Borrow), and for the no repeat problems without a borrow operation (NR-No Borrow) and the no repeat problems with a borrow operation (NR-borrow) in Experiment 4. Error bars represent standard errors.

Thus, as can be seen from Table 2 and Figure 6, while the multiple repeat problems did not differ under pressure as function of problem difficulty, no repeat problems exhibited a different pattern of results. Here the more difficult borrow problems showed performance decrements under pressure, while the less difficult no borrow problems did not. This pattern of results supports distraction theories of choking in that

only t
be sus

for the
decren
probler
evidenc
may be
equatio
difficul
less diff
the simp
opportur
and Figu
showed
not. In fa
not signi

2.4.3 Di

E
pressure-
modular a
theories o
once or nc

only those modular arithmetic problems with heavy working memory demands appear to be susceptible to performance decrements under pressure.

In contrast, explicit monitoring theories do not appear to be the best explanation for the pattern of results outlined above. If pressure-induced attention causes performance decrements in novel problems based on practiced algorithmic procedures, then all novel problems governed by such algorithms, regardless of difficulty level, should show some evidence of skill failure under pressure. One might assert that more difficult problems may be more prone to attention-induced performance decrements than less difficult equations, as there are more opportunities to induce maladaptive step-by-step control in difficult, multi-step equations. However, if explicit monitoring theories are correct, the less difficult no repeats should still show some evidence of choking under pressure. Even the simplest modular arithmetic equations require multiple steps and thus present the opportunity for some form of attention-induced error under pressure. As seen in Table 2 and Figure 6, however, this was not the case. While the difficult no repeat problems showed significant performance declines under pressure, the less difficult no repeats did not. In fact, these latter problems improved under pressure, although this difference was not significant.

2.4.3 Discussion

Experiment 4 was designed to further examine the mechanisms responsible for pressure-induced performance decrements in the working-memory-intensive task of modular arithmetic. In line with the first three experiments, support for distraction theories of choking was found for this task. While modular arithmetic problems repeated once or not at all during practice showed performance decrements under pressure,

probl
Furth
perfor
results
perfor
line w
Experi
earmar
specific

of incre
manifes
common
expressio
practice i
experime
arithmetic
problem i
the algorit
not suffici

It i
acquire ge
maintain a

problems practiced 50 times each (multiple repeats) did not show this decline. Furthermore, only those no repeat problems involving a borrow operation were performed at a lower level in the high pressure test. Similar to Experiments 2 and 3, these results parallel findings in the test anxiety literature demonstrating the most pronounced performance decrements in capacity-limiting situations for those problems with large on-line working memory demands (Ashcraft & Kirk, 2001; Darke, 1988). The results of Experiment 4 also suggest that it is not merely an issue of general task practice that earmarks those skills most susceptible to choking, but rather the structural demands of specific problems.

It should be noted that while choking in the first three experiments took the form of increased error rates, performance decrements under pressure in Experiment 4 were manifested in terms of increased reaction times. Reaction times and error rates are commonly used as interchangeable indexes of performance. However, differences in the expression of choking in the present experiments may be a function of the point in task practice in which each of these performance declines occurred. In the first three experiments, pressure-induced performance failures occurred prior to substantial modular arithmetic exposure. In Experiment 4, participants had been exposed to over 800 modular problem instances before the high pressure situation. Thus there may be some changes in the algorithm with practice that are sufficient to shift choking from accuracy to speed, but not sufficient to prevent choking from occurring.

It may be that with extensive modular arithmetic training, individuals are able to acquire general problem solving strategies that increase the efficiency with which they maintain and manipulate problem information in working memory (Charness &

Campbell, 1988; Wenger & Carlson, 1996). As a result, pressure-induced distraction does not interfere with the execution of the intermediate steps needed to successfully solve modular arithmetic equations. Instead, such distraction leads to increases in the time required to maintain, rehearse, or act on these problem representations. This is in contrast to modular arithmetic problem solving before practice in which individuals should not have yet developed problem-general strategies that facilitate the management of information in working memory. In this pre-practice situation, the distracting environment created by performance pressure leads to the loss or alteration of the intermediate steps needed to accurately solve modular arithmetic equations.

If the algorithm does change somewhat with practice, then it is possible that pressure-induced failures following general task practice may be better explained by explicit monitoring rather than distraction theories of choking. It may be that pressure-induced attention causes skill failure in novel problems based on practiced procedures by serving to slow down or disrupt efficient algorithmic computations. However, two pieces of evidence work against this notion. First, under this view, problems governed by practiced algorithms, regardless of difficulty level, should show some evidence of skill failure under pressure, as presumably all practiced algorithms are susceptible to attentional-control-induced performance decrements. This was clearly not the case in Experiment 4. While the difficult no repeat problems showed significant performance declines under pressure, the less difficult no repeats did not. Secondly, research examining the acquisition of goal-directed sequences of cognitive steps, similar to those underlying the performance of the modular arithmetic task in the present work, suggests that although high levels of general task practice may increase the efficiency with which working memory is utilized, the temporary storage and manipulation of information in working memory for novel sequences still occurs (Charness & Campbell, 1988; Wenger & Carlson, 1996). Such on-line processes almost certainly impose capacity demands that should be susceptible to pressure-induced performance decrements according to distraction theories.

Nonetheless, it remains an open possibility that practiced algorithms may fail under pressure via the mechanisms proposed by explicit monitoring theories of choking. Future research specifically targeting differences in susceptibility to choking as a function of the type and amount of practice will help to elucidate this notion. In subsequent work I intend to systematically explore how practice changes the algorithm governing modular arithmetic performance, and how such changes affect pressure-induced skill failures.

One way to achieve this goal might be to explore the impact of both explicit attention to execution and dual-task performance on modular arithmetic performance prior to and following general algorithmic practice. Given the results of the first three experiments presented above, early in learning, dual-task performance should impinge on the resources needed to successfully solve unpracticed modular arithmetic equations. This may also be the case following general practice of the algorithm, provided that the algorithm continues to rely heavily on working memory at all levels of practice. In contrast, it may be that practice serves to proceduralize the algorithm in much the same way as sensorimotor skills are learned. If so, dual-task performance should not significantly harm execution, as attention once devoted to step-by-step algorithmic computation may now be available for processing secondary task demands. In terms of conditions that prompt attention to execution, early in learning, attending to performance should not significantly harm execution as the novel algorithmic procedure is presumably already explicitly monitored in real time. This may also hold for performance following practice, provided as Logan's (1988) theory would propose, the algorithm continues to be governed by step-by-step control. However, if the algorithm does become more efficient with practice, then similar to the sensorimotor work presented in the introduction (see also Appendix B), explicit attention to performance may serve to harm execution that is not normally based on step-by-step control. Such a finding would be consistent with explicit monitoring theories of choking.

A second study designed to shed light on how novel problems based on practiced algorithms function under pressure might explore the execution of such problems under simultaneous pressure and dual task conditions. If dual-task demands and a high pressure environment result in a "catastrophic" breakdown in performance, then it is likely that the combination of dual-task execution and performance pressure served to increase working memory capacity demands beyond a manageable limit (Baddeley, 1986). In this case, capacity limitations may be implicated in the failure of practiced algorithms under pressure. On the other hand, if dual-task conditions and pressure do not produce such an overadditive interaction, or if the instantiation of a dual-task condition alleviates the negative effects of performance pressure, support for explicit monitoring theories in the domain of mathematical problem solving will occur. Here, such an underadditive interaction would suggest that performance pressure and dual-task demands are not acting on the same performance processes in modular arithmetic.

3 GENERAL DISCUSSION

The purpose of the present work was to explore performance under pressure in a task with a control structure that might make it susceptible to choking via distraction, at least at low levels of practice; with a possible shift of mechanisms to choking via explicit monitoring at high levels of practice. Explicit monitoring theories suggest that performance pressure prompts attention to skill processes and their step-by-step control. Attention to execution at this component level is thought to disrupt the proceduralized or automated processes of high level skills that are normally run off without such explicit attention (Baumeister, 1984; Beilock & Carr, 2001; Butler & Baumeister, 1998; Kimble & Perlmuter, 1970; Langer & Imber, 1979; Lewis & Linder, 1997; Marchant & Wang, in press; Masters, 1992). On the other hand, distraction theories propose that pressure serves to create a dual-task environment in which controlling execution of the task at hand and performance worries divide the attentional capacity once devoted solely to primary task performance (Wine, 1971). While explicit monitoring theories have received substantial support in accounting for the choking phenomenon (see Appendix A and Appendix B for more detail), most of this evidence has been derived from well-learned sensorimotor tasks that automate via proceduralization with extended practice (Beilock & Carr, 2001; Marchant & Wang, in press). These skills may not be an adequate domain in which to test the predictions of distraction theories.

Experiment 1 examined performance under pressure in difficult modular arithmetic equations – problems that should be susceptible to choking as a result of distraction, but not explicit monitoring theories at low levels of practice. The introduction of a pressure scenario increased participants' perceptions of performance pressure and decreased their performance accuracy in comparison to individuals who were not exposed to a high pressure situation.

In Experiment 2, individuals assigned to either a low or high pressure group again performed novel modular arithmetic problems. However, in contrast to the first experiment, the modular arithmetic equations utilized in Experiment 2 varied as a function of problem difficulty. Individuals in the high pressure group demonstrated increased levels of state anxiety and perceptions of performance pressure following the introduction of a high pressure scenario in comparison to their low pressure counterparts. Additionally, high pressure group participants performed at a significantly lower accuracy level on the modular arithmetic equations after receiving the pressure scenario in comparison to those in the low pressure group. These pressure-induced performance decrements were limited to the most difficult and capacity demanding equations requiring a borrow operation.

Experiment 3 extended the examination of performance under pressure in the modular arithmetic task to include highly practiced problems. Similar to Experiment 2, only the most difficult, capacity demanding modular arithmetic problems (i.e., those that required a borrow operation) showed performance decrements under pressure. Furthermore, these pressure-induced failures were limited to low levels of practice when problem solutions were based on the explicit application of task-relevant solution algorithms that required the on-line maintenance of intermediate results in working memory.

In Experiment 4, individuals performed over 800 modular arithmetic practice problems (presented either once, twice, or 50 times during practice) and were then exposed to a high pressure test environment. Again, only difficult modular arithmetic problems that had not been highly practiced, and hence relied on algorithmic computation rather than one-step answer retrieval from long term memory, showed performance decrements under pressure.

These findings are consistent with distraction theories of choking and suggest that pressure-induced capacity limitations may result in performance decrements in those tasks that have the right properties to be harmed by such constraints. That is, choking via distraction may occur in tasks whose successful performance requires a sequence of mental operations with interdependent demands on storage and processing, rather than direct retrieval of an answer from long term memory (Klapp et al., 1991; Logan, 1988; Zbrodoff & Logan, 1986), or the execution of a motor program, as in previous studies of sensorimotor skills (Beilock & Carr, 2001; Beilock et al., 2002; Brown & Carr, 1989; Keele, 1986; Keele & Summers, 1976).

While finding support for distraction theories in the domain of mathematical problem solving sheds new light on the phenomenon of choking under pressure, it also begs additional questions. Namely, given the extensive support for explicit monitoring theories outlined in the introduction (see also, Appendix A and Appendix B), how can distraction and explicit monitoring both be viable explanations for choking?

3.1 Working Memory Intensive Tasks vs. Automated Skills

It may be that a simple dichotomy is sufficient to solve this problem: Distraction is responsible for performance decrements under pressure in tasks that engage attention and are capacity demanding, while explicit monitoring is better able to explain choking in skills that have become automatic. This hypothesis would coincide with previous findings in the literature on choking regarding the susceptibility of proceduralized sensorimotor skills to performance decrements under pressure, as well as the present work demonstrating pressure-induced failures in working-memory-intensive math tasks. However, while this notion is at first glance appealing, two pieces of evidence work against it.

First, if distraction leads to performance decrements in all tasks that rely on explicit attentional control mechanisms during on-line performance, then one would expect novice sensorimotor skill performance, which is thought to be based on declaratively accessible performance rules (Proctor & Dutta, 1995) and has been shown to be harmed by dual-task distracting manipulations (Beilock et al., 2002; Leavitt, 1979; Smith & Chamberlin, 1992), to show signs of choking under pressure. As outlined in the introduction, Beilock and Carr (2001) have addressed this issue. Participants learned a golf putting skill to a high level and were exposed to a high pressure situation both early and late in practice. Early in practice, pressure to do well actually facilitated execution. At later stages of learning however, performance decrements under pressure emerged.

In contrast to the predictions of distraction theories then, pressure does not appear to induce the type of distraction that harms performance at low levels of practice in sensorimotor skills such as golf putting. For if this were the case, novel sensorimotor skills should show performance decrements under pressure. Rather, explicit monitoring seems to be better able to account for the role of pressure in the motor skill domain at all levels of learning. The unpracticed performances of novices are thought to be controlled by declarative knowledge that is held in working memory and attended in a step-by-step fashion (Anderson, 1983, 1993; Fitts & Posner, 1967; Proctor & Dutta, 1995). As a result, pressure-induced attention to task processes and procedures actually benefits skill execution. Highly practiced or overlearned performances however, supported by procedural knowledge in the form of motor programs that operate without the need for explicit or attended monitoring (Fitts & Posner, 1967; Keele, 1986; Keele & Summers, 1976; Proctor & Dutta, 1995), decline under this type of explicit attentional control.

The notion that distraction may be responsible for choking in capacity-demanding tasks, while explicit monitoring may better explain performance decrements in automated skills is also problematic because automated skills often do not show signs of pressure-induced failure. Previous examinations of well-learned alphabet arithmetic performance under pressure (Beilock & Carr, 2001), as well as the experiments in the present work, demonstrate that some tasks that become automated with extended practice are relatively robust to pressure-induced performance decrements. In particular, this applies to tasks that are thought to automate via a shift to direct retrieval of answers from memory.

3.2 Cognitive vs. Sensorimotor Skills

While a working memory versus automated skill distinction does not appear to adequately explain the choking under pressure phenomenon, a sensorimotor versus cognitive task division may achieve this goal. Distraction theories may be able to explain the performance of predominantly cognitive-based tasks in high pressure environments, while explicit monitoring theories may successfully account for sensorimotor skill performance under pressure. If so, this would elucidate why most of the support for explicit monitoring theories has originated in sensorimotor skills, while support for performance decrements as a result of distracting environments has been found in working-memory-intensive cognitive tasks such as mental arithmetic (Ashcraft & Kirk,

2001) and analogical reasoning (Tohill & Holyoak, 2000). This distinction also suggests that there are fundamental differences in the control structures governing cognitive and sensorimotor skill performance. While this is most certainly an issue for future research, we offer some tentative ideas below.

It may be the case that cognitive tasks that show performance decrements under pressure rely on working memory in a very different way than either novice or well-learned sensorimotor skills. That is, skills such as modular arithmetic appear to be based on a hierarchical and sequentially dependent task representation in which initial steps are used to generate subsequent processes and final solutions. In modular arithmetic problems such as "$72 \equiv 39 \pmod 4$" for example, the derivation of a final problem solution is dependent on the correct answer to the first subtraction operation, which is in turn reliant on the successful maintenance of the intermediate steps necessary to produce the borrow operation.

Well-learned sensorimotor skills that operate largely outside of working memory are almost certainly not based on such a representation (Fitts & Posner, 1967; Proctor & Dutta, 1995). Furthermore, novice sensorimotor skills do not appear to depend on this type of task representation either. Despite the fact that unpracticed motor skills may be based, in part, on explicitly accessible declarative knowledge (Beilock, Wierenga, and Carr, 2002), this knowledge is not organized in such a fashion that the execution of each element of performance is dependent on the maintenance of every prior step. Novice golfers may have explicit access to such skill rules as "keep knees bent." However, subsequent steps in performance such as, "bring club back straight," are not dependent on this knowledge in the same way as a borrow operation in modular arithmetic is dependent on the maintenance of the specific numbers necessary to carry out this operation. Hence it may be the sequentially dependent interweaving of processing and information storage demands that makes a complex cognitive task susceptible to choking via distraction. This idea is similar to Humphreys and Revelle's (1984) suggestion that the impact of variables such as motivation on execution depend on the characteristics of the task being performed. Future research examining the idea that the manner in which a skill fails varies as a function of the specific composition of the control structures supporting performance will certainly further our knowledge of the choking under pressure phenomenon.

3.3 Future Directions

3.3.1 Choking via Explicit Monitoring and Distraction Mechanisms in one Task

The findings of the present work lend support to the notion that both explicit monitoring and distraction are viable explanations for the choking phenomenon, with different domains of applicability. While it may be that distraction and explicit monitoring theories are limited to different task domains, it may also be possible to find performance decrements under pressure in one skill via both choking mechanisms, but at different levels of task experience or practice. One such task that I am interested in exploring as a possible skill that may exemplify both forms of performance decrements under pressure is computer programming.

Skilled computer programmers implement both simple procedures that have been executed hundreds of times, as well as more complex novel procedures that require the integration and on-line maintenance of a number of different sources of information. The former type of programming may become automated in much the same way as a well-learned sensorimotor skill. If so, performance decrements under pressure should occur as a result of over attention to proceduralized performance processes – the mechanisms that explicit monitoring theories would predict cause choking under pressure. In contrast, novel code that must be written while maintaining multiple execution goals on-line, may be more susceptible to performance decrements under pressure as a result of pressure-induced limitations on working memory capacity – the mechanisms that distraction theories predict cause choking.

More research in this area is needed to identify the properties of skills that demonstrate performance patterns under pressure consistent with each type of theory. One product of such research will be an understanding of the nature and representation of the specific task control structures necessary to find pressure-induced performance decrements via distraction and explicit monitoring mechanisms.

3.3.2 Individual Differences in Choking under Pressure

It may also be beneficial to explore the role of individual difference variables in the choking phenomenon. One individual difference that I am currently investigating is working memory capacity.

Working memory and capacity-demanding cognitive skill performance under pressure. It may be that the degree of performance decrements under pressure in working-memory-intensive skills such as mathematical problem solving depends on individual differences in working memory capacity or span. For example, individuals with low working memory span may be more prone to choke under pressure. Low span individuals have limited on-line resources to compute problem solutions. Hence, pressure-induced distraction may reduce low spans' available capacity below the minimum needed to successfully solve modular arithmetic problems.

In contrast, it may instead be the case that individuals with higher working memory span are more prone to performance decrements under pressure. High span individuals depend more so on working memory for problem solutions than their low span counterparts. Thus, these individuals may be harmed to a greater extent than low span individuals when working memory capacity is limited through pressure-induced distraction. Kane and Engle (2000) have made a similar argument with respect to high and low span working memory capacity differences in susceptibility to proactive interference following the instantiation of dual-task demands. Specifically, Kane and Engle have suggested that adding secondary task demands to primary skill execution essentially makes low span individuals out of high span individuals by reducing the capacity that high spans normally rely on for primary task performance (Kane & Engle, 2000, in press). Under this view, if pressure serves to create a distracting environment in mathematical problem solving, then it is conceivable that high spans' performance may be affected to a greater extent than low spans' performance for novel problems that rely on real-time algorithmic computation.

The rate of learning for low and high span individuals may differ as well. Kyllonen and Stephens (1990) have suggested that individual differences in working memory capacity are related to learning rates in complex cognitive skills. If the rate at which modular arithmetic problems become instance-based varies as a function of working memory capacity, then working memory span will have implications for performance decrements under pressure at high levels of practice. Specifically, to the extent that low spans' modular arithmetic performance does not become reliant on automatic answer retrieval as quickly as individuals with higher working memory spans, low capacity individuals should be more prone to performance decrements under pressure according to distraction theories than their high span counterparts.

Working memory capacity and proceduralized sensorimotor skill performance under pressure. It may seem rather intuitive to examine the link between individual differences in working memory capacity and susceptibility to performance decrements under pressure in working-memory-intensive tasks such as modular arithmetic. However, it may also be beneficial to our understanding of the choking phenomenon to explore the relationship between working memory capacity and performance decrements under pressure in other types of tasks – for example, sensorimotor skills that automate via proceduralization with extended practice.

As outlined in the introduction, explicit monitoring theories of choking propose that performance pressure increases anxiety and self-consciousness about performing correctly, which in turn enhances the attention paid to skill processes and their step-by-step control. Attention to performance at this component level is thought to disrupt the proceduralized or automated processes of high-level proceduralized sensorimotor skills (Baumeister, 1984; Beilock & Carr, 2001; Lewis & Linder, 1997). Under this view, it may be that individuals with low working memory capacity will be less prone to performance decrements under pressure at high skill levels than their higher span counterparts. If low spans have less attentional resources to devote to the explicit guidance, on-line maintenance, and step-by-step control of proceduralized skills in comparison to individuals with higher working memory capacity, then the type of pressure-induced attentional processes that explicit monitoring theories propose to be detrimental to sensorimotor skill execution may be less likely to occur.

3.3.3 Stereotype Threat as a Form of Choking under Pressure

In addition to exploring task-type and individual differences in susceptibility to performance decrements under pressure, theories of choking will also benefit from the examination of other performance phenomena that demonstrate unwanted skill decrements. One such phenomenon that my colleagues and I are currently exploring is stereotype threat. Theories of stereotype threat suggest that a negative stereotype about a social group in a particular domain can adversely affect performance by members of that group (Steele, 1997). In studies from golf putting (Stone, Lynch, Sjomeling, & Darley, 1999) to math problem solving (Aronson et al., 1999; Spencer, Steele, & Quinn, 1999), it has been demonstrated that the introduction of a negative stereotype concerning how an individual should perform leads to skill decrements on subsequent tests of the stereotyped ability, independent of an individual's actual task proficiency. In math performance for example, Spencer, Steele, and Quinn (1999) have demonstrated that introducing a negative stereotype about women's math ability (e.g., women are generally worse at math than their male counterparts) leads to decrements in women's performance on difficult math tests in comparison to their male counterparts and also in comparison to tests in which the stereotype was not thought to be applicable (e.g., an English test).

However, while the stereotype threat phenomenon has been demonstrated across a wide range of social groups and skill areas (Shih, Pittinsky, & Ambady, 1999; Spencer, Steele, & Quinn, 1999; Steele, Spencer, & Aronson, 2002; Stone, Lynch, Sjomeling, & Darley, 1999), there is a paucity of research examining the underlying causal mechanisms of such performance failures (Wheeler & Petty, 2001). My colleagues and I are currently investigating the cognitive processes underlying the stereotype threat phenomenon in sensorimotor skills such as golf putting that become proceduralized with extended practice (Beilock et al., 2003), and in more working-memory-intensive tasks such as modular arithmetic (Beilock, Rydell, McConnell, & Carr, 2003). We are interested in determining whether the processes underlying the stereotype threat phenomenon look similar to the mechanisms governing performance decrements under pressure. Preliminary findings suggest that this is indeed the case.

Stereotype threat in proceduralized sensorimotor skills. In the sensorimotor skill of golf putting for example, Beilock et al. (2003) have found that the putting accuracy of expert male golfers is adversely impacted by the instantiation of a negative stereotype about performance (i.e., men are generally poorer putters than women). In contrast, the putting performance of novice male golfers is not harmed by the introduction of the same negative performance stereotype. This pattern of stereotype threat coincides with explicit monitoring theories of "choking under pressure," that suggest that pressure heightens self-awareness and anxiety about performing correctly, which in turn increases the attention paid to skill processes and their step-by-step control (Lewis & Linder, 1997). Although this attention may help novice performance explicitly monitored in real time, attention to execution at this component level disrupts the proceduralized performances of experts (Beilock & Carr, 2001). Negative stereotypes about performance appear to induce explicit monitoring of skill execution that does not harm novice execution, but induces a form of "choking" in experts.

Thus, in sensorimotor skill domains, performance decrements following the introduction of a negative stereotype appear to result from stereotype-induced attention to step-by-step execution – a finding consistent with explicit monitoring theories of choking under pressure. Do the same form of performance decrements characterize the stereotype threat phenomenon across all tasks?

Stereotype threat in working-memory-intensive cognitive skills. The work of my colleagues and I has shown that very different mechanisms may govern stereotype-induced failures in working-memory-intensive tasks such as math problem solving, in comparison to sensorimotor skills such as the golf putting task described above. Specifically, it appears that stereotype threat-induced performance decrements in tasks such as modular arithmetic are consistent with distraction theories of choking.

In a preliminary study, Beilock, Rydell, McConnell, and Carr (2003) had female undergraduate students at Miami University perform a series of easy and difficult unpracticed modular arithmetic problems both prior to and following the introduction of a negative stereotype about their performance (i.e., females are generally worse at math than males). Similar to the impact of pressure on modular arithmetic performance, skill decrements following the instantiation of a negative stereotype about women's math ability only occurred on the most difficult, capacity-demanding problems. This finding, similar to distraction theories of choking under pressure, suggests that stereotype threat creates a capacity-limited environment that impinges on the resources necessary for the performance of highly demanding tasks.

3.4 Conclusion

The present work examines unwanted performance decrements in situations where the desire to perform at an optimal level is at a maximum – situations thought to create performance pressure (Baumeister, 1984, Hardy, Mullen, & Jones, 1996). While early work investigating these skill failures or "choking under pressure" found support for explicit monitoring theories, more recent research suggests that this evidence may be dependent on the type of task and the control structures under investigation.

It appears that both explicit monitoring and distraction are viable explanations for the choking phenomenon. More research in this area is needed to identify the precise mechanisms of the skills that demonstrate performance patterns under pressure consistent with each type of theory. One product of such research will be a taxonomy of skills based on a clear understanding of the nature and representation of their control structures at different levels of expertise. Continued exploration of the choking phenomenon across diverse skill domains will speak to task-type, skill-level, and individual differences in susceptibility to performance failures, and ultimately to means of engineering training regimens to diminish such susceptibility – knowledge that will benefit researchers, practitioners, and performers alike.

REFERENCES

Allard, F. & Starkes, J. L. (1991). Motor-skill experts in sports, dance and other domains. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise* (pp. 126-152). Cambridge: Cambridge University.

Allport, D. A., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single channel hypothesis. *Quarterly Journal of Experimental Psychology, 24,* 225-235.

Anderson, J. R. (1982). Acquisition of a cognitive skill. *Psychological Review, 89,* 369-406.

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University.

Anderson, J. R. (1993). *Rules of mind.* Hillsdale, NJ: Erlbaum.

Anderson, J. R. & Lebiere, C. (1998). *The atomic components of thought.* Mahwah, NJ: Erlbaum.

Aronson, J., Lustina, M., Good, C., Keough, K., Steele, C., & Brown, J. (1999). When white men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35,* 29-46.

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44,* 75-106.

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General, 130,* 224-237.

Baddeley, A. (1986). *Working memory.* Oxford, England: Clarendon Press.

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology, 46,* 610-620.

Beilock, S. L., Bertenthal, B. I., McCoy, A. M., & Carr, T. H. (in press). Haste does not always make waste: Expertise, direction of attention, and speed versus accuracy in performing sensorimotor skills. *Psychonomic Bulletin & Review.*

Beilock, S. L. & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General, 130,* 701-725.

Beilock, S. L., Carr, T. H., MacMahon, C., & Starkes, J. L. (2002). When paying attention becomes counterproductive: Impact of divided versus skill-focused attention on novice and experienced performance of sensorimotor skills. *Journal of Experimental Psychology: Applied, 8,* 6-16.

Beilock, S. L., Jellison, W. A., Noechel, L., Furrow, N., McConnell, A. R., & Carr, T. H. (February, 2003). On the causal mechanisms underlying stereotype threat: Is stereotype threat a form of "choking under pressure?" Poster presented at the Annual Meeting of the Society for Personality and Social Psychology. Los Angeles, California.

Beilock, S. L., Rydell, R. J., McConnell, A. R., & Carr, T. H. (2003) Stereotype threat: A threat on working memory. Manuscript in preparation.

Beilock, S. L., Wierenga, S. A., & Carr, T. H. (2002). Expertise, attention, and memory in sensorimotor skill execution: Impact of novel task constraints on dual-task performance and episodic memory. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 55,* 1211-1240.

Beilock, S. L., Wierenga, S. A., & Carr, T. H. (in press). Memory and expertise: What do experienced athletes remember? In J. L. Starkes and K. A. Ericsson (Eds.), *Expert performance in sports: Advances in research on sport expertise.* Champaign, IL: Human Kinetics.

Bogomolny, A. (1996). Modular arithmetic. Retrieved March 1, 2000 from http://www.cut-the-knot.com/blue/Modulo.shtml

Brown, T. L., & Carr, T. H. (1989). Automaticity in skill acquisition: Mechanisms for reducing interference in concurrent performance. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 686-700.

Butler, J. L., & Baumeister, R. F. (1998). The trouble with friendly faces: Skilled performance with a supportive audience. *Journal of Personality and Social Psychology, 75,* 1213-1230.

Charness, N., & Campbell, J. I. D. (1988). Acquiring skill at mental calculation in adulthood: A task decomposition. *Journal of Experimental Psychology: General, 117,* 115-129.

Darke, S. (1988). Effects of anxiety on inferential reasoning task performance. *Journal of Personality and Social Psychology, 55,* 499-505.

Ericsson, K. A., & Charness, N. (1994). Expert performance - Its structure and acquisition. *American Psychologist, 49,* 725-747.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100,* 363-406.

Eysenck, M. W. (1979). Anxiety learning and memory: A reconceptualization. *Journal of Research in Personality, 13,* 363-385.

Eysenck, M. W. (1992). *Anxiety: The cognitive perspective.* Hillsdale, NJ: Erlbaum.

Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition and Emotion, 6,* 409-434.

Eysenck, M. W. & Keane, M. T. (1990). *Cognitive psychology: A student's handbook.* Hillsdale, NJ: Erlbaum.

Fitts, P. M. (1964). Perceptual-motor skill learning. In A. W. Melton (Ed.). *Categories of human learning.* New York: Academic Press.

Fitts, P. M., & Posner, M. I. (1967). *Human performance.* Belmont, CA: Brooks/Cole.

Halford, G. S. (1993). *Children's understanding: The development of mental models.* Hillsdale, NJ: Erlbaum

Hardy, L., Mullen, R., & Jones, G. (1996). Knowledge and conscious control of motor actions under stress. *British Journal of Psychology, 87,* 621-636.

Humphreys, M. S., & Revelle, W. (1984). Personality, motivation, and performance: A theory of the relationship between individual differences and information processing. *Psychological Review, 91,* 153-184.

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 26,* 336-358.

Kane, M. J., & Engle, R. W. (in press). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review.*

Keele, S. W. (1986). Motor control. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 7). New York: John Wiley.

Keele, S. W., & Summers, J. J. (1976). The structure of motor programs. In G. E. Stelmach (Ed.), *Motor control: Issues and trends* (pp. 109-142). New York: Academic press.

Kimble, G. A., & Perlmuter, L. C. (1970). The problem of volition. *Psychological Review, 77,* 361-384.

Klapp, S. T., Boches, C. A., Trabert, M. L., & Logan, G. D. (1991). Automatizing alphabet arithmetic: II. Are there practice effects after automaticity is achieved? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 196-209.

Kyllonen, P. C., & Stephens, D. L. (1990). Cognitive abilities as determinants of success in acquiring logic skill. *Learning and Individual Differences, 2,* 129-150.

Lachman, R., Lachman, J. L., & Butterfield, E. C. (1979). *Cognitive psychology and information processing: An introduction.* Hillsdale, NJ: Lawrence Erlbaum.

Langer, E., & Imber, G. (1979). When practice makes imperfect: Debilitating effects of overlearning. *Journal of Personality and Social Psychology, 37,* 2014-2024.

Leavitt, J. (1979) Cognitive demands of skating and stick handling in ice hockey. *Canadian Journal of Applied Sport Sciences, 4,* 46-55.

Leon, M. R. (1989). Anxiety and the inclusiveness of information processing. *Journal of Research in Personality, 23,* 85-98.

Lewis, B., & Linder, D. (1997). Thinking about choking? Attentional processes and paradoxical performance. *Personality and Social Psychology Bulletin, 23,* 937-944.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492-527.

Marchant, D. B., & Wang, J. (in press). Predictors of Choking: I. A Quantitative Investigation. *Journal of Sport and Exercise Psychology.*

Masters, R. S. W. (1992). Knowledge, knerves and know-how: The role of explicit versus implicit knowledge in the breakdown of a complex motor skill under pressure. *British Journal of Psychology, 83,* 343-358.

Masters, R. S. W., Polman, R. C. J., & Hammond, N. V. (1993). 'Reinvestment': A dimension of personality implicated in skill breakdown under pressure. *Personality and Individual Differences, 14,* 655-666.

Mullen, R., & Hardy, L. (2000). State anxiety and motor performance: Testing the conscious processing hypothesis. *Journal of Sports Sciences, 18,* 785-799.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Erlbaum.

Pachella, R. G. (1974). The interpretation of reaction time information-processing research. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition.* Hillsdale, NJ: Erlbaum.

Proctor, R. W., & Dutta, A. (1995). *Skill acquisition and human performance.* Thousand Oaks, CA: Sage.

Reimann, P., & Chi, M. T. (1989). Human expertise in complex problem solving. In K. J. Gilhooly (Ed.), *Human machine and problem-solving* (pp. 161-189). New York: Plenum.

Rosenbaum, D. A., Carlson, R. A., & Gilmore, R. O. (2001). Acquisition of intellectual and perceptual-motor skills. *Annual Review of Psychology, 52,* 453-470.

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience ad shifts in quantitative performance. *Psychological Science, 10,* 80-83

Smith, M. D., & Chamberlin, C. J. (1992). Effect of adding cognitively demanding tasks on soccer skill performance. *Perceptual and Motor Skills, 75,* 955-961.

Sorg, B. A., & Whitney, P. (1992). The effect of trait anxiety and situational stress on working memory capacity. *Journal of Research in Personality, 26,* 235-241.

Spencer, S. J., Steele, C. M., and Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35,* 4-28.

Spielberger, C. C., Gorsuch, R. L., & Lushene, R. (1970). *State-Trait Anxiety Inventory.* Palo Alto, CA: Consulting Psychology Press.

Staszewski, J. J. (1988). Skilled memory and expert mental calculation. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.). *The nature of expertise* (pp. 71-128). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Sternberg, S. (1969) Memory-scanning: Mental processes revealed by reaction-time experiments. *American scientist, 57,* 421-457.

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.). *Advances in experimental social psychology* (pp. 379-440). Amsterdam: Academic Press.

Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on Black and White athletic performance. *Journal of Personality and Social Psychology, 77,* 1213-1227.

Tohill, J. M., & Holyoak, K. (2000). The impact of anxiety on analogical reasoning. *Thinking and Reasoning, 6,* 27-40.

Touron, D. R., Hoyer, W. J., & Cerella, J. (2001). Cognitive skill acquisition and transfer in younger and older adults. *Psychology & Aging, 16,* 555-563.

Wenger, J. L., & Carlson, R. A. (1996). Cognitive sequence knowledge: What is learned? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 599-619.

Wheeler, S. C., & Petty, R. E. (2001). The effects of stereotype activation on behavior: A review of possible mechanisms. *Psychological Bulletin, 127,* 797-826.

Wine, J. (1971). Test anxiety and direction of attention. *Psychological Bulletin, 76,* 92-104.

Wulf, G., & Prinz, W. (2001). Directing attention to movement effects enhances learning: A review. *Psychonomic Bulletin & Review, 8,* 648-660.

Zbrodoff, N. J., & Logan, G. D. (1986). On the autonomy of mental processes: A case of study arithmetic. *Journal of Experimental Psychology: General, 115,* 118-130.

APPENDIX A

Beilock, S. L. & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General, 130,* 701-725.

Journal of Experimental Psychology: General
2001, Vol. 130, No. 4, 701–725

0096-3445/01/$5.00 DOI: 10.1037//0096-3445.130.4.701

On the Fragility of Skilled Performance: What Governs Choking Under Pressure?

Sian L. Beilock and Thomas H. Carr
Michigan State University

Experiments 1–2 examined generic knowledge and episodic memories of putting in novice and expert golfers. Impoverished episodic recollection of specific putts among experts indicated that skilled putting is encoded in a procedural form that supports performance without the need for step-by-step attentional control. According to *explicit monitoring* theories of choking, such proceduralization makes putting vulnerable to decrements under pressure. Experiments 3–4 examined choking and the ability of training conditions to ameliorate it in putting and a nonproceduralized *alphabet arithmetic* skill analogous to mental arithmetic. Choking occurred in putting but not alphabet arithmetic. In putting, choking was unchanged by dual-task training but eliminated by self-consciousness training. These findings support explicit monitoring theories of choking and the popular but infrequently tested belief that attending to proceduralized skills hurts performance.

Why does the execution of a well-learned skill fail under pressure? Research investigating skill and expertise has produced a number of important findings regarding the variables that mediate optimal skill performance (Allard & Starkes, 1991; Anderson, 1982, 1987; Ericsson, Krampe, & Tesch-Romer, 1993; Keele, 1986; Logan, 1988; Reimann & Chi, 1989). Nevertheless, the phenomenon of "choking under pressure" remains unexplained—and feared by skilled performers across many domains. *Performance pressure* has been defined as an anxious desire to perform at a high level in a given situation (Hardy, Mullen, & Jones, 1996) and is thought to vary as a function of the personally felt importance of a situation (Baumeister, 1984). *Choking*, or performing more poorly than expected given one's level of skill, tends to occur in situations fraught with performance pressure. This phenomenon seems particularly visible in sensorimotor or action-based skills, where it has garnered interest in both experimental and real-world settings. People often speak of the "bricks" in basketball free throw shooting or the "yips" in golf putting, and a majority of the experimental research on choking done to date has used sensorimotor tasks of one kind or another (Baumeister, 1984; Lewis & Linder, 1997; Masters, 1992).

Two competing theories have been proposed to account for decrements in skilled performance under pressure. ***Distraction*** theories propose that pressure creates a distracting environment that shifts attentional focus to task-irrelevant cues, such as worries about the situation and its consequences (Wine, 1971). This shift of focus changes what was single-task performance into a dual-task situation in which controlling execution of the task at hand and worrying about the situation compete for attention. *Self-focus* theories (perhaps more appropriately termed ***explicit monitoring*** or *execution focus* theories, as they are concerned with attention to skill execution) suggest that pressure raises self-consciousness and anxiety about performing correctly, which increases the attention paid to skill processes and their step-by-step control (Baumeister, 1984; Lewis & Linder, 1997). Attention to execution at this step-by-step level is thought to disrupt well-learned or proceduralized performances (Kimble & Perlmuter, 1970; Langer & Imber, 1979; Lewis & Linder, 1997; Masters, 1992).

Distraction and explicit monitoring theories appear to be competing alternatives—indeed, they are complete opposites in their proposed mechanisms. However, it is important to note that they may have different domains of applicability and hence could turn out to be complementary rather than mutually exclusive. Distraction theory holds that the mechanisms of choking operate on task control structures that are attended during performance. Thus, under distraction theory, breakdowns under pressure are most likely in skills that rely on working memory for storage of decision and action-relevant information that might be vulnerable to corruption or forgetting as a result of dual-task interference. This calls to mind skills based on fact retrieval as possible test cases. In contrast, explicit monitoring theory suggests that the mechanisms of choking operate on task control structures that are proceduralized—based on mental or motor programs that run largely unattended, without the services of working memory, and might best remain outside the scrutiny of introspection. This calls to mind sensorimotor skills as test cases.

Given these differences in potential domain of applicability, our first two experiments were aimed at identifying a particular skill that had the right properties to be susceptible to choking according to one of these theories but not the other. We chose golf putting, which is a complex sensorimotor task that is thought to become proceduralized with practice and hence falls into the domain of explicit monitoring theory. Because a proceduralized skill ought not to require constant on-line attentional control (e.g., Fitts & Posner, 1967; Proctor & Dutta, 1995), it should be relatively robust

---

Sian L. Beilock, Departments of Psychology and Kinesiology, Michigan State University; Thomas H. Carr, Department of Psychology, Michigan State University.

Correspondence concerning this article should be addressed to Sian L. Beilock, Department of Psychology, Michigan State University, Psychology Research Building, Room 236, East Lansing, Michigan 48824. Electronic mail may be sent to beilocks@pilot.msu.edu.

against conditions that draw attention away from the primary task as in distraction theory. However, this type of skill should be sensitive to the kind of attention-induced disruptions of fluent execution envisioned by explicit monitoring theory.

To confirm the proceduralized status of golf putting, in Experiments 1 and 2 we compared reports of generic, schematic, or prescriptive knowledge about putting with episodic memories of particular putts in expert and novice golfers. The goal was to document a particular property of the cognitive substrate of this sensorimotor skill—the declarative accessibility, or openness to introspection, recollection, and report, of the skill's processes and procedures at different levels of expertise. In particular, we sought to use as a diagnostic tool the well-documented dependence of explicit episodic memory on the presence of attention (e.g., Craik, Govini, Naveh-Benjamin, & Anderson, 1996). If golf putting becomes proceduralized with practice and, as a consequence, task control structures are largely unattended during skill execution, then episodic memory for the step-by-step unfolding of particular instances of performance should be impoverished. Observing such a pattern would earmark practiced golf putting as a skill that should be susceptible to choking under pressure according to explicit monitoring theory.

After documenting the proceduralized status of practiced golf putting, in Experiment 3 we trained novices to an asymptotic level of achievement and then created a high-pressure test situation intended to induce choking. Participants performed either our chosen sensorimotor task of golf putting or a comparison task whose practiced control structure has already been shown to depend on fact retrieval rather than proceduralized motor programs. The comparison task was Zbrodoff and Logan's (1986) *alphabet arithmetic* task.

Training took place under one of three different regimens. Choice of regimens followed what we called a "vaccination strategy," intended to test the theories of choking by determining whether practice at dealing with the particular causal mechanism proposed by each theory would reduce choking in performers who would have been likely to choke had the training not been received. One regimen was ordinary single-task practice, which provided a baseline measure of the occurrence of choking. The other two regimens exposed performers to the particular aspects of high-pressure situations that have been proposed by the two theories of choking to cause performance decrements. In the "dual-task distraction" regimen, practice took place under dual-task conditions (while monitoring an auditory word list for a target word) in order to expose performers to being distracted from the primary task by execution-irrelevant activity in working memory. In the "execution-oriented self-consciousness" regimen, practice took place while being videotaped for subsequent analysis by experts in order to expose performers to having attention called to themselves and their performance in a way intended to induce explicit monitoring of skill execution.

Finally, Experiment 4 replicated and extended Experiment 3. The goal was to test the distraction and explicit monitoring theories' predictions at lower levels of practice in the golf putting task. Most conceptions of skill acquisition, including motor program theories applicable to golf putting, propose that early in learning, performance is supported by unintegrated control structures that are held a few steps at a time in working memory (Anderson, 1987, 1993; Fitts & Posner, 1967; Keele, 1986; Keele & Summers, 1976; Proctor & Dutta, 1995). With practice, control is thought to evolve toward the integrated procedures that are the objects of the explicit monitoring theory. One might imagine, then, that in tasks that follow such a developmental trajectory, choking due to distraction might be observed early in learning whereas choking due to explicit monitoring would occur in a more practiced state. Experiment 4 explored this idea by imposing high-pressure tests on golf putting performance at two different points in practice.

We now turn to a more detailed discussion of differences in the knowledge representations controlling the execution of various skills at different levels of expertise and of how these differences may aid in our examination and understanding of the two competing theories of choking under pressure.

Generic, Episodic, and Procedural Skill Representations

Skill knowledge can be declaratively accessible in two different forms. *Generic* knowledge captures schema-like or prescriptive information about how a skill is typically done. *Episodic* knowledge, on the other hand, captures a specific memory—an autobiographical record of a particular performance. According to current theories of skill acquisition and automaticity, changes in expertise should affect these two types of declaratively accessible representations very differently.

First, declaratively accessible generic knowledge should increase with increasing expertise—experts have more explicitly available general knowledge about the domain in which they are skilled than do their novice counterparts (for reviews, see Proctor & Dutta, 1995; Van Lehn, 1989). It would be shocking to discover that experts could not describe the dos and don'ts of their skill in as much detail or explain its ideal execution as competently as novices. Thus, experts' off-line generic or prescriptive accounts of their skill should provide a more extensive and systematic chronicle of how, in general, performance should be accomplished than the generic accounts of novices.

Second, declaratively accessible episodic memories of any particular performance should decrease with increasing expertise. Why should this be? It is widely believed that highly practiced, overlearned performances are *automated*—meaning they are controlled in real time by procedural knowledge that requires little attention, operates largely outside of working memory, and is substantially closed to introspection (Anderson, 1987, 1993; Fitts & Posner, 1967; Keele & Summers, 1976; Kimble & Perlmuter, 1970; Langer & Imber, 1979; Proctor & Dutta, 1995; Squire & Knowlton, 1994). Because of the well-established relation between attention and episodic memory, this belief carries implications for recollecting one's performances. In both short-term memory (Daneman & Carpenter, 1980; Muter, 1980; Peterson & Peterson, 1959; Posner & Rossman, 1965) and long-term memory (Craik et al., 1996; Naveh-Benjamin, Craik, Guez, & Dori, 1998), diverting or reducing the amount of attention paid to material being encoded for storage reduces subsequent explicit memory for that material. The impact of reducing attention is greatest in recall and is present in cued recall and recognition as well. To the extent that practiced tasks are indeed carried out with less attention to processes, procedures, and the control structures that govern them, real-time performance ought to leave impoverished episodic memories of the performance's execution.

In contrast, the relatively unpracticed performances of novices are thought to be controlled by declarative knowledge that is held in working memory and attended to step-by-step during performance (Anderson, 1987, 1993; Fitts & Posner, 1967; Kimble & Perlmuter, 1970; Proctor & Dutta, 1995; Squire & Knowlton, 1994). Attending to such knowledge should leave an explicitly retrievable episodic record of task execution—a declaratively accessible memory of the performance as an autobiographical experience that includes the step-by-step operations by which the performance was implemented.

"Expertise-Induced Amnesia"

Thus, current theories of skill acquisition and automaticity suggest that increasing expertise through practice will create a kind of domain-specific amnesia. If a skill is controlled by declarative knowledge that is attended to during performance, episodic memory for skill execution processes should be explicitly retrievable. However, if a skill is supported by procedural knowledge that automates real-time performance, then episodic memory for this performance should be minimized.

The idea of "expertise-induced amnesia" may seem uncontroversial to some investigators of skilled performance. To others, however, a problem will come immediately to mind: There is well-known evidence suggesting that expertise serves to enhance episodic recollection, not degrade it. For example, in their classic chess study, Chase and Simon (1973) found that chess masters were better able to recall briefly presented chess positions than were less experienced players (for confirmatory data, see De Groot, 1946/1978; and for similar results from computer programmers, see Soloway & Ehrlich, 1984). Analogous evidence comes from studies of reading. Real-time deployment of world knowledge, creating superior comprehension of a situation described in a narrative text, leads to better recall of the text's wording (Bransford & Johnson, 1972; Dooling & Christiaansen, 1977; McCandliss & Carr, 1994, 1996). In light of such evidence, one anonymous reviewer of an earlier attempt to report this work called the idea of expertise-induced amnesia "otherworldly" and "patently false," claiming instead that "experts have exquisite episodic recall of the most arcane minutiae in their area of competence" (personal communication, August 24, 1999).

However, problems exist in using results such as those mentioned above as evidence against the prediction of expertise-induced amnesia. The chess studies focused on memory for the kinds of stimuli that are operated on by chess players, not memory for the operations themselves. That is, experts were asked to recreate the positions of specific pieces on the board. Experts were not asked for the steps or processes by which the situation was assessed, how a move appropriate to that stimulus configuration was chosen, or how a chosen move was physically implemented. The same applies to studies of reading, where people able to deploy greater world knowledge were asked to remember the stimulus material they read, not the sequence of reading operations that took place. Thus, the above-mentioned studies can be taken to support the notion that experts have better episodic recollection for the stimuli to which they apply their knowledge. However, these studies do not demonstrate that experts have superior recollection for the sequence of cognitive processes involved in formulating specific plans of action or the sequence of cognitive processes by which actions are implemented in real time. For this reason, it remains a reasonable idea, despite the existing literature on experts' episodic memories, that because expert knowledge runs automatically during real-time skill execution, experts may neither attend to nor later remember the step-by-step unfolding of their performances.

Consistent with this possibility is the finding from the chess literature that in both on-line and retrospective verbal-report protocols, experts report having explicitly considered fewer alternatives in making any given move than do novices. The experts report that the best move or a small number of good moves just popped into their heads, whereas the novices report a serial process of generating and evaluating several possible moves in succession (Ericsson & Smith, 1991).

Furthermore, it is not always true that highly practiced experts demonstrate superior episodic memory for the stimuli on which they have operated. Fisk and Schneider (1984) studied the acquisition of expertise at searching through arrays of visually presented words for members of a target category. They found that after a great deal of practice, during which speed and accuracy of finding targets greatly increased and sensitivity of performance to the number of words in each array greatly decreased, recognition memory for the words that had been searched through was markedly worse than it had been at lower levels of practice. Fisk and Schneider argued that practice automated performance and that automating performance increased real-time skill but decreased subsequent episodic memory. In light of Fisk and Schneider's finding, it should be noted that the literature on skill acquisition and automaticity from which we derived the prediction of expertise-induced amnesia has been dominated by studies of the speeded performance of reaction time tasks with significant sensorimotor components. In contrast, much of the work on expertise that suggests good rather than poor episodic memory has focused on cognitive tasks that are based on a great deal of factual knowledge and whose real-time sensorimotor demands are minimal (e.g., chess, computer programming, physics problem solving). The memorial results of performing such fact-reliant cognitive tasks may be different from those of tasks that rely more on sensorimotor knowledge. More generally, fact-reliant tasks and sensorimotor tasks may diverge in the nature of their underlying representations and control structures and hence may differ in many ways, not just memorial consequences. An argument for such differences in underlying representations and control structures has been made by Klapp, Boches, Trabert, and Logan (1991), to which we return later in the article.

Experiments 1 and 2: Declaratively Accessible Knowledge of Golf Putting

This brings us to the first two experiments in the present study, which document the prediction of expertise-induced amnesia in the sensorimotor task of golf putting. Putting was chosen because it is a complex task in which considerable time and effort is required to become an expert performer. Even at the highest levels, putting is not easy and success depends heavily on extensive past experience. This is in some ways similar to chess, where expert chess masters hone their skills over a long period of time, developing a large knowledge base consisting of many relevant chess piece configurations and game scenarios (De Groot, 1946/1978). Nevertheless, putting is a sensorimotor task in a way that chess is not. In addition, putting's discrete nature enables straightforward trial-by-trial measurement of accuracy, so that differences in expertise can be readily verified.

Experiment 1

In Experiment 1 expert golfers' generic knowledge of golf putting and episodic recollection of specific putts were compared to the generic knowledge and episodic recollection of novice golfers within the context of a laboratory golf putting task. If on-line well-learned golf putting is supported by procedural knowledge, as theories of sensorimotor skill acquisition would predict, then expert golfers should give longer, more detailed generic descriptions of the steps involved in a typical putt compared with the accounts given by novices, but shorter, less detailed episodic recollections of a particular putt. Because proceduralization reduces the need to attend to the specific processes by which skill execution unfolds, experts' episodic recollections of step-by-step real-time performance should be impoverished.

Method

Participants

Participants (N = 48) were undergraduate students enrolled at Michigan State University and consisted of intercollegiate golf team members (n = 16), intercollegiate athletes with no golf experience (n = 16), and introductory psychology students with no golf experience (n = 16). An equal number of male and female participants were recruited from each of the three populations. The two groups of novices were included in order to examine the possibility that the intense training and practice engaged in by elite athletes may alter their strategic approach to skill acquisition, even in a new domain outside their already-acquired expertise, causing differences in performance and knowledge representation in comparison with nonathletes.

Procedure

After giving informed consent and filling out a demographic sheet concerning previous golf experiences, participants were told that the purpose of the study was to examine the accuracy of golf putting over several trials of practice. Participants were instructed that the object of the task was to putt a golf ball as accurately as possible, making it stop at a target located 1.5 m away, marked by a square of red tape on a carpeted indoor putting green (3 × 3.7 m). A standard golf putter and golf ball were supplied. All groups participated in identical pretest, practice, and posttest conditions, though the participants were not made aware of the separate conditions. To the participant, the golf putting task appeared to involve three blocks of putts with a short break after each block during which a questionnaire was filled out.

Pretest condition. Participants were set up 1.5 m from the target. They were asked whether they preferred to putt right-handed or left-handed and were then given the appropriate putter. Participants took a series of 20 putts. After completing the putts, participants filled out a questionnaire eliciting a description of the steps involved in a typical golf putt (Appendix A, first paragraph).

Practice condition. Participants were again set up 1.5 m from the target. Participants took a series of 30 putts. After completing the putts, participants filled out an identical questionnaire to the one that they had previously filled out in the pretest condition eliciting a description of the steps involved in a typical golf putt (Appendix A, first paragraph).

Posttest condition. Participants were set up 1.5 m from the target. Participants then took a series of 20 putts. Immediately following the posttest condition, participants filled out a questionnaire designed to access their episodic recollection of the last putt they had just taken (Appendix A, second paragraph).

Results

Putting Performance

Accuracy of putting was measured by the distance (in centimeters) away from the center of the target at which the ball stopped after each putt. The mean distance from the target of the last 10 putts in the pretest condition was used as a measure of pretest golf putting skill. The mean distance from the target of the middle 10 putts in the practice condition was used as a measure of practice putting skill. The mean distance from the target of the last 10 putts in the posttest condition was used as a measure of posttest golf putting skill. Means and standard errors for putting performance appear in Figure 1.

As Figure 1 makes clear, golf team members showed superior putting performance in comparison to the two novice groups, who did not differ. This was true both before and after the practice phase, during which the three groups all improved by approximately the same amount. This pattern was confirmed by a 3 (undergraduate, athlete, golf team) × 2 (pretest, posttest) analysis of variance (ANOVA), which revealed significant main effects of experience, $F(2, 45) = 16.23$, $p < .001$, $MSE = 56.88$, and test, $F(1, 45) = 29.21$, $p < .001$, $MSE = 14.91$, with no interaction ($F < 1$).

Thus the expertise of the golf team members transferred substantially to the somewhat novel task demands of making the ball stop on a target rather than drop into a hole. However, the golfers did improve with practice, indicating that there were still some elements of the present task left for them to learn. In contrast to the skill displayed by the golfers, the nongolf sensorimotor expertise of the athlete group did not transfer to putting whatsoever. Athletes enjoyed no advantage in putting accuracy over the nonathlete undergraduates at any point in practice.

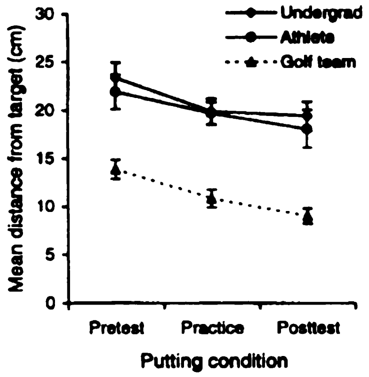

Figure 1. Mean (± SE) distance from the target at which the ball stopped after each putt in the pretest, practice, and posttest conditions for each group. Undergrad = undergraduate.

Generic and Episodic Memory Protocols

Questionnaire responses were analyzed quantitatively, in terms of the number of golf putting steps included in each type of protocol, and qualitatively, in terms of the relative frequencies of different categories of steps.

Quantitative analysis. Three expert golfers and the how-to golf putting book *Classic Instruction in Golf* (Jones, Davis, Crenshaw, Behar, & Davis, 1998) were used in establishing a master list of steps involved in a successful golf putt (Appendix B). The statements in each participant's protocol were compared with this master list. If a step given by a participant referred to the same action or the same biomechanism as a step on the master list, it was counted as one step. For example, the step given by a participant, "I swung the club back behind me," and Step 13 on the expert golfer list, "Backswing—swing the club straight back," were coded as a match because they both refer to the same action (i.e., taking a backswing). Similarly, the step given by a participant, "I kept my hips still," and Step 21 on the list developed by the expert golfers, "Head/trunk/hips/legs—should remain still during the stroke," were deemed a match because they both refer to the same biomechanism (i.e., motion of the hips). If two steps given by participants both described one step on the list developed by the expert golfers, they were combined and counted as one step. For example, if a participant reported the two steps "I held the putter with two hands" and "My right hand was above my left hand," these steps were combined to match the step on the list developed by the expert golfers that referred to the grip of the putter. If a step given by a participant did not match a step on the master list, yet did refer to a necessary part of the participant's putting process (e.g., "I brushed my hair out of my face so I could see the target"), it was counted as a step. Because the master list was thorough and detailed, these "nonmatch steps" were quite rare. Finally, if a step given by a participant did not match a step on the master list, and was not part of the putting process itself (e.g., "I thought about the fact that I needed golf lessons"), it was not counted as a step. Although such nonprocess commentary is legitimately part of the autobiographical record, it is not part of the specific object of prediction in testing for expertise-induced amnesia. However, nonprocess commentary was also quite rare and, if included, would not have changed the results in any way.

The order in which the steps were recorded was not taken into account in determining the number of steps given by participants. Two experimenters independently coded the data. Interexperimenter reliability was extremely high (r = .97). Table 1 gives representative generic and episodic golf putting protocols for all three participant groups.

Table 2 and Figure 2 present the results. Mean number of steps did not differ significantly between the two generic protocols in any of the groups, as confirmed by a 3 (undergraduate, athlete, golf team) × 2 (first generic protocol, second generic protocol) ANOVA in which the main effect of test and the interaction of Expertise × Test produced $Fs < 1$. The second generic protocol was used in a 3 (undergraduate, athlete, golf team) × 2 (second generic protocol, episodic protocol) ANOVA to compare the lengths of the generic and episodic protocols produced at each level of expertise. The analysis revealed an interaction of Expertise × Protocol type, $F(2, 45) = 24.30$, $p < .001$, $MSE = 2.10$. Direct comparisons of the number of generic versus episodic steps within each group showed that the undergraduates gave significantly more steps in their episodic than in their generic protocols, $t(15) = 4.29$, $p < .001$. The athletes produced a difference in the same direction as the undergraduates, but it was not significant, $t(15) = 0.29$, $p < .78$. In contrast, golf team members gave significantly more steps in their generic than in their episodic protocols, $t(15) = 4.70$, $p < .001$. Thus, as can be seen in Figure 2, golfing expertise was associated with longer generic descriptions and shorter episodic recollections.

Qualitative analysis: Types of steps. The first qualitative analysis divided steps into three categories (see Table 3). Assessment or planning referred to deciding how to take a particular putt and what properties the putt ought to have. Examples are "read the green," "read the line" (from the ball to the hole or target), "focus on the line," and "visualize the force needed to hit the ball." Mechanics or execution referred to the components of the mechanical act that implements the putt. Examples are "grip the putter with your right hand on top of your left," "bring the club straight back," and "accelerate through the ball," all of which deal with the effectors and the kinesthetic movements of the effectors required to implement a putt. Ball destinations or outcomes referred to where the ball stopped or landed and hence to degree of success. A 3 (undergraduate, athlete, golf team) × 2 (generic protocol, episodic protocol) ANOVA was conducted on the number of steps given in each of these three categories.

The analysis of assessment produced a significant interaction between expertise and type of protocol, $F(2, 45) = 14.56$, $p < .001$, $MSE = 1.07$, which is displayed in the left panel of Figure 3. Assessment steps appeared more often in the generic descriptions of golf team members than anywhere else. A simple effects test confirmed a difference among groups in the generic protocol, $F(2, 45) = 13.75$, $p < .001$, $MSE = 2.14$, and Fisher's least significant difference (LSD) test showed that the golf team gave significantly more assessment steps in their generic descriptions than did either the undergraduates or the athletes, who did not differ. Furthermore, the golf team gave more assessment steps in their generic descriptions than they did in their episodic recollections, $t(15) = 4.90$, $p < .001$, whereas the undergraduate and athlete groups did not differ in the number of assessment steps included in the two kinds of protocols, $t(15) = 0.64$ and $t(15) = 0.00$, respectively ($ps > .10$).

As an adjunct to the analysis of assessment, those steps that involved mental imagery (i.e., imagining some aspect of how a putt ought to look or feel before executing the action) were counted. Mental imagery is a topic of considerable interest in sports psychology and has been defined in that literature as "the imagined rehearsal of skill processes, procedures, and possible outcomes prior to task performance" (Woolfolk, Murphy, Gottesfeld, & Aitken, 1985). In the undergraduate group, 0.0% of generic steps and 0.7% of episodic steps referred to mental imagery. In the athlete group, 0.7% of the generic steps and 0.0% of the episodic steps referred to imagery. In the golf team group, 7.0% of the generic steps and 2.0% of the episodic steps referred to imagery. Thus almost all of the reports of imagery were from golfers, and most of these were part of the generic descriptions.

One might worry that the experts' exclusion of assessment steps from their episodic recollections was merely an artifact of our very simple and highly repetitive situation, in which assessment was not much needed by the time episodic memory was measured, which

Table 1
Representative Generic and Episodic Putting Descriptions

Generic putting description	Episodic putting description
Undergraduates	
1. Feet apart	1. Feet apart
2. Lean forward	2. Knees not locked
3. Aim ball	3. Leaning forward
4. Swing	4. Positioning hands
	5. Lining putter up with the ball
	6. Look at the hole
	7. Aim ball
	8. Swing
	9. Follow through
Athletes	
1. Estimating distance	1. Estimate distance to target
2. Bending knees	2. I placed my feet a comfortable distance apart
3. Looking back at target	3. Bent my knees
4. Relaxed backswing	4. Line up the putter with the target
5. Follow through	5. Slowly pulled the putter back
	6. Follow through lightly
	7. Using straight arms
Golf team members	
1. Walk behind the ball and look at the putt	1. Look up at putt
2. Read the green from behind the ball	2. Place putter behind ball with the head square at the target
3. Make sure nothing is in its path	3. Look at target
4. Look at distance of putt	4. Look at putter and ball
5. Pick a target to aim at	5. Take putter back
6. Place putter behind ball lined up with the target	6. Swing through ball
7. Move putter closer to you of the ball and line up at target	7. Look up at target
8. Take a practice swing	
9. Move putter back to behind the ball	
10. Line up squarely with target	
11. Move feet and body square with putter head	
12. Look at target	
13. Look down at the ball	
14. Swing the putter head straight back	
15. And straight through	
16. Look up at ball	

was after the 70th putt. To guard against this alternative explanation, we performed a reanalysis of the golf team members' protocols, dropping from each generic protocol all assessment steps that (a) did not appear in the corresponding episodic protocol and (b) were likely to be unnecessary once 69 putts had been taken in our laboratory situation. Excluded were steps such as "read the green" and "read the lie of the ball," because neither the green nor the lie of the ball changed during the experiment. Steps such as "taking aim," that would always be necessary in order to execute a putt, were maintained. This reanalysis of assessment produced the same-shaped interaction between expertise and type of protocol as the original, $F(2, 45) = 3.34$, $p < .05$, $MSE = 0.68$.

Turning to mechanics, this analysis also produced an interaction between expertise and protocol type, $F(2, 45) = 7.96$, $p < .001$, $MSE = 1.68$, but of a very different nature, as can be seen in the middle panel of Figure 3. Undergraduates gave significantly more mechanics steps in their episodic descriptions than in their generic descriptions, $t(15) = 3.34$, $p < .005$, and athletes produced a nonsignificant difference in the same direction, $t(15) = 0.36$. In contrast, the golfers gave more mechanics steps in their generic descriptions than in their episodic descriptions, though the difference was only marginally significant, $t(15) = 1.75$, $p < .10$. The greater number of mechanics steps in the episodic protocols of undergraduates compared with golfers was significant, $t(30) = 2.13$, $p < .05$. In sum, mechanics was a category of steps that for experts tended to appear more often in generic descriptions than in episodic descriptions, but for novices appeared more often

Table 2
Questionnaire Responses: Number of Steps (Experiment 1)

	Generic 1		Generic 2		Episodic	
Group	*M*	*SE*	*M*	*SE*	*M*	*SE*
Undergraduate	5.19	0.39	5.63	0.38	7.69	0.58
Athlete	5.94	0.54	6.25	0.55	6.75	0.77
Golf team	8.63	0.94	8.44	0.97	5.56	0.60

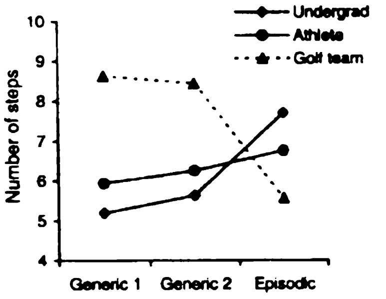

Figure 2. Mean number of steps for the first and second generic questionnaires and the episodic questionnaire for each group. Undergrad = undergraduate.

in episodic descriptions than in generic descriptions. Athletes were intermediate.

The analysis of ball destinations produced two main effects but no interaction. Overall, more ball destinations were included in the episodic protocols than in the generic protocols, $F(1, 45) = 9.36$, $p < .004$, $MSE = 0.22$, and the golf team included more destination information than either the undergraduates or the athletes, $F(2, 45) = 3.98$, $p < .026$, $MSE = 0.21$. Thus, as shown in the right panel of Figure 3, ball destinations were more likely to appear in the episodic recollections of experts than anywhere else, though even there they were relatively infrequent, accounting for only 13% of the steps that were included.

A second qualitative analysis looked for steps present in both protocols that referred to the same action or biomechanism but provided more detail in one type of protocol than in the other. For instance, a step in the episodic description of one participant was stated as "I positioned my feet so that they were shoulder length apart." This was scored as an elaboration of a step in the same participant's generic description that was stated as "feet positioning." Overall, elaborations were more likely to occur in episodic descriptions relative to generic descriptions than vice versa. Therefore, greater detail in the episodic description was scored as a "positive" elaboration whereas greater detail in the generic description was scored as a "negative" elaboration. In the undergraduate group, 14% of the steps in the episodic descriptions were elaborations of steps in the generic descriptions. In the athlete group, −1% of the episodic steps were elaborations of generic steps. In the golf team group, 5% of the episodic steps were elaborations of generic steps. A one-way ANOVA on these data produced a significant effect of expertise, $F(2, 45) = 3.53$, $p < .038$, $MSE = 1.23$. Fisher's LSD test showed that undergraduates elaborated their episodic recollections relative to their generic descriptions significantly more often than the athletes and marginally more often than the golf team ($p < .06$). Athletes and golfers did not significantly differ from one another.

Although the athlete group consisted of novice golfers, their elaborations were more similar to the golf teams' than to the undergraduates'. Similar to the athletes' pattern of mechanics steps, the athletes' pattern of elaborations suggests that sport training and participation lead athletes to approach novel skill situations in certain ways that resemble the approach of more experienced performers. This occurs despite the fact that the athletes' measured achievements in golf putting performance are no better than those of other novices.

Discussion

The results of Experiment 1 demonstrate an effect of level of expertise on the content of generic knowledge and episodic memories of golf putting. Experts gave longer, more detailed generic descriptions of the steps involved in a typical putt compared with the accounts given by novices and shorter, less extensive episodic recollections of a particular putt. These quantitative differences were accompanied by qualitative differences between experts and novices in the nature of the steps included in each type of description.

Expert golfers' generic descriptions dealt considerably more with assessing and planning a putt than did novices'. This finding is consistent with research on expert performers across a wide range of task domains (Chi, Feltovitch, & Glaser, 1981; Lesgold et al., 1988; Priest & Lindsay, 1992; Proctor & Dutta, 1995; Voss & Post, 1988). In areas as diverse as physics problem solving and radiological X-ray diagnosis, experts spend more time evaluating a situation and deciding how to approach or formulate a problem before they actually begin to work on it than do novices.

Expert golfers' episodic recollections included fewer assessment steps than did their generic descriptions. Expert golfers also made fewer references to putting mechanics in their episodic recollections than did novices. This pattern follows the prediction of expertise-induced amnesia derived from current theories of skill acquisition and automaticity. According to this idea, experts' extensive generic knowledge of putting is declaratively accessible during off-line reflection, but it is not used during real-time per-

Table 3
Assessment, Mechanic, and Destination Descriptions by Questionnaire Type—Experiment 1

	Generic								Episodic							
	Assessment steps		Mechanics steps		Destination description		Total steps		Assessment steps		Mechanics steps		Destination description		Total steps	
Group	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Undergraduate	1.44	0.32	4.19	0.56	0.00	0.00	5.63	0.38	1.62	0.24	5.88	0.74	0.19	0.14	7.69	0.58
Athlete	1.25	0.27	5.00	0.58	0.00	0.00	6.25	0.55	1.25	0.27	5.12	0.63	0.38	0.15	6.75	0.77
Golf team	3.69	0.48	4.50	0.84	0.25	0.11	8.44	0.97	1.37	0.30	3.63	0.75	0.56	0.16	5.56	0.60

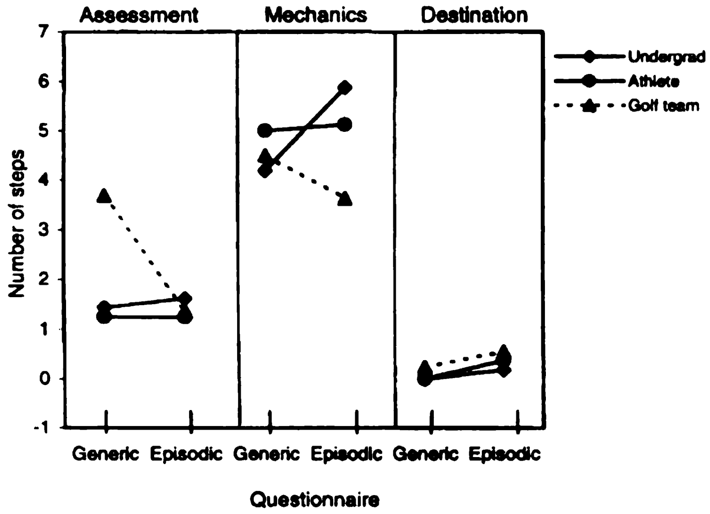

Figure 3. Mean number of steps in each category for the second generic questionnaire and the episodic questionnaire for each group. Undergrad = undergraduate.

formance, which is controlled by automated procedural knowledge. Because proceduralization reduces the need to attend to the processes by which skill execution unfolds, episodic recollection of step-by-step real-time performance is impoverished.

How are the details of these declarative reports related to the accuracy of performance? A significant negative correlation was found between the length of the undergraduates' generic descriptions and their pretest putting accuracy ($r = -.52, p < .03$). Because the measure of golf putting accuracy in the present study was an error score (i.e., mean distance from the target), it appears that the more detailed the generic descriptions supplied by the undergraduate novices early in practice, the better they performed. This correlation is consistent with an additional idea derived from current theories of skill and automaticity stipulating that novices' real-time performance is controlled by declaratively accessible knowledge concerning skill execution. Furthermore, this correlation was the only significant individual-differences relationship found between the contents of the declarative protocols and the accuracy of putting within any of the groups. This pattern suggests that a more extensive generic representation aids putting performance in the very earliest stages of skill learning but loses its impact as practice proceeds, once again consistent with expectations generated from theories of skill and automaticity. The disappearance of the correlation between undergraduates' generic knowledge and their performance accuracy appears to have occurred rapidly in the present situation, disappearing by 60–70 putts in the posttest scores. Thus procedural control structures may be established and begin to come to the fore quite quickly, at least in certain task domains (see Brown & Carr, 1989; Klapp et al., 1991; Raichle et al., 1994).

Although experts gave less elaborate episodic recollections of putting mechanics than did their novice counterparts, they gave more extensive recollections of ball destinations. This result suggests that performance outcomes are more salient to expert golfers, paralleling findings in other, more cognitive domains. It has been shown that expert physicists allocate more attentional resources to assessing and monitoring specific goal outcomes during problem solving than do less experienced physicists (Voss & Post, 1988). Of course, in our simple and repetitive situation, outcomes were generally similar to one another in both form and importance. The very low rate of inclusion of outcome information, even by experts, should increase as competitive motivations and consequences of success or failure become greater. We now turn to the second experiment in the present study, which was designed to replicate and extend the findings of Experiment 1.

Experiment 2

In Experiment 2 expert golfers' generic knowledge of golf putting and episodic recollection of specific putts were again compared with the generic knowledge and episodic recollection of novice golfers. Knowledge and recollection were assessed during either a standard golf putting task using a normal putter (i.e., the same task as in Experiment 1) or an altered putting task using a "funny putter" that consisted of a regular putter head attached to an S-shaped curved and arbitrarily weighted putter shaft. The design of the funny putter required experienced golfers to alter their well-practiced putting form in order to compensate for the distorted club, forcing them to allocate attention to the new skill execution processes. If experts' golf putting skill is proceduralized, then the disruption caused by the novel putter should not only lead to a lower level of performance in comparison to regular putter use but should also produce more elaborated episodic memory protocols—possibly similar to those of the novice golfers—as a result of the need to attend to the specific processes of skill execution under the constraints of the new putter. However, novice performers should not be affected by the funny putter in the same way as more experienced golfers. Because novices have not yet adapted to putting under normal putter constraints, performance should not depend as heavily on the type of putter used. Furthermore, according to the theories of skill acquisition we have reviewed, novices' on-line representations of golf putting are explicitly monitored in real time. Therefore, attending to novel putter constraints should not produce different episodic memory protocols in comparison with regular putter use, because in both cases novices attend to their performances in a way that should support explicit episodic memory.[1]

The design of Experiment 2 was similar to that of Experiment 1, with three exceptions. First, in order to ensure that individuals were not adapting to the highly repetitive task of putting from one specific spot on the green, all participants alternately putted from nine different spots, located at varying angles and distances from the target. Second, the experienced golfers in Experiment 2 were university students with 2 or more years of high school varsity golf experience rather than intercollegiate golf team members. Last, in Experiment 2 participants filled out two episodic protocols. As in Experiment 1, the first episodic questionnaire was unexpected. Prior to the last putt taken before the second episodic questionnaire, however, individuals were instructed to monitor their performance carefully for later recall.

Method

Participants

Participants (N = 72) were undergraduate students enrolled at Michigan State University and consisted of experienced golfers with 2 or more years of high school varsity golf experience (n = 36) and introductory psychology students with no golf experience (n = 36). Participants were randomly assigned within skill level to either a regular putter or funny putter condition in a 2 (novice golfer, experienced golfer) × 2 (regular putter, funny putter) experimental design with 18 participants in each group.

Procedure

After giving informed consent and filling out a demographic sheet concerning previous golf experiences, participants were told that the purpose of the study was to examine the accuracy of golf putting over several trials of practice. Participants were instructed that the object of the task was to putt a golf ball as accurately as possible from nine locations on a carpeted indoor putting green (3 × 3.7 m) that were either 1.2, 1.4, or 1.5 m away from a target, marked by a square of red tape, on which the ball was supposed to stop. All participants followed the same random alternation of putting from the nine different locations. A standard golf putter and golf ball were supplied for those participants who took part in the regular putter condition, and the funny putter and a standard golf ball were supplied for those participants in the funny putter condition.

All groups participated in identical pretest, practice, and posttest conditions, though the participants were not made aware of the separate conditions. To the participant, the golf putting task appeared to involve four blocks of putts with a short break after each block during which a questionnaire was filled out.

Pretest condition. Participants were set up at the first putting spot. They were asked whether they preferred to putt right-handed or left-handed and were given the appropriate putter. Participants were then informed that they would be putting from nine different locations on the green, each with a corresponding number. The experimenter reviewed the numbers associated with each putting location and asked participants to repeat back the numbers corresponding to each putting spot. Participants were informed that the experimenter would call out a number corresponding to a particular spot on the green from which they were to execute their next putt. Participants then took a series of 20 putts. After completing the putts, participants filled out a questionnaire eliciting a description of the steps involved in a typical golf putt (Appendix A, first paragraph).

Practice condition. Participants were again set up at the first putting spot. Participants took a series of 30 putts. After completing the putts, participants filled out an identical questionnaire to the one that they had previously filled out in the pretest condition eliciting a description of the steps involved in a typical golf putt (Appendix A, first paragraph).

Posttest 1 condition. Participants were set up at the first putting spot. Participants then took a series of 20 putts. Immediately following the first posttest condition, participants filled out a questionnaire designed to access their episodic recollection of the last putt they had just taken (Appendix A, third paragraph).

Posttest 2 condition. Participants were again set up at the first putting spot. Participants then took a series of 10 putts. Immediately prior to the 10th putt in the trial block, the experimenter instructed participants that they should pay close attention to the processes involved in their next putt because after it was complete, they would be asked to fill out another questionnaire, identical to the one they had just filled out, regarding their memories of this next putt. Immediately following the second posttest condition, participants filled out a questionnaire designed to access their episodic recollection of the last putt they had just taken (Appendix A, third paragraph).

Results

Putting Performance

Accuracy of putting was measured by the distance (in centimeters) away from the center of the target at which the ball stopped after each putt. As in Experiment 1, the mean distance from the target of the last 10 putts in the pretest condition was used as a measure of pretest golf putting skill. The mean distance from the target of the middle 10 putts in the practice condition was used as a measure of practice putting skill. The mean distance from the target of the last 10 putts in the first posttest condition was used as a measure of Posttest 1 golf putting skill. The mean distance from the target of the 10 putts in the second posttest condition was used as a measure of Posttest 2 golf putting skill. Means and standard errors for putting performance appear in Figure 4.

As can be seen from Figure 4, the experienced golfers showed superior putting performance in comparison with the novice golfers, regardless of type of putter used. This was true both before and after the practice phase. This pattern was confirmed by a 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) × 2 (pretest, Posttest 1) ANOVA, which revealed significant main effects of experience, $F(1, 68) = 42.73$, $p < .001$, $MSE = 51.55$, and test, $F(1, 68) = 4.04$, $p < .048$, $MSE \div 25.25$; no significant main effect of putter, $F(1, 68) = 1.47$, $p < .229$, $MSE = 51.55$; and no interaction of Test × Experience × Putter ($F < 1$).

In order to assess putting performance from the pretest condition to the second posttest condition, a three-way ANOVA similar to the one reported above was computed using the mean distance from the target of the last 10 putts in the pretest condition as a measure of pretest skill and the mean distance from the target of the 10 putts in the second posttest condition as a measure of

---

[1] We thank Claudia Carello for suggesting the funny putter as a diagnostic tool.

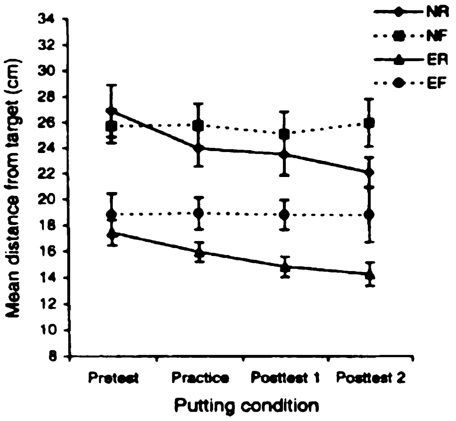

Figure 4. Mean (± *SE*) distance from the target at which the ball stopped after each putt in the pretest, practice condition, Posttest 1, and Posttest 2. NR = novice golfer–regular putter; NF = novice golfer–funny putter; ER = experienced golfer–regular putter; EF = experienced golfer–funny putter.

Posttest 2 golf putting skill. The results of this analysis did not differ from those reported above.

Thus, as can be seen from Figure 4, the experienced golfers, regardless of type of putter used, outperformed the novice golfers at all stages of practice. In addition, experienced golfers using the funny putter were less accurate than the regular putter–experienced golfers—especially during the practice condition and posttests. Independent sample *t* tests within the experienced golfers revealed no significant differences between putter type during the pretest, $t(34) = 0.74$, $p > .47$, but significant differences during the practice condition, $t(34) = 2.08$, $p < .05$, and the first posttest, $t(34) = 2.87$, $p < .007$, and marginally significant differences during the second posttest, $t(34) = 2.0$, $p < .054$. In contrast, the novice golfers did not significantly differ by putter type at any point in the experiment, although novices using the funny putter generally performed at a slightly lower level than their regular putter counterparts. Thus, although the funny putter produced differences in performance within higher levels of experience, it did not significantly affect the less experienced golfers. It should be noted that although experienced golfers using the funny putter performed at a lower level than regular putter experts during the pretest, this difference was not statistically significant. It may be that in the pretest condition, expert golfers—regardless of putter type—were adjusting to the novel experimental demands of having to land the ball on the target rather than in a hole. Thus, regular putter experts were not performing up to their potential in the pretest. The difference between the regular and funny putter experts widened quickly, however, appearing as early as the practice condition. Because experienced golfers often encounter novel putting green environments and must adapt to these situations in order to maintain a low handicap, it is not surprising the regular putter experts were able to rapidly adjust to our indoor green. In fact, several of the experienced golfers mentioned adjusting to the "fast green" or having to "land the ball on the tape" in their episodic protocols, suggesting that these individuals were able to identify and adapt to our somewhat irregular putting environment. In contrast, as can be seen from Figure 4, those experts using the funny putter were unable to adapt to the demands of the new putter within the time frame of the experiment, performing at a similar level of accuracy across experimental conditions.

Generic and Episodic Memory Protocols

As in Experiment 1, questionnaire responses were analyzed quantitatively, in terms of the number of golf putting steps included in each type of protocol, and qualitatively, in terms of the relative frequencies of different categories of steps.

Quantitative analysis. Analysis of number of golf putting steps given by participants was performed in the exact same manner as in Experiment 1. Two experimenters independently coded the data. Interexperimenter reliability was extremely high ($r = .95$).

Table 4 and Figure 5 present the results. A 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) × 2 (first generic protocol, second generic protocol) ANOVA on the two generic protocols revealed a marginally significant main effect of test, $F(1, 68) = 3.03$, $p < .086$, $MSE = 1.67$, and no interaction of Expertise × Putter × Test ($F < 1$). Thus, as in Experiment 1, the second generic protocol was used in a 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) × 2 (second generic protocol, first episodic protocol) ANOVA to compare the lengths of the generic and episodic protocols produced at each level of expertise. This analysis revealed an interaction of Experience × Putter × Questionnaire, $F(1, 68) = 9.63$, $p < .003$, $MSE = 2.77$.

A 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) general factorial ANOVA on the second generic

Table 4
Questionnaire Responses: Number of Steps (Experiment 2)

	Generic 1		Generic 2		Episodic 1		Episodic 2	
Group	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
NR	6.39	0.51	6.69	0.59	9.28	0.94	9.78	0.96
NF	6.11	0.65	6.67	0.75	9.11	0.72	9.83	0.81
ER	8.17	0.81	8.79	0.76	7.11	0.57	8.60	0.72
EF	10.22	0.81	10.30	0.85	11.89	0.75	11.78	0.89

Note. NR = novice golfer–regular putter; NF = novice golfer–funny putter; ER = experienced golfer–regular putter; EF = experienced golfer–funny putter.

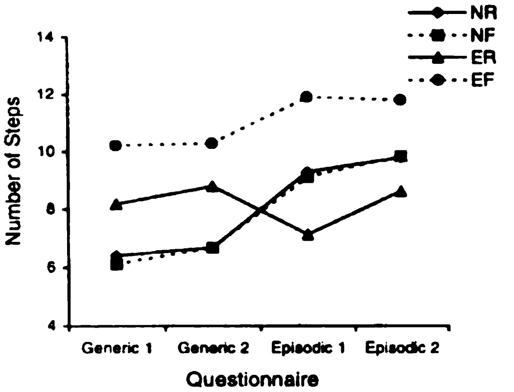

Figure 5. Mean number of steps for the first and second generic questionnaires and the first and second episodic questionnaires for each group. NR = novice golfer–regular putter; NF = novice golfer–funny putter; ER = experienced golfer–regular putter; EF = experienced golfer–funny putter.

protocol produced a main effect of expertise, $F(1, 68) = 14.72$, $p < .001$, $MSE = 10.01$, with the experienced performers giving longer generic protocols than the novices; no main effect of putter, $F(1, 68) = 1.01$, $p > .318$, $MSE = 10.01$; and no Experience × Putter interaction, $F(1, 68) = 1.01$, $p > .318$, $MSE = 10.01$. In contrast, a 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) general factorial ANOVA on the first episodic questionnaire produced an Experience × Putter interaction, $F(1, 68) = 10.70$, $p < .002$, $MSE = 10.28$. Independent sample t tests revealed that whereas the two novice groups did not differ in terms of the number of steps given in their episodic protocols, $t(34) = 0.14$, $p > .89$, the funny putter–experienced golfers gave significantly more steps in their episodic protocol than the regular putter–experienced golfers, $t(34) = 5.09$, $p < .001$. In addition, both novice groups and the funny putter–experienced group gave significantly more steps than the regular putter–experienced golfers in the episodic questionnaire ($ps < .05$).

Direct comparisons of the number of generic versus episodic steps within each group showed that, similar to Experiment 1, both the regular and funny putter novices gave significantly more steps in their episodic than their generic protocols, $t(17) = 4.10$, $p < .001$, and $t(17) = 6.27$, $p < .001$, respectively. In addition, the experienced golfers using the funny putter gave significantly more steps in their episodic than in their generic protocols, $t(17) = 2.64$, $p < .017$. In contrast, the experienced golfers using the regular putter gave significantly more steps in their generic than in their episodic protocols, $t(17) = 3.04$, $p < .007$. Furthermore, as can be seen from Figure 5, the experienced golfers using the funny putter gave longer generic and episodic putting descriptions than any other group. Increased attention to the novel constraints of the funny putter most likely prompted these golfers to allocate more attention to skill execution processes, enhancing generic descriptions and leaving explicit episodic memory traces of performance.

If the funny putter–experienced golfers gave more elaborate episodic descriptions as a result of increased attention to the specific processes involved in novel skill execution, then instructing these individuals to pay close attention to a particular instance of a putt, as did the instructions given prior to filling out the second episodic questionnaire, should not significantly change episodic descriptions in comparison to the first unexpected episodic questionnaire. That is, if the constraints of the funny putter serve to increase attention to skill execution, instructing experienced golfers to explicitly monitor performance should not alter attentional allocation and thus should not affect episodic memory protocols. In contrast, if those experts using the regular putter are asked to monitor performance for a later recall test, their episodic descriptions should increase in comparison to their first episodic protocol—especially if the first recollection was truly based on an unmonitored proceduralized instance of performance.

A 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) × 2 (first episodic protocol, second episodic protocol) ANOVA was performed, producing a significant Protocol × Experience × Putter interaction, $F(1, 68) = 4.08$, $p < .047$, $MSE = 1.86$. As can be seen in Figure 5, the novices, regardless of putter type, gave marginally longer putting descriptions in the second episodic questionnaire than in the first, $t(17) = 1.7$, $p < .108$, and $t(17) = 1.83$, $p < .085$, respectively. The experienced golfers using the funny putter did not differ in putting description length from the first to second episodic questionnaire, $t(17) = 0.11$, $p > .92$. In contrast, the experienced golfers using the regular putter gave longer protocols in the second episodic questionnaire in comparison to the first, $t(17) = 2.82$, $p < .012$. Furthermore, a 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) general factorial ANOVA on the second episodic questionnaire produced a marginally significant Experience × Putter interaction, $F(1, 68) = 3.35$, $p < .071$, $MSE = 12.99$. Thus, instructing participants to monitor skill execution did not affect the funny putter experts' episodic recollections and only marginally influenced the novice golfers' episodic descriptions. However, although instructing regular putter experts to monitor performance did increase their episodic recollections, they still did not reach the level of either the novice group or the funny putter–experienced golfers.

Qualitative analysis: Types of steps. The qualitative analysis was performed in the same manner as in Experiment 1. Steps were divided into three categories (assessment, mechanics, and ball destinations), and a 2 (experienced golfers, novice golfers) × 2 (funny putter, regular putter) × 2 (second generic protocol, first episodic protocol) ANOVA was conducted on the number of steps given in each of these three categories (see Table 5).

The analysis of assessment steps produced a significant interaction of expertise and type of protocol, $F(1, 68) = 14.53$, $p < .001$, $MSE = 1.2$, which is displayed in the left panel of Figure 6, along with a nonsignificant interaction of Expertise × Protocol × Putter ($F < 1$). A one-way ANOVA on the generic protocol with putter collapsed within skill level produced a main effect of experience, $F(1, 70) = 23.47$, $p < .001$, $MSE = 2.98$. Assessment steps appeared more often in the generic descriptions of experienced golfers, regardless of putter type, than anywhere else. In terms of the episodic protocol, a one-way ANOVA with putter collapsed within skill level produced a marginally significant main effect of experience, $F(1, 70) = 3.58$, $p < .063$, $MSE = 1.71$, with experienced golfers continuing to give more assessment steps in their episodic recollections than the novices. Paired sample t tests further revealed that the regular putter–experienced golfers gave

Table 5
Assessment, Mechanic, and Destination Descriptions by Questionnaire Type—Experiment 2

	Assessment		Mechanics		Destination		Total	
Group	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
				Generic				
NR	1.70	0.32	5.00	0.54	0.00	0.00	6.69	0.59
NF	1.56	0.29	5.11	0.74	0.00	0.00	6.67	0.75
ER	3.54	0.41	5.26	0.63	0.00	0.00	8.79	0.76
EF	3.65	0.57	6.66	0.82	0.00	0.00	10.30	0.85
				Episodic 1				
NR	1.72	0.37	7.33	0.71	0.22	0.10	9.28	0.94
NF	1.94	0.26	7.00	0.80	0.17	0.09	9.11	0.72
ER	2.06	0.19	4.94	0.58	0.11	0.08	7.11	0.57
EF	2.78	0.37	8.72	0.66	0.39	0.16	11.89	0.75
				Episodic 2				
NR	1.89	0.25	7.61	0.88	0.28	0.14	9.78	0.96
NF	1.67	0.26	7.89	0.90	0.28	0.11	9.83	0.81
ER	2.35	0.23	5.96	0.61	0.28	0.11	8.60	0.72
EF	2.56	0.35	9.06	0.70	0.17	0.09	11.78	0.89

Note. NR = novice golfer–regular putter; NF = novice golfer–funny putter; ER = experienced golfer–regular putter; EF = experienced golfer–funny putter.

significantly more assessment steps in their generic descriptions than in their episodic recollections, $t(17) = 4.03, p < .001$, and the funny putter–experienced golfers gave somewhat more assessment steps in their generic descriptions, though the difference was not significant, $t(17) = 1.72, p < .104$. In contrast, the regular putter novices did not differ in terms of the number of assessment steps given in the generic and episodic protocols, $t(17) = 0.00$, while the funny putter–novice group gave more assessment steps in their episodic recollections than in their generic protocols, $t(17) = 2.12$, $p < .049$. Thus, similar to Experiment 1, assessment steps decreased in number from the generic to episodic protocol for the two experienced groups, regardless of type of putter used, whereas the opposite pattern occurred in the two novice groups.

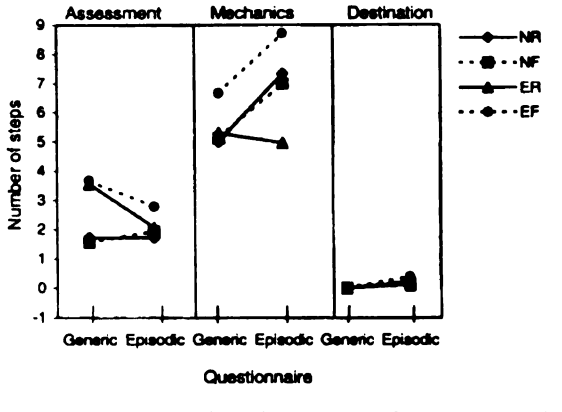

Figure 6. Mean number of steps in each category for the second generic questionnaire and the first episodic questionnaire for each group. NR = novice golfer–regular putter; NF = novice golfer–funny putter; ER = experienced golfer–regular putter; EF = experienced golfer–funny putter.

As an adjunct to the analysis of assessment, those steps that involved mental imagery (i.e., imagining some aspect of how a putt ought to look or feel before executing the action) were counted. In the regular putter–novice group, 2.7% of generic steps and 1.5% of episodic steps referred to imagery. In the funny putter–novice group, 0.6% of generic steps and 1.9% of episodic steps referred to imagery. In the regular putter–expert group, 2.2% of generic steps and 5.0% of episodic steps involved imagery. Finally, in the funny putter–expert group, 1.2% of generic steps and 1.4% of episodic steps referred to imagery. Thus, as in Experiment 1, golf putting steps involving imagery were predominantly reported by experienced golfers using the regular putter.

Turning to mechanics, this analysis produced an interaction of Experience × Protocol × Putter, $F(1, 68) = 5.26, p < .025$, $MSE = 3.43$, as can be seen in the middle panel of Figure 6. A 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) general factorial ANOVA on the generic protocol produced a nonsignificant main effect of experience, $F(1, 68) = 1.77, p < .188$, $MSE = 8.55$ (though experienced golfers did give more mechanics steps in their generic protocols than novices in terms of absolute number); no main effect of putter, $F(1, 68) = 1.18, p > .280$, $MSE = 8.55$; and no interaction of Experience × Putter ($F < 1$). In the episodic protocol, a 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) general factorial ANOVA produced an Experience × Putter interaction, $F(1, 68) = 8.82, p < .004$, $MSE = 8.63$. As can be seen from the middle panel of Figure 6, the experienced golfers using the funny putter gave more mechanics steps than any other group, while the regular putter–experienced golfers gave fewer mechanics steps than the other three groups. The two novice groups did not differ.

In addition, both regular and funny putter novices gave significantly more mechanics steps in their episodic recollections than in their generic protocols, $t(17) = 2.20$, $p < .001$, and $t(17) = 1.84$, $p < .001$, respectively. Similarly, the experienced golfers using the funny putter gave significantly more mechanics steps in their episodic recollections as compared with their generic protocols, $t(17) = 4.27$, $p < .001$. In contrast, the regular putter experienced golfers gave fewer mechanics steps in their episodic recollections than in their generic descriptions, although this difference was not significant, $t(17) = 0.556$, $p < .585$.

The analysis of ball destinations produced a main effect of protocol type, $F(1, 68) = 15.43$, $p < .001$, $MSE = 0.12$; no main effect of experience or putter ($Fs < 1$); and no interaction of Protocol × Experience × Putter, $F(1, 68) = 2.17$, $p > .145$, $MSE = 0.12$. Thus, as shown in the right panel of Figure 6, regardless of putter type or expertise, ball destinations were more likely to appear in episodic recollections than generic protocols.

As in Experiment 1, a second qualitative analysis looked for elaborations—steps present in both protocols that referred to the same action or biomechanism but provided more detail in one type of protocol than in the other. Because elaborations were more likely to occur in episodic descriptions relative to generic descriptions than vice versa, greater detail in the episodic description was scored as a positive elaboration whereas greater detail in the generic description was scored as a negative elaboration. In the regular putter–novice group, 11.1% of the steps in the episodic description were elaborations of steps in the generic descriptions. In the funny putter–novice group, 7.5% of the episodic steps were elaborations of generic steps. In the regular putter–experienced group, −0.3% of episodic descriptions were elaborations of generic steps. Finally, in the funny putter–experienced group, 3.5% of episodic recollections were elaborations of generic steps. A 2 (experienced golfer, novice golfer) × 2 (regular putter, funny putter) general factorial ANOVA on these data produced a significant main effect of experience, $F(1, 68) = 8.33$, $p < .005$, $MSE = 1.28$; a nonsignificant effect of putter ($F < 1$); and no Experience × Putter interaction, $F(1, 68) = 1.97$, $p > .165$, $MSE = 1.28$. Regardless of putter used, the novice golfers gave more elaborations in their episodic protocols than the more experienced golfers. Furthermore, all groups gave more elaborations in their episodic descriptions than in their generic protocols, with the exception of the regular putter–experienced golfers, who gave fewer elaborations in their episodic recollections.

In order to assess qualitative differences between the first and second episodic protocols, a 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) × 2 (first episodic protocol, second episodic protocol) ANOVA was also performed on each of the three categories of steps (assessment, mechanics, destination; see Table 5). The analysis of assessment steps produced a main effect of experience, $F(1, 68) = 6.00$, $p < .017$, $MSE = 2.34$; no main effect of protocol or putter; and no Experience × Putter × Protocol interaction ($Fs < 1$). Thus, as can be seen from the left panel of Figure 7, the experienced golfers gave more assessment steps in both episodic questionnaires than did either group of novices, who did not differ.

As an addition to assessment steps, the percentage of second episodic protocol steps involving references to mental imagery was also assessed. In the regular putter–novice group, 1.3% of second episodic steps referred to imagery. In the funny putter–novice group, 2.2% of second episodic steps referred to imagery. In the regular putter–expert group, 4.8% of second episodic steps involved imagery. Finally, in the funny putter–expert group, 1.3% of second episodic steps referred to imagery. Thus, as in the first episodic questionnaire, golf putting steps involving mental imagery were more likely to be found in the regular putter–experienced golfers' protocols than anywhere else.

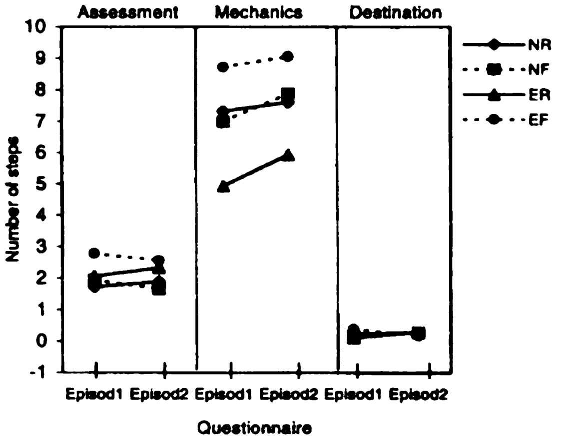

Figure 7. Mean number of steps in each category for the first and second episodic questionnaires for each group. Episod = Episodic; NR = novice golfer–regular putter; NF = novice golfer–funny putter; ER = experienced golfer–regular putter; EF = experienced golfer–funny putter.

The analysis of mechanics produced a main effect of protocol, $F(1, 68) = 8.13$, $p < .006$, $MSE = 1.73$, and an Experience × Putter interaction, $F(1, 68) = 6.07$, $p < .016$, $MSE = 17.87$. A 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) general factorial ANOVA was then performed on the combined average of mechanics steps involved in the first episodic questionnaire and second episodic questionnaire, revealing an Experience × Putter interaction, $F(1, 68) = 6.07$, $p < .016$, $MSE = 8.93$. As can be seen in the middle panel of Figure 7, mechanics steps increased from the first to second episodic protocol across all groups. Furthermore, the funny putter–experienced golfers gave more mechanics steps than any other group in either episodic questionnaire, whereas the regular putter–experienced golfers gave fewer mechanics steps than any other group. The two novice groups did not differ. Thus, although the regular putter–experienced golfers knew that they would be asked to give an episodic recollection of the last putt they took in the second episodic questionnaire, these golfers still gave fewer mechanics steps than both groups of inexperienced golfers and the experienced golfers using the funny putter.

Analysis of ball destinations was not interpretable because of inhomogeneity of variance across groups for destination steps reported in the episodic questionnaires. However, as can be seen from the right panel of Figure 7, all groups gave similar absolute numbers of ball destination steps in both the first ($M = 0.22$, $SD = 0.48$) and the second ($M = 0.25$, $SD = 0.47$) episodic questionnaires.

Discussion

Experiment 2 yielded results similar to those of Experiment 1. Experts using the regular putter gave more elaborate generic representations of putting than novices using either type of putter, in parallel with diminished episodic accounts of particular instances of performance. Those experts who were asked to use the funny putter also gave more detailed generic representations than did novices. However, in contrast to experts using the regular putter in both experiments, experts using the funny putter did not show diminished episodic memories for specific performances as would be expected if on-line performance was executed automatically and without leaving an explicit memory trace. In fact, funny putter experts gave more elaborate episodic descriptions of particular instances of skill execution than did the experts using the regular putter and both novice groups. The results of Experiment 2 suggest that real-time putting performance for experienced golfers is supported by proceduralized knowledge that may be disrupted through the addition of novel task constraints. When this disruption occurs and experts are forced to attend to step-by-step performance in the same way as novices, their expertise allows them to remember more of what they have attended to than do less skilled performers. This outcome resembles findings of superior episodic memory in chess and computer programming experts for the stimuli to which they have applied their knowledge (Chase & Simon, 1973; Soloway & Ehrlich, 1984). Furthermore, regardless of the type of putter used, novice golfers in the present study produced similar putting performances and generic and episodic putting descriptions, thus suggesting that in contrast to experienced golfers, novel skill performance in novice golfers is not based on a proceduralized, practice-specific skill representation.

The results of Experiments 1 and 2 speak to the nature of skill representations at various levels of expertise. With respect to the sensorimotor skill of golf putting, it appears that highly practiced, well-learned task components are encoded in a procedural form that supports effective real-time performance without requiring step-by-step attentional control. Reduced attention leads to a reduction in declaratively accessible episodic memory for details of the performance. However, if task constraints (e.g., a funny putter) are imposed that force experienced golfers to alter execution processes in order to adjust to the novel environment, the proceduralized skill knowledge that once drove normal execution is disrupted. The consequence is a reduction in putting accuracy, an extended period of adaptation in which learning appears to proceed rather slowly, and a more detailed episodic memory trace for specific instances of skill execution. The notion that well-learned sensorimotor skill performance is governed by a proceduralized representation carries implications for how this type of skill will behave in pressure or attention-demanding situations. Specifically, the two main theories that have been proposed to account for choking make different predictions concerning the types of skills that will be susceptible to performance decrements under pressure. Next we review these theories and how the cognitive mechanisms hypothesized by each theory to account for choking under pressure may be related to the type of knowledge representation governing task performance.

Experiments 3 and 4: Choking Under Pressure

As outlined in the introduction, two types of theories have attempted to explain choking under pressure. Distraction theory proposes that pressure influences task performance by creating a distracting environment. Distraction-based accounts of suboptimal performance propose that performance pressure shifts attentional focus to task-irrelevant cues—such as worries about the situation and its consequences. In essence, this shift of focus changes what was single-task performance into a dual-task situation in which controlling the task at hand and worrying about the situation compete for attention. The most notable arguments for the distraction hypothesis come from research involving academic test anxiety (Eysenck, 1979; Kahneman, 1973; Wine, 1971). Individuals who become highly anxious during test situations, and consequently perform at a suboptimal level, are thought to divide their attention between task-relevant and task-irrelevant thoughts more so than those who do not become overly anxious in high-pressure situations (Wine, 1971).

The distraction explanation for performance decrements under pressure is consistent with the idea that complex performances are attention or capacity demanding and that removing attention will disrupt performance (Humphreys & Revelle, 1984; Norman & Bobrow, 1975; Proctor & Dutta, 1995). However, as demonstrated in Experiments 1 and 2, there are skills that become automated or proceduralized with extended practice and thus may not require constant on-line attentional control during execution (Anderson, 1987, 1993; Fitts & Posner, 1967; Kimble & Perlmuter, 1970; Proctor & Dutta, 1995; Squire & Knowlton, 1994). Such skills should be able to withstand the attentional demands of a dual-task environment in that explicit attention to step-by-step skill procedures is not mandatory for successful performance. The well-learned sensorimotor task of golf putting may be one such task. However, skills that rely on declaratively accessible knowledge even in their practiced state may behave quite differently under pressure. A potential example is Zbrodoff and Logan's (1986) alphabet arithmetic task.

Alphabet arithmetic is a laboratory task analogous to mental arithmetic in which skilled performance is thought to be supported by the retrieval of stored instances of particular equations to which the performer has been exposed. Answers to an alphabet arithmetic problem such as "A + 2 = ?," whereby individuals must count two units down the alphabet to obtain the answer "C," may be achieved in two ways: either by using a rule-based system or algorithm to solve the equation or by the stimulus-driven retrieval of past instances of the problem from memory. Logan (1988) assumed that solutions are derived by a race between these two processes. As exposure to examples of problems increases, instances stored in memory increase as well. The larger the base of instances stored in memory, the higher the probability that memory retrieval will provide an answer to the problem before the rule-based algorithm reaches a solution. In either case, the answer enters working memory and hence is declaratively accessible—what differs is how it gets there (Klapp et al., 1991). Logan's model is supported by changes in speed and accuracy of performance with practice on the alphabet arithmetic task and other rapidly performed tasks involving judgment and choice, such as lexical decision and semantic categorization.

Answers to Logan's (1988) alphabet arithmetic task are thought to be declaratively accessible at all stages of skill learning. If choking is due to distraction of attention, one might imagine that choking would be a more imminent danger in tasks based on declarative knowledge that often enters working memory during the course of performance, because distraction of attention is a primary antecedent of corruption of information and forgetting in working memory (Daneman & Carpenter, 1980; Muter, 1980; Peterson & Peterson, 1959; Posner & Rossman, 1965). Furthermore, if the distraction hypothesis is valid, then it is possible that training in a dual-task environment would enable performers to adapt to distraction and the concurrent allocation of attention to something other than the primary task, alleviating the negative impact of pressure (Hirst, Neisser, & Spelke, 1978; Spelke, Hirst, & Neisser, 1976).

It has also been proposed that pressure situations raise anxiety and self-consciousness about performing correctly and successfully. The resulting focus on the self prompts individuals to turn their attention inward on the specific processes of performance in an attempt to exert more explicit monitoring and control than would be applied when a high-achievement outcome is less desired and its consequences are less important (Baumeister, 1984; Lewis & Linder, 1997). Note that the essence of this proposal is exactly the opposite of the distraction hypothesis. The main idea behind the self-focus, or what we would like to term the explicit monitoring hypothesis, is that although close attention and control may benefit novice performers in the initial stages of task learning, it will become counterproductive as practice builds a more and more automated performance repertoire. This is due to the fact that explicit monitoring of step-by-step skill processes and procedures is thought to disrupt well-learned or proceduralized skill execution processes (Kimble & Perlmuter, 1970; Langer & Imber, 1979; Lewis & Linder, 1997). Masters (1992) proposed that performance disruption occurs when an integrated or compiled real-time control structure that can run as an uninterrupted unit is broken back down into a sequence of smaller, separate, independent units—similar to how the performance was organized early in learning. Once broken down, each unit must be activated and run separately, which slows performance and, at each transition between units, creates an opportunity for error that was not present in the integrated control structure.

In addition to the differences in the types of knowledge that may govern task performance, variations in complexity of skills may also mediate the pressure–performance relationship. That is, it may be that complex skills, involving the integration and sequencing of multiple steps or parts, are more prone to breakdowns and performance deficits than less complex one-step skills. Certainly the skill of golf putting involves such complexity.

According to the explicit monitoring theory, then, the complex, proceduralized sensorimotor skill of golf putting analyzed in Experiments 1 and 2 should be extremely susceptible to performance decrements under pressure, as it is unaccustomed to being explicitly attended in real time. However, alphabet arithmetic, in which answers to particular problems are thought to be declaratively accessible at all stages of skill learning, should not be negatively affected by pressure-induced attention to performance processes. Furthermore, if the explicit monitoring hypothesis is valid, then training in an environment that heightens self-consciousness and achievement anxiety is likely to alleviate the negative impact of pressure, by adapting performers to conditions that entice them to pay too much attention to step-by-step execution.

Experiment 3

As a first step toward determining whether type of task knowledge and/or complexity might influence susceptibility to choking, we conducted Experiment 3. In this experiment participants learned either the sensorimotor task of golf putting or Zbrodoff and Logan's (1986) more declaratively based alphabet arithmetic task under single-task, dual-task, or self-consciousness-raising training conditions. Following training, participants were exposed to single-task low-pressure and high-pressure posttest situations in their task.

Testing the hypotheses concerning the distraction and explicit monitoring theories of choking under pressure requires control over the training environment. To ensure that our manipulation was the major source of each participant's golf putting or alphabet arithmetic experience, we recruited novice golfers and individuals with no exposure to alphabet arithmetic and taught them these tasks in the laboratory. However, despite the predictions we have made with respect to novice performance, choking as a concept is primarily aimed at individuals who can be expected to perform at a relatively accomplished level. Therefore, in order to examine performance at the later stages of skill acquisition, we trained participants rather heavily. Participants performed more than 280 golf putts or alphabet arithmetic trials in our laboratory prior to being exposed to a high-pressure situation. This number of task repetitions was chosen because pilot testing revealed a leveling off in performance with this amount of practice, suggesting that performance on the practiced putts or alphabet arithmetic problems was reaching asymptote.

Method

Participants

Undergraduate students ($N = 108$) with little or no golf experience who were enrolled in an introductory psychology class at Michigan State University served as participants. Participants were randomly assigned to either a single-task, self-consciousness, or dual-task distraction training group in either the golf putting or alphabet arithmetic task. Eighteen participants took part in each training group.

Procedure: Golf Putting Task

After giving informed consent and filling out a demographic sheet concerning previous golf experiences, participants were told that the purpose of the study was to examine the accuracy of golf putting over several trials of practice. Participants were instructed that the object of the task was to putt a golf ball as accurately as possible from nine locations on a carpeted indoor putting green (3 × 3.7 m) that were either 1.2, 1.4, or 1.5 m away from a target, marked by a square of red tape, on which the ball was supposed to stop. All participants followed the same random alternation of putting from the nine different locations. A standard golf putter and golf ball were supplied. Participants took part in a 270-putt training condition followed by an 18-putt low-pressure posttest and an 18-putt high-pressure posttest described below.

Single-task group. Participants were set up at the first putting spot. They were asked whether they preferred to putt right-handed or left-handed and were given the appropriate putter. Participants were then informed that they would be putting from nine different locations on the green, each with a corresponding number. The experimenter reviewed the numbers associated with each putting location and asked participants to repeat back the numbers corresponding to each putting spot. Participants were informed that the experimenter would call out a number corresponding to a particular spot on the green from which they were to execute their next putt. Participants then completed a total of 270 putts consisting of three training blocks of 90 putts each, with a short break after each set of putts. On completion of the training condition, participants were given a short break during which the experimenter computed the mean distance from the target of their last 18 putts.

Participants then completed an 18-putt single-task low-pressure posttest, though they were not made aware of the test situation. To the participant, the low-pressure posttest appeared to be just another series of putts. Participants were then informed of their mean putting performance for the last 18 putts in the training condition and given a scenario designed to create a high-pressure situation. Specifically, participants were told that if they could improve their accuracy by 20% in the next set of putts, they would receive $5. However, participants were also informed that this monetary award was a "team effort." Participants were told that they had been randomly paired with another participant, and in order to receive their $5, not only did they themselves have to improve by 20% but the participant that they had been paired with had to improve by 20% as well. Next, participants were informed that the individual they had been paired with had already completed the experiment and had improved by 20%. Therefore, if the present participant improved by 20%, both participants would receive $5. However, if the present participant did not improve by the required amount, neither participant would receive the money. Participants then took another 18 putts constituting the high-pressure posttest. Following these putts, the experimenter computed the participants' putting average and informed them of their performance. Finally, participants were fully debriefed and given the monetary award regardless of their performance.

Distraction group. Participants were set up at the first putting spot. They were asked whether they preferred to putt right-handed or left-handed and were given the appropriate putter. The experimenter informed participants that they would be putting from nine locations on the green. The experimenter then directed participants' attention to a tiny light that had been set up next to each putting spot. Participants were informed that the lights were hooked up to a switchboard controlled by the experimenter. Participants were told that before every putt, a light would illuminate beside the location from which they were to take their next putt. The experimenter then explained to participants that while they were putting they would be listening to a recorded list of spoken words being played from a tape recorder. Participants were told to monitor the words carefully, and, each time they heard the word *cognition*, to repeat it back to the experimenter. Words were played at the rate of one every 2 s. The target word occurred randomly once every four words. Participants then completed a total of 270 putts consisting of three training blocks of 90 putts each, with a short break after each set of putts during which the tape recorder was turned off. When participants completed the training condition, the tape recorder was turned off and participants were given a short break during which the experimenter computed the mean distance from the target of their last 18 putts. Participants then took part in an 18-putt low-pressure posttest and an 18-putt high-pressure posttest identical to that of the single-task group.

Self-consciousness group. Participants were set up at the first putting location. They were asked whether they preferred to putt right-handed or left-handed and were given the appropriate putter. The experimenter then explained that participants would be putting from nine different locations on the green, each with a corresponding number. Once the participants understood the number–putting spot relationships, the experimenter informed participants that they would be filmed by a video camera while putting. The video camera was set up on a tripod that stood on a table directly in front of participants, approximately 1.8 m away. Participants were told that they would be videotaped so that a number of golf teachers and coaches at Michigan State University could review the tapes in order to gain a better understanding of how individuals learn a golf putting skill. The experimenter adjusted the camera and turned it on. Participants then completed a total of 270 putts consisting of three training blocks of 90 putts each, with a short break after each set of putts during which the video camera was turned off. When participants completed the training condition, the video camera was turned off and faced away. Participants were then given a short break in which the experimenter computed the mean distance from the target of their last 18 putts. The participants then took part in an 18-putt low-pressure posttest and an 18-putt high-pressure posttest identical to that of the single-task and distraction groups.

Procedure: Alphabet Arithmetic Task

After giving informed consent and filling out a demographic sheet concerning previous golf and alphabet arithmetic experiences, participants were told that the purpose of the study was to examine how individuals learned the alphabet arithmetic task. Participants were set up in front of a monitor controlled by a standard laboratory computer. Participants were informed that they would be solving alphabet arithmetic equations such as "A + 2 = C" by counting two units down the alphabet to C. Next, the experimenter verbally presented three alphabet arithmetic equations to participants and instructed them to solve the equations out loud in order to ensure proper understanding. Participants were then shown a small keyboard containing two buttons labeled *True* and *False* and told to press the appropriate button when they derived the answer to an equation presented on the screen. Participants were instructed to try to judge the validity of the equations as quickly as possible without sacrificing accuracy.

The stimuli were capital letters, digits, the plus symbol (+), and the equal sign (=). All participants were presented with the same random order of nine different equations, consisting of the letters (A, G, and S) and the digits (2, 3, and 4) that were equally randomly repeated. Each trial began with a fixation point exposed for 500 ms in the center of the screen. The fixation point was immediately replaced by an equation, which remained on the screen until the participant pressed the *True* or *False* button on the keyboard. When the participant responded, the equation was extinguished, and the screen remained blank for a 1.5-s intertrial interval. All participants took part in a 270-equation training condition in which each equation was randomly repeated 30 times, followed by an 18-equation low-pressure posttest and an 18-equation high-pressure posttest, to be described below, in which each equation appeared twice in a random order.

Single-task group. Participants were set up in front of the monitor and given the alphabet arithmetic instructions. Participants then completed a total of 270 equations consisting of three training blocks of 90 equations each, with a short break after each set.

Following completion of the training condition, participants took part in an 18-equation single-task low-pressure posttest, similar to the low-pressure posttest in the golf putting task. Participants then took a short break and were given a scenario designed to create a high-pressure situation. Participants were told that the computer used a formula that equally took into account reaction time and accuracy in computing an "alphabet arithmetic performance score." The experimenter then described the same high-pressure scenario used in the golf putting task. Participants next completed the 18 equations constituting the high-pressure posttest, were fully debriefed, and were given the monetary award regardless of their performance.

Distraction group. Participants were set up in front of the monitor and given the alphabet arithmetic instructions. Participants were also told that while they were performing the arithmetic task they would be listening to a series of words being played through a headset. Participants were instructed to monitor the words carefully and, each time they heard the word *cognition*, to press a foot pedal that was located near their feet. The experimenter then instructed participants to put on the headphones, move the foot pedal to a comfortable location, and practice pressing it a few times. Participants then completed a total of 270 equations consisting of three training blocks of 90 equations each, with a short break after each set during which the headset was taken off.

Following completion of the training condition, participants were instructed to remove the headset and move the foot pedal away from their feet. Participants then took part in an 18-equation low-pressure posttest and 18-equation high-pressure posttest identical to that of the single-task group.

Self-consciousness group. Participants were set up in front of the monitor and given the alphabet arithmetic instructions. Participants were also told that they would be filmed by a video camera while solving the equations. The video camera was set up on a tripod directly to the left of participants, approximately 0.9 m away. Participants were told that they would be videotaped so that a number of math teachers at Michigan State University could review the tapes in order to determine how quickly and accurately individuals learn the arithmetic skill. The experimenter adjusted the camera and turned it on. Participants then completed a total of 270 equations consisting of three training blocks of 90 equations each, with a short break after each set during which the video camera was turned off.

Following completion of the training condition, the video camera was turned off and faced away from the participants. Participants then took part in an 18-equation low-pressure posttest and 18-equation high-pressure posttest identical to that of the single-task and distraction groups.

Results

Putting Performance

Accuracy of putting was measured by the distance (in centimeters) away from the center of the target that the ball stopped after each putt. All three groups improved significantly with practice as demonstrated by a 3 (single-task, distraction, self-consciousness) × 2 (mean distance from target of first 18 putts in training condition, mean distance from target of last 18 putts in training condition) ANOVA revealing a main effect of practice, $F(1, 51) = 85.03$, $p < .001$, $MSE = 27.57$; a nonsignificant main effect of training group, $F(2, 51) = 0.658$, $p > .522$, $MSE = 90.65$; and no interaction, $F(2, 51) = 0.214$, $p > .808$, $MSE = 27.57$. As can be seen in Figure 8, although there was not a significant effect of training group, the distraction group's performance was slightly, but not significantly, degraded in comparison to the single-task and the self-consciousness groups both early and late in the training condition. These results coincide with research in the skill performance literature demonstrating that dual-task performance may lead to a decrement in the performance of a primary task that has not become fully automatized (Proctor & Dutta, 1995).

In the low-pressure posttest (measured by the mean distance from the target of the 18 putts in the low-pressure test), putting accuracy was nearly identical across groups ($F < 1$). However, this homogeneity disappeared in the high-pressure posttest (measured by the mean distance from the target of the 18 putts in the high-pressure test) as confirmed by a one-way ANOVA revealing a significant difference between the three groups, $F(2, 51) = 4.57$, $p < .015$, $MSE = 23.57$. The above results are further supported by a 3 (single-task, distraction, self-consciousness) × 2 (low-pressure posttest, high-pressure posttest) ANOVA revealing a significant interaction of Training Group × Posttest, $F(2, 51) = 7.37$, $p < .002$, $MSE = 9.71$. Direct comparisons of putting performance within each group showed that both the single-task group and the distraction group significantly declined in putting accuracy from the low-pressure posttest to the high-pressure posttest, $t(17) = -2.21$, $p < .04$, and $t(17) = -3.24$, $p < .005$, respectively. In contrast, the self-consciousness group improved in putting accuracy from the low-pressure posttest to the high-pressure posttest, although this improvement was only marginally significant, $t(17) = 1.81$, $p < .09$. Thus, as can be seen in Figure 8, whereas both the single-task and distraction groups were adversely affected by the high-pressure situation, the self-consciousness group actually improved.

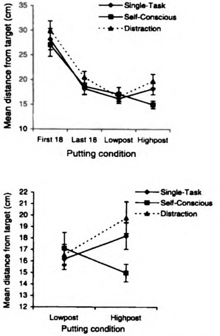

Figure 8. Mean (± *SE*) distance from the target at which the ball stopped after each putt in the training and posttest conditions for each group in the golf putting task (top) and posttest performance only (bottom). Lowpost = low-pressure posttest; Highpost = high-pressure posttest.

Alphabet Arithmetic Performance

Accuracy (shown in Table 6) was relatively high and did not differ significantly between groups both early (as measured by the mean number of correct judgments of the first 18 equations in the training condition), $F(2, 51) = 0.178$, $p > .838$, $MSE = 5.94$, and late (as measured by the mean number of correct judgments of the last 18 equations in the training condition), $F(2, 51) = 0.735$, $p > .485$, $MSE = 2.80$, in the training phase. Early in training, the single-task, distraction, and self-consciousness groups' accuracy was 90%, 89%, and 87% correct, respectively, and late in training the single-task and distraction groups' accuracy was 93% correct and the self-consciousness group's was 96% correct. Similarly, there were no significant differences in accuracy between groups during either the low-pressure posttest (as measured by the mean

Table 6
Mean Accuracy (Percentage Correct) in Training and Posttest Conditions of Alphabet Arithmetic Task

	Training 1 (first 18 equations)		Training 2 (last 18 equations)		Low-pressure posttest		High-pressure posttest	
Group	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Single task	89.51	2.00	93.21	2.96	96.30	1.56	95.06	1.61
Self-consciousness	87.04	4.42	96.30	1.19	96.60	1.20	94.75	1.38
Distraction	89.20	2.64	92.90	2.05	94.14	1.52	94.75	1.05

number of correct judgments of the 18 equations in the low-pressure test), $F(2, 51) = 0.879$, $p > .421$, $MSE = 1.2$, or the high-pressure posttest (as measured by the mean number of correct judgments of the 18 equations in the high-pressure test), $F(2, 51) = 0.017$, $p > .983$, $MSE = 1.09$. During the low-pressure posttest, accuracy for the single-task, distraction, and self-consciousness groups was 96%, 94%, and 97% correct, respectively, and accuracy during the high-pressure posttest was 95% correct for all three groups.

Reaction times were computed for only those equations that were answered correctly. Mean reaction times and standard errors both early and late in the training condition, as well as for the low- and high-pressure posttests, are illustrated in Figure 9. All three groups significantly decreased their reaction times across the training condition as shown by a 3 (single-task, distraction, self-consciousness) × 2 (mean reaction time of first 18 equations in the training condition, mean reaction time of last 18 equations in the training condition) ANOVA revealing main effects of practice, $F(1, 51) = 171.63$, $p < .001$, $MSE = 6.2 \times 10^5$, and training group, $F(2, 51) = 5.91$, $p < .005$, $MSE = 9.6 \times 10^5$, and a marginally significant interaction, $F(2, 51) = 2.88$, $p < .066$, $MSE = 6.2 \times 10^5$ (see Figure 9). Tukey's honestly significant difference (HSD) tests on the main effect of training group further revealed that the distraction group performed significantly worse than both the single-task and self-consciousness groups (who did not differ) early in the training condition ($p < .035$ and $p < .017$, respectively) and significantly worse than the self-consciousness group ($p < .05$) and nonsignificantly worse than the single-task group ($p < .227$) late in the training condition. This pattern did not change in either the low-pressure or high-pressure posttests as shown by a 3 (single-task, distraction, self-consciousness) × 2 (low-pressure posttest, high-pressure posttest) ANOVA that revealed main effects of both training group, $F(2, 51) = 4.41$, $p < .017$, $MSE = 2.4 \times 10^5$, and posttest, $F(1, 51) = 43.42$, $p < .001$, $MSE = 1.3 \times 10^4$, with no interaction, $F(2, 51) = 2.27$, $p > .114$, $MSE = 1.3 \times 10^4$. Tukey's HSD tests on the main effect of training group further revealed that in the low-pressure posttest, the distraction group produced significantly slower reaction times than the self-consciousness group ($p < .04$) and nonsignificantly slower reaction times than the single-task group ($p < .293$). In the high-pressure posttest, Tukey's HSD tests revealed that the distraction group produced significantly slower reaction times than both the self-consciousness ($p < .008$) and single-task ($p < .04$) groups. Thus, similar to golf putting accuracy, all three groups in the alphabet arithmetic task significantly improved their reaction times with practice. However, in contrast to the golf putting task, the dual-task alphabet arithmetic group performed substantially worse than the single-task and self-consciousness groups both early and late in the training condition, as well as in the posttest situations. The main effect of low-pressure versus high-pressure posttests observed in alphabet arithmetic also contrasts with putting. All three training groups improved somewhat in the high-pressure posttest, showing no signs of choking under pressure in the alphabet arithmetic task.

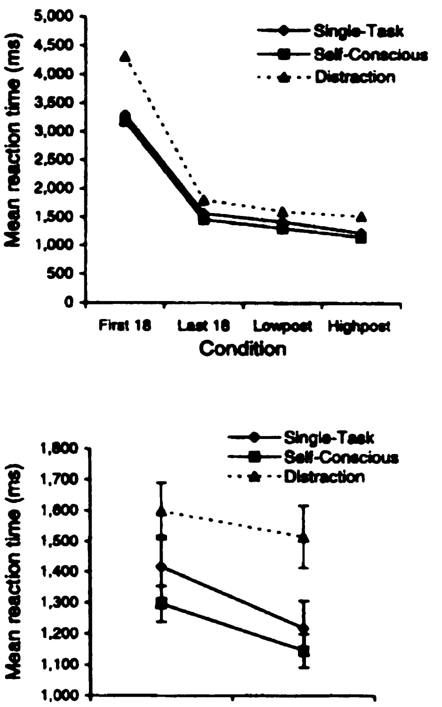

Figure 9. Mean reaction times (± *SE*) for arithmetic equations in the training and posttest conditions for each group in the alphabet arithmetic task (top) and posttest performance only (bottom). Lowpost = low-pressure posttest; Highpost = high-pressure posttest.

As an addendum, we should note that we believe the absence of choking in alphabet arithmetic to be a result of the fact that this type of skill does not become proceduralized with practice but moves from a declaratively accessible algorithm to retrieval of declaratively accessible facts into working memory. However, it may be that at higher levels of practice than we have examined, pressure-induced decrements in alphabet arithmetic performance could appear. Significant differences in performance between the dual-task training group and the other two groups in the alphabet arithmetic task at later stages of practice indicate that alphabet arithmetic performance was not yet fully automatized during the high-pressure situation (see Klapp et al., 1991). Thus, it may be that once differences in reaction times between these groups disappear, indicating full automatization, alphabet arithmetic will more closely resemble golf putting performance under pressure.[2]

To pursue this possibility, we conducted an analysis of alphabet arithmetic reaction times as a function of digit addend—the number of counting steps up the alphabet required by each equation. An effect of this variable has been used to diagnose the extent to which the control structure of alphabet arithmetic has shifted from the counting algorithm (which produces a significant effect of addend) to memory retrieval (which is independent of addend). This analysis produced a significant interaction of digit addend by pretest versus high-pressure posttest, $F(2, 104) = 7.06$, $p < .01$. Reaction time averaged across training groups increased markedly as a function of digit addend in the pretest ($M = 3{,}240$ ms for two-digit addend, $M = 3{,}690$ ms for three-digit addend, $M = 3{,}880$ ms for four-digit addend). Reaction time in the high-pressure posttest flattened considerably ($M = 1{,}185$ ms for two-digit addend, $M = 1{,}417$ ms for three-digit addend, $M = 1{,}310$ ms for four-digit addend), though there was still a significant effect of digit addend averaged across training groups. Further analysis of individual training groups showed that the significant effect of digit addend during the high-pressure posttest was a result of the single-task and distraction groups ($M = 1{,}084$ ms for two-digit addend, $M = 1{,}286$ ms for three-digit addend, $M = 1{,}268$ ms for four-digit addend and $M = 1{,}324$ ms for two-digit addend, $M = 1{,}679$ ms for three-digit addend, $M = 1{,}533$ ms for four-digit addend, respectively). Reaction time did not differ significantly as a function of addend for the self-consciousness group, $F(2, 34) = 2.73$, $p > .1$ ($M = 1{,}098$ ms for two-digit addend, $M = 1{,}211$ ms for three-digit addend, $M = 1{,}125$ ms for four-digit addend).

These data indicate that the self-consciousness training group achieved the most automated alphabet arithmetic performance, diagnosed by the relative independence they showed between alphabet arithmetic reaction time and digit addend. This is consistent with the prediction that increased monitoring of task components enhances skill acquisition among novices undergoing training (Anderson, 1987, 1993). If, similar to golf putting performance, those individuals trained under self-consciousness-raising conditions are immune to the detrimental effects of performance pressure, whereas those trained in single-task or distraction conditions are not, and the likelihood of choking increases as performance becomes more automated, then differences in high-pressure posttest reaction times should be apparent between the single-task and distraction groups on the one hand, and the self-consciousness group on the other. However, as can be seen in Figure 9, reaction time shows the same pattern across all three training groups. Thus it would appear that neither degree of automatization as measured by the effect of digit addend nor the differential impact of training condition bears on whether choking is observed in alphabet arithmetic. Increasing the amount of practice so that susceptibility to choking could be assessed in a completely automatized alphabet arithmetic skill would serve to further clarify this issue.

Putting Versus Alphabet Arithmetic

In order to further verify the differences in performance across posttests in the golf putting and alphabet arithmetic tasks, measurements taken in the putting and alphabet arithmetic tasks were converted into z scores. A 3 (single-task, distraction, self-consciousness) × 2 (low-pressure posttest, high-pressure posttest) × 2 (putting task, alphabet arithmetic) ANOVA was then performed, revealing a significant three-way interaction, $F(2, 102) = 6.5$, $p < .002$, $MSE = .08$. This confirms the pattern of data obtained above demonstrating that performance across the golf putting and alphabet arithmetic posttest conditions is different.

Discussion

Experiment 3 yielded three main results. First, following single-task practice, choking under pressure occurred in golf putting but not in alphabet arithmetic. Second, practice under dual-task conditions reduced performance in both tasks and altered practice benefits in alphabet arithmetic but did not alter either task's susceptibility to choking. Finally, practice under conditions intended to raise self-consciousness and execution-oriented achievement anxiety did not harm performance or change practice benefits relative to single-task practice in either skill but did inoculate putters against choking. Thus, at least at the levels of practice examined in the present study, choking arises in a task whose underlying knowledge base is thought to be procedural, but not one in which the underlying knowledge base is assumed to be more explicitly accessible. Furthermore, in terms of the effects of the two training regimens in the proceduralized task, it appears that when choking occurs, it results from explicit monitoring in response to self-consciousness and achievement anxiety. Performance pressure appears to elicit maladaptive efforts to impose step-by-step monitoring and control on complex, procedural knowledge that would have run off more automatically and efficiently had such monitoring not intervened. Practice at dealing with self-consciousness-raising situations counteracts this tendency.

We now turn to Experiment 4 in which we sought to replicate and extend Experiment 3's findings concerning the choking under pressure phenomenon. Because we were interested in the mechanisms governing choking, and alphabet arithmetic did not appear to show decrements in performance under pressure, in Experiment 4 we only examined the sensorimotor task of golf putting.

Experiment 4

In Experiment 4, the two possible sources of choking were examined at different stages of practice. It has been proposed that

---

[2] We thank Stuart Klapp for suggesting this possibility.

early in skill acquisition performance is supported by unintegrated control structures that are held in working memory and attended to in a step-by-step fashion (Anderson, 1987, 1993; Fitts & Posner, 1967). With practice, however, control evolves toward the type of integrated procedures assumed by explicit monitoring theory. According to explicit monitoring, tasks that follow this developmental trajectory should benefit from performance pressure early in learning yet be susceptible to choking at later stages of practice. Attention to task components is thought to be an integral part of novel skill performance (Proctor & Dutta, 1995). The explicit monitoring theory predicts that performance pressure prompts individuals to attend to skill execution processes. Thus, at low levels of practice, performance pressure should facilitate skill execution by prompting novice performers to allocate more attention to the task at hand.

According to distraction theory, however, performance pressure serves to create a dual-task environment. If individuals are attending to step-by-step execution processes in the early stages of skill learning, a distracting environment that draws attention away from the task at hand may harm performance. One could infer from the distraction hypothesis, then, that novice performers with little or no practice under divided attention conditions would be negatively affected by performance pressure, whereas those trained to a high skill level in a divided attention environment would not.

In Experiment 4 participants learned a golf putting task to a high level of skill under dual-task or self-consciousness-raising training conditions and were subjected to identical single-task low- and high-pressure situations both early and late in the training phase. If distraction is the reason for suboptimal performance under pressure, then individuals trained in either a dual-task or self-consciousness-raising environment should show performance decrements in pressure situations early in skill learning because, at this point, individuals in either training condition have not adapted to performing under divided attention conditions and do not possess a proceduralized skill response. Later in learning, however, those individuals trained in a dual-task environment will presumably be accustomed to performing under divided attention conditions and thus will not be affected by pressure, whereas the performance of those trained under conditions designed to increase anxiety and self-consciousness should decline. In contrast, if explicit monitoring is the reason for skill decrements under pressure, than at low levels of practice individuals trained under either distraction or self-consciousness-raising conditions should improve under pressure. If, as the explicit monitoring hypothesis predicts, pressure induces attention and control to skill performance, then novice performers may benefit from performance pressure in the initial stages of task learning. However, once the golf putting skill has become proceduralized later in practice, only those individuals who have adapted to performance anxiety and the demands to explicitly monitor skill performance (i.e., those trained under self-consciousness-raising conditions) should improve under pressure.

Method

Participants

Undergraduate students (N = 32) with little or no golf experience who were enrolled in an introductory psychology class at Michigan State University served as participants. Participants were randomly assigned to either a self-consciousness (n = 16) or dual-task distraction training group (n = 16).

Procedure

Participants completed the same golf putting task as in Experiment 3. Individuals took part in a 27-putt training condition followed by the first 18-putt single-task low-pressure posttest and 18-putt single-task high-pressure posttest. Participants then took part in a 225-putt training condition followed by a second 18-putt single-task low-pressure posttest and a second 18-putt single-task high-pressure posttest.

Distraction group. Participants completed the putting task in the same distraction training environment used in Experiment 3. Participants completed the first training condition of 27 putts, after which the tape recorder was turned off. Participants were then given a short break during which the experimenter computed the mean distance from the target of their last 18 putts.

Following the first training condition, participants completed the first 18-putt single-task low-pressure posttest, though they were not made aware of the test situation. Participants were then informed of their mean putting performance for the last 18 putts of the first training condition and given a scenario designed to create a high-pressure situation. Specifically, participants were told there would be two test situations in the present experiment. Participants were informed that they were about to take part in the first test situation and that the second test situation would take place toward the end of the experiment. Participants were given the same high-pressure scenario used in the golf putting task in Experiment 3, with the exception that they were told that they needed to improve their putting accuracy in two test situations to receive the monetary award. Participants then took the 18 putts constituting the first single-task high-pressure posttest. Following these putts, the experimenter informed participants that their putting average would be computed at the end of the experiment after both test situations had been completed.

Participants were then told that they would be taking another series of practice putts under the same dual-task conditions. The experimenter turned on the tape recorder, and participants completed the second training condition consisting of a total of 225 putts broken down into one training block of 72 putts, a training block of 81 putts, and another training block of 72 putts, with a short break after each block in which the tape recorder was turned off. When participants completed the second training condition, the tape recorder was turned off. Participants were then given a short break during which the experimenter computed the mean distance from the target of the last 18 putts in the training condition.

Participants then took part in the second 18-putt single-task low-pressure posttest. As in the first low-pressure posttest, participants were not made aware of the test situation. The experimenter then informed participants that they were about to take part in the second test situation and repeated the high-pressure scenario. Participants completed the second 18-putt single-task high-pressure posttest, were fully debriefed, and were given the monetary award regardless of their performance.

Self-consciousness group. Participants were set up at the first putting location and given instructions similar to those given to the distraction group regarding putting from the nine different locations on the green. Participants completed the putting task in the same self-consciousness-training environment used in Experiment 3. Participants completed the first training condition of 27 putts, after which the video camera was turned off and faced away from participants. Following the first training condition, participants were given a short break during which the experimenter computed the mean distance from the target of their last 18 putts. Participants then completed the first 18-putt single-task low-pressure posttest and high-pressure posttest identical to that of the distraction group.

Participants were then informed that they were going to complete another series of practice putts, again while being filmed by the video camera. The experimenter turned on the video camera, and participants completed the second training condition consisting of 225 total putts broken down into one training block of 72 putts, a training block of 81 putts, and another training block of 72 putts, with a short break after each block during which the video camera was turned off. When participants completed the second training condition, the camera was turned off and faced away from participants. Participants next completed the second 18-putt single-task low-pressure posttest and high-pressure posttest identical to that of the distraction group. Following the second high-pressure posttest, participants were fully debriefed and were given the monetary award regardless of their performance.

Results

Accuracy of putting was measured by the distance (in centimeters) away from the center of the target at which the ball stopped after each putt. Both groups improved significantly with practice as demonstrated by a 2 (distraction, self-consciousness) × 2 (mean distance from target of first 18 putts in first training condition, mean distance from target of last 18 putts in second training condition) ANOVA revealing a main effect of practice, $F(1, 30) = 58.63$, $p < .001$, $MSE = 31.73$; a nonsignificant main effect of training group ($F < 1$); and no interaction ($F < 1$; see Figure 10).

In the first low-pressure posttest (measured by the mean distance from the target of the 18 putts in the first low-pressure test), putting accuracy was similar across groups, $F(1, 30) = 1.40$, $p > .245$, $MSE = 26.15$. This homogeneity continued in the first high-pressure posttest (measured by the mean distance from the target of the 18 putts in the first high-pressure test) as confirmed by a one-way ANOVA again revealing no significant difference between groups ($F < 1$). These results are further supported by a 2 (distraction, self-consciousness) × 2 (first low-pressure posttest, first high-pressure posttest) ANOVA revealing a significant effect of test, $F(1, 30) = 17.73$, $p < .001$, $MSE = 12.38$; no significant effect of group, $F(1, 30) = 1.00$, $p > .323$, $MSE = 34.96$; and no interaction ($F < 1$). Direct comparisons of putting performance within each group showed that both the distraction and the self-consciousness groups significantly improved in putting accuracy from the first low-pressure posttest to the first high-pressure posttest, $t(15) = 3.76$, $p < .002$, and $t(15) = 2.30$, $p < .036$, respectively.

As can be seen in Figure 10, following the first high-pressure posttest, both training groups' performance accuracy decreased. This was confirmed by a 2 (distraction, self-consciousness) × 2 (first high-pressure posttest, first 18 putts of second training condition) ANOVA revealing a significant effect of condition, $F(1, 30) = 4.92$, $p < .034$, $MSE = 10.83$; no main effect of training group ($F < 1$); and no interaction ($F < 1$). Thus, whereas the first high-pressure posttest led to an increase in golf putting accuracy in comparison to the first low-pressure posttest in both the distraction and self-consciousness training groups, both groups showed performance decrements in the initial putts following the high-pressure situation. The explicit monitoring hypothesis suggests that performance pressure prompts individuals to explicitly monitor skill execution. Under this hypothesis, one would expect individuals in the initial stages of skill learning to improve under pressure as a result of increased attention to the novel demands of skill execution. However, once performance pressure and increased monitoring of performance are alleviated, a reduction in accuracy should occur.

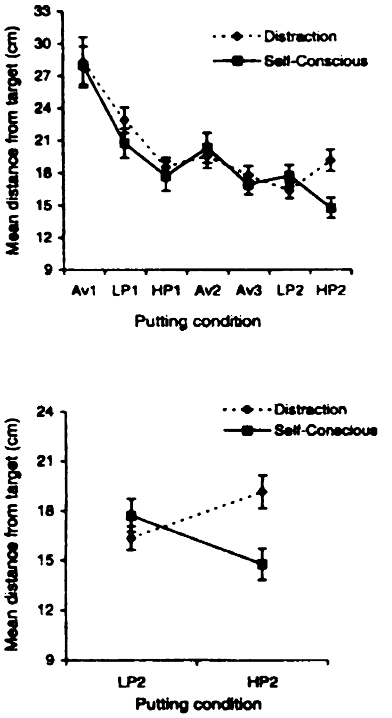

Figure 10. Top: Mean (± *SE*) distance from the target at which the ball stopped after each putt for the first 18 putts in the first training condition (Av1), the first low-pressure posttest (LP1), the first high-pressure posttest (HP1), the first 18 putts of the second training condition (Av2), the last 18 putts of the second training condition (Av3), the second low-pressure posttest (LP2), and the second high-pressure posttest (HP2). Bottom: Second posttest performance only.

In the second low-pressure posttest (measured by the mean distance from the target of the 18 putts in the second low-pressure test), putting accuracy was similar across groups, $F(1, 30) = 1.27$, $p > .269$, $MSE = 12.11$. This homogeneity disappeared in the second high-pressure posttest (measured by the mean distance from the target of the 18 putts in the second high-pressure test) as confirmed by a one-way ANOVA revealing a significant difference between groups, $F(1, 30) = 10.43$, $p < .003$, $MSE = 14.87$. The above results are further supported by a 2 (distraction, self-consciousness) × 2 (second low-pressure posttest, second high-pressure posttest) ANOVA revealing a significant interaction of Training Group × Second Posttest, $F(1, 30) = 24.16$, $p < .001$, $MSE = 5.55$. Direct comparisons of putting performance within each group showed that the distraction group significantly declined in putting accuracy from the second low-pressure posttest to the second high-pressure posttest, $t(15) = -2.79$, $p < .014$. In contrast, the self-consciousness group improved in putting accuracy from the second low-pressure posttest to the second high-pressure posttest, $t(15) = 4.84$, $p < .001$. Thus, as can be seen in Figure 10, both the distraction and self-consciousness groups improved from the first low- to the first high-pressure posttest. However, later in learning, those individuals trained in a dual-task environment showed decrements in performance under pressure, whereas those who learned the golf putting task under conditions designed to foster adaptation to a self-consciousness-raising environment that would increase achievement anxiety and explicit monitoring actually improved.

Discussion

The results of Experiment 4 once again are consistent with the predictions of the explicit monitoring theory of choking under pressure. Early in practice, regardless of training environment, performance pressure facilitated skill acquisition. However, as the golf putting skill became more proceduralized at later stages of practice, only those individuals who were accustomed to performing under conditions that heightened performance anxiety and the explicit monitoring of task processes and procedures were inoculated against the detrimental effects of performance pressure. These findings lend support to the notion that increased attention to the execution of a well-learned, complex skill may disrupt skill execution.

General Discussion

The purpose of the present study was to explore the cognitive mechanisms responsible for the disruption in the execution of a well-learned skill under pressure. Experiments 1 and 2 assessed the declarative accessibility of the knowledge representations governing real-time performance of golf putting at various levels of expertise. Results conformed in remarkable detail to predictions derived from current theories of automaticity and proceduralization of task performance as a function of practice. In Experiments 3 and 4 we looked at the phenomenon of choking under pressure in two very different tasks: the complex, sensorimotor skill of golf putting and a simpler, declaratively based alphabet arithmetic task. Results showed that choking under pressure occurred in golf putting but not in alphabet arithmetic, which demonstrates that sequential complexity, proceduralization, or both determine susceptibility to choking at the levels of practice we have examined. Furthermore, it was found that a particular training environment can eliminate choking when it does occur. Whereas single-task and dual-task practice left individuals in the golf putting task susceptible to performance decrements under pressure, self-consciousness training eliminated choking completely. Indeed, performers who experienced self-consciousness training actually improved under pressure—a highly desirable result. In addition to supporting the explicit monitoring hypothesis about why choking occurs, these experiments lead immediately to very practical ideas about training for real-world tasks in which serious consequences depend on good or poor performance in relatively public or consequential circumstances.

Properties of Tasks That are Susceptible to Choking

The pattern of results found in the present study speaks to the kinds of task properties that should be considered in the investigation of the pressure–performance relationship. Evidence of choking in the complex, proceduralized sensorimotor skill of golf putting but not in the simpler, declaratively based alphabet arithmetic task suggests at least three task properties that may be involved in choking. The first is task complexity. Masters (1992) argued that performance pressure prompts attention to skill execution, which results in the breakdown of task components. Well-learned complex skills may possess on-line control structures that run off as uninterrupted units. When attended to, these units may be broken down into a sequence of smaller, independent units, each of which must be run separately. As a result, performance slows and the transition between units creates an opportunity for error that was not present in the integrated control structure. Although skill breakdown may occur in complex, multistep tasks, this may not be the case in simple, one-step retrieval tasks. One-step retrieval tasks are not thought to consist of multiple integrated units and thus may not be susceptible to dismantling in the event of performance pressure. According to current theory, the automated form of alphabet arithmetic is such a one-step task (Klapp et al., 1991; Logan, 1988; Logan & Klapp, 1991).

However, our data indicate that alphabet arithmetic was not fully automated among our participants and hence was supported by some mixture of one-step fact retrieval and the multistep algorithm based on counting through the alphabet. Nevertheless, no hint of choking was observed in alphabet arithmetic. If task complexity is involved in choking, then performance decrements might have been expected at least on those trials supported by the multistep algorithm. In contrast, if task complexity does not affect susceptibility to choking, and instead choking occurs for alphabet arithmetic equations based completely on fact retrieval, then one might expect to see some indication of performance decrements for the portion of alphabet arithmetic equations that have switched to a fact retrieval mechanism. Either the former or latter of these possibilities should affect overall reaction time. However, as can be seen in Figure 9, alphabet arithmetic reaction time shows the same pattern both across training groups and between the low- and high-pressure conditions.

This leads to the second task property that may be involved in mediating the pressure–performance relationship. This is the degree to which task components become proceduralized with practice. Attention to the explicit processes involved in skill execution is thought to decrease as a function of skill level (Anderson, 1987, 1993; Fitts & Posner, 1967). As a result, skilled performances (e.g., complex sensorimotor tasks) are thought to operate largely outside of working memory. However, there may be certain skill types that rely on working memory for storage of control-relevant information during all stages of skill acquisition. Alphabet arithmetic is one such task. There is a substantial body of evidence demonstrating that performance on practiced alphabet arithmetic problems is not based on the establishment of a proceduralized version of the algorithm that controls action directly with no involvement from working memory. Instead, practice results in a shift from running through the steps of the algorithm in working memory to retrieving the answer into working memory from episodic memory (Klapp et al., 1991; Logan, 1988; Logan & Klapp, 1991). In either case the answer enters working memory, from where it controls the choice of an overt response. If choking is due to explicit monitoring, such skills should not be susceptible to decrements in performance because the practiced version of this task does not rely on the right kind of complex but proceduralized control structure. First, as already discussed, the control structure is too simple, and, second, control-relevant information always enters working memory and hence is always declaratively accessible. The alphabet arithmetic equation as a perceptual stimulus retrieves a single piece of information from episodic memory as the answer, and the elements of the control structure, the perceived equation and the retrieved answer, enter working memory rather than remaining outside the scope of attention, as does the relatively encapsulated procedure or motor program.

Finally, cognitive and motor tasks may differ in their susceptibility to breakdowns under pressure. The present study demonstrated that the sensorimotor task of golf putting, but not the cognitive task of alphabet arithmetic, was negatively affected by performance pressure. From these results it is tempting to conclude that choking may be confined to sensorimotor skills. However, such a conclusion is problematic in that it does not speak to the specific task characteristics that make a skill vulnerable to breakdowns under pressure. As mentioned in the introduction to this study, sensorimotor skills in both real-world and experimental settings are often associated with the choking phenomenon. Yet, the apparent prevalence of choking in sensorimotor domains may be not a function of sensorimotor skills per se but instead a result of specific task characteristics embedded in sensorimotor tasks that are susceptible to performance pressure (e.g., complexity and/or proceduralization). Furthermore, the notion that choking is limited to sensorimotor skills contrasts with research in the educational psychology literature demonstrating decrements in academic test performance under pressure in highly anxious individuals (Eysenck, 1979; Kahneman, 1973; Wine, 1971). Clearly such academic test performances do not have a large sensorimotor component, yet evidence of choking under pressure still emerges. Distraction theorists have suggested that suboptimal academic test performance results from the creation of a dual-task, distracting environment in which attention is divided between the task at hand and worries about the situation and its consequences. Thus, it remains a possibility that distraction as a mechanism for choking does hold for certain task types. It may be that pressure-induced distraction is detrimental to performance in tasks in which a large amount of information must be held in working memory and is susceptible to interference when attention is allocated to secondary sources (see, e.g., Tohill & Holyoak, 2000). This is a notion that is open to exploration in future work. However, it may also be the case that the types of problems encountered in these cognitive-based academic test situations have characteristics in common with many sensorimotor skills (e.g., complexity and/or proceduralizability) and are thus vulnerable to the same type of negative performance effects associated with the explicit monitoring of task execution processes. For example, Anderson (1993) suggested that complex cognitive skills such as algebra, geometry, and computer programming may become largely proceduralized in experts. We are pursuing this possibility in our laboratory.

Choking Research in Social Psychology

The present findings accord with research in the social psychology literature concerning the relationship between arousal, attention, and performance. It has been demonstrated that heightened anxiety and/or arousal levels induce self-focused attention (Fenigstein & Carver, 1978; Wegner & Giuliano, 1980). Wegner and Giuliano postulated that increments in arousal prompt individuals to turn their attention inward on themselves and current task performance in an attempt to seek out an explanation for their aroused state. Similar results have been found for skill execution in the presence of an audience. Butler and Baumeister (1998) recently demonstrated that supportive audiences were associated with unexpected performance decrements in the execution of complex, procedurally based tasks. The authors proposed that supportive audiences may increase attention to the processes involved in well-learned task performance, thus disrupting performance processes. And, finally, in a recent study investigating the effects of pressure on golf putting performance, Lewis and Linder (1997) found that pressure caused choking when participants had not adapted to performing in self-awareness-heightened environments—results similar to the present study. Furthermore, Lewis and Linder also found that decrements in performance could be alleviated through the use of a distractor (in this case, counting backward from 100) during real-time performance. Lewis and Linder suggested that attending to the distractor during on-line performance under pressure prevented participants from focusing attention inward on skill execution processes, thus alleviating the possibility of choking due to the "self-focus mediated misregulation" (p. 937) of performance. As can be seen from the literature just described, the notion that performance pressure induces self-focused attention, which in turn may lead to decrements in skill execution, is now a reasonably well-supported concept for proceduralized skills.

In conclusion, the findings of the four experiments in the present study lend support to the notion that pressure-induced attention to the well-learned components of a complex, proceduralized skill disrupts execution. Future research in this area is needed in order to illustrate the precise nature of the control structures that lead to decrements in performance under pressure. In addition, the generalizability of the present results to other task types and to different levels of practice must also be assessed. Further exploration of both the task and learning environments that mediate the pressure–performance relationship will serve to enhance our understanding of the choking under pressure phenomenon, which stands as an intriguing exception to the general rule that well-learned skills are robust and resistant to deterioration across a wide range of conditions.

References

Allard, F., & Starkes, J. L. (1991). Motor-skill experts in sports, dance and other domains. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise* (pp. 126–152). Cambridge, England: Cambridge University Press.

Anderson, J. R. (1982). Acquisition of a cognitive skill. *Psychological Review, 89,* 369–406.

Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review, 94,* 192–210.

Anderson, J. R. (1993). *Rules of mind.* Hillsdale, NJ: Erlbaum.

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology, 46,* 610–620.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11,* 717–726.

Brown, T. L., & Carr, T. H. (1989). Automaticity in skill acquisition: Mechanisms for reducing interference in concurrent performance. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 686–700.

Butler, J. L., & Baumeister, R. F. (1998). The trouble with friendly faces: Skilled performance with a supportive audience. *Journal of Personality and Social Psychology, 75,* 1213–1230.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4,* 55–81.

Chi, M. T., Feltovitch, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5,* 121–152.

Craik, F. M., Govini, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General, 125,* 159–180.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466.

De Groot, A. (1978). *Thought and choice in chess.* The Hague, the Netherlands: Mouton. (Original work published 1946)

Dooling, D. J., & Christiaansen, R. E. (1977). Episodic and semantic aspects of memory for prose. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 428–436.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100,* 363–406.

Ericsson, K. A., & Smith, J. (1991). Prospects and limits of the empirical study of expertise: An introduction. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise* (pp. 1–38). Cambridge, England: Cambridge University Press.

Eysenck, M. W. (1979). Anxiety learning and memory: A reconceptualization. *Journal of Research in Personality, 13,* 363–385.

Fenigstein, A., & Carver, C. S. (1978). Self-focusing effects of heartbeat feedback. *Journal of Personality and Social Psychology, 36,* 1241–1250.

Fisk, A. D., & Schneider, W. (1984). Memory as a function of attention, level of processing, and automatization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 181–197.

Fitts, P. M., & Posner, M. I. (1967). *Human performance.* Belmont, CA: Brooks/Cole.

Hardy, L., Mullen, R., & Jones, G. (1996). Knowledge and conscious control of motor actions under stress. *British Journal of Psychology, 87,* 621–636.

Hirst, W., Neisser, U., & Spelke, E. (1978). Divided attention. *Human Nature, 1,* 54–61.

Humphreys, M. S., & Revelle, W. (1984). Personality, motivation, and performance: A theory of the relationship between individual differences and information processing. *Psychological Review, 91,* 153–184.

Jones, B. T., Davis, M., Crenshaw, B., Behar, T., & Davis, M. (1998). *Classic instruction in golf.* New York: Broadway.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice Hall.

Keele, S. W. (1986). Motor control. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 2, pp. 1–60). New York: Wiley.

Keele, S. W., & Summers, J. J. (1976). The structure of motor programs. In G. E. Stelmach (Ed.), *Motor control: Issues and trends* (pp. 109–142). New York: Academic Press.

Kimble, G. A., & Perlmuter, L. C. (1970). The problem of volition. *Psychological Review, 77,* 361–384.

Klapp, S. T., Boches, C. A., Trabert, M. L., & Logan, G. D. (1991). Automatizing alphabet arithmetic: II. Are there practice effects after automaticity is achieved? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 196–209.

Langer, E., & Imber, G. (1979). When practice makes imperfect: Debilitating effects of overlearning. *Journal of Personality and Social Psychology, 37,* 2014–2024.

Lesgold, A., Robinson, H., Feltovitch, P., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise in a complex skill: Diagnosing X-ray pictures. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise* (pp. 311–342). Hillsdale, NJ: Erlbaum.

Lewis, B., & Linder, D. (1997). Thinking about choking? Attentional processes and paradoxical performance. *Personality and Social Psychology Bulletin, 23,* 937–944.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527.

Logan, G. D., & Klapp, S. T. (1991). Automatizing alphabet arithmetic: I. Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 179-195.

Masters, R. S. (1992). Knowledge, knerves and know-how: The role of explicit versus implicit knowledge in the breakdown of a complex motor skill under pressure. *British Journal of Psychology, 83,* 343–358.

McCandliss, B., & Carr, T. H. (1994, November). *Washing clothes twice: Do repetition and referential interpretation influence the same processes during reading?* Paper presented at the annual meeting of the Psychonomic Society, St. Louis, MO.

McCandliss, B., & Carr, T. H. (1996, November). *Repetition, comprehension, and intention to learn as sources of memory for text.* Paper presented at the annual meeting of the Psychonomic Society, Chicago.

Muter, P. (1980). Very rapid forgetting. *Memory & Cognition, 8,* 174–179.

Naveh-Benjamin, M., Craik, F. I., Guez, J., & Dori, H. (1998). Effects of divided attention on encoding and retrieval processes in human memory: Further support for an asymmetry. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1091–1104.

Norman, D. A., & Bobrow, D. J. (1975). On data-limited and resource-limited processes. *Cognitive Psychology, 7,* 44–64.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58,* 193–198.

Posner, M. I., & Rossman, E. (1965). Effect of size and location of informational transforms upon short-term retention. *Journal of Experimental Psychology, 70,* 496–505.

Priest, A. G., & Lindsay, R. O. (1992). New light on novice–expert differences in physics problem solving. *British Journal of Psychology, 83,* 389–405.

Proctor, R. W., & Dutta, A. (1995). *Skill acqusition and human performance.* Thousand Oaks, CA: Sage.

Raichle, M. E., Fiez, J. A., Videen, T. O., Macleod, A. M. K., Pardo, J. V., Fox, P. T., & Petersen, S. E. (1994). Practice-related changes in human brain functional-anatomy during nonmotor learning. *Cerebral Cortex, 4,* 8–26.

Reimann, P., & Chi, M. T. (1989). Human expertise in complex problem solving. In K. J. Gilhooly (Ed.), *Human machine and problem-solving* (pp. 161–189). New York: Plenum.

Soloway, E., & Ehrlich, K. (1984). Empirical studies of programming knowledge. *IEEE Transactions on Software Engineering, SE-10,* 595–609.

Spelke, E., Hirst, W., & Neisser, U. (1976). Skills of divided attention. *Cognition, 4,* 215–230.

Squire, L. R., & Knowlton, B. J. (1994). The organization of memory. In H. Morowitz & J. L. Singer (Eds.), *The mind, the brain, and complex adaptive systems* (Vol. 22, pp. 63–97). Reading, MA: Addison-Wesley.

Tohill, J. M., & Holyoak, K. (2000). The impact of anxiety on analogical reasoning. *Thinking and Reasoning, 6,* 27–40.

Van Lehn, K. (1989). Problem solving and cognitive skill acquisition. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 527–580). Cambridge, MA: MIT Press.

Voss, J. F., & Post, T. A. (1988). On the solving of ill-structured problems. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise* (pp. 261-285). Hillsdale, NJ: Erlbaum.

Wegner, D. M., & Giuliano, T. (1980). Arousal-induced attention to self. *Journal of Personality and Social Psychology, 38*, 719-726.

Wine, J. (1971). Test anxiety and direction of attention. *Psychological Bulletin, 76*, 92-104.

Woolfolk, R. L., Murphy, S. M., Gottesfeld, D., & Aitken, D. (1985) Effects of mental rehearsal of task motor activity and mental depiction of task outcome on motor skill performance. *Journal of Sport Psychology, 7*, 191-197.

Zbrodoff, N. J., & Logan, G. D. (1986). On the autonomy of mental processes: A case study of arithmetic. *Journal of Experimental Psychology: General, 115*, 118-130.

Appendix A

Questionnaires

Generic Questionnaire—Experiments 1 and 2

Certain steps are involved in executing a golf putt. Please list as many steps that you can think of, in the right order, which are involved in a typical golf putt.

Episodic Questionnaire—Experiment 1[a]

Pretend that your friend just walked into the room. Describe the last putt you took, in enough detail so that your friend could perform the same putt you just took.

Episodic Questionnaire—Experiment 2[b]

Pretend that your friend just walked into the room. Describe the last putt you took, in enough detail so that your friend could duplicate that last putt you just took in detail, doing it just like you did.

---

[a] Additional explanation was given in order to make it clear that what was being asked for was a "recipe" or "set of instructions" that would allow the putt to be duplicated in all its details by someone who had not seen it. Golf team members were told that the friend was not another golf team member but someone with an ordinary knowledge of the game. This was done to prevent excessive use of jargon or "in-group" shorthand, in an attempt to equate the need for knowledge that would be assumed by the describers across groups.

[b] This episodic questionnaire was changed slightly from Experiment 1 in an attempt to elicit the most detailed episodic descriptions possible from participants.

Appendix B

Steps Involved in a Typical Golf Putt

1. Judge the line of the ball.
2. Judge the grain of the turf.
3. Judge the distance and angle to the hole.
4. Image the ball going into the hole.
5. Position the ball somewhere between the center of your feet. You should be able to look straight down on top of the ball.
6. Align shoulders, hips, knees, and feet parallel and to the left of the target (e.g., image railroad tracks from the ball to the cup—feet outside the tracks, the ball in the middle).
7. Grip—thumbs should be pointed straight down, palms facing each other, a light grip.
8. Posture—stand tall enough so that if you were to practice putting for 30 minutes you would not experience a stiff or sore back.
9. Arms—should hang naturally and be relaxed.
10. Hands—should be relative to ball position. Hands should be slightly in front of the ball.
11. Head position—eyes should be positioned directly over the ball.
12. Weight—distribute weight evenly, about 50-50, or with a little more weight on the left foot.
13. Backswing—swing the club straight back. The distance back that the club goes must equal the through stroke distance.
14. Stroke—the club must accelerate through the ball. Finish with the "face" of the club head pointing directly at the target.
15. Length of the stroke—it is better to err to a shorter more compact stroke rather than a longer stroke.
16. Stroke direction—straight back and straight through.
17. Stroke rhythm—not too fast and not too slow.
18. Keep head and lower body stationary throughout stroke and swing with the arms.
19. Wrists—should not break during the stroke.
20. Arms and shoulders—should do most of the work.
21. Head/trunk/hips/legs—should remain still during the stroke.
22. Watch the ball go into the hole.

Received February 11, 2000
Revision received June 28, 2000
Accepted January 29, 2001 ■

APPENDIX B

Beilock, S. L., Carr, T. H., MacMahon, C., & Starkes, J. L. (2002). When paying attention becomes counterproductive: Impact of divided versus skill-focused attention on novice and experienced performance of sensorimotor skills. *Journal of Experimental Psychology: Applied, 8,* 6-16.

Journal of Experimental Psychology: Applied
2002, Vol. 8, No. 1, 6–16

1076-898X/02/$5.00 DOI: 10.1037//1076-898X.8.1.6

When Paying Attention Becomes Counterproductive: Impact of Divided Versus Skill-Focused Attention on Novice and Experienced Performance of Sensorimotor Skills

Sian L. Beilock and Thomas H. Carr
Michigan State University

Clare MacMahon and Janet L. Starkes
McMaster University

Two experiments examined the impact of attention on sensorimotor skills. In Experiment 1, experienced golfers putted under dual-task conditions designed to distract attention from putting and under skill-focused conditions that prompted attention to step-by-step putting performance. Dual-task condition putting was more accurate. In Experiment 2, right-footed novice and experienced soccer players dribbled through a slalom course under dual-task or skill-focused conditions. When using their dominant right foot, experts again performed better in the dual-task condition. However, when using their less proficient left foot, experts performed better in the skill-focused condition. Novices performed better under skill-focus regardless of foot. Whereas novices and the less-proficient performances of experts benefit from online attentional monitoring of step-by-step performance, high-level skill execution is harmed.

What drives the performance of a well-learned skill? Knowledge structures (Chi, Feltovich, & Glaser, 1981), memory capacities (Chase & Simon, 1973; de Groot, 1978; Starkes & Deakin, 1984), problem-solving abilities (Priest & Lindsay, 1992; Tenenbaum & Bar-Eli, 1993), and individual differences (Ackerman, 1987; Ackerman & Cianciolo, 2000; R. Kanfer & Ackerman, 1989) involved in high-level performance have been extensively examined, as well as compared across skill levels, in an attempt to shed light on the variables mediating exceptional task execution. Although work in this area has produced a number of important findings in both cognitive and sensorimotor skill domains (see Ericsson & Lehmann, 1996), there remain aspects of high-level performance that have not yet received adequate analysis.

One such area centers around the attentional mechanisms supporting skill execution in real time. That is, the manner in which experienced performers allocate attention to skill processes and procedures as actual skill execution unfolds, as well as differences in the attentional requirements of low- and high-level performances, are not yet fully understood. The purpose of the present study was twofold: (a) to assess the attentional mechanisms supporting performance of two sensorimotor skills in real time and (b) to explore the relationship between the attentional demands of online skill execution and degree of task proficiency. This knowledge will not only aid in developing the most appropriate techniques for optimal skill acquisition (Singer, Lidor, & Cauraugh, 1993; Wulf, Hob, & Prinz, 1998; Wulf, McNevin, Fuchs, Ritter, & Toole, 2000) but may also help to explain the suboptimal performance of well-learned skills in situations such as high-pressure environments thought to stress attentional capacity or interfere with its effective deployment (Baumeister, 1984; Beilock & Carr, 2001; Lewis & Linder, 1997).

Attention and the Proceduralization of Skill

Skill acquisition is believed to progress through distinct phases characterized by both qualitative differences in the cognitive structures supporting performance and differences in performance itself. Researchers have proposed that early in learning, skill execution is supported by a set of unintegrated control structures that are held in working memory and attended to one-by-one in a step-by-step fashion (Anderson, 1983, 1993; Fitts & Posner, 1967; Proctor & Dutta, 1995). As a result, attention is committed to controlling task performance and hence is largely unavailable for the interpretation or processing of nontask-related stimuli. With practice, however, procedural knowledge specific to the task at hand develops. Procedural knowledge does not require constant control and operates largely outside of working memory (Anderson, 1993; Fitts & Posner, 1967; Keele & Summers, 1976; Kimble & Perlmuter, 1970; Langer & Imber, 1979). Thus, in contrast to earlier stages of performance, once a skill becomes relatively well learned, attention may not be needed for the step-by-step control of execution and may be available for the processing of extraneous stimuli.

The impact of secondary task demands on novel skill acquisition and performance, as well as the ability of high-level performers to successfully operate in environments with substantial attentional load, have been widely demonstrated. Nissen and Bullemer (1987) examined individuals' performance on a novel sequence-learning task under both single-task and dual-task conditions. The sequence task was a four-choice reaction time task in which stimuli were

---

Sian L. Beilock, Departments of Psychology and Kinesiology, Michigan State University; Thomas H. Carr, Department of Psychology, Michigan State University; Clare MacMahon and Janet L. Starkes, Department of Kinesiology, McMaster University.

Correspondence concerning this article should be addressed to Sian L. Beilock, Department of Psychology, Michigan State University, Room 236, Psychology Research Building, East Lansing, Michigan 48824. E-mail: beilocks@msu.edu

presented in a consistent pattern in one of four locations on a computer screen. Participants were instructed to respond to the presentation of each stimulus by pressing a key at a corresponding spatial location. In the single-task condition, individuals practiced the sequence in isolation. In the dual-task condition, participants practiced the sequence while performing an attention-demanding, secondary auditory-monitoring task. In contrast to the single-task sequence condition, individuals performing under added secondary task demands showed no evidence of sequence learning during performance in the dual-task situation or during a transfer test in which individuals received the same sequence in isolation. Subsequent work by A. Cohen, Ivry, and Keele (1990) showed that the severity of the dual-task decrement in sequence learning varied with the complexity of the sequential pattern. These findings suggest that at the initial stages of performance, attention may be a necessary ingredient of skill learning, and the more so the more complicated the skill.

Although attention to task components may be important for novel skill execution and learning, this may not be the case at higher levels of practice. Allport, Antonis, and Reynolds (1972) found support for this notion through the examination of skilled pianists' ability to sight-read music while performing a secondary auditory-monitoring task. When skilled pianists were asked to sight-read in addition to shadowing a series of auditorially presented words, their sight-reading performance was not significantly altered in comparison with sight-reading in isolation. Allport et al. suggested that well-learned sight-reading does not demand constant attentional control. As a result, attention is available to devote to secondary task demands without significantly disrupting primary task performance.

Fisk and Schneider (1984) have also argued that well-learned performances are not based on explicit attentional control mechanisms. Individuals were trained on a visual search task in which they learned to search through arrays of visually presented words for members of a target category. With practice, both reaction time and accuracy in finding targets increased, whereas recognition memory for the words that had been searched through declined in comparison with memory at lower levels of practice. Fisk and Schneider suggested that practice automated performance, and automating performance improved real-time skill execution while decreasing attention to specific task components.

Thus, novel and well-learned performances appear to require different levels of attentional resources for successful execution. This experience by attentional demand interaction has also been demonstrated in complex sensorimotor skills drawn from the real world. For example, Leavitt (1979) examined novice and experienced ice hockey players' ability to complete a hockey task while performing a secondary visual shape-identification task. Individuals were required to skate and stick-handle a puck through a slalom course of pylons in isolation and while performing a monitoring task in which they identified geometric shapes projected onto a screen they could see from the ice. Leavitt found that the addition of the secondary visual shape-identification task to the primary skating-and-stick-handling task did not affect experienced hockey players' skating and stick-handling performance. However, when novices were required to perform the primary skating-and-stick-handling task in addition to the secondary monitoring task, their performance on the primary task declined markedly in comparison with skating and stick handling in isolation. In a similar study, Smith and Chamberlin (1992) also found differences across skill levels in performers' abilities to attend to multiple tasks simultaneously. Experienced and less skilled soccer players dribbled a soccer ball through a series of cones set up on a gymnasium floor. Individuals dribbled in isolation or while performing a secondary visual-monitoring task similar to that used in Leavitt's study mentioned above. Adding the secondary task harmed the dribbling of less skilled players in comparison with dribbling in isolation but did not significantly affect experienced soccer players' dribbling performance.

The results of Leavitt (1979) and Smith and Chamberlin (1992) support the notion that well-learned skill performance does not require constant online attentional control. However, it should be noted that these findings are potentially confounded. The secondary tasks used in the hockey and soccer tasks were both in the visual modality. Low-skill players often spend a considerable amount of time looking down at the puck or the ball while performing these skills. Thus, skill level differences in these studies may not have been due to experts' more automated control processes per se but instead may have been the result of less skilled individuals' higher need for visual information and feedback from the objects they were attempting to manipulate (i.e., a hockey puck or soccer ball). In essence, differences in "structural interference" (Kahneman, 1973; Wickens, 1980, 1984), rather than differential demands of novice and experienced skill performance on attention, may have led to the above mentioned findings. For novices, both the primary task of stick handling or dribbling and the secondary visual-monitoring task require visual input, and the information-gathering structures of the visual system cannot be directed toward both the primary and secondary tasks' stimuli simultaneously. Hence, the two tasks compete for visual information. This means that novice task performance under dual-task conditions may have suffered as a result of structural interference, whereas experienced individuals' performance, which presumably does not demand constant visual contact with the objects under control, did not.

Recently, Beilock, Wierenga, and Carr (in press a, b) addressed this confound in the exploration of the attentional mechanisms governing novice and well-learned golf putting. Novice and experienced golfers took a series of golf putts in a single-task putting condition (involving putting in a quiet, isolated environment) and in a dual-task condition. In the dual-task condition, individuals putted while monitoring verbally presented words for a specified target word. Because auditory capacity should not differ as a function of golf-putting experience, and auditory input should not create structural interference with the visual input of putting, this secondary task was designed to be free of the confound present in Leavitt (1979) and Smith and Chamberlin's (1992) work. Results demonstrated that experienced golfers' putting accuracy was not affected by the addition of the secondary monitoring task, in comparison with single-task putting. Furthermore, when experienced golfers were given an unexpected recognition memory test for a subset of the words contained in the monitoring task, their performance did not differ from a single-task word recognition test given as a baseline measure. Because attention is known to influence recognition memory (Craik, Govini, Naveh-Benjamin, & Anderson, 1996), this result indicates that experienced golfers were able to pay just as much attention to the words while putting as when attending to the words was their only task, and that doing so did not harm their putting performance. In contrast, novice golfers showed both putting and word recognition decrements from single- to dual-task conditions. Consistent with Leavitt and Smith and Chamberlin, Beilock et al. (in press a, b) concluded that expertise leads to the encoding of task components in a proceduralized form that supports effective real-time performance, without the need for constant online attentional control.

When Attention to Performance May Be Counterproductive

Well-learned skills do not appear to require constant attentional control during execution. However, the notion that these skills are based on a proceduralized or "automated" representation carries even stronger implications for attending to practiced performances. Researchers have proposed that attending to the step-by-step component processes of a proceduralized skill may actually disrupt execution (Baumeister, 1984; Beilock & Carr, 2001; Kimble & Perlmuter, 1970; Langer & Imber, 1979; Lewis & Linder, 1997). Therefore, attention to performance may become counterproductive as practice builds an increasingly automated performance repertoire. Masters and colleagues (Masters, 1992; Masters, Polman, & Hammond, 1993) proposed that attention to high-level skills results in their "breakdown," in which the compiled real-time control structure of a skill is broken down into a sequence of smaller, separate, independent units—similar to how performance may have been organized early in learning. Once broken down, each unit must be activated and run separately, which slows performance and, at each transition between units, creates an opportunity for error that was not present in the "chunked" control structure. Researchers have proposed that this process of breakdown contributes to the suboptimal performance of well-learned skills in high-pressure situations (Baumeister, 1984; Beilock & Carr, 2001; Lewis & Linder, 1997). This brings us to Experiment 1, in which we attempted to assess the attentional mechanisms governing the real-time execution of a well-learned golf-putting skill.

Experiment 1

If high-level skill execution is supported by procedural knowledge that does not mandate and, furthermore, may be harmed by continuous online control, then, as demonstrated by Leavitt (1979), Smith and Chamberlin (1992), and Beilock et al. (in press a; b), experienced performers should not be negatively affected by a dual-task environment that draws attention away from the task at hand. In contrast, attending to an explicit component of a well-learned skill may actually serve to disrupt or degrade automated performance procedures. In Experiment 1, experienced golfers performed a golf-putting task in a skill-focused attention condition in which individuals were prompted to attend to a specific component of their performance (i.e., the exact moment that their club head stopped its follow-through) and a dual-task attention condition in which experienced golfers executed the putting task while performing a secondary auditory-tone-monitoring task.

Method

Participants

Participants ($N = 21$) were undergraduate students (7 women, 14 men), ages 18–22 ($M = 19.86$ years, $SD = 0.96$ years), who were enrolled at Michigan State University with 2 or more years of high school varsity golf experience or a Professional Golfers' Association (PGA) handicap less than 8.

Task

Individuals performed the golf-putting task on a carpeted indoor putting green (3 m × 3.7 m). The task required participants to putt a golf ball as accurately as possible from nine different locations marked by squares of red tape. Three of the locations were 1.2 m, three locations were 1.4 m, and three locations were 1.5 m away from a target, also marked by a square of red tape, on which the ball was supposed to stop. All participants followed the same random alternation of putting from the nine different locations. A standard golf putter and golf ball were supplied. Participants took part in both the skill-focused and the dual-task attention condition.

Skill-focused condition. In the skill-focused condition, participants were instructed to attend to a particular component of their golf-putting swing. Specifically, individuals were instructed to monitor the swing of their club and at the exact moment they finished the follow-through of their swing, bringing the club head to a stop, to say the word "stop" out loud.

Dual-task condition. The dual-task attention condition involved putting while listening to a series of recorded tones being played from a tape recorder. Participants were instructed to monitor the tones carefully and each time they heard a specified target tone to say the word "tone" out loud. The target tone was played three times prior to the start of the dual-task condition to ensure that participants were familiar with this tone. Tones (500 ms each) occurred at a random time period once within every 2-s time interval. The target tone occurred randomly once every four tones. The random placement of the tones within the 2-s time intervals, as well as the random embedding of the target tone within the filler tones, was designed to prevent the golfers from anticipating secondary task tone presentation.

Studies in our laboratory revealed that experienced golfers perform a golf putt (from beginning the initial putt assessment to completing the actual mechanical act of implementing the putt) in roughly 10 s ($M = 10.40$ s, $SD = 1.69$ s, for 420 putts taken by 21 experienced participants). Tones occurred on average once every 2 s, and target tones occurred once every four tone presentations. Thus, individuals received about five tone presentations per putt, including a minimum of one target tone presentation.

Procedure

After giving consent and filling out a demographic sheet concerning previous golf experiences, participants were instructed that the purpose of the study was to examine the accuracy of golf putting over several trials of practice. Participants were set up at the first putting spot and asked whether they preferred to putt right-handed or left-handed. The experimenter informed participants that they would be putting from nine locations on the green. The experimenter then directed participants' attention to a tiny light that had been set up next to each putting spot. Participants were informed that the lights were hooked up to a switchboard controlled by the experimenter. Participants were told that before every putt, a light would illuminate beside the location from which they were to take their next putt. Individuals then performed one set of 20 putts. These putts constituted the practice trials.

The order of the attention conditions was counterbalanced between participants. Individuals performed one set of 20 putts in the dual-task condition and one set of 20 putts in the skill-focused attention condition. Participants were given a short break in between the two attention conditions. Accuracy of putting was measured by the distance in centimeters away from the center of the target that the ball stopped after each putt. The measurement was made by the experimenter while the participant was setting up for the next putt.

Results

Putting Performance

The mean distance from the target of the 20 putts in each condition was used as the measure of performance. In the practice condition, M = 15.09 cm (SD = 3.27); in the dual-task condition, M = 13.74 cm (SD = 2.65), and in the skill-focused condition, M = 19.44 cm (SD = 5.42).

A Bonferroni adjustment was performed on the critical p value of the following comparisons of experienced performers' putting accuracy to guard against inflation of Type I error rates as a result of multiple comparisons. The resulting critical p value was .017. Cohen's d was used as the measure of effect size (for equation, see, Dunlap, Cortina, Vaslow, & Burke, 1996). J. Cohen (1992) suggested that 0.20 is a small effect size, 0.50 is a medium effect size, and 0.80 is a large effect size.

Experienced golfers performed significantly better during the dual-task condition in comparison with the skill-focused condition, $t(20) = 5.22$, $p < .01$, $d = 1.26$. Additionally, putting in the skill-focused condition was significantly less accurate than in the single-task practice condition, $t(20) = 3.94$, $p < .01$, $d = 0.94$, whereas performance in the dual-task condition did not significantly differ from the practice condition, $t(20) = 1.71$, $d = 0.45$. This pattern of results coincides with predictions derived from the skill acquisition and automaticity literature. Specifically, high-level skill execution is thought to be governed by proceduralized knowledge that does not require explicit monitoring and control. Thus, a dual-task environment should not degrade performance in comparison with skill execution under single-task conditions, as attention should be available to allocate to secondary task demands if necessary without detracting from control of the primary skill. The above results demonstrate precisely this notion. It should be noted, however, that these findings may not hold true for dual-task environments in which the tasks draw on similar processes and hence create structural interference (e.g., looking at a golf ball while lining up a putt and visually monitoring a screen for a target object).

Attention Condition Secondary Task Performance

Skill-focused condition. Each trial in which individuals failed to say "stop" at the cessation of their golf swing was recorded. On average, failure to follow instructions occurred in 2.9% of the 20 putts taken in the skill-focused condition by each individual (M = 0.57 "stops," SD = 0.87 "stops"), or 0.14% of the 420 skill-focused putting trials across all participants.

Dual-task condition. Each instance in which individuals failed to identify a target tone was recorded. Failure to identify target tones occurred infrequently (M = 0.62 target tones, SD = 0.86 target tones). On average, individuals received 1.25 target tone presentations per putt for a total of 25 target tones per 20 putts taken in the dual-task condition. This led to an error rate of .025 per tone. Given that each participant's errors were distributed over 20 putts, failure to identify a target tone occurred in 3.1% of the 20 putts taken in the dual-task condition by each individual.

Additionally, a significant positive correlation was found between the number of target tone identification errors and our measure of putting accuracy (i.e., mean distance from the target that the ball landed after each putt) in the dual-task condition ($r = .47$, $p < .05$). That is, individuals who performed at a lower accuracy level in the putting task were also more likely to miss identifying a target tone. The fact that individuals with poorer putting accuracy were also less able to identify target tones suggests that putting performance under dual-task conditions was not the result of a simple trade-off between primary putting and secondary tone-monitoring performance. It should be noted that there may be individual differences in performance ability such that certain individuals are less accurate at both putting and secondary task target tone detection. Therefore, the possibility of a more complex trade-off cannot be completely ruled out. However, the positive correlation between putting accuracy error scores and tone detection error rates does suggest that even though average performance was at least as good in the dual-task condition as in the single-task practice condition, there remained a degree of variation in how much attention experienced golfers paid to their putting, which could be detected in accuracy of tone detection. The less accurate the putting was, the greater the amount of attention was paid to it as indexed by decreased accuracy of tone detection.

Finally, a comparison of target tone detection errors between the skill-focused and dual-task condition illustrates that secondary task performance was not significantly affected by our manipulations of attention, $t(20) = 0.21$, $d = 0.05$, *ns*. This observed effect size is substantially smaller than the standard for a small effect size ($d = 0.20$; J. Cohen, 1992), suggesting that if there is a difference in secondary task performance across conditions, it is trivial.

Discussion

The results of Experiment 1 demonstrate that well-learned golf putting does not require constant online control. As a result, attention is available for the processing of secondary task information if necessary (such as monitoring a series of auditory tones). However, when prompted to attend to a specific component of the golf swing, experienced performance degrades in comparison with both single-task practice performance and dual-task conditions. Although the negligible effects of divided attention on well-learned performance have been previously demonstrated (Beilock et al., in press a; b; Leavitt, 1979; Smith & Chamberlin, 1992), the consequences of explicitly attending to automated or proceduralized performance processes have not received so much investigation. The present findings suggest that well-learned performance may actually be compromised by attending to skill execution. This result complements recent evidence on "choking under pressure." Researchers have proposed that pressure to perform at a high level prompts attention to the step-by-step components of a well-learned skill. This attention is thought to disrupt or slow down skill execution, resulting in a less than optimal performance outcome (Baumeister, 1984; Beilock & Carr, 2001; Lewis & Linder, 1997). In the present study, directly instructing experienced golfers to attend to a specific component of their swing produced just this result—a less than optimal performance.

Experiment 2

Experiment 2 was designed to replicate the results of Experiment 1 in a movement skill that uses different effectors and imposes different temporal demands, as well as to examine the effects of dual-task and skill-focused attention on performance at differing levels of skill proficiency. Experiment 1 demonstrated that it can be disadvantageous to explicitly attend to a specific component of an automated or proceduralized well-learned skill performance. However, by analogy to ballistic versus nonballistic movements (Banich, 1997), it might be that explicit attention plays a different role in continuous tasks extended in time, such as soccer dribbling, than in the discrete golf-putting task with a defined beginning and ending point just reported. Furthermore, in addition to task differences, there also may be expertise differences in the impact of attention. Researchers have suggested that close attentional monitoring and attentional control benefits novice performers in the initial stages of task learning (Anderson, 1983; Fitts & Posner, 1967). This notion has received both empirical and anecdotal support (see Curran & Keele, 1993; Proctor & Dutta, 1995). Recently, however, the benefits of attention to specific task components in novel sensorimotor skill performance have been challenged (Singer et al., 1993; Wulf et al., 1998, 2000). Instead, researchers have suggested that instructing novices to attend to task properties during online motor skill performance may actually hinder skill acquisition. Wulf et al. (1998) examined the effects of both an internal focus of attention (defined as attention to specific body movements, much like our skill-focused condition) and an external focus of attention (defined as attention to the effects or outcomes of body movements) on the learning and retention of a ski simulator and stabilometer task. Results demonstrated that an external focus of attention led to more effective learning than an internal focus. Wulf and colleagues proposed that explicitly attending to skill execution at the initial stages of skill learning may actually hinder performance. Thus, a controversy remains over the types of attentional mechanisms thought to support less experienced or less practiced performance processes.

Experiment 2 was designed to assess the attentional mechanisms supporting soccer dribbling performance at different levels of skill proficiency. This was accomplished in two ways: First, the influence of dual-task and skill-focused attention on both novice and experienced soccer dribbling performance was examined. Second, the effects of these attentional manipulations on dominant and nondominant foot performance within soccer skill level were assessed.

If attention to well-learned skill execution disrupts performance, then one might expect explicit attention to experienced performers' dominant right-foot dribbling skill to compromise performance in comparison with dual-task conditions. However, this may not be the case for experienced players' nondominant left foot. That is, although soccer players must be skilled with both feet to compete at a high level, these athletes admit foot preferences and are often more skilled with one foot than with the other (Helsen & Starkes, 1999; Peters, 1981, 1988). Comments from the experienced soccer players in the present study concerning foot preference are consistent with this evidence. For example, one experienced participant stated that in comparison with right-foot dribbling, "dribbling with my left foot is the worst," and another stated that "when I use my left foot performance suffers." As with all introspective reports about task performance, these comments must be viewed with caution. Nevertheless, these comments do indicate that experienced players may not perceive their dominant right-foot and nondominant left-foot dribbling skills as equivalent. If it is true that experienced performers' right- and left-foot dribbling skills do not support the same level of task proficiency, then current theories of automaticity in skilled performance predict that these skills are likely not to be supported by the same attentional mechanisms (Anderson, 1983; Fitts & Posner, 1967; Logan, 1990). Therefore, right- and left-foot performance may be differentially affected by the skill-focused and dual-task attention manipulations in the present study. Put another way, if experienced performers' nondominant foot is not supported by a proceduralized knowledge structure, and explicit attention to less practiced performances serves to enhance skill execution, then in contrast to dominant right-foot dribbling, left-foot performance may actually benefit from explicit attention to skill execution. Similarly, skill-focused attention may lead to a higher level of performance than the dual-task condition for novices, regardless of foot—as novices should not be skilled dribblers with either foot.

In Experiment 2, right-foot dominant novice and experienced soccer players performed a dribbling task in which they dribbled a soccer ball through a slalom course made up of a series of pylons. Individuals performed the task under a dual-task condition involving an auditory word-monitoring task (dual-task condition) and a condition in which individuals were prompted to focus on a specific component of the dribbling task—the side of the foot that last made contact with the ball (skill-focused condition). As with golf putting in Experiment 1, the combination of auditory word monitoring and soccer dribbling should not create structural interference.

Participants took part in both attention conditions while dribbling with their dominant right foot and again while dribbling with their nondominant left foot. The attention and foot manipulations afforded the comparison of dribbling performance between novice and experienced soccer players under the different attentional manipulations in the soccer-dribbling task, as well as within-individual comparisons of dominant and nondominant foot performance.

Method

Participants

Participants ($N = 20$) were self-proclaimed right-handed and right-footed undergraduate students at McMaster University, ages 18–26 ($M = 20.20$ years, $SD = 1.85$ years). The novice participants (8 women, 2 men) had less than 2 years of organized soccer experience ($M = 1.10$ years, $SD = 0.74$ years). The experienced participants (8 women, 2 men) had 8 or more years of competitive soccer experience ($M = 13.30$ years, $SD = 2.75$ years).

Task

Individuals performed the soccer-dribbling task on an indoor gymnasium-type surface. The task required participants to dribble a soccer ball as rapidly as possible through a slalom course that consisted of six cones set 1.5 m apart for a total of 10.5 m from start to finish. Prior to each dribbling trial, participants were instructed to dribble the ball through the cones with either their right foot or their left foot. Individuals were also given instructions concerning the skill-focused and dual-task attention manipulations.

Skill-focused condition. In the skill-focused attention condition, individuals dribbled through the slalom course while a single tone occurred at a random time period on a blank tape once during every 6-s interval. The tone was temporally aligned with the occurrence of the target word in the dual-task condition so that the tone in the skill-focused condition appeared at the same rate as the target word in the dual-task condition. Individuals were instructed to attend to the side of their foot that was in contact with the ball throughout the dribbling trial, so that upon hearing the tone, individuals could verbally indicate whether they had just touched the ball with the outside or inside of their foot. The random placement of the tone was designed to prevent participants from anticipating its occurrence.

Dual-task condition. The dual-task condition involved dribbling through the slalom course while performing a secondary auditory-word-monitoring task. Individuals heard a series of single-syllable concrete nouns spoken from a tape recorder. Words were presented at a random time period once within every 2-s time interval. The target word, *thorn*, occurred randomly, averaging once every three words (6 s). Participants were instructed to monitor the list of words and to repeat the target word out loud every time it was played. The random placement of the words within the 2-s time intervals, as well as the random embedding of the target word within the filler words, was designed to prevent participants from anticipating secondary task word presentation.

Procedure

Participants completed a consent form and demographic sheet detailing previous soccer experience. Individuals were also asked to report their dominant hand and foot. The experimenter further explored individuals' foot preference by asking participants "Which foot would you normally kick a ball with?" This specific question was asked because it is relevant to the predictions made in the present study and is included on several measures of footedness (Day & MacNeilage, 1996; Searleman, 1980). Only those individuals who were self-proclaimed right-handed and right-footed were used.

Participants were instructed that the purpose of the task was to dribble a soccer ball as quickly and accurately as possible through the series of cones set up in front of them. Individuals were also informed that prior to each dribbling attempt, the experimenter would instruct them as to which foot to use. Finally, participants were told that each dribbling trial would be timed by the experimenter. If an error in dribbling performance occurred or the proper foot was not used, the dribbling trial was repeated. This was done to ensure that participants completed the entire slalom course with the specified foot. Because we were interested in making specific predictions concerning the dribbling performance of each foot under the various attentional manipulations, it was extremely important to ensure that participants were solely using the correct foot for each attention condition. Thus, trials containing dribbling errors were repeated. However, errors as a result of failure to use the specified foot were quite infrequent and did not significantly differ across the attention or foot conditions (novice right-foot practice: M = 0.05, SD = 0.22 errors; both skill-focus and dual-task: M = 0 errors; novice left-foot practice, skill-focus, and dual-task: M = 0 errors; experienced right-foot practice, skill-focus, and dual-task: M = 0 errors; experienced left-foot practice: M = 0.05, SD = 0.22 errors; skill-focus and dual-task: M = 0 errors).

The dependent measure was the time taken to complete each error-free trial, measured with a stopwatch to the nearest tenth of a second. Participants performed two dribbling trials with their right foot only and two dribbling trials with their left foot only. These four dribbling trials constituted the practice trials.

The order of the remaining dribbling trials was counterbalanced between participants. Individuals performed four sets of two dribbling trials (8 total dribbling trials), alternating feet (i.e., right foot only, left foot only) and attentional focus manipulations (i.e., dual-task or skill-focused attention) every two trials. All participants performed the dribbling task with all possible foot and attentional focus combinations. After every two trials in a specific attention condition had been completed, individuals were given a short break during which time they were asked to verbally count backward from 100 by 7s. This manipulation was designed to limit the influence of persisting thoughts about the previous attention condition on subsequent skill performance.

Results

Dribbling Performance

We used the mean of the two error-free dribbling trials performed with each foot under each condition as a measure of dribbling performance for that specific foot and condition. Table 1 presents means and standard deviations for left- and right-foot dribbling performance in the practice, skill-focused, and dual-task attention conditions for both novice and experienced participants. Bonferroni adjustments on the critical p value of dribbling time comparisons in the practice condition were performed to control for the inflation of Type I error rate as a result of multiple between-skill-level and within-skill-level comparisons. The resulting critical p value was .025. The experienced soccer players were significantly faster than novices during practice when instructed to dribble with either their right foot, $F(1, 18) = 52.54$, $p < .01$, $MSE = 1.11$, $d = 3.24$, or their left foot, $F(1, 18) = 13.47$, $p < .01$, $MSE = 3.55$, $d = 1.64$.

Direct comparisons within skill level demonstrated that the novices did not significantly differ in dribbling time between their right and left feet during the practice condition, $t(9) = 1.16$, $d = 0.35$, *ns*. However, this null effect exceeds J. Cohen's (1992) criterion for a small effect size, and thus this nonsignificant result most likely reflects the fact that we do not have adequate power to detect an effect this small. With a medium effect size of 0.50, power is equal to .18 (J. Cohen, 1988). Thus, given the low power of this comparison, it may be unwise to conclude from the lack of a significant difference that the null hypothesis is true.

However, it should be noted that the similarity in dribbling times between novices' right and left feet in the practice condition parallels other findings in our laboratory concerning novel skill performance. In golf putting, for example, novices have been found to putt at a similar accuracy level while using a standard golf putter or an S-shaped and arbitrarily weighted "funny putter" (Beilock & Carr, 2001; Beilock et al., in press a; b). Because novices are not accustomed to performing with either type of putter, the distorted funny putter does not significantly alter their putting accuracy. This is in contrast to experienced golfers, whose performance is degraded by the altered golf putter. In the present dribbling task, novices should not have been accustomed to dribbling with either foot. Thus, despite expressed foot preferences, it may not be too surprising that novices were not significantly more

Table 1
Novice and Experienced Participants' Mean Dribbling Times and Standard Deviations Across Conditions for Both Right and Left Feet

	Condition					
	Practice		Skill-focused		Dual-task	
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*
Novice						
Right foot	10.26	1.29	8.81	1.23	9.70	1.91
Left foot	11.02	2.48	9.30	1.90	10.47	2.04
Experienced						
Right foot	6.85	0.74	8.38	1.23	6.55	0.88
Left foot	7.93	0.97	7.01	0.85	8.21	1.01

skilled with one foot in comparison with the other. However, this was not the case for the experienced soccer players in the present study. Experienced soccer players were significantly faster with their right foot in comparison to their left foot during practice, $t(9) = 2.90$, $p < .03$, $d = 1.25$. This result is consistent with the earlier documented notion that high-level soccer players are often more skilled with one foot than with the other.

Turning to the attention conditions, we performed a 2 (novice, experienced) × 2 (right foot, left foot) × 2 (skill-focused attention, dual-task attention) repeated measures analysis of variance (Table 2). This analysis revealed a main effect of experience, in which the experienced participants dribbled faster than the novice participants across all foot and attention conditions. Furthermore, there was a significant Attention × Expertise interaction and a significant Attention × Foot interaction. However, these two-way interactions are qualified by a significant Experience × Foot × Attention Condition interaction.

In terms of right-foot dribbling, shown in the upper panel of Figure 1, experienced performers were faster than the novices during the dual-task condition ($d = 2.12$). In contrast, experienced and novice participants dribbled at a more similar speed in the skill-focused attention condition ($d = 0.35$). It should be noted, however, that this effect does exceed J. Cohen's (1992) criterion for a small effect size, and thus the similarity in novice and experienced players' right-foot dribbling speed in the skill-focused attention condition should be interpreted with caution. Thus, it is clear that experienced performers were markedly faster than novices in the dual-task condition, whereas in the skill-focused condition their advantage was substantially reduced. Furthermore, experienced soccer players dribbled faster in the dual-task condition in comparison with the skill-focused condition ($d = 1.62$), whereas a tendency toward the opposite pattern occurred in novices, who dribbled faster in the skill-focused condition than in the dual-task condition ($d = 0.56$).

Table 2
Analysis of Variance for Mean Dribbling Times in the Attention Conditions

Source	*df*	*MS*	*F*	*f*
Between subject				
Experience (E)	1	82.51	17.09**	.558
Error	18	4.83		
Within subject				
Attention condition (A)	1	2.61	1.94	.098
A × E	1	9.02	6.71*	.282
Error (A)	18	1.34		
Foot (F)	1	2.99	2.80	.135
F × E	1	1.92	1.12	.058
Error (F)	18	1.07		
A × F	1	13.67	13.82**	.482
A × F × E	1	9.46	9.56**	.370
Error (A × F)	18	0.99		

Note. Cohen's *f* was used as the measure of effect size (for equation, see J. Cohen, 1988). Cohen suggested that 0.10 is a small effect size, 0.25 is a medium effect size, and 0.40 is a large effect size.
* $p < .05$. ** $p < .01$.

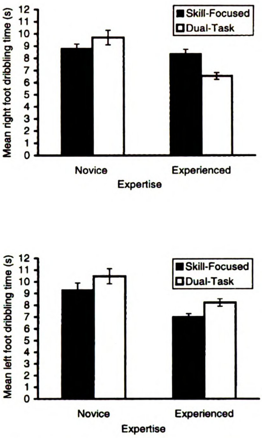

Figure 1. Mean right and left foot dribbling times in the skill-focused and dual-task attention conditions for novice and experienced performers. Error bars represent standard errors.

In terms of left-foot dribbling performance, the lower panel of Figure 1 illustrates that experienced performers were faster than novices during both the dual-task condition ($d = 1.40$) and skill-focused condition ($d = 1.56$). Additionally, novice and experienced soccer players performed better in the skill-focused condition than in the dual-task condition ($d = 0.59$ and $d = 1.25$, respectively). Thus, regardless of skill level, in left-foot dribbling, a higher level of performance occurred in the skill-focused condition, designed to draw attention to skill execution, than in the dual-task condition, designed to distract attention away from skill execution. This is in contrast to dominant right-foot dribbling, in which experienced and novice soccer players were differentially affected by the skill-focused and dual-task attention manipulations.

Finally, separate post hoc comparisons of novice and experienced dribbling performance in the attention conditions in contrast to dribbling in the practice condition were performed. Novices' right-foot dribbling during the skill-focused condition was faster than their right-foot practice condition performance (*d* = 1.19). In contrast, experienced soccer players' right-foot dribbling during the skill-focused condition was slower than their right-foot practice condition performance (*d* = 1.44). In terms of left-foot performance, both novice and experienced participants dribbled faster during the skill-focused condition in comparison with their respective left-foot practice condition performances (*d* = 0.76 and *d* = 1.01, respectively). Thus, although attention to dominant right-foot performance in the skill-focused condition led to an improvement in dribbling speed in comparison with the practice condition for novices, this same condition led to a decrement in experienced soccer players' dribbling skill. With the nondominant left foot, however, both novice and experienced performers improved in dribbling speed from the practice to skill-focused condition.

Attention Condition Secondary Task Performance

Skill-focused condition. On average, novice participants dribbling with their right foot heard 2.80 tones (*SD* = 0.42), whereas experienced participants heard 2.60 tones (*SD* = 0.52). During left-foot dribbling, novices heard an average of 2.80 tones (*SD* = 0.42), whereas experienced players heard 2.10 (*SD* = 0.32). Each instance in which individuals failed to verbalize the side of the foot that had just touched the ball following tone presentation was recorded. These errors occurred infrequently across both foot and experience level (*M* = 0.10 foot identifications, *SD* = 0.31 foot identifications). Overall, there were just two instances of participants failing to verbalize the side of the foot after tone presentation (1 novice and 1 experienced participant in the right-foot dribbling condition). Therefore, analysis of failures to identify the foot that had just touched the ball following tone presentations across foot and experience level was not interpretable because of the infrequency of these errors. Finally, if a skill-focused dribbling condition was completed in which no tones were heard, this trial was counted as an error and repeated. However, this occurred only on one trial across all participants (a novice right-foot dribbling trial).

Dual-task condition. On average, novice participants heard 5.30 words (*SD* = 0.82; *M* = 2.90 target words, *SD* = 0.32) while dribbling with their right foot, whereas experienced participants heard 3.80 words (*SD* = 0.63; *M* = 2.10 target words, *SD* = 0.32). In terms of left-foot dribbling, novices heard an average of 5.6 words (*SD* = 1.08; *M* = 2.90 target words, *SD* = 0.32), whereas experienced players heard 4.60 words (*SD* = 0.70; *M* = 2.50 target words, *SD* = 0.53). Each instance in which individuals failed to identify a target word was recorded. As in the skill-focused condition, errors were infrequent (*M* = 0.20 target words, *SD* = 0.41 target words). There were five instances of failure to identify a target word across both foot and expertise level (three target word identification failures in novice right-foot dribbling, one target identification failure in experienced right-foot dribbling, and one target identification failure in novice left-foot dribbling). Analysis of target identification differences across foot and experience level was not interpretable because of the infrequency of these errors. Thus, similar to Experiment 1, errors in secondary task performance were infrequent across both attention condition and level of expertise.

Discussion

The purpose of Experiment 2 was to explore differences in the attentional mechanisms supporting online sensorimotor skill execution in novice and experienced soccer players, as well as to assess differences in the attentional requirements of dominant and nondominant foot performance within level of expertise. Theories of skill acquisition have proposed that distinct cognitive processes are involved at different stages of skill execution. Early in learning, individuals are thought to attend to the step-by-step processes of performance. However, once a high level of performance has been achieved, constant online attentional control may not be necessary (Anderson, 1983, 1993; Fitts & Posner, 1967; Logan, 1988). One could infer from this framework of skill acquisition that novices might benefit from conditions that prompt attention to task properties yet not profit to the same extent in environments that divert attention away from the primary task at hand. In contrast, experienced performers may be harmed by explicit attention to skill processes that normally run as uninterrupted programs or procedures, yet they may not be adversely affected by conditions that draw attention away from performance. However, this may hold true only for those aspects of an experienced performer's skill execution repertoire that are indeed governed by a proceduralized or automated knowledge representation. If particular aspects of a skill are not as well learned or as highly accomplished, then experienced individuals—like less practiced performers—may benefit more from conditions that prompt attention to the task at hand rather than take it away.

The results of the present study conform quite well to these predictions derived from theories of skill acquisition. For right-foot dribbling, novices performed at a lower level in the dual-task condition, designed to distract attention from task performance, in comparison with the skill-focused manipulation, designed to draw attention toward the task at hand. Furthermore, novices substantially improved in dribbling speed from the single-task practice condition to the skill-focused condition. Experienced soccer players showed an opposite pattern of results. Experienced individuals performed at a lower level in the skill-focused condition compared with either the dual-task or practice condition. These results coincide with those of Experiment 1 and, as mentioned above, are consistent with current theories of choking under pressure (Beilock & Carr, 2001).

Performance with the left foot differed. In contrast to right-foot dribbling, novice and experienced soccer players alike performed better in the skill-focused condition than in the dual-task or practice condition. In the present study, there were significant differences in experienced performers' right- and left-foot dribbling speed in the practice trials. This pattern of results suggests that experienced players' left-foot dribbling skill was not at the same performance level as their dominant right-foot skill. The fact that the differential impact of the attentional manipulations in the present study was evident not only between skill levels but within experienced performers' dominant and nondominant feet performance as well speaks to the robust nature of the impact of attention on skill performance.

General Discussion

When Attention to Performance Becomes Counterproductive

The findings of Experiments 1 and 2 demonstrate that skill-focused attention benefits less practiced and less proficient performances yet hinders performance at higher levels of skill execution. High-level skills are thought to become proceduralized or automated with extended practice. The encoding of task components in a proceduralized form supports effective real-time performance, without the need for constant online control. As a result, skill performance decrements occur in conditions that impose step-by-step monitoring and control on complex, procedural knowledge that would have operated more automatically and efficiently had such monitoring not intervened. Therefore, experienced performers suffer more than novices from conditions that call their attention to individual task components or elicit step-by-step monitoring and control. However, experienced performers are better able than novices to spare a portion of their attention for other stimuli and task demands, and hence are better able than novices to deal with conditions that create dual-task environments (e.g., taking a series of golf putts or dribbling a soccer ball while performing an auditory-monitoring task). As shown by the contrast between right-foot and left-foot dribbling, however, this may hold only for that portion of an experienced performers' skill repertoire that is supported heavily by proceduralized knowledge structures.

The findings of the present study confirm results generated in Leavitt's (1979) hockey study, Smith and Chamberlin's (1992) soccer-dribbling task, and Beilock et al.'s (in press a; b) golf-putting study. Furthermore, the present findings expand previous results by examining the consequences of explicitly attending to both novel and well-learned performances. Researchers have recently suggested that attention to the step-by-step components of a novel skill may be detrimental to performance (Singer et al., 1993; Wulf et al., 1998, 2000). However, the present findings demonstrate that attention benefits both novel skill performance and performance that is not based on a heavily proceduralized knowledge representation, even though carried out by an experienced performer (e.g., experienced soccer players' nondominant foot performance). In contrast, at higher levels of learning and proficiency, increased attention to the step-by-step execution of a well-learned skill appears to have the opposite effect—disrupting skill execution processes.

It should be noted that the present study examined the performance of a golf-putting task and a soccer-dribbling skill under different attentional manipulations at approximately constant levels of performance, rather than examining the learning or transfer of these skills to novel task situations. For this reason it remains possible that under conditions commonly used to assess skill learning (e.g., transfer tests), a different pattern of performance may arise. Future research in this area would serve to shed light on this issue.

Not All Forms of Attention Are Counterproductive to Well-Learned Skill Performance

The above mentioned results indicate that attention to step-by-step skill execution—what we term *skill-focused attention*—may benefit novel performances yet hinder well-learned and highly proficient task execution. However, this relationship may not extend to other forms of attention to task-related information. R. Kanfer and Ackerman (1989) demonstrated that "self-regulatory" activities, including the allocation of attention to performance outcomes and goal attainment, self-evaluation, and self-reactions, detract from the lower level performances of novices yet enhance skill execution at later stages of learning and higher levels of proficiency. Self-regulatory activities are thought to require attentional capacity for successful initiation and implementation. Thus, self-regulation may disrupt novel skill execution by recruiting attentional resources needed for control of task performance (F. H. Kanfer & Stevenson, 1985). However, this may not be the case for more experienced performance that does not rely on constant attentional control. Instead, self-regulatory functions may be implementable in parallel with proceduralized control processes, serving to store information about the outcomes and evaluations of performance (rather than the unfolding of their step-by-step components) that is needed for subsequent cognitions about ones' abilities, effort, and strategies for task control (R. Kanfer & Ackerman, 1989; Kluwe, 1987). We propose, then, that self-regulatory attention and skill-focused attention differ in a crucial way: Self-regulatory attention is metacognitive and aimed at the plans that precede skill execution and the products that follow skill execution (Brown, 1987), whereas skill-focused attention is cognitive and aimed at the component steps that constitute execution itself (Beilock & Carr, 2001).

If attention devoted to self-regulation is different from skill focus yet depends on some of the same attentional resources, then self-regulatory activities may provide a secondary, unintended benefit to experienced skill execution: Specifically, self-regulation applied to the plans, the outcomes, and the feelings accompanying performance may prevent individuals from paying too much attention to the step-by-step control of that performance as it unfolds in real time. It may be that individuals involved in self-regulatory functions do not have the resources available to explicitly attend to, monitor, or try to control particular steps or components of online performance—a practice that was shown in the present study to disrupt high-level execution. Thus, although explicit attention to component steps of proceduralized performances may disrupt or dismantle optimal task execution at high levels of learning and proficiency, attentional processes that serve higher level, more metacognitive roles may instead promote optimal performance, both by focusing attention on plans and outcomes and also by preventing attention from focusing on step-by-step control of execution.

Furthermore, skill-focused attention may not always be detrimental to well-learned performances. The present study demonstrated that skill-focused attention applied to current real-time performance disrupts execution. However, if applied in other circumstances, such as practice situations, in which performers are consciously attempting to dismantle their skill and modify certain parts in accord with data collected by self-regulatory activities such as those mentioned above, skill-focused attention may actually be helpful. That is, when the goal is not to maximize real-time performance but instead to explicitly alter or change performance processes to achieve a different outcome, skill-focused attention may be beneficial. In this manner, skill-focused attention may become embedded in the metacognitive activities of self-regulation. Specifically, individuals may attend to specific components of their skill (i.e., implement skill-focused attention) to alter control strategies and execution processes that, through self-regulatory actions, have been deemed unproductive or maladaptive to progress toward a desired goal state. Although this monitoring of performance may be temporarily detrimental to skill execution, as performers will most likely have to slow down and break down previous execution procedures to attend to and alter these processes, and then readapt to and proceduralize the new execution parameters, ultimately these changes should produce performance benefits as skill execution becomes more closely aligned with desired outcomes.

Implications for Skill Training and Performance

Coaches and teachers have long believed that different teaching styles are required at various stages of learning to address the changing attentional mechanisms of the performer. The findings of the present study begin to lend empirical support to this notion. For example, the results of the present study suggest that it may be beneficial to direct performers' attention to step-by-step components of a skill in the early stages of acquisition. This might be achieved through instructions that draw learners' attention to task-relevant kinesthetic or perceptual cues. However, at later stages of performance, this type of attentional control may be detrimental, at least in situations where maximum performance is the desired real-time outcome. In the present study, experienced golfers and soccer players showed decrements in performance under conditions designed to prompt attention to step-by-step execution. Thus, it may be beneficial for experienced individuals to allocate attention to aspects of performance that are not directly involved in the online control of skill execution. McPherson (2000) demonstrated that successful, experienced tennis players spend a significant amount of time examining their own performance outcomes, as well as those of their opponents, as a tool for diagnosing and updating performance strategies and maintaining focus on the task at hand. At higher levels of practice, then, when attention may not be necessary for the online control of performance, such functions as strategizing about choices of actions not only may help to achieve desired goal-states but also may prevent individuals from over attending to well-learned performance processes and procedures.

References

Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin, 102*, 3–27.

Ackerman, P. L., & Cianciolo, A. T. (2000). Cognitive, perceptual-speed, and psychomotor determinants of individual differences during skill acquisition. *Journal of Experimental Psychology: Applied, 6*, 259–290.

Allport, D. A., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single channel hypothesis. *Quarterly Journal of Experimental Psychology, 24*, 225–235.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R. (1993). *Rules of mind*. Hillsdale, NJ: Erlbaum.

Banich, M. T. (1997). *Neuropsychology*. New York: Houghton Mifflin.

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology, 46*, 610–620.

Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General, 130*, 701–725.

Beilock, S. L., Wierenga, S. A., & Carr, T. H. (in press-a). Expertise, attention, and memory in sensorimotor skill execution: Impact of novel task constraints on dual-task performance and episodic memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology.*

Beilock, S. L., Wierenga, S. A., & Carr, T. H. (in press-b). Memory and expertise: What do experienced athletes remember? In J. L. Starkes & K. A. Ericsson (Eds.), *Recent advances in research on sport expertise*. Champaign, IL: Human Kinetics.

Brown, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 65–116). Hillsdale, NJ: Erlbaum.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55–81.

Chi, M. T., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121–152.

Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 17–30.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York. Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159.

Craik, F. M., Govini, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General, 125*, 159–180.

Curran, T., & Keele, S. W. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 189–202.

Day, L. B., & MacNeilage, P. F. (1996). Postural asymmetries and language lateralization in humans (*Homo sapiens*). *Journal of Comparative Psychology, 110*, 88–96.

de Groot, A. (1978). *Thought and choice in chess*. The Hague, the Netherlands: Mouton. (Original work published 1946)

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods, 1*, 170–177.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology, 47*, 273–305.

Fisk, A. D., & Schneider, W. (1984). Memory as a function of attention, level of processing, and automatization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 181–197.

Fitts, P. M., & Posner, M. I. (1967). *Human performance*. Monterey, CA: Brooks/Cole.

Helsen, W. F., & Starkes, J. L. (1999). A multidimensional approach to skilled perception and performance in sport. *Applied Cognitive Psychology, 13*, 1–27.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.

Kanfer, F. H., & Stevenson, M. K. (1985). The effects of self-regulation on concurrent cognitive processing. *Cognitive Therapy and Research, 6*, 667–684.

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude—treatment interaction approach to skill acquisition. *Journal of Applied Psychology, 74*, 657–690.

Keele, S. W., & Summers, J. J. (1976). The structure of motor programs. In G. E. Stelmach (Ed.), *Motor control: Issues and trends* (pp. 109–142). New York: Academic Press.

Kimble, G. A., & Perlmuter, L. C. (1970). The problem of volition. *Psychological Review, 77,* 361–384.

Kluwe, R. H. (1987). Executive decisions and regulation of problem-solving behavior. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 31–64). Hillsdale, NJ: Erlbaum.

Langer, E., & Imber, G. (1979). When practice makes imperfect: Debilitating effects of overlearning. *Journal of Personality and Social Psychology, 37,* 2014–2024.

Leavitt, J. (1979). Cognitive demands of skating and stick handling in ice hockey. *Canadian Journal of Applied Sport Sciences, 4,* 46–55.

Lewis, B., & Linder, D. (1997). Thinking about choking? Attentional processes and paradoxical performance. *Personality and Social Psychology Bulletin, 23,* 937–944.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527.

Logan, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms. *Cognitive Psychology, 22,* 1–35.

Masters, R. S. W. (1992). Knowledge, knerves, and know-how: The role of explicit versus implicit knowledge in the breakdown of a complex motor skill under pressure. *British Journal of Psychology, 83,* 343–358.

Masters, R. S. W., Polman, R. C. J., & Hammond, N. V. (1993). "Reinvestment:" A dimension of personality implicated in skill breakdown under pressure. *Personality & Individual Differences, 14,* 655–666.

McPherson, S. L. (2000). Expert-novice differences in planning strategies during collegiate singles tennis competition. *Journal of Sport and Exercise Psychology, 22,* 39–62.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19,* 1–32.

Peters, M. (1981). Handedness: Effect of prolonged practice on between-hand performance differences. *Neuropsychologia, 19,* 587–590.

Peters, M. (1988). Footedness: Asymmetries in foot preference and skill and neuropsychological assessment of foot movement. *Psychological Bulletin, 103,* 179–192.

Priest, A. G., & Lindsay, R. O. (1992). New light on novice-expert differences in physics problem solving. *British Journal of Psychology, 83,* 389–405.

Proctor, R. W., & Dutta, A. (1995). *Skill acquisition and human performance.* Thousand Oaks, CA: Sage.

Searleman, A. (1980). Subject variables and cerebral organization for language. *Cortex, 16,* 239–254.

Singer, R. N., Lidor, R., & Cauraugh, J. H. (1993). To be aware or not aware? What to think about while learning and performing a motor skill. *The Sport Psychologist, 7,* 19–30.

Smith, M. D., & Chamberlin, C. J. (1992). Effect of adding cognitively demanding tasks on soccer skill performance. *Perceptual and Motor Skills, 75,* 955–961.

Starkes, J. L., & Deakin, J. (1984). Perception in sport: A cognitive approach to skilled performance. In W. F. Straub & J. M. Williams (Eds.), *Cognitive sport psychology* (pp. 115–128). Lansing, MI: Sport Science Associates.

Tenenbaum, G., & Bar-Eli, M. (1993). Decision-making in sport: A cognitive perspective. In R. N. Singer, M. Murphey, & L. K. Tennant (Eds.), *Handbook of research on sport psychology* (pp. 171–192). New York: Macmillan.

Wickens, C. D. (1980). The structure of processing resources. In R. Nickerson (Ed.), *Attention and performance VII* (pp. 239–257). Hillsdale, NJ: Erlbaum.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 63–102). Orlando, FL: Academic Press.

Wulf, G., Hob, M., & Prinz, W. (1998). Instructions for motor learning: Differential effects of internal versus external focus of attention. *Journal of Motor Behavior, 30,* 169–179.

Wulf, G., McNevin, N. H., Fuchs, T., Ritter, F., & Toole, T. (2000). Attentional focus in complex skill learning. *Research Quarterly for Exercise and Sport, 71,* 229–239.

Received March 2, 2001
Revision received August 15, 2001
Accepted August 20, 2001 ■