

THE ROLE OF EVIDENCE IN ANIMAL WELFARE SCIENCE AND STANDARDS:
AN ETHICAL ANALYSIS

By

Monica List

A DISSERTATION

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

Philosophy- Doctor of Philosophy

2019

ABSTRACT

THE ROLE OF EVIDENCE IN ANIMAL WELFARE SCIENCE AND STANDARDS: AN ETHICAL ANALYSIS

By

Monica List

The development of the field of animal welfare science in recent decades, together with growth in public interest and ethical concerns for the welfare of animals has arguably led to an increase in regulation, both public and private, in the production and use of farmed animals. Animal welfare regulation, in the form of legislation, directives, industry guidelines and private standards is characterized by its reliance on scientific evidence to justify the conditions under which farmed animals should be bred, raised, transported, and slaughtered. One of the important roles scientific evidence plays in discussions around the regulation of farmed animal welfare is providing a seemingly ethically neutral understanding of how animals should be treated, in other words, it sidesteps ethical arguments for the treatment of animals, using science as a justification. However, per the dominant philosophy of science discourse on the role of values in science, no science can be considered value-free, and furthermore, there are acceptable roles for social and ethical values in scientific practice. These roles are not just acceptable, but necessary for the direction, interpretation, and application of science. This work argues that: a) given the broad range of ethical views regarding how we should treat animals, animal welfare science provides robust and credible guidance; b) furthermore, that animal welfare science is not only inspired or informed by animal ethics concepts and frameworks, but also has embedded social/ethical and cognitive values throughout; c) thus, in order to fulfil its purpose as a socially mandated science, animal welfare science should engage in intentional processes to determine adequate roles for

various kinds of values underlying all stages of the scientific process and the interpretation and implementation of findings. This intentional examination of the role of values can be supported by more effective interdisciplinary collaboration. While animal welfare science is characterized in part by its interdisciplinary nature, it is important to question to what extent the research is truly interdisciplinary in the sense of fostering epistemic integration. Philosophical tools and analyses, beyond the typical uses of ethical frameworks as a starting point, can be valuable in facilitating effective interdisciplinary work, leading to a better understanding of the normative dimensions of animal welfare science.

Copyright by
MONICA LIST
2019

This work is dedicated to the women in my life who inspired me to think that I could do this, and to Pablo, for always reminding me that I should.

ACKNOWLEDGMENTS

I have many people to thank for their support throughout the many years it took me to complete this process. To my family: Pablo, Beatriz, Max, and Santi, as well as my parents; I am grateful for your willingness to literally follow me (across many miles!) into this process, and for your continued confidence in me even when circumstances were less than ideal. Thank you for believing in me; I have no words to express how happy I am to be able to share this success with you.

I was fortunate to have a wonderful, interdisciplinary dissertation committee. My dissertation director, Paul Thompson, made it possible for me to come to Michigan State University, and was always supportive in ways that went beyond his academic responsibilities; I am forever grateful. I have Linda Kalof to thank for my participation in the Animal Studies graduate specialization, which provided a unique academic home and support network during my time at Michigan State. I am also grateful to committee members Janice Siegford, Kyle Whyte, and Fred Gifford; thank you for your mentorship, and especially for being open-minded and always supportive of my interdisciplinary interests.

To mentors, friends, and colleagues at the Center for Ethics and Humanities in the Life Sciences, especially Tom Tomlinson and Libby Bogdan-Lovis, thank you for providing me with a rich intellectual environment, a much-valued work space that I enjoyed for several years, and many wonderful opportunities to develop skills beyond the scope of my doctoral program. I am also grateful to many other mentors and colleagues who made it possible for me to partake in unique interdisciplinary projects that informed this project; many thanks in particular to Laurie Thorp, Julie Funk, Maria Lapinski, and Michael O'Rourke.

I am also grateful to fellow graduate students who inspired me, supported me, made me laugh and sometimes cried with me throughout these years. In particular, I wish to thank my Philosophy Department colleagues Samantha Noll and Esme Murdock, and my Animal Studies cohort. Finally, I wish to thank my colleagues at Compassion in World Farming and World Animal Protection, for allowing me the space to complete this work while holding full-time jobs with them, and for helping to make sure that my work was always grounded in the goal of making the world better for animals and humans alike.

TABLE OF CONTENTS

Chapter 1: Introduction: definitions, historical context and outline.....	1
1.1 Introduction and definitions	1
1.2 Historical context	7
1.2.1 Animal ethics	8
1.2.2 Animal welfare science	17
1.3 Contribution to the literature and scope of the project.....	26
1.4 Project outline	27
Chapter 2: Animal welfare as science: conceptual frameworks, research areas, and evidence	30
2.1 Introduction.....	30
2.1.1 Functional areas in animal welfare science.....	31
2.2 Conceptual frameworks	40
2.2.1 The three pillars framework	43
2.2.2 Needs and wants framework	45
2.2.3 The five domains model.....	48
2.3 Types of evidence and indicators of animal welfare.....	50
2.3.1 Physiological measures	50
2.3.2 Ethological evidence	53
2.4 Interpreting and applying evidence	56
2.5 Conclusions.....	59
Chapter 3: Evidence and values in animal welfare science	62
3.1 Introduction.....	62
3.2 Evidence: definitions and uses	63
3.3 Limitations of evidence: underdetermination	66
3.4 Limitations of evidence: inductive risk.....	67
3.5 The role of values in science: an overview	69
3.5.1 Types of values	69
3.5.2 Epistemic, cognitive, non-epistemic, and social/ethical values	70
3.5.3 Roles of values in science	71
3.6 Values in animal welfare science	73
3.7 Conclusion	75
Chapter 4: Standards, science, and ethics	76
4.1 Introduction.....	76
4.2 Classification of standards	77
4.3 Standards and their ecosystems.....	81
4.4 Private farm animal welfare standards in the United States: motivations and development	85
4.5 The role of animal welfare concepts in the development of standards	87
4.6 Ethical and scientific justifications for animal welfare standards.....	89
4.7 Conclusions.....	90
Chapter 5: Practical perspectives on the development and implementation of animal welfare standards..	92
5.1 Introduction.....	92
5.2 Methods.....	93
5.3 Examples of private animal welfare standards in the US	95
5.3.1 United Egg Producers Guidelines for Commercial Laying Hens	95

5.3.2 Global Animal Partnership Standard for Chickens Raised for Meat	97
5.4 Interviews.....	100
5.4.1 Background of interviewees and organizations	100
5.4.2 Personal experience and views on farm animal welfare standards	103
5.4.3 Perceptions on the role of standards and animal welfare certifications	105
5.4.4 Perceptions of animal welfare science	111
5.4.5 Science, standards, and public trust	113
5.5 Discussion and conclusions	115
Chapter 6: Conclusions and recommendations	118
6.1 Conclusions.....	118
6.2 Practical recommendations and future research directions	121
APPENDICES	124
Appendix I: Interview guide	125
Appendix II: Informed consent template	127
BIBLIOGRAPHY	130

Chapter 1: Introduction: definitions, historical context and outline

1.1 Introduction and definitions

This project presents a critical analysis of the definition, scope, and uses of scientific evidence in the development and application of animal welfare standards and policies. It examines the specific kinds of evidence used in animal welfare science and traces the path of this evidence from laboratories and field trials to supermarket shelves, and ultimately, the plates of consumers. The project draws from the field of philosophy of science in order to examine the concept of evidence, including definitions, theoretical frameworks, interpretation, and roles of evidence in the broader context of contemporary Western science. However, the ultimate goal of this work is to draw conclusions and make recommendations on the use of scientific evidence in the justification of animal welfare standards, policies, and practices, thus it should be understood as a contribution to the field of practical ethics. The objectives of this dissertation are the following: 1) to conduct a detailed documentation and analysis of the history and roles of evidence in animal welfare science; 2) to critically examine these uses of evidence through the lens of contemporary views on this topic from the field of philosophy of science as well as ethical frameworks; 3) to describe how this evidence is used to substantiate standards and policy, and answer whether or not its use as an ethical justification or motivation is sufficient; 4) to draw on personal experiences that relate to the previous points in order to support a discussion of practical implications of the central issues addressed in this work.

Building on these objectives, a central premise of this thesis is that while scientific evidence plays an important role in justifying ethical views and guiding normative action, in most cases it is a necessary but not sufficient condition of such justifications. On its own, this premise may seem unsurprising and unsophisticated to philosophers, who would generally agree

that you cannot derive “ought” from “is”, in other words, that a satisfactory factual description of something (such as a robust body of evidence) is not enough to morally justify a certain course of action or state of affairs.¹ At the same time, scientists who uphold some version of the value-free ideal of science would push back against the notion that robust evidence is insufficient to justify normative action and tools such as standards. Thus, a more well-rounded argument attending to both of these views is the following: a) that given the broad range of ethical views on the treatment of animals, together with a lack of regulatory mechanisms, animal welfare science provides robust and credible guidance; b) that animal welfare science is not only inspired or informed by animal ethics concepts and frameworks, but also has embedded social/ethical and cognitive values throughout; c) thus, in order to fulfil its purpose as a socially mandated science, animal welfare science should engage in intentional processes to determine adequate roles for various kinds of values underlying all stages of the scientific process and the interpretation and implementation of findings. Premise “a” is developed in sections 1.2.1 and 1.2.2 of this chapter, which provide a broad overview of common animal ethics frameworks, and of the role that animal welfare science plays in providing justification for better treatment of animals in the face of divergent ethical views. Premise “b” is expanded upon in Chapter 3, which provides theoretical background on the roles of values in science generally, and in animal welfare science particularly. Regarding the conclusion, “c”, the intentional processes by which the role of values in animal welfare science are should be acknowledged are illustrated by the interviews reported

¹ In contemporary Western philosophy, the “ought-is fallacy”, also known as Hume’s law or Hume’s guillotine derives from David Hume’s *Treatise of Human Nature* (Book III, part I, section I), where he observes that the distinction between vice and virtue (or what is right and wrong) should not be founded simply on reason, or observed relations (e.g. causal relationships) between objects. A standard contemporary interpretation of this is that non-moral premises cannot entail a moral conclusion (MacIntyre 1959).

in Chapter 5, and discussed in Chapters 5 and 6. Specific recommendations for improving these processes are provided in Chapter 6.

Although this work analyzes the use of evidence in animal welfare science and standards generally, the focus of discussion and most examples used pertain to farmed animals. This is because farmed animals make up the majority of land animals used by humans, and accordingly, many animal welfare standards developed and used thus far apply to animals used for food or fiber. In the United States alone, a record production of beef and pork was reported for 2016, with 30.6 million head of cattle and 118.2 million pigs slaughtered at federally inspected plants (USDA 2017). Meanwhile, the availability and popularity of chicken has consistently increased in the United States since 1970; according to the National Chicken Council (NCC), in 2018, Americans consumed on average 93.5 pounds of chicken per capita, versus 56.5 pounds of beef, and 50.1 pounds of pork (Bentley 2017; NCC 2019). In addition to sheer numbers, farmed animals also constitute a priority because they are increasingly at the center of public controversies regarding whether and how we continue to raise them for food, and if so, how to do so in socially and environmentally responsible ways. Finally, regardless of particular moral stances on the production and consumption of animals, there are practical questions and concerns regarding their welfare in industrialized farming systems. Evidence produced by animal welfare science and its interpretation and application through policy and standards contributes to answering these questions.

While concerns over the welfare of farmed animals often make their appearance in the public sphere as dramatic media features portraying undercover investigations by animal activist groups, they are also one of the main drivers of animal welfare research, which is, in turn, one of the sources of information used by standards bodies and policymakers to define animal welfare

priorities, targets, and the methods to measure and track them. A preliminary search of academic publications featuring articles within the scope of animal welfare science conducted for the purposes of this research found twenty-two journals, some of which are specifically focused on animal welfare, while others regularly include publications on variants of the topic but have a broader scope. Within these journals, publications on farmed animal welfare are by far more frequent than publications on companion animals or other species. For example, in 2017, the Journal of Applied Animal Behaviour Science, official journal of the International Society for Applied Ethology (ISAE), published 73 articles pertaining to species commonly known as farmed or food animals, and only 37 articles on companion animal welfare^{2,3}. This proportion of articles focused on animal welfare is hardly surprising because of factors previously mentioned, as well as the economic importance of animal agriculture. Overall, agriculture and its related industries provide 11% of jobs in the US; in the manufacturing sector, 31% of jobs are in the meat and poultry sectors (USDA ERS 2017). In 2017, the US poultry industry alone grossed an approximate \$26 billion in sales (this included fresh, frozen, and processed chicken and other poultry) (Mintel Group Ltd. 2017). In a report focusing on economic impact, the National Chicken Council of the United States estimates the total economic impact of the US poultry industry at close to 347 billion US dollars, from which wages account for 76.5 billion US dollars (National Chicken Council and U.S. Poultry and Egg Association 2019).

However, scientific interest in animal welfare is not only driven by economic motivations. Social interest in the welfare of animals, fueled by an increasing media presence of

² These results were obtained from a Web of Science query using the following keywords and operators: Query set 1: (PY=(2017) AND TS=(poultry OR chicken OR turkey OR pig OR pork OR cattle OR beef OR dairy OR hen OR egg OR pullet OR lamb OR goat OR calf OR calves) AND IS=(01681591)) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article). Query set 2: (PY=(2017) AND TS=(dog OR cat OR companion) AND IS=(01681591)) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article)

³ It is also true that fewer species are considered companion animals; this is generally limited to dogs and cats.

animal welfare topics is also an important influence on research agendas. Media coverage and other publicly available information play a central role in shaping public attitudes toward farmed animal welfare specifically, especially since the public is increasingly disconnected from food production and farming. The physical and epistemic distancing of consumers from the sources of their food results in animal welfare being a credence attribute, in other words, an attribute that consumers cannot directly assess when forming an opinion of an animal product or practice, but rather one that depends on proxies such as labeling and certifications (Tonsor and Olynk 2011). Thus, increased media presence of animal welfare, as well as this inability to directly assess the welfare of animals used in food production influenced consumers' growing interest in these topics, which in turn motivates researchers in both industry and academia to produce evidence to support, question, or improve practices and methods of production.

As regards definitions, throughout this work, I will be frequently using several terms, defined as follows. First, as is customary with works on the topic of animals, it is necessary to clarify the use of “animal(s)” as an abbreviation of “non-human animals”; furthermore, due to the scope of this work, “animals” will mainly refer to higher order vertebrates, and more specifically, those commonly used for food and fiber production in the West. Second, “animal welfare”, refers specifically to the accepted scientific understanding of the concept, in summary, “the state of an individual as regards its attempts to cope with its environment”, where “coping” involves having control of mental and body stability, and “state” refers generally to the physical and mental state of the animal (Broom 1986; Broom and Johnson 1993).

I also refer to “the animal welfare movement”, broadly understood as individuals and organizations (including charities, government organizations and professional bodies) who have an interest in advancing the cause of animal welfare. These individuals and organizations operate

within the movement with different objectives, ranging from observational and advisory roles to lobbying and grassroots activism. Often, those who consider themselves part of the animal welfare movement are keen to make a distinction between animal welfare and animal rights activism. In some respects, it is an important distinction to make; from a philosophical perspective, there are relevant differences between the two camps in terms of what moral motivations and ethical theories substantiate each one, with animal rights generally characterized as relying on deontological and contractarian views, and animal welfare being markedly utilitarian⁴ (Sumner 1988). In practice, many organizations (and individuals) who identify with animal rights do indeed hold more radical views with respect to our moral obligations to animals, and the rights that they have (or should have) in our societies. However, it is also true that many individuals who identify one way or another may have little or no knowledge of these theoretical differences, and often alternate between rights views and welfarism, depending on the issue at hand. Furthermore, organizations who consider themselves to be in distinct camps will often work together to put pressure on governments and/or industry in order to effect change for animals. Having said this, within the scope of this work, I will generally consider “the animal welfare movement” as a social movement with the goal of improving the lives of animals, while maintaining open the possibility of using them in a responsible and ethical way for the benefit of human populations, be it for companionship, food, fiber, or otherwise. Finally, I also use the term “animal welfare science”, which refers to the community of scientists doing research to advance scientific knowledge of animal welfare, and to the body of work resulting from this.

⁴ These ethical theories are explained in further detail in section 1.2.1

1.2 Historical context

The histories of animal ethics and animal welfare are intertwined; the view that we have ever expanding moral obligations toward animals has been an important driver for the emergence of animal welfare science as a distinct area of scientific inquiry. Likewise, the rapid growth and new findings of animal welfare science are constantly refreshing and refining our ethical perspectives on animals, as well as other areas of animal philosophy, such as metaphysics and epistemology.

The idea that animals are, or should be, morally considerable is at the heart of animal ethics. It is important to clarify, however, that the discussion of whether animals have moral rights is a different question. In order to assert that we have moral obligations toward animals (or any other entity), it is only necessary to make the case that they are worthy of moral consideration. As Goodpaster (1978) clarifies, *moral considerability* is distinct from *moral significance*, as the latter entails an understanding of moral worth either in a certain context, or in comparison to another being or category of beings. Moral considerability, on the other hand, entails only that the issue raises moral questions. Moral considerability is, in this broad sense, arguably one of the drivers of a socially mandated science, such as animal welfare. While science motivated by curiosity or by practical needs may be seemingly detached from value considerations, a mandated science in fact emerges *from* these value considerations, and is framed and driven by public concerns (Fraser 2008). These public concerns are grounded in moral considerability, but not necessarily in beliefs that animals have a particular moral standing, or moral rights. Regardless, the different ethical theories arguing for the moral significance or moral rights of animals often inform and permeate the kinds of questions that animal welfare science attempts to answer, and as such, it is useful to provide a summary.

1.2.1 Animal ethics

Animal ethics theories can be broadly classified into two groups based on the kind of ethical justifications that ground them. The first group of theories I will refer to as “principle-centered”. This group includes theories grounded in categorical ethical principles that dictate morally required actions regarding a particular group of beings (in this case, groups of animals delimited by specific criteria). Well-known examples of this group of theories are Singer’s utilitarian account, Regan’s animal rights position, and Nussbaum’s capabilities approach. The second group of theories I will refer to as “contextual”; representative examples include Clare Palmer’s contextual animal ethics, Josephine Donovan and Carol Adams’ feminist care ethics, and pragmatist animal ethics as described by Light, McKenna and Thompson. While principle-centered theories entail *categorical* moral obligations to animals, theories in the second group generally entail *conditional* moral obligations.

For principle-centered theories, the moral obligation to apply universal principles arises from one or more morally relevant inherent attributes of individuals included in a category to which the principle applies. In this sense, it can be said that this set of theories is not only principle-centered, but individualist. Individual attributes, such as consciousness, the ability to reason, self-motivated activity, a capacity for communication analogous to human language, and evidence for a concept of self and self-awareness have been used to ground and advance theories in this group (Bekoff and Gruen 1993). These individual attributes are often referred to as *anthropocentric*, because they have been extensively studied and used in humans to justify their intellectual, moral, and even legal status. However, non-anthropocentric attributes, such as genetically and environmentally evolved capacities unique to certain species, have also been used to justify the moral standing of some animals (Thompson 2003). Non-anthropocentric attributes are those that signal that animals have interests and have species-specific

characteristics or behaviors geared toward the fulfilment of needs and realization of their individual potential, without regard for the existence of analogous characteristics in humans (Rutgers and Heeger 1999). Examples of this may be forms of complex communication that are non-analogous to human concepts of language, such as communication involving pheromones, ultrasonic waves, and body placement (for instance, starling murmurations).

Peter Singer's utilitarian proposal is perhaps the most widely known ethical theory in Western contemporary animal ethics, and is broadly representative of other utilitarian views in animal ethics. It is generally accepted that within utilitarian frameworks like Singer's, moral traction comes from the weighing of actions, their consequences, and utility, where the interests of all morally considerable parties have equal weight. Since this notion of equal interests is a condition for moral considerability, it is there that the normative principle of the framework lies. Singer claims that the reason why we should be morally compelled to act in a way that maximizes good outcomes for humans and (some) animals alike is because we are equal in a way that is morally relevant, that is, we are both the kinds of beings that have commensurable interests (Singer 2009, 2; 5). These commensurable interests, in turn, entail moral obligations; we should ensure that whatever particular interests are central to the good life of an individual are equally considered (Singer 2009, 6). To clarify, the theory does not entail that all beings have the same kinds of interests, and therefore require equal treatment; it simply states that morally relevant interests have equal weight and we should therefore strive to meet them without regard for what Singer claims are trivial attributes such as species, race, or gender, among others (Singer 2009, 2;6). It is important to point out that although Singer proposes a universal application of this principle of interests-based equality, he limits the scope of his moral circle to those beings considered sentient (possessing the capacity to feel pain and pleasure) (Singer 2009,

7). He is somewhat vague about where the line for sentience is drawn; while he has previously claimed that perhaps this is “somewhere between a shrimp and an oyster”, he also defends that since we really have no need to eat or otherwise harm these animals, it is best to exercise caution (Singer 2009, 174).

A second group of principle-centered theories relies on Kantian or deontological principles, in which moral considerability, grounded in the inherent worth of individuals, entails imperative moral obligations or duties. While animal ethicists in this group would agree that both humans and animals have the capacity to have good lives, they assert that what is good for one individual may be very different from what is good for another (even within the same species), so not only is it impossible to weigh different goods (or interests) against each other, but it is morally unacceptable to try to aggregate goods and weigh them against harms, as utilitarians would (Korsgaard 2011, 95-96). Therefore, while Kantian or deontological positions also rely on universal principles, they reject the claim that the grounds for equality between humans and animals (and corresponding moral obligations) is that we possess a comparable set of interests.

Tom Regan, whose rights-based theory is grounded in deontological principles, proposes that the moral obligation to treat animals as ends in themselves is based on the *subject of a life* criterion (emphasis mine) (Regan 2004, 243). To be a subject of a life is not merely to be alive and conscious, but to have, among other things: beliefs⁵ and desires, memory, a sense of the future, a rich emotional life, a psychophysical identity over time, and individual welfare (Regan 2004, 243). Note that, according to this description not all animals would be considered subjects

⁵ Although Regan does not define “beliefs” in this context, it is plausible he is referring to basic beliefs, which entail having a basic concept of something (for instance, a tree), and then, based on observation or perception, being able to identify instances of trees. More complex forms of belief (higher order) are also possible in animals with complex cognitive skills; higher order beliefs involve, for instance, making inferences about something based on contextual or historical information. In animals, complex puzzle solving could be considered an example of this.

of a life (Regan 2004, 245). Like Singer, Regan draws a tentative line; he considers that all normal mammals over the age of one⁶, as well as humans who cannot be considered moral agents but who are like these animals in all relevant aspects should be considered *moral patients* (Regan 2004, 239). The concept of moral patient is central to Regan's argument. He sees humans as part of a broader category of beings with whom we share something morally relevant: inherent value (Regan 2004, 236). In this case, then, it is inherent value, ascribed to all moral patients and agents, and not reason, which gives normative traction to the Kantian mandate to treat all who fulfill the relevant criteria described by Regan as ends in themselves.

A third example of principled-centered theories is Martha Nussbaum's capabilities approach, which shares some elements with the deontological position, such as the obligation to ensure the dignity of forms of life with relevant needs and abilities by properly enabling them through a variety of activities that lead to a worthy life (Nussbaum 2004, 305). Despite these commonalities, Nussbaum clearly distances her account from the deontological position, claiming that moral worth should not *only* be determined by morally relevant criteria or attributes (be it humanity, rationality, or being a subject of a life), but that it is also important to consider the subject's potential to flourish; in other words, it's not just being a subject of a life, but being a subject of a life *potentially* worth living that matters (Nussbaum 2004, 306). Nussbaum's theory is also different from others in this group in that in addition to advancing a universal moral principle, it cashes out that principle through a series of political commitments based on the list of capabilities that she believes are central to flourishing (Nussbaum 2004, 313-317). For

⁶ This distinction is arbitrary in the sense that Regan provides no convincing argument for drawing the line at "all mammals over 1". Arguably, he could be hinting at certain levels of cognitive complexity or cognitive and social skills present in most mammals at this age. However, it is still a very broad generalization considered all of the species included in the class Mammalia, and the different rates at which they mature and develop cognitive skills, not to mention the range of possible skills across different mammalian species.

animals, this list of basic capabilities (which she extrapolates from her prior work on capabilities in humans) includes: an entitlement to live out a natural life-span, bodily health, bodily integrity, facilitating the use of the senses, imagination and thought⁷ (where these capabilities exist and are a conscious or apparent interest), a satisfactory emotional life, affiliation (not limited to a social life with conspecifics, but to be integrated and seriously considered within the broader social, cultural, and legal landscape), and control over one's environment (which entails choice of environment) (Nussbaum 2004, 314-317). Although she does not elaborate on how capabilities and their associated entitlements are ultimately assessed, behavioral and other observable indicators could plausibly be developed for each.

In contrast to principle-centered theories, contextual animal theories do not rely on a singular moral principle. In fact, many try to incorporate a variety of principles in order to better navigate the moral complexities of human-animal relations. A second important feature of this group is, as the name suggests, the importance of context. Here I refer to context in a broad sense that includes physical as well as social and cultural aspects of the environments in which humans and animals interact. Generally speaking, these theories posit that humans have a moral responsibility toward animals because of the kinds of relationships that we have with them; however, this claim is not meant to be understood as a universal principle of the sort "humans have moral obligations to all animals that they have X kind of relationships with", precisely because context may, in some cases, overrule that sort of principle. A relationship in this case refers not only to affective or emotional relationships (such as one would have with a companion animal), but any kind of interaction that is likely to have a morally significant effect on any of the parties involved, whether reciprocal or not (Palmer 2010, 49). A third feature that makes

⁷ Nussbaum lists play as a separate entitlement, but it could arguably be included within this category.

these theories philosophically distinct from the first group is that there is no singular source or locus of normative traction, furthermore, these theories may rely on a mix of intrinsic and instrumental values or eschew this distinction altogether.

Palmer's contextual ethics position is based on the moral relevance of our relations with animals, including causal relations (where human actions have caused animals to be in a certain situation, for example relocated wild animals), and pseudo-contractual relations, such as those we have with companion and farmed animals (2010, 55-57). Given the many difficulties of developing a relational ethics account based on contract-driven obligations in the case of human-animal relations, Palmer opts to develop a view based on causal relations. In order to argue for this view, she compares two groups of animals; domesticated animals, to whom we owe care and assistance, and wild animals, to whom she claims we have no such obligations (2010, 63). Although Palmer does not elaborate on this, the manner in which an animal is domesticated, or our purpose for doing so *may* matter, as the domestication process ultimately shapes the relationships that we have with domestic animals, and therefore our potential obligations. She discusses in detail the different kinds of duties that these moral obligations entail, for example, positive duties (duties to assist) versus negative duties (duties to not harm or not interfere) (2010, 68). The distinction between these two kinds of duties, and questions regarding which groups of animals we owe these duties to lead Palmer to develop several versions of an argument which she calls the *laissez-faire intuition* (LFI), which in summary proposes that we have positive duties toward some animals but not others, based on the kinds of relationships we have with them (2010, 63;68). A refined iteration of the LFI argument is the no-contact LFI, which entails positive duties to animals with whom we maintain morally relevant forms of contact, but only negative duties to animals with whom we have limited or no contact, and who are better off

without any form of human intervention (thus the “no contact” qualification) (2010, 90).

Furthermore, obligations to fulfill both negative and positive duties may be informed by past actions or relations that have either violated negative duties or created vulnerabilities (domestication and relocation of wild animals are examples) (Palmer 2010, 90).

A second cluster of theories within the contextual group falls under the label of Western feminist care ethics. Initially developed in the mid-1980s, its origins are attributed to Carol Gilligan’s *In a Different Voice*, published in 1982 (Donovan and Adams 2007, 2). This group of theories emerged as a rejection of patriarchal concepts of morality based on rights, rules, agency, and rationality (Donovan and Adams 2007, 2). Consistent with other contextual theories, feminist care animal ethics rejects abstract or rule- based principles, and is rather guided by a series of key concepts and research directions (Donovan and Adams 2007, 2). Among these are the attention to individual suffering in the context of the structural underpinnings that perpetuate it, demanding not only moral commitments but political action, and a need to open different forms of dialogue between humans and animals in order to shift the epistemic source of theorizing about animals to the animals themselves (Donovan 2007 and Adams, 3; 361). As far as specific moral commitments of feminist care ethics, Palmer points out that while emotional closeness and a particular kind of attentiveness can tell us that these beings matter morally, it gives little guidance as to specific moral obligations (2010, 52). Donovan admits that the specific moral obligations entailed by a feminist care ethics can be somewhat vague, nevertheless she also claims that at least for her work, the purpose is not to lay out a detailed practical ethic, but rather to reorient how we think (and know) about the morality of human-animal relationships (2007, 76). Nevertheless, I believe it is safe to say that most scholars in this vein would agree with Palmer that some *prima facie* principles, such as the duty to cause no harm must be upheld;

but in this case, the moral reasons for doing so are found in an improved understanding of the moral worth of animals qua animals, and not necessarily by way of expanding the human moral circle, or extrapolating the foundations of human moral considerability to animals. As regards the use of animals for food or fiber, feminist care ethics does not broadly or categorically rule out the permissibility of using animals in these ways under certain circumstances, however, this group of theories generally makes strong associations between the oppression of women (and other marginalized humans) and animals, where consumption in the form of meat or fiber is considered an ultimate form of objectification and violence.⁸

The final example I will consider in this group, pragmatist animal ethics, is perhaps the most theoretically distant from the principle-centered theories, as it expressly rejects the foundation of moral commitments on universal principles (in contrast to contextual theories that admit a limited role for *prima facie* moral commitments), and does not aim to propose a set of normative premises that can be uniformly applied to either kinds of animals or situations involving animals (Thompson 2004, 147). Pragmatism shares some common points with principle-centered theories, for example a rejection of the privileging of human interests and suffering, but it does so based on a practical, experience-centered understanding of why we should not readily accept human privilege (McKenna and Light 2004, 10). A second important aspect that informs the moral commitments of pragmatism is the rejection of a metaphysical/ethical divide, in other words, the tendency to consider ontological and ethical questions separately, and often inconsistently (McKenna and Light 2004, 10). Pragmatism, as a pluralistic,

⁸ Although I claim that feminist ethics as a whole does not categorically prohibit the killing and consumption of animals for human benefit, some of the most influential works in Western feminist animal ethics are explicitly against these practices. Notably, the work of Carol Adams and Josephine Donovan makes strong connections between violence against women and violence against animals (including their consumption), condemning both. Well known Western ecofeminists such as Marti Kheel and Val Plumwood hold similar views against the use of animals solely for human benefit.

fallibilistic, and flexible philosophical approach has the potential to better adapt the ethical questions we ask to the rapidly changing circumstances of our relationships and interactions with animals (McKenna and Light 2004, 9); many of these changes are related to the expansion of our moral consideration to species other than humans, or to our changing relationships. McKenna and Light point out that while there has been a progressive expansion of the human moral circle, it is implausible to think that this is some sort of natural progression; rather, they claim, it has expanded as we have gained a better understanding of others' interests, and how and why they matter (2004, 2-3).

As a decidedly contextual theory, pragmatist ethics also pays attention to the power dynamics of the systems that ethical decision-making takes place in. For example, Thompson points to the role that science and technology have played (and continue to play) in shaping our views of animals and their moral standing by reflecting the interests of actors or networks of actors that are more powerful or successful, such as the industrial farming complex (2004,150). In this sense, a pragmatist animal ethics aims to center the role of specific public debates over more general social issues, and to use this information to propose specific ethical solutions in response to problems rather than propose principle-based frameworks (McKenna and Light, 2004,2).

While some theories in both the principle-centered and contextual groups incorporate empirical evidence of the attributes used to justify moral obligations, others propose that even in the absence of evidence, a reasonable doubt with regard to the fact that these animals may possess these attributes is enough to justify moral obligations towards them (Bekoff and Gruen 1993; Singer 2009, 174). This mirrors a tension that also exists in the realm of behavioral science, which, not surprisingly, is the source for much of the evidence for morally relevant

attributes in animals. The principle of parsimony, which is at the root of well-known scientific canons such as Occam's razor and Morgan's canon, demands that natural phenomena, including animal behavior, be explained in the simplest terms possible, and making as few assumptions as possible (de Waal 2006). For animal behavior scientists, however, the principle can be interpreted in at least two different ways; *cognitive parsimony* involves favoring explanations of animal behavior that rely on lower order mental capacities, while *evolutionary parsimony* considers shared phylogeny in crafting these explanations, seeking consistency in explaining similar behaviors in different species (de Waal 2006, 62).

1.2.2 Animal welfare science

Regardless of whether they incorporate empirical evidence for moral considerability, moral arguments alone may not have enough traction to convince those who have little moral consideration for animals to begin with, and because of this, appeals for the improved treatment of animals have long relied on a variety tactics, which include both moral arguments and scientific opinions. In this sense, philosophy and science play distinct and complementary roles. However, using both ethical theories and scientific evidence to make arguments for social change can lead to confusion about the scope and roles of ethics and science, and ultimately skepticism regarding the capacity of animal welfare science to deliver robust scientific evidence that can be used to substantiate policy.

A relevant difference between scientific and philosophical views of animal welfare is that the former are often characterized as strictly descriptive (although this can be debated), while the latter are decidedly normative. For example, one of the first scientific definitions of animal welfare is the following: "the welfare of an individual is its state as regards its attempts to cope with its environment" (Broom 1986). Per this definition, coping includes behavioral,

immunological, and physiological mechanisms, all of which can be scientifically assessed to some extent. However, the term “cope”, specifically in the context of animal welfare science has no precise definition, and as such, understandings of how well we can assess coping hinge on particular definitions. Definitions that align with Broom’s remain true to the claim that animal welfare is a strictly objective scientific concept, and as such, its assessment should not be influenced by moral considerations (although the outcomes of its assessment can be morally relevant) (Fraser 2008, 261; Broom 2014, 27-28).

Broom’s definition can be characterized as a functional definition of animal welfare; in other words, animal welfare understood as a function of an organism’s ability to cope. However, definitions of animal welfare have evolved to include emotional or affective factors that are also central to the coping response, and to welfare as a whole. If we consider scientific definitions of animal welfare as a spectrum, definitions akin to Broom’s would be at one end of a spectrum favoring function, while at the opposite end we would have definitions that describe welfare as a state mainly defined by an animal’s affective states or feelings (such as pain, frustration, fear, and distress, among others). Marian Dawkins and Ian Duncan both contend that affect plays a major role in defining an animal’s welfare, and may even be the most important element (Keeling et al 2011, 18). Regardless of where scientific definitions fall along this spectrum, there is general agreement among animal welfare scientists that animal welfare is a complex concept informed by objective measures of health and functioning, and comparatively more subjective assessments of affective states and behaviors (Fraser 2008, 222).

The most common origin stories of animal welfare science in the West refer to the 1964 publication of Ruth Harrison’s *Animal Machines* as a pivotal moment at which social concern for animals, long relegated to small groups of concerned private citizens and occasional public

outcry over extreme cruelty, finally made it to influential political and scientific circles.

Harrison's work was indeed influential; it is believed that her book and its public reception are what prompted the British government to establish the Brambell Committee in 1965 (Broom 2014). The Committee was so named after its chairperson, Professor F. W. Rogers Brambell, an immunologist; however, most influential in developing the foundations that would continue to shape animal welfare science and policies for years to come was Dr. William Thorpe, an animal behaviorist (Fraser 2008).

The Brambell Committee was tasked with investigating the conditions of animals in intensive husbandry systems, and in 1965 delivered on this task in the report entitled *Report of the Technical Committee to enquire into the Welfare of Animals kept under Intensive Livestock Husbandry Systems* (henceforth "the Brambell Report"). In addition to providing the groundwork for what would later be known as "The Five Freedoms" (described later in this chapter), the Brambell Report was perhaps the first widely known public document to use the term *animal welfare* in its contemporary sense. Interestingly, the Brambell Report provides no definition of the concept, which, according to Broom (2011) would not be defined in the academic literature until the early 1980s. At around the same time, animal welfare started to be thought of as a distinct field of scientific inquiry, as evidenced by Marian Stamp Dawkins' 1980 publication *Animal Suffering: The Science of Animal Welfare*.

Dawkins' 1980 publication is one of the earliest works addressing a topic that would preoccupy animal welfare scientists for years to come: developing methods for the objective observation and assessment of animal welfare. She recognizes that one of the main challenges of the emerging science of animal welfare is to move beyond its origins in social concerns for animals, and establish itself as a field of inquiry grounded in the methods of natural science

(Dawkins 1980). Per Jeremy Bentham's well-known quote⁹, the notion that animals are capable of suffering should be enough to motivate us to give them moral consideration, but, as Dawkins points out, there is much disagreement about what suffering is, which animals are capable of it, how much suffering is tolerable, and a myriad of related questions that can only be consistently and meaningfully answered if there is also agreement on the methods used to answer them (Dawkins 1980). These same concerns were shared by William Thorpe in the appendix to the Brambell report, where he writes "(...) If we consider that certain treatments may cause pain or suffering, we then have to examine more carefully what we mean by these words" (Brambell 1965).

The debatable meanings of suffering, sentience, pain, naturalness, and many other terms, as well as the evolving meaning of animal welfare as a whole have continued to influence the direction of animal welfare as a science. Words and their different meanings, cautions Dawkins, are also a power influence on the public's interpretation of animal welfare science and standards (Dawkins 2012). For example, the term "free-range", which generally refers to access to an outdoor area featuring natural elements has a positive connotation, whereas "exposed to the elements", which may be used to refer to a similar scenario, would be considered negative (Dawkins 2012). One strategy to avoid ambiguous terminology that results in biased understandings of animal welfare is to focus on aspects of welfare that pertain to the animals themselves, rather than to inputs provided to them (such as enrichments), or outputs (such as

⁹ "Other animals, which, on account of their interests having been neglected by the insensibility of the ancient jurists, stand degraded into the class of things (...) It may come one day to be recognized, that the number of legs, the villosity of the skin, or the termination of the os sacrum, are reasons equally insufficient for abandoning a sensitive being to the same fate. What else is it that should trace the insuperable line? Is it the faculty of reason, or perhaps, the faculty for discourse?...the question is not, Can they reason? nor, Can they talk? but, Can they suffer? (...)"

fertility or productivity). These animal-centered definitions of welfare are commonly referred to as animal-based measures or welfare outcomes.

Beyond definitions of welfare and the factors that influence it, there are also frameworks that flesh out definitions in order to set the groundwork for the assessment of welfare, moving from the conceptual to the practical. A common foundation for these frameworks are the widely known “Five Freedoms”, a list of principles for the welfare of farm animals that was implicit in the Brambell Report and was later published by the UK Farm Animal Welfare Council (FAWC-replaced by the Farm Animal Welfare Committee in 2011) as a list similar to the updated version used today, listed below (Webster 2005b,12).

1. Freedom from hunger and thirst, by ready access to water and a diet to maintain health and vigor.
2. Freedom from discomfort, by providing an appropriate environment.
3. Freedom from pain, injury and disease, by prevention or rapid diagnosis and treatment.
4. Freedom to express normal behavior, by providing sufficient space, proper facilities and appropriate company of the animal’s own kind.
5. Freedom from fear and distress, by ensuring conditions and treatment, which avoid mental suffering.

Despite their significant and continued influence on the development of animal welfare science and the animal welfare movement, the Five Freedoms have been criticized from several perspectives, including the problematic interpretation of the term “freedom”, and their focus on the avoidance of negative experiences rather than the promotion of positive welfare states (Mellor 2016a). In addition to these shortcomings, the Five Freedoms are limited in terms of providing concrete guidance for the practical assessment of animal welfare. In recent years, several applied frameworks based on the Five Freedoms have been developed in order to

overcome these limitations. An example is the “Five Opportunities to Thrive” framework, developed to assess positive welfare states in animals kept in zoos (Vicino and Miller 2015).¹⁰

A number of these frameworks rely on three criteria; basic health and functioning, affective states, and natural living (Fraser et al 1997; Fraser 2008, 229). These criteria are often understood and visually represented as three overlapping circles, in the form of a Venn diagram (Fraser 2008, 230). The overlap between the three criteria represents not only the complexity of the concept of animal welfare and the related difficulties of assessing it, but also the possibility of shifting the focus from one area to another depending on the specific question about welfare that is being asked, and of considering synergies and dependencies among the criteria. For instance, one could ask whether outdoor access improves the welfare of laying hens. From the natural living perspective, the answer might be positive; hens with access to an outdoor area designed with their behavioral needs in mind may be able to display behaviors such as dustbathing and foraging. The ability to perform these behaviors may, in turn, have a positive effect on the affective state of the animal; however, it is also possible that hens who have not previously experienced outdoor environments would be fearful or stressed. The same is true of effects on basic health and functioning: while access to natural light, forage, and more space to range may have a positive effect, access to the outdoors will expose the hen to the elements, and possibly parasites and predators that could have detrimental effects. The issue of disagreement on what leads to good welfare depending on the angle from which the question is being asked is one that is carried over from the original Five Freedoms concept, where the fulfilment of one freedom can compromise the fulfilment of others.

¹⁰ The Five Opportunities to Thrive are: 1) Opportunity for a well-balanced diet, 2) Opportunity for animals to look after themselves (self-care), 3) Opportunity for optimal health, 4) Opportunity to express species-specific behavior, and 5) Opportunity for choice and control (Vicino and Miller 2015).

Some positions argue for a hierarchical understanding of the three criteria; Ian Duncan, for instance, proposes that affective states are the lead criterion by which welfare should be defined and assessed, because the ability to experience these affective states is what allows the animal to have a state of welfare in the first place (Fraser et al 1997; Fraser 2008, 231). Meanwhile, for Bernard Rollin, the affective states of particular species depend more deeply on the unique set of activities, interests, and functions that make up the nature of that species, for which he uses the Aristotelian concept of *telos* (Rollin 2006). Rollin's view prioritizes the set of unique needs that make up a species' *telos*, although he considers that in some cases, it is acceptable to alter this *telos* in order to reduce suffering.

The Five Domains model builds on the original Five Freedoms and the related provisions in order to advance a proposal that reorganizes the three original criteria (affective states, basic health and function, and naturalness) into five domains (Mellor 2016b). The physical and functional domains, which dominate the model, include: nutrition, environment, health, and behavior; the first three classified as “survival related factors”, and the latter as a “situation related factor” (Mellor 2016b). Mental state is the fifth domain, and like the four others, it includes a detailed list of examples of positive and negative experiences and situations; for example, in the mental state domain, pleasure related to drinking and eating, and goal-directed activity are examples of positive outcomes, while breathlessness, boredom, and pain are regarded as negative (Mellor 2016b).

Other than their common origin in the Five Freedoms, the models discussed thus far have in common a reliance on a concept central to many discussions of animal welfare: sentience. While sentience generally implies a capacity to feel pain and pleasure, it also suggests the ability to be self-aware and to evaluate the actions of others, some capacity for memory, and the ability

to assess actions and their consequences in terms of risks and benefits (Broom 2014). To some extent, these capacities hinge on the broader concept of consciousness, which involves subjective experience and its more objective, observable expressions. As part of the conscious experience, sentience requires complex higher order neurophysiological and cognitive functions, many of which cannot be directly studied. Thus, study of sentience has relied on observable expressions of sentience and related phenomena such as pain, learning, play, and memory. Due to its subjective nature, and despite advances in the fields of animal cognition and ethology, the issue of animal consciousness remains contentious, and to some extent is thought to have hindered the progress of animal welfare science.

For instance, Marian Dawkins proposes that while animal consciousness is a worthwhile issue, it is not, at this point, one able to be explained by science (Dawkins 2012, 48). While she acknowledges that some questions surrounding consciousness may be scientifically answered- for instance, how animals discriminate between stimuli or recognize objects- the truly “hard problem” of consciousness, that of subjective experience, cannot (Dawkins 2012, 46). Furthermore, she claims that continuing to use consciousness, and phenomena for which consciousness is a necessary condition such as pain, pleasure, or sentience to ground the science of animal welfare may result in the questioning of animal welfare as a science (2012, 48).

Given the problems with using these concepts as part of the groundwork and evidence for animal welfare science, Dawkins suggests bracketing the problem of consciousness, and instead focusing on what we already know about animals with enough certainty, simply from observing their behavior. For Dawkins, there are two main kinds of provisions that determine the state of welfare of an animal: needs and wants (2012,139). Needs are defined as inputs or provisions that an animal requires in order to sustain its basic biological functions (food, water, shelter); wants

are the mechanisms that evolved in order to trigger the satisfaction of those needs (Dawkins 2012, 139). Hunger, for instance, is a state in which an animal both needs and wants food; the need arises from physiological mechanisms that trigger the state of hunger, resulting in motivation for the animal to make food-seeking a priority. The physiological state of hunger (low blood sugar, for instance) constitutes the “need”, while the motivation to seek food in response to the need would be considered the "want". For domestic animals, especially those who are genetically far removed from their wild ancestors, such as the domestic chicken (*Gallus gallus domesticus*), needs and wants are often distanced or even decoupled, such that the mechanisms triggered by a need (hunger), often result in behaviors (associated with the state of wanting) that are really no longer necessary in order to satisfy the need, for example, foraging, ground-pecking or scratching. However, even chickens whose appetites are satisfied by a continuous availability of nutritionally adequate food will perform these "want" behaviors, signaling that although the direct link between the want and the need has been broken, the motivation to satisfy both still persists. Furthermore, while it is not surprising that the deprivation of a need (food) will result in poor health and ultimately death, the deprivation of wants, even when they do not directly affect an animal's ability to survive, has also been observed to have negative effects, which can compromise overall fitness and welfare (Dawkins 2012,127-139).

Dawkins’ needs/wants definition of welfare, in its simplicity, is very useful in grounding arguments in favor of animal welfare in that it does not require one to hold particular beliefs regarding the mental lives of animals and their significance for our moral obligations toward them (2012, 143). To state that animals have needs and wants, and that they should be satisfied such that their health and welfare are not compromised, especially when under human care, is

not a radical ethical claim, and one that is perhaps better suited to be the subject of scientific inquiry.

1.3 Contribution to the literature and scope of the project

Despite its exponential growth and increasing presence in public discussions, it can be argued that within the realm of social causes, animal welfare remains marginal or secondary, especially when considered within the broader landscape of socially relevant causes, including poverty, food security, water conservation, and a many others which are considered a priority because of their importance to the survival and well-being of human populations. The science of animal welfare is at least in part a response to social concern for the well-being of animals, and contributes to the continued presence of animal welfare on the map of social issues and its consideration by those who may not be swayed by emotional arguments for animal advocacy. However, animal welfare science continues to be challenged in its mission to objectively study subjective states in animals. The question of what constitutes an objective definition and assessment of welfare, including the types and uses of evidence to substantiate this objectivity, as well as serious and systematic consideration of robust subjective assessment are areas worthy of further inquiry that philosophy may contribute to.

While animal ethics and animal welfare science are said to have developed somewhat simultaneously and synergically, the bridging of the literature of these two fields is largely limited to analyses of how ethical views have contributed to, challenged, or complemented the development of animal welfare science and vice versa. Much of this work is, in a way, retrospective, and fails to point out how, moving forward, animal philosophy and animal welfare science can work together to tackle practical problems that hinder the development of knowledge in both fields, as well as *de facto* improvement of animal welfare. This project specifically aims

to address a gap in philosophical discussion regarding the characterization and merits of animal welfare as a science, and its uses in shaping and supporting the development of standards and policies.

More narrowly, the project focuses on the issue of evidence, which in the case of animal welfare is transformed from research outcome to an information product that transcends the scientific and ethical spheres and permeates public discussions of animal welfare. This evidence, produced largely within the bounds of animal welfare science, is used to justify the creation and application of animal welfare standards (public and private), which are ultimately mechanisms used to draw the line between acceptable and unacceptable ways of treating animals. Thus, animal welfare science is not only informed and directed by social concerns regarding the treatment of animals, but also feeds back into public conversations by confirming or challenging public perceptions through the provision of evidence. This work attempts to put into conversation the normative dimensions of animal welfare science, using insights from philosophy of science regarding the role of evidence in scientific inquiry and its applications.

1.4 Project outline

Chapter 2 of this work builds on the initial description of different models and frameworks for animal welfare in order to describe in detail the particular kinds of evidence used for each of the domains of animal welfare, namely, affective states, basic health and functioning, and naturalness. It will address the two main kinds of animal welfare evidence (physiological and ethological evidence), and describe the kinds of observations commonly used, as well as their theoretical underpinnings and justifications.

Chapter 3 offers a broad overview of definitions and uses of evidence from the perspective of philosophy of science, also touching on extensively discussed limitations of

evidence, including underdetermination and inductive risk. This background is used to frame a discussion of the role of evidence in animal welfare science through the lens of the substantial body of literature on the role values in science. The ultimate purpose of this background and discussion is to provide the necessary setup for an analysis of the ethical dimensions of animal welfare science at various stages of the scientific process, starting with the influential role of societal values in research agendas, and going more in depth into the less transparent and more contested roles of values within the research process itself. The uses of evidence in animal welfare science are contrasted with examples from other applied fields where the use of evidence has been more thoroughly discussed, for instance the use of evidence-based practice in human medicine.

Chapter 4 considers evidence specifically in its role in the creation and implementation of animal welfare standards. The chapter starts with a brief overview of definitions of standards, focusing then on their philosophical underpinnings and ethical ramifications. Discussion on the philosophy of standards includes two subtopics; first from a philosophy of science perspective, it addresses the role of standards as classification tools, which requires the assignment of entities to particular kinds. The second subtopic deals with the normative implications of standards, which foster what Busch (2011) refers to as "the moral economy of standards". The theoretical discussion is then applied to an analysis examples of animal welfare standards in the US.

Chapter 5 provides an empirical perspective of the topics of science and standards, by way of insights obtained through semi-structured interviews of four people representing different perspectives from the field of animal welfare standards: an animal welfare researcher, an industry leader, and two people who direct private farm animal standards organizations in the US. Topics addressed by these interviews include general perceptions on the process of standard-

making, perceived roles of animal welfare science in the creation and application of standards, the role of science and standards in generating public trust around animal industries, and the perceived role of the public in pushing for the implementation of animal welfare standards.

Chapter 6 offers a synthesis of the central issues examined in this work, as well as conclusions and recommendations regarding areas for improvement and future research regarding the roles of science and philosophy in guiding animal welfare policy and standards. Recommendations also include specific insights on how philosophers can contribute to the improvement of interdisciplinary collaboration between stakeholders who influence the standards-making process, and who are more generally drivers of change for animal industries.

Chapter 2: Animal welfare as science: conceptual frameworks, research areas, and evidence

2.1 Introduction

Animal welfare science encompasses several disciplines within the biological and biomedical sciences, including ethology, animal science, veterinary medicine, and physiology, among others. Because of this multidisciplinary nature, and the fact that animal welfare involves objective and subjective aspects of an animal's physical and mental states, animal welfare science relies on a multidisciplinary mixed methods approach to produce comprehensive assessments of welfare. In its complexity, animal welfare eludes precise quantification. While exact measurement is not possible, animal welfare can be assessed by determining and measuring indicators that allow the observer to make an inference about the welfare state of an animal (Fraser 2008). The selection of indicators and thresholds is generally based on natural history, prior observations and background information about the health and behavior of a particular species or individual. Two common types of indicators used in animal welfare assessment are resource-based measures or inputs, which include genetic makeup, housing, enrichment, food and water, veterinary care and husbandry, and animal-based measures or outcomes, which rely on direct observation of the physical state, physiological measures and/or behavior of an individual animal or a representative sample within a larger population (Webster 2005a). While it is generally agreed that these indicators are acceptable indirect measures of welfare, there is less agreement about what kind and amount of evidence is enough to substantiate the use of these indicators, and their relative contribution to the comprehensive assessment of welfare. In an attempt to address this problem, conceptual frameworks of animal welfare aim to break down definitions of welfare into functional areas, common indicators for

each area, and methodological guidance on how they should be used in order to produce reliable assessments of welfare.

This chapter describes common conceptual frameworks used in animal welfare to explain, organize, and analyze the types of scientific evidence commonly used in welfare assessment. It also provides a more detailed description of three broad research areas of animal welfare science introduced in Chapter 1; health, affective states, and naturalness. These three areas are common elements of the conceptual frameworks discussed in this chapter. In section 2, the topic of conceptual frameworks is introduced by touching briefly on theoretical concepts, providing definitions for conceptual frameworks and scientific strategies. Then, three common frameworks or models for the definition and assessment of animal welfare are described and analyzed: the three pillars framework, the needs and wants model, and the five domains model. Finally, this chapter looks in detail at the kinds of evidence used in animal welfare science. This informs further analysis of the three conceptual frameworks as scientific strategies, and sets the stage for a deeper exploration of theoretical aspects, uses and limitations of evidence within scientific frameworks and strategies. The discussion of evidence is also relevant to the examination of the roles of values, both social and cognitive, at different stages of the scientific process, a topic introduced in this chapter in the description of scientific strategies, and further developed in Chapter 3.

2.1.1 Functional areas in animal welfare science

Like health, animal welfare is a concept that reflects the functioning of a system. As briefly introduced in Chapter 1, definitions of animal welfare largely agree on the systemic nature of the concept, and range from purely descriptive, focusing on the functional balance of the system (e.g. Broom's definition), to more normative, where the system is judged to be well

balanced if it is aligned with the goals or purpose of the organism in question (teleological definitions, e.g. Dawkins and Rollin). Before discussing instances of how these definitions are fleshed out in conceptual frameworks, three main functional areas which are present across most definitions and frameworks are described in more detail in this section. The objective is to provide further insights into how each of these areas are defined and their role within systemic understandings of animal welfare, and to contribute to the later discussion of different scientific strategies in animal welfare science.

Biological function, health, or “fitness” is perhaps considered the most basic requirement for good welfare, and is included in most if not all conceptual frameworks and assessment protocols for animal welfare. Arguably, it is also the functional area of welfare that has been most extensively researched; animal health, growth, and reproduction were studied by veterinarians, animal scientists, and other agricultural professionals long before the concept of animal welfare emerged (Fraser 2008, 84). Regardless of this long-standing scientific interest in animal health, precise definitions of animal health are scarce. For example, an analysis of 500 veterinary textbooks from the library collection at the Swedish University of Agricultural Sciences conducted in 2003 found that only 8 percent of those textbooks included explicit definitions of health and disease (Gunnarsson 2006). This study grouped definitions of health into five categories: normality, biological function, homeostasis, physical and psychological well-being, and growth and productivity (Gunnarsson 2006). Animal welfare definitions and frameworks often define health, either explicitly or implicitly, with an emphasis on one or more of these categories.

From the perspective of normality, biological function, and well-being, states of health and illness matter with regard to their effects on an animal’s ability to function within an

expected range of normality (to note, what is considered normal can vary significantly even within one species, depending on context). Consider, for example, a modern broiler chicken at six weeks, who spends most of its time sitting or eating, versus a modern laying hen of the same age, who spends proportionately more time engaging in active behaviors such as perching, dustbathing, or foraging (when the environment allows such activities). Although these animals are of the same species and age, genetic selection and specialization for different functions (weight gain versus egg-laying) radically changes what is considered normal and healthy in each case.

Conversely, disease is defined as a condition that disrupts normal physical and/or mental function, sickness is the state of being ill, and health the absence of illness or injury, as well as a state of soundness, fitness, or vigor (Cockram and Hughes 2011). Growth and productivity are more relevant to welfare assessment in farmed animals, where good welfare is considered instrumental, rather than a goal in of itself, and is often defined and assessed in correlation to production practices and productivity thresholds. This is because productivity (and its attached economic value) is of course the main purpose for the lives of these animals, but also because in intensive production settings, the demands of production, the scale of operations, and the density of populations are the main factors influencing health outcomes. For instance, some specific health issues in laying hens (foot lesions, feather damage, keel-bone fractures) are specific to the kind of system they are housed in (systems designed to increase productivity and profitability), and may not be as prevalent when animals are housed in different systems.

Arguably, in farmed animals, poor health outcomes are a concern in relation to their effects on production, and consistently, not only welfare, but also health tends to be defined and valued as a condition for production, and not as a goal in of itself. For example, health issues

likely to cause pain and suffering, such as lameness in broilers, are a concern not only because pain is considered detrimental to the welfare of the animal, but because of the effects that it may ultimately have on production parameters, such as decreased appetite and slower weight gain. However, while health and productivity are often correlated in farmed animals, productivity is not considered a reliable indicator of health and welfare within most current frameworks. The reason for this is that some animals have been so heavily selected for production traits, that they will continue to produce even at the expense of their health. A classic example is that of laying hens who continue to produce eggs in spite of low calcium levels; since calcium is required to make the shell of the egg, rather than stop producing eggs, the hen's body will utilize calcium stores from its bones, resulting in brittle bones and fractures, all with the purpose of maintaining high egg production (Fraser 2008, 100). If health is to be assessed from a welfare perspective, or as part of a comprehensive welfare evaluation, its assessment must include indicators that go beyond conventional veterinary diagnostics, considering cross-relations between disease processes, affective states, and behavioral outcomes. According to Broom (2014), clinical assessment of welfare is an emerging area of medical and veterinary science.

A second functional area encompasses affective states or emotions. The term “affective states” was first used by physiologist Walter Cannon to refer to emotions and feelings that are either pleasant or unpleasant rather than neutral (Fraser 2008). In the context of animal welfare, affect is a term that includes emotions, feelings, and moods; these terms are often used interchangeably and generally refer to the mental state of an animal (Broom 2014). However, there are subtle but relevant differences among these terms that should be kept in mind when articulating how affective states fit within the broader scheme of animal welfare science, and what kind of evidence is available to support their relevance to animal welfare. Contemporary

research on affective states proposes that they are defined by physiological, behavioral, cognitive, and subjective components (Paul et al 2005). This means that, whereas emotions have a subjective component, they are not simply a subjective or experiential state.

Broom provides a markedly biological definition of emotions as “physiologically describable conditions characterized by electrical and neurochemical activity of the nervous system” (Broom and Fraser 2007). In later work, he acknowledges that they are also characterized as a conscious state, and are considered a kind of elicited response, attached to an object (Broom 2014). Feelings can also be used to refer to an emotional state, but the term is sometimes used to describe sensory experiences; a feeling is a brain construct, requires perceptual awareness (but not necessarily conscious experience), is associated with a life-regulating system, and may result in an observable change in behavior (Broom 2014). Thus, according to these definitions, emotions entail a more complex interaction of feelings and their physiological correlates, while feelings do not necessarily involve physiological mechanisms or changes in behavior. Moods are generally understood to be longer lasting and may persist for hours or even days after the stimulus that prompted them has disappeared (Dawkins 2012).

From the perspective of welfare assessment, these conceptual differences may seem irrelevant; however, the subtle differences in the language that animal welfare scientists use to describe them may serve as clues regarding underlying assumptions about the animals in question, the kinds of evidence that can be used to confirm and assess the occurrence of affective states, and the weight assigned to them in the overall evaluation of welfare. For instance, believing that an animal can experience something such as a “mood”, as opposed to a feeling, presupposes the existence of certain cognitive capacities, and could also influence research at

various stages (framing of questions, method selection) as well as the reporting and interpretation of results.

Conceptual differences aside, affective states are arguably the functional area that carries the most weight with regard to the actual state of welfare of an animal, and are the most difficult to assess objectively. This difficulty was acknowledged even prior to the emergence of animal welfare science as such; classical ethologists recognized affect as relevant to animal behavior, even when they considered it futile to attempt to study affective states scientifically. Niko Tinbergen, one of the founding figures of modern ethology, points to this in *The Study of Instinct* (first published in 1951), where he writes “because subjective phenomena cannot be observed objectively in animals, it is idle to either claim or deny their existence” (Tinbergen 1989). Later publications, such as *The Question of Animal Awareness* (Griffin 1976), made headway on the topic of scientifically assessing affective states in animals (Keeling et al 2011).

The assessment of affective states is still a contentious topic in animal welfare research, but one that continues to develop, especially with the increased availability of technologies such as brain imaging that enable researchers to establish neurological correlates between different affective states and their corresponding stimuli. It is reasonable to expect growing interest in research of affective states and their relevance to welfare, as well as the shaping of this research by the increasing sophistication and availability of technologies such as brain imaging. For example, a scientific literature review of articles on the topic of animal sentience published between 1990 and 2012 found 2,562 publications that included one or more of forty-three keywords related to animal sentience (Proctor et al 2013). This review also notes that the majority of this research focused on negative affective states; only 154 of articles reviewed referred to positive states or experiences in animals, compared to 2,359 articles which referred to

negative keywords, and 31 articles considered to have no valence, but discussed issues related to sentience, such as theory of mind or consciousness (Proctor et al 2013). A significant proportion of the work on negative affective states focuses on the concept of suffering, a fuzzy notion which encompasses fear, pain and boredom, and can presumably be recognized through observation (Dawkins 2008). Within the area of suffering, research on pain has been extensive, and continues to grow. A review of publications in the Journal of Dairy Science (JDS), spanning a hundred year period between 1917 and 2017 notes that the first paper in JDS using the word “pain” was published in 1999, in an article related to mastitis (for reference, the first paper in JDS using the words “animal” and “welfare” or “animal” and “well-being” was published in 1998) (von Keyserlingk and Weary 2017). Since 1999 more than 70 papers using the word “pain” (in reference to dairy cattle and dairy production) have been published in JDS (von Keyserlingk and Weary 2017).

Scientific assessment of affective states can range from simply establishing the presence or absence of a state based on behavioral and physiological data, to attempts to qualitatively or quantitatively assess these states. For example, assessing valence, which is the extent to which affective states are pleasant or unpleasant, can be done through preference testing, or by observing the frequency or strength of behavioral signs correlated with specific affective states (Broom 2014). Needless to say, regardless of the development of increasingly complex and reliable indices for the objective assessment of affect, its defining subjectivity remains.

The bias toward research of negative affective states may be related to practical limitations in the research of positive affective states, as negative emotions are often expressed more strongly and consistently, in ways that can be readily observed and correlated to the originating stimuli (Proctor et al 2013). However, these perceived limitations may also stem

from assumptions about the role of positive emotions in survival and reproduction (Fraser 2008, 159). This relates to the question of whether or not animals are able to experience pleasure at all, and even if they are, how much energy they are willing to expend toward behaviors that do not directly improve their fitness or likelihood of survival.

Recent research on play behaviors has provided some plausible explanations. In animals, play is not an affective state, but a set of behaviors that have been deemed “useful” to animals in terms of learning survival skills and strengthening social bonds (Fraser 2008, 159-160). Positive affective states (such as joy, contentment, comfort) and their behavioral cues are believed to play a role in the initiation of play and other behaviors that do not have immediate effects on preventing harm or poor welfare, but may contribute to long term fitness, and thus represent a worthy energy investment for animals (Fraser 2008, 160-161). As argued by Mellor and Beausoleil, effective conceptual frameworks and assessment tools for animal welfare need to evolve in line with our expanding knowledge of positive affective states and their role in defining thresholds for good welfare (2015, 211).

A third functional area, naturalness, is more loosely defined than health and affect, an issue perhaps related to the contested meaning of “nature” or “natural” in general. One interpretation of these terms alludes to the idea of natural environments, and the assumption that these environments (outdoor access, pasture, vegetation, natural light) are without a doubt better for animals than intensive, indoor systems. The notion that a natural life, understood as that of a traditional, pre-industrial farm setting, is a good life for animals may be related to the idea of such a life as wholesome because it is part of a broader natural landscape and its natural cycles (Fraser 2008, 48). However, in the context of animal welfare science, this somewhat romantic notion of a naturalness as good is considered insufficient, largely because its effects on welfare

cannot be reduced to or robustly correlated with objective indicators. An alternative interpretation of nature or naturalness refers to the nature of the animal, including instinctive behaviors and other mechanisms that assist in coping and survival (Broom 2014).

In a strictly scientific sense, natural conditions are good (or better than non-natural conditions) only when they can be demonstrably linked to better welfare outcomes (Dawkins 2012, 147). Unlike health, where objective evidence may be more readily available (for example, absence of lesions, good feather condition) and can be linked in a fairly straightforward way to welfare outcomes (skin integrity is a barrier against infection, good feather condition is related to birds being more active, and to thermoregulation, both indicators and/or conditions for health), arguing that a natural environment, or even elements of a natural environment, is conducive to better welfare is more difficult. Furthermore, because natural or extensive production systems are generally less efficient and productive than intensive systems, and scaling them to meet demand is not yet considered feasible given demand for animal products, interest and funding for research on the welfare aspects of these systems is limited. For example, despite extensive research on the effect of different lighting programs on broiler productivity, the majority of these studies focus on the effects of different artificial lighting programs on production parameters, and not more broadly on welfare outcomes (Olanrewaju et al 2006). An even smaller amount of this research addresses the effect of natural light on broiler welfare.

Despite these gaps in the research, standards, policies, legislation and other regulatory documents often make references to naturalness or natural behavior in definitions of animal welfare. For example, the OIE Terrestrial Animal Health Code, in its section on animal welfare (Volume I, Chapter 7), refers to “the ability to express innate behavior” as one of the requirements for a good state of welfare (OIE 2017). Naturalness is also an important part of

many animal welfare assessment schemes, not as blanket requirement, but rather by including elements of the natural environments of a particular species (or its wild ancestor), based on the assumption that they help to satisfy needs and wants and therefore influence welfare. Thus, from a scientific assessment perspective, “naturalness” should not be evaluated with regard of what environmental provisions are in place, or whether or not they fulfil the vision of “natural” environments for these animals, but rather by observing whether animals are able to express innate (natural) behaviors that demonstrably correlate to better states of welfare.

2.2 Conceptual frameworks

A conceptual framework is, in simple terms, a schematic representation of basic concepts and their relationships within a particular area of knowledge. Jabareen (2009) defines conceptual frameworks as “a network, or “a plane,” of interlinked concepts that together provide a comprehensive understanding of a phenomenon or phenomena”. Furthermore, these frameworks are not simply descriptive tools or heuristic devices, but contain ontological, epistemological, and methodological dimensions that filter and shape the kinds of data that flows through them, and the kind of information and assessments generated from them (Jabareen 2009). In addition to the practical function of structuring and organizing information in order to produce assessments, these frameworks are also an instance of what Lacey (2003) refers to as a *scientific strategy*, which he defines as an ideology or paradigm that provides direction to a field of inquiry, and constrains the kinds of acceptable theories, and the empirical data and descriptive categories that support such theories.¹¹ Lacey proposes that most modern natural science adopts a materialist scientific strategy, which in essence limits acceptable theories to those related to structures,

¹¹ Lacey’s use of “strategy” is, per his own assessment, consistent with Kuhn’s (1962) use of “paradigm” and Laudan’s (1977) “research tradition” (Lacey and Staddon, 2003).

processes, interactions, or laws, and acceptable empirical data to that which is intersubjective, quantitative, and measurable (Lacey and Staddon 2003, 212). One of the main objectives of materialistic strategies is to maintain the impartiality and neutrality of the research process through the exclusion of social values (Lacey and Staddon 2003, 211). However, there are phenomena within the natural sciences that cannot be adequately explained or applied within the constraints of materialistic strategies; these phenomena are complex systems whose explanation requires the investigation of tendencies, capacities, interactions, among other attributes (Lacey and Staddon 2003, 213). Examples of these include agroecological systems (in which, Lacey and Staddon claim, the farmer's observations and experience are essential to the research), and cognitive and behavioral research (Lacey and Staddon 2003). The selection of strategies constitutes the first of three moments in scientific research, followed by the acceptance of data, hypotheses and theories, and finally, the application of scientific knowledge (Lacey and Staddon 2003, 211). The selection of a strategy, which in this case corresponds to the selection of a framework, should help answer questions regarding problems to solve, what kinds of explanations to explore, what to observe, measure, and experiment upon and what procedures to use (Lacey and Staddon 2003, 212).

As previously stated, although it is generally agreed that animal welfare pertains to an animal's state, rather than to environmental factors or provisions that may affect that state, there is no singular definition of welfare. Accordingly, multiple conceptual frameworks have been developed in order to explain welfare as a holistic concept, to map out the functional areas that define welfare and how they relate to each other, and ultimately, propose how this understanding can be applied to the assessment of welfare.

At first glance, the dominant conceptual frameworks of animal welfare science, seem to follow a materialistic strategy; they rely on categories of observable phenomena all of which meet the criteria of being quantitative, measurable, and intersubjective. These categories include biological function (including health, physiology, and productivity), affect and feelings (measured by observable behavioral expressions), naturalness, and needs and wants (observed as preferences and motivations) (Webster 2005a; Nordenfeldt 2006; Fraser 2008; Appleby et al 2011; Broom 2014). For the purposes of this work, I selected three well-known conceptual frameworks or models that are representative of the main schools of thought in animal welfare science. The first, which I refer to as the three pillars framework, includes three commonly identified domains of animal welfare, the physical, mental, and natural aspects, which rely mainly on animal-based measures (Fraser et al 1997; Appleby 1999). The second is the five domains model, a comprehensive model that rather focuses on mapping negative and positive subjective states, and correlating them to environmental factors that influence them (Mellor and Beausoleil 2015). Third, the needs and wants view largely bypasses the distinction between intrinsic (animal-based) and extrinsic (environmental) determinants, and rather views welfare as the result of an animal's process of adaptation to its circumstances, manifested in terms of how well its needs and wants are met (Dawkins 2012). While the different frameworks propose varying combinations and interpretations of information about animals and their environments in order to produce an assessment of welfare, the kinds of information or data collected in order to make these assessments ultimately fall into one of the three functional areas of welfare: biological function, affect and feelings, and naturalness. Each of these areas is discussed in more detail, and examples of the indicators used to measure and quantify them are provided.

2.2.1 The three pillars framework

The conceptualization of animal welfare as a result of the combined effects of biological function, mental states or affect, and naturalness was arguably influenced by the disciplinary backgrounds of those who took part in early discussions of animal welfare as a science.

Veterinary and agricultural scientists had long dominated research in the areas of animal health and productivity, but the emergence of animal welfare, and the idea that animals might suffer within industrial production systems invited the views of ethologists and other behavioral scientists, who exerted significant influence on the development of the new field (Broom 2014, 24). For example, the work of ethologists and psychologists including Miller (1959), Hinde (1970), and McFarland and Sibly (1975) provided insights into control systems and decision-making in animals; these foundations were then developed by researchers like Ian Duncan, David Wood-Gush (1971; 1972), and others who developed key theories on motivation, frustration, and the biological basis of needs (Broom 2014, 24). Thus, behavioral research became one of the methodological pillars of animal welfare science, regardless of which of the three main areas of inquiry, (biological function, affective states, or naturalness) was favored. For example, Broom, in one of the earliest works defining animal welfare and its indicators, notes that physiological indicators such as heart rate and adrenal enzyme activity are central to welfare assessment, but abnormal behaviors may also indicate poor welfare (Broom 1986). On the other hand, Duncan, who claimed that welfare should be defined strictly in terms of how an animal feels, also finds behavioral research to provide suitable indirect evidence for affective states (Duncan 2005).

As illustrated by these examples, even early contributors to the development of the concept of animal welfare were amenable to the idea that all three functional areas were relevant to animal welfare, but disagreed on the priority of functional areas within these frameworks. The

challenges of integrating views, methods, and evidence from multiple disciplines into coherent conceptual frameworks remain to this day, however, initial efforts to produce a synthetic framework can be attributed to Fraser et al (1997), Appleby (1999) and Lund (2006), who argued for the overlapping and complementary nature of the three functional areas. Fraser (2008) analyzes these challenges in greater detail, discussing the practical implications of integrating the different vocabularies, methods, and criteria for proof, as well as the challenge of integrating theoretical constructs. With regard to the difficulty of navigating and integrating different theoretical constructs, Rushen (2003) raises concerns about the lack of applicability of experimental research in animal welfare to practical welfare concerns; he blames this in part on the inconsistencies between conceptual definitions of animal welfare that cater to research needs, and practical definitions that reflect public concerns.

In this case, an initial analysis of the three pillars framework, given its origins in sciences that largely follow a materialistic strategy (per Lacey's definition), would indicate that the framework as a whole is also an instance of a materialistic strategy. However, the weighting and integration of different kinds of evidence from each of the three domains would arguably require a value judgment. Recalling Lacey's three logically distinct moments within the research process: 1) adoption of strategy, 2) acceptance of data, hypotheses and theories, and 3) application of scientific knowledge, the integration of evidence within the three pillars framework arguably occurs in the second moment (2005). Furthermore, the kind of judgment required to weigh and synthesize evidence from all three domains would arguably require researchers to determine what is most important to the animal in that instance, based not only on their scientific knowledge of the species, but also on a more normative notion of what constitutes good welfare or a good life for that animal. Thus, it is possible that this framework is rather

included in what Lacey calls alternative strategies that allow for the use of value judgments at all stages of the scientific process without necessarily compromising the scientific (cognitive) values of impartiality, neutrality, and autonomy (2005).

2.2.2 Needs and wants framework

In contrast to views included in the three pillars framework, the needs and wants approach admits the complexity of animal welfare, but proposes that this complexity should not translate to how we define or measure welfare. Marian Dawkins, one of the strongest proponents of the needs and wants view, states that welfare can be both defined and assessed by answering two questions: 1) Is the animal healthy? 2) Does the animal have what it wants? (Dawkins 2003). In this sense, the needs and wants approach strips away much of the conceptual baggage of the three pillars framework in order to focus on that which can be measured by simply assessing whether an animal is healthy, and if it has what it needs to meet its basic needs, and perhaps even thrive. But, despite the apparent simplicity of this approach, health itself is a complex concept, and one that is strongly reliant on social context, for humans and animals alike. Additionally, assessing whether an animal has what it needs and wants and what that means from a welfare perspective demands a deep examination of the origin of those wants, how they are expressed through different behaviors, and whether or not they have been decoupled from needs in the processes of evolution or domestication. For Dawkins, understanding the evolutionary roots of needs and wants, and how they are expressed in modern domestic animals is key to formulating the kinds of research questions that lead to a better understanding of animal welfare. Dawkins makes a distinction between evolutionarily driven mechanisms, calling them “needs” and the more proximate mechanisms that animals have evolved to fulfill those needs, referring to them as “wants”. Together, needs and wants constitute coping mechanisms, which allow not only

individuals, but species as whole, to survive by getting better at predicting and anticipating threats to survival (Dawkins 2012). Some anticipatory mechanisms are related to individual, immediate needs, such as the activation of physiological thirst mechanisms in response to changes in hydration, while others, such as seasonal migration, have evolved in a species over time, and in response to complex environmental and internal cues (Dawkins 2012).

Even when these behavioral mechanisms generally increase an animal's or species' odds of survival, there are often conflicts between the satisfaction of needs and wants, and the risks that arise in the process, which can result in harm or even death. For instance, a thirsty animal visiting a water hole, driven by its "want" to drink is also exposed to potential predators because of that action (Dawkins 2012). In domestic animals, the situation is complicated by the fact that although these animals have "needs" passed down through evolution, the "wants" related to those needs are often already satisfied in domestic environments, such that the want mechanism no longer serves its original purpose (Dawkins 2012). For example, broiler chickens are motivated to scratch the ground and forage, a behavior related to food-seeking in a natural or undomesticated setting. In commercial production, broilers are provided food and water ad libitum in trays or automatic feeders, and chickens have to do no more than approach these containers in order to consume food and water. But, according to Dawkins, animals can "want" the things they no longer "need" just as strongly, if not more so, than the things they really need (Dawkins 2012, 137). In domestic animal welfare research, the tensions and discrepancies between needs and wants, and the balance between these is one of the main drivers of research (2012, 143). Because these tensions exist, Dawkins' argues, the goal of animal welfare should not be to reach a state where needs and wants are perfectly satisfied, but where they exist in an acceptable balance within a particular environment (2012, 142-143).

While the needs and wants approach offers several advantages, as described above, it also has limitations. One limitation is related to research design; for example, in designing preference studies, presenting sets of choices that do not reflect choices that an animal might have to make in a “natural” setting may produce results that do not really address the question of what animals need and want. For example, one such study attempted to answer the question of whether access to a dustbathing substrate or food was more important to laying hens, in situations where they were deprived of one of those resources (Petherick et al 1993). The study, which used a Y maze setup in order to assess preference, reported that food-deprived birds consistently chose food, while substrate deprived birds often chose food although they were not food deprived (Petherick et al 1993). Arguably, choosing between food and the opportunity to dustbathe is not a choice that red junglefowl (the wild ancestors of domestic chickens) would have to make, but a study like this could help determine which wants are a higher priority for domestic laying hens. In this case, determining whether food or dustbathing are a higher priority is scientifically interesting, but given that both of those resources can be provided without conflict in a well-designed system, these results would not necessarily contribute to building systems where needs and wants are in balance, one of Dawkins’ stated goals for animal welfare science.

From the perspective of scientific strategies, this framework does not differ significantly from the three pillars framework. Like the three pillars, it seems to follow a largely materialistic strategy, relying on measurable, quantitative data. One interesting difference occurs in the set-up of the strategy, which, as Lacey states, is informed by questions such as what problems to solve, what kinds of explanations to explore, what to observe, measure, and experiment upon and what procedures to use. Recalling the notion that conceptual frameworks include ontological, epistemological, and methodological dimensions, in the three-pillars framework ontological

concerns seem to precede epistemic and methodological questions. On a general reading, the starting point for many of its proponents is the definition of animal welfare based on the three areas central to its determination. Dawkins' starting point, on the other hand, can said to be driven by epistemic and methodological questions: what is it that we need to know in order to solve problems, and how can we best gain this knowledge.

2.2.3 The five domains model

The five domains model, proposed by Mellor and Reid (1994) focuses on the dynamics between the factors that affect welfare and their expression as mental or affective states (Mellor and Beausoleil 2015). Much like the three pillars framework, this model divides welfare into physical/functional domains (nutrition, environment, health, and behavior), and then determines their consequences in a fifth domain, mental state (Mellor and Beausoleil 2015). Thus, although the model recognizes the relevance of these four functional domains to overall welfare (i.e., if any of these is compromised, this will in turn compromise the outcome), it centers the animal's mental state as the locus for the expression of welfare. In practice, then, welfare assessment using the Five Domains model involves first, accounting for all of the factors in domains 1-4, and then assessing their impact on the mental state of the animal, through the use of indicators assessed as present or absent. For example, the nutrition domain is assessed in terms of opportunities or restrictions to readily access water and sufficient, nutritionally adequate, and varied food, while the behavioral domain evaluates opportunities or restrictions in terms of exercising agency in a variety of ways, including movement, engagement, exploration, play, and shelter (Mellor and Beausoleil 2015).

The conceptualization of the model, on the other hand, begins with mental states, expressed as affective states or emotions, and follows with proposed correlations with the other

four domains. The original Five Domains model proposed by Mellor and Reid mirrors the five affective states contained in the Five Freedoms statement: thirst, hunger, anxiety, fear, pain and distress (Mellor and Reid 1994). This list was later updated in order to break down the content of “distress” in more detail, as well as expand the list of affective states considered, reflecting updated research in ethology, physiology, and neuroscience (Mellor and Beausoleil 2015). The updated list includes: breathlessness, thirst, pain, hunger, nausea, dizziness, debility, weakness and sickness, anxiety, fear, frustration, anger, helplessness, loneliness and boredom (Mellor and Beausoleil 2015). From a conceptual perspective, this model integrates all of the three functional areas included in the three pillars approach, while at the same time evaluating them essentially in terms of whether or not the animal’s needs and wants are satisfied. Thus, the Five Domains model is not merely a conceptual model, but also provides a tool for practical assessment, based on a five-tier grading system; the four functional domains are graded first, and their aggregated scores are then used to calculate overall impact in the fifth domain, mental health (Mellor and Beausoleil 2015).

In terms of strategy, it is useful to examine the authors’ motivation for proposing this revised model. The authors state that given the expansion of knowledge in the field of animal welfare, available frameworks are no longer able to capture the breadth and depth of knowledge of biological processes, and the translation of this knowledge to updated codes and standards (Mellor 2016a). More specifically, Mellor and Beausoleil (2015, 241) note a shift in the goals of animal welfare science, from a focus on diagnosing and correcting poor welfare, to an emphasis on the promotion of positive welfare. From this, it is reasonable to infer that epistemic concerns precede ontological and methodological concerns, although methods are also relevant if understood as a tool meant to be used not only in the conduct of research but in its application.

2.3 Types of evidence and indicators of animal welfare

When stripped of most of the conceptual baggage and theoretical assumptions of conceptual frameworks, evidence used for welfare assessment falls into the two categories originally suggested by William Thorpe in the Brambell Report: physiological and ethological. Thorpe writes: "(...) There are two main kinds of scientific evidence which bear closely on the problem of cruelty to animals, namely physiological and ethological (...) when we come to interpretation, we shall find that the two lines of evidence act and interact upon each other" (Brambell, 1965). The following section describes each type of evidence, and then provides examples of their application within each of the three functional areas previously described.

2.3.1 Physiological measures

The prominence of physiological evidence can be traced back to early work on the concept of animal welfare. In Broom's original definition of welfare as "the state of an individual as regards its attempts to cope with the environment", the mechanism of coping involves having control over the stability of mind and body (Broom 1986; Broom and Johnson 1993). In order to cope, in other words to be able to maintain a state of physical and mental balance in the face of an environmental challenge or disturbance, an individual needs to adapt; this can be done through behavior (removing the challenge or removing oneself from its effect), through physiological mechanisms, or a combination of these. A related concept, stress, is also involved in the mechanisms of coping and adaptation, and has been prominent in animal welfare research, especially as regards its physiological mechanisms, and outputs of these mechanisms that can be measured and used as welfare indicators. According to Broom, stress occurs when stressors (factors that challenge or upset a state of balance) overtax the mechanisms by which an

individual maintains balance, thus leading to reduced fitness and poor welfare (Broom and Johnson 1993).

In physiological terms, the state of balance maintained by coping and adaptation is known as homeostasis. Homeostasis entails a balance between an individual and its surroundings, which is internally reflected by physiological variables kept at a “set point” (Korte et al 2007). Physiological variables can include a variety of measures that are used as indicators for physiological activity or mechanisms; some examples are heart rate, respiratory rate, blood pressure, and concentration of glucocorticoids (commonly known as “stress hormones”) in blood, saliva, or feces. The “set point” for each of these is the value or range of values considered normal for an individual of a particular species, sex, and age. As regards welfare assessment, one of the advantages of physiological measures is their degree of objectivity. For most of these measures, a range of normal or desirable values (set point) has been established, and the mechanisms used to measure them are standardized, and often mechanized, such that human bias in the recording of these measures is considered minimal.

However, in terms of how accurately they portray welfare, physiological measures have at least three limitations. The first is that many require invasive procedures or extensive handling of the animal (for example, taking blood samples, or placing probes for EEG or EKG procedures), which is bad from a welfare perspective given that they are stressful, but also, since stress indicators are what is often being measured, results may be skewed. Second, physiological measures often lack specificity. For example, concentrations of glucocorticoids in blood are a frequently used measure of stress. When an animal is stressed, this activates the hypothalamus/pituitary/adrenal (HPA) axis (a physiological mechanism stress response mechanism involving nervous system and vascular structures), and as a result, increased levels of

glucocorticoids are released into the bloodstream in order to prepare the body for whatever environmental challenge triggered the response. These triggers are not always related to potentially harmful or dangerous situations; for example, the proximity of a viable sexual partner may trigger a response just as much as the proximity of a predator. Thus, unless the animal is in a highly controlled experimental situation (which might itself trigger a response), establishing correlations between potential triggers and observed measures is difficult. The third limitation is that individuals can adapt to stress, therefore, a situation that triggered a strong response at first (for instance, close confinement) may eventually not cause a significant response, thus, observations would seem to indicate that the situation is not stressful. This adaptation, however, often comes at a price, especially when stressors are persistent, and can lead to long term effects (for example suppression of the immune system) that could also be used as measures of poor welfare (Webster 2005b).

Recent work on the use of physiological measures as indicators of welfare proposes that the concept of physiological “set points” representing an ideal state of balance- homeostasis- may be inaccurate. Korte et al (2007) suggest the alternative concept of allostasis- the maintenance of stability through change- which implies that physiological set points change through time (for example, in response to natural selection), to enable the individual to adjust their balance point according to long term changes in their environment. According to this concept, unusual trends in physiological parameters do not necessarily signal a failure to reestablish balance, but an adjustment in anticipation of future demands (Korte et al 2007). The concept of allostasis as dynamic balance, versus homeostasis as a set of predetermined balance points may align better with conceptual frameworks like the needs and wants model, which acknowledges that circumstances like domestication and intensive rearing may significantly

change the way animals respond to their environments, and therefore the environmental circumstances necessary for good welfare in those animals.

2.3.2 Ethological evidence

Ethological or behavioral evidence, as the name suggests, includes observable behaviors, which can be part of the normal behavioral repertoire of an animal, or occur in response to internal (physiological or pathological) or external (environmental) challenges. These behavioral responses are shaped by genetics, context, experience, and perception, and delimited by physical capabilities (Webster 2005b). Categories of behavior that can be observed to make inferences about the state of welfare of an animal include maintenance behaviors, orientation responses and postures, defense and avoidance behaviors, injurious behaviors (directed at self or others), stereotypies (repetitive behavior without a purpose), redirected behaviors, and suppression of normal behavior (Webster 2005b). In terms of their correlation to welfare, these behaviors can be classified according to whether they indicate good welfare, effective coping, or failure to cope, which results in poor welfare (Webster 2005b).

Methods for collecting behavioral evidence are based on observation, either directly, or with the assistance of technology, for instance video-recording. A common starting point for behavioral observation is to conduct an ethogram, which is an inventory of all observed behaviors in an individual or representative sample in a given environment (Webster 2005b). These observed behaviors can be considered “events” (a relatively quick action or series of actions), or “states” (a position or attitude adopted by an animal over a longer period) (Altmann 1974). For example, Altmann (1974) distinguishes between sitting (an event), and being or remaining seated (a state). An ethogram can serve as a baseline assessment to study the change

of behavior over time, but it does not necessarily provide insights into what should be the normal behavior pattern for an individual or population; in other words, it is merely descriptive.

In some cases, ethograms conducted in “natural” environments can be used as a baseline against which observations in experimental or commercial production conditions are compared. For example, Dawkins (1989) produced an ethogram for red junglefowl, wild ancestor to the domestic chicken, in order to provide a baseline for the assessment of domestic poultry. This ethogram determined the seven most common activities in red junglefowl (walking, pecking, looking, scratching, preening, roosting, and standing), as well as the percentage of time spent on each, with walking and ground pecking being the most frequent activities, occupying up to 60% of the time (Dawkins 1989). In comparison, modern meat chickens (broilers) spend more than 80% of their time sitting, which has been correlated to genetic selection for fast growth and related metabolic demand, as well their predisposition to lameness (Weeks et al 2000). Dawkins’ work on red junglefowl behavior has been used to argue that the low activity levels and poor mobility of modern broilers is an indicator of compromised welfare.

As illustrated by the example above, behavioral observations and their interpretations are shaped by background knowledge, conceptual frameworks, and assumptions. The outcomes and interpretation of a behavioral observations are also influenced by sampling methods. In one of the most comprehensive and cited works on the topic, behavioral ecologist Jeanne Altmann describes behavioral sampling methods, and addresses the question of if and how they restrict the kinds of behavior processes that can be studied (Altmann 1974). Altmann distinguishes between manipulative and non-manipulative methods, where the former entail controlled laboratory conditions, and the latter observation in natural settings. In either case, another methodological choice determined by research questions and practical limitations involves the scope of the

observation (1974). The most common include *ad libitum* observations, which involve observing and recording everything that occurs within a given time frame, focal observations, which involve only a target individual or small sample, observation of focal behaviors, and scan sampling, which involves recording observed states in an individual or group at established intervals (Altmann 1974). While it can be argued that these methodological choices are informed mainly by the kinds of research questions being addressed, Altmann notes that researchers often select methods without considering how this choice directly limits the kinds of questions that can be answered (1974). In addition to limiting the kinds of questions that can be answered, methods, as well as settings, can limit the degree of objectivity with which those questions can be answered, and the conclusions that can be drawn from them (Altmann 1974).

In addition to methodological limitations, behavioral studies have drawbacks related to the management of confounding factors. In natural settings, controlling confounding factors may prove difficult if not impossible. In experimental settings, this is more feasible, with studies controlling not only environments, and husbandry practices, but also, genotypes. Interestingly, the high levels of environmental and genetic standardization in behavioral studies of laboratory animals (rodents in particular) have been claimed to interfere with brain development and behavior, thus affecting the meaningfulness and applicability of study findings (Olsson et al 2003). An additional limitation to consider, related to those previously discussed, refers to the usefulness of behavioral study results for practical applications. While well designed studies may generate results that establish reliable correlations between observed behaviors and welfare outcomes, their internal validity does not necessarily translate to external validity, especially when experimental conditions are relevantly different from non-experimental settings.

2.4 Interpreting and applying evidence

In animal welfare research, research questions and hypotheses presuppose correlations between observations and welfare outcomes from three main functional areas that are the foundation of the conceptual frameworks discussed: biological function (including health or fitness), affective states, and naturalness. For example, research to determine whether higher stocking density (defined as a ratio of number of animals or total weight of animals per area unit, e.g. 30 lb/ ft², or 5 chickens/ ft²) results in higher prevalence of footpad dermatitis¹² in broiler chickens would presumably have one of three outcomes: inconclusive, negative correlation, or positive correlation (either weak, moderate, or strong). However, the interpretation of these results and their application in practice arguably depends on the scientific strategy and conceptual framework adopted at the outset of the research process. A purely materialistic strategy would demand that any data used as evidence be measurable, quantifiable, and intersubjective. For instance, the prevalence and severity of lesions can be quantified using objective indices and measures, as can the stocking density. As such, methods such as randomized controlled trials can be used to conclude whether higher stocking densities result in higher prevalence (and/or severity) of footpad dermatitis. Arguably, while the immediate goal of this particular study may be to answer that specific research question, the question is being framed by one or more scientific strategies or conceptual frameworks. Given the focus of the research on establishing a correlation between a health indicator (footpad dermatitis) and one related to housing or husbandry conditions (stocking density), it is likely that such a question is

¹² Footpad dermatitis (FPD) is a condition that causes lesions on the plantar surface of birds' feet (ranging in severity, from thickening of the external layers of the skin to ulceration and tissue death). In broilers chickens, the interest in research of FPD allegedly stemmed from the emergence of an export market for chicken feet in the early 1980s. Prevalence and severity of FPD is now an accepted and widely used criterion for the assessment of broiler and turkey welfare in the US, Europe, and other geographies (Shepherd and Fairchild, 2010).

informed by a framework based on the three pillars. Interestingly, while the three pillars approach does consider affective states, a reliance on materialistic strategies means that it cannot admit evidence that falls outside of the bounds of what is objectively measurable and quantifiable. Thus, within research informed by this framework, conclusions suggesting correlations between health indicators and affective states (such as pain or suffering) are rare. While knowing what conceptual framework is used has no impact whatsoever on the conclusions of the research itself (and as such is only implicit in published research), its contribution to a body of evidence that favors certain frameworks over others and thus drives a research direction for the field is relevant. The boundaries imposed on research conclusions not only influence the research agenda, but can have direct effects on actions to correct or mitigate situations which perpetuate less than ideal living conditions for animals.

No matter what definition of welfare, conceptual framework, or research method is used to gather evidence about the welfare of an animal, there comes a point at which the available evidence must be comprehensively considered in order to make an overall assessment of welfare. Most scientific studies of animal welfare do not engage in the aggregation or synthesis of different kinds of evidence, but rather make conservative conclusions about discreet correlations between well defined, measurable and quantifiable variables. However, the application of this research, either through husbandry protocols, standards, or policies does involve making sense of multiple sets of evidence in order to establish thresholds for acceptable welfare outcomes and corresponding environmental inputs. This process might include deciding which evidence or results will be used (based on scientific criteria as well as its applicability to the assessment being made), which aspects will be given priority, and how the assessment will be reported (for

instance, using a descriptive or numerical scale, or a mathematical calculation based on scores and weights assigned to each measure).

An example of a numerical scoring system based on weighted scores is the Animal Needs Index (ANI) developed by Austrian animal scientist Helmut Bartussek (Fraser 2008, 255). This index scores animal housing in relation to welfare, and is based on the adequacy of the housing system with regard to five criteria: the animal's ability to move, to have social contact with other animals, the quality of the floor for lying, standing, and walking, environmental conditions (light, air quality), and quality of human care (Fraser 2008, 255). Note that although these are all provisions (and therefore are not a direct measure of welfare), correlations between the criteria being assessed and welfare outcomes have presumably been established beforehand. Each criterion is assigned a point value, which varies based on the relative importance of said criterion; then, the scores are added up to obtain a final score that is considered an overall measure of the suitability of that housing system in terms of welfare (Fraser 2008, 255).

While the relative scores assigned to each criterion in the ANI are presumably based on evidence as regards their relative importance to the welfare of the animal, it also illustrates one of the major challenges in aggregating and synthesizing welfare measures in order to make a comprehensive assessment: establishing commensurability. Commensuration is defined as the transformation of different qualities into a common metric (Espeland and Stevens 1998). Commensurability is a relative process; it relies on establishing relations between attributes or measurements, and expresses that relativity in the form of a value (Espeland and Stevens 1998). It is also an abstractive and reductive process, and information is necessarily lost, or intentionally discarded as a result (Espeland and Stevens 1998). Of course, commensuration is a value-laden process, not only because it requires us to choose what information is ultimately relevant, but

because it is often motivated by goals or concerns which are also value-laden. For instance, it may be used as a strategy to rationalize something, or conform to powerful expectations or assumptions (Espeland and Stevens 1998). In animal welfare assessment, it is also a way to manage uncertainty, and fulfill a need for the objective assessment of what are ultimately subjective states.

2.5 Conclusions

The complexity of the concept of animal welfare inevitably involves competing definitions and conceptual frameworks that inform the agenda and direction of the field as a whole. Arguably, this is driven not only by what is considered the dominant discourse within the field, but also by broader scientific paradigms, which continue to favor strategies deemed objective based on the relative absence of social and ethical value judgments. Although the notion of value-free science has been broadly contested by philosophers of science in recent decades, proposing acceptable roles for social values at different stages of the scientific process, the matter of the study of subjective states in animals, which may require value judgments beyond what is considered scientifically acceptable. Wemelsfelder (1997) identifies and characterizes this problem in detail, noting that scientific models of animal behavior are limited to describing behavior as regards its physical constituents and their causal relations, rather than as subjective, experiential phenomena. This, she claims, is an attempt to address a problem from a theoretical perspective that does not reflect the inherent nature of the problem (Wemelsfelder 1997). Wemelsfelder suggests starting points for reliable and systematic investigation of subjective experiences, including the creation of revised terminology that more accurately captures the expressive repertoire of animals, leveraging the methods of scientific disciplines such as psychology and psychiatry, which have developed nuanced methods and scales to

characterize subjective phenomena such as depression (1997). Based on the discussion presented in this chapter, I argue that the issue needs to be addressed at a more fundamental level, namely that of scientific strategies or conceptual frameworks. Wemelsfelder's description of an adequate framework for the assessment of subjective experiences is one that 1) centers subjective experience, 2) is not mechanistic in nature, 3) presents behavioral organization as an explanation itself, rather than as an explanandum fragmented into underlying causal factors (1997). On initial examination, none of the conceptual frameworks discussed in this chapter fulfil these conditions, thus the difficulties in researching subjective phenomena such as consciousness lie not only in the technological limitations of science, as Dawkins proposes, but in the very definition and frameworks delimiting animal welfare science.

The difficulty of providing acceptable evidence for affective/subjective states, either as part of a framework that merely includes them or one that centers them is not new. The Brambell report warns that in interpreting observations of animal behavior and other welfare indicators, it is justifiable to assume that animal suffering is not identical to that of humans, but it is equally justifiable to assume that they suffer in (relevantly) similar ways; the balance between these two extremes is therefore a matter of judgment (Brambell 1965).

However, as Fraser states, animal welfare research, as socially mandated research, stems from ethical concerns regarding our use, treatment of, and relationships with animals (2008). As such, it requires collaboration and feedback between the society that frames the issues, and the science that poses questions and proposes solutions (von Keyserlingk et al 2017). Ongoing discussions over definitions of welfare are pointless if the goal is to reach consensus, but these discussions are fruitful in the sense of bringing to the surface disagreements that may be rooted in differences in values and assumptions that are ultimately relevant to the framing, conduct, and

interpretation of research. Farmed animal welfare presents an especially challenging case, as the economization of animals necessarily implies that their welfare also becomes an economic good, adding to the complexity of the web of values underlying the nature, interpretation, and communication of animal welfare evidence (Bueller and Roe 2014).

Chapter 3: Evidence and values in animal welfare science

3.1 Introduction

Having covered the descriptive and practical aspects of the conceptual frameworks and kinds of evidence used in animal welfare science and their limitations, this chapter turns to a more theoretical discussion of evidence, its limitations, how cognitive and social values inform what we consider acceptable evidence, and how it is interpreted and applied to scientific judgments, in this case, assessments of animal welfare. The first section of this chapter is a broad overview of the nature, definitions, and roles of evidence in the scientific process. The discussion then turns to specific limitations of the use of evidence within the scientific process, specifically, underdetermination and inductive risk; these issues are illustrated with specific examples of evidence used in animal welfare science described in the previous chapter. Finally, this chapter addresses the role of values in science, and particularly in animal welfare science. Because of its interdisciplinary nature, animal welfare science produces heterogeneous evidence, requiring synthesis and interpretation that often incorporates social and moral concerns and values at different stages of the scientific process. As such, it constitutes an interesting example of how values play different roles in socially mandated science. These values transcend the sphere of scientific research, and also influence the development of standards and policy, which have direct consequences on the lives of animals, humans, and the environment.

3.2 Evidence: definitions and uses

Theoretically, evidence is data or information that supports a theory or hypothesis, where the relationship between the data (evidence) and the hypothesis is logically sound (Kincaid 2012). But in practice, what exactly constitutes evidence, and the conditions for a significant, logically sound relationship between data and the hypothesis are not easy to define. Among scientists, there is frequent disagreement about whether, or to what extent, some set of data or observational results constitute evidence for a scientific hypothesis (Achinstein 2003). These disagreements can be empirical, for example regarding methods, assumptions, or data accuracy, but they can also be due to incompatible concepts of evidence (Achinstein 2003).

Achinstein (2003) distinguishes between two concepts of evidence; veridical evidence, which he considers strong, and potential evidence, the weaker of the two. Veridical evidence is that which provides sufficient reason to believe that a hypothesis is true, where the objective truth of both the evidence and the hypothesis is independent from anyone's particular beliefs on whether they are true or not (Achinstein 2003). Veridical evidence is considered conclusive when it establishes, with absolute certainty, that the hypothesis is true; to put it in terms of probability, if the evidence is conclusive, the probability that the hypothesis is true is 1 (where 1 is the maximum possible value) (Achinstein 2003). Potential evidence also provides good reason to believe that a hypothesis is true, but it does not presuppose that the hypothesis is true, and is empirically incomplete (Achinstein 2003). Given these definitions, it would seem that in no case would scientific evidence be veridical, as science is considered inherently uncertain, and related to this uncertainty, hypotheses can be supported by evidence to a greater or lesser extent, but can never be true.

Weight of evidence is a concept that is often used to indicate how strongly a body of evidence supports a hypothesis or theory (Weed 2005). According to Weed (2005), within the

biomedical literature, it is used in at least three different senses. The first is used synonymously with “summary interpretation of the evidence”, or “synthesis of the evidence”; the second use involves the assumption that all available evidence for a given issue was systematically examined; a third use refers to a methodological approach by which the evidence is systematically weighted and analyzed (Weed 2005). This third interpretation involves examining available studies and weighing them in the context of well-accepted criteria for causation (Weed 2005). Although the terms weight of evidence and body of evidence are not as frequently (or explicitly) used in the animal welfare literature, all three concepts are common within the assessment of evidence in animal welfare science. For example, the Animal Needs Index described in Chapter 2 involves the methodological approach of weighing criteria and assigning a value to them according to their correlation with different welfare outcomes (Fraser 2008, 255).

Another common use of “evidence” in the biomedical literature is in reference to evidence-based practices. The concept of evidence-based medicine, which emerged in the 1990s promoted the use of evidence from clinical research as the gold-standard for medical decision-making (Goldenberg 2014). The motivation was to make medical reasoning as impersonal and unbiased as possible, relying on systematic reviews of the literature and application of the formal rules of evidence in order to make diagnoses and determine treatment based on the best observed results in practice (Goldenberg 2014). Critics of evidence-based medicine note that the push to move away from possibly biased views resulted in removing the social and ethical values central to healthcare practice (Goldenberg 2014). An additional critique addresses the limits of generalizing the knowledge obtained through evidence analyses beyond the confines of randomized controlled trials (Goldenberg 2014).

Similar issues arise in the practice and application of animal welfare science. The push to establish animal welfare science as a science has long favored more objective methods (as described in Chapter 2); this is not undesirable per se, but may carry the unintended consequence of ignoring or excluding more subjective methods that may produce more comprehensive assessments of welfare in favor of objective methods that may be too narrow and reductionist. The issue of internal versus external validity also arises in the application of animal welfare science. Experimental conditions often differ in significant ways from commercial settings, and thus, regardless of how well a study was designed and the quality of evidence produced, its broader applicability to commercial operations is limited. This is arguably one of the limitations of methods that are otherwise scientifically robust and widely used in animal welfare science, such as randomized controlled trials.

Another use of evidence in the biomedical field involves the collaboration of scientists and ethicists in a relatively new area of research called “evidence-based ethics”. Empirical ethics involves the application of methods from the social sciences to ethical issues; the results of this are generally descriptive rather than prescriptive or normative (Goldenberg 2005). Evidence-based ethics, however, takes the empirical approach one step further, using evidence regarding the best care and prognosis of the patient to promote better informed and justified ethical decision-making (Goldenberg 2005). On the face of it, this seems innocuous, and even laudable; however, as Goldenberg (2005) points out, the approach is problematic in several ways, including the limits it imposes on the kinds of experiences that can be explored and quantified, and that would therefore be considered adequate justification within ethical analysis and decision-making.

3.3 Limitations of evidence: underdetermination

In addition to the issues described above regarding the uses and application of evidence in decision-making, evidence presents some inherent limitations that must be considered, one of which is underdetermination. Underdetermination refers to evidence being insufficient to warrant acceptance or rejection of a theory (Turnbull 2017). Specific forms of underdetermination include equivalence, holist, transient, and practical underdetermination (Turnbull 2017). Equivalence underdetermination is the notion that for every theory there is at least one other rival theory that is empirically equivalent, in other words, all empirical evidence that supports a given theory also supports its competing theory or theories, even when it is impossible for more than one theory to be valid at the same time (Turnbull 2017). Holist underdetermination, also known as the Duhem-Quine thesis, posits that evidence for one hypothesis alone is insufficient to accept or reject any theory, and therefore related hypotheses and background assumptions must be tested together (Turnbull 2017). Given this condition, even when there is data that appears to falsify the theory, a scientist can maintain that the theory is valid, and that the conflicting data is due to an error in either background assumptions, or the design and execution of the experiment (Turnbull 2017). Transient underdetermination also asserts that evidence alone is insufficient to confirm a theory, but keeps open the possibility that in the future further evidence may become available to confirm a superior version of this theory that is yet unavailable (Turnbull 2017). Some views conflate this version with a fourth type, practical underdetermination, which also refers to a state in which evidence is insufficient to choose one theory over another (Turnbull, 2017). However, Turnbull argues that unlike transient underdetermination, practical underdetermination does not assume anything about future evidence being able to confirm particular theories; it simply observes that for a given theory or set of theories, at a particular point in time, there is insufficient evidence to accept one over another (Turnbull 2017). In the

empirical sciences, and especially in the applied sciences, underdetermination is, in most cases, considered transient, as it is commonly assumed that while theories can never be proved with certainty, more and better evidence will become available with the development of better methods and technologies.

3.4 Limitations of evidence: inductive risk

According to Douglas (2017), science is endemically uncertain because of its inductive and ampliative nature; thus, although evidence is never complete, we are still able to make generalizations that transcend this incompleteness. It is in this process of generalizing in the face of uncertainty that inductive risk arises. Inductive risk involves having to make a choice about whether a theory, hypothesis, or claim is true even when the evidence is not conclusive, especially when making the wrong choice entails serious, morally relevant consequences. Birch (2017) argues that because animal welfare scientists often play the role of both scientist and policy advisor, they are especially vulnerable to the problem of inductive risk.

Inductive risk involves the possibility of making two kinds of errors. Type I errors (false positives) are the incorrect rejection of a true null hypothesis, while Type II errors (false negatives), which involves a failure to reject a false null hypothesis (Birch 2017). While these terms were originally applied in the context of statistical significance, the terms false positive and false negative are commonly applied more broadly to the acceptance of hypotheses as either true or false when the opposite is true (Birch 2017). According to Birch (2017), among scientists, false positives are considered worse, as they assert something that is not true. Additionally, Birch points out in animal welfare science, false positives generally involve the over-attribution of phenomena or attributes that cannot be studied objectively or quantified, such as pain.

Birch presents the case of cognitive enrichment as an example of common over-attribution errors in animal welfare research. In Birch's example, cognitive enrichments are provided as a way to stimulate the expression of natural behaviors that require cognitive engagement, thus avoiding boredom, and possibly eliciting positive affective states, such as joy. Birch contends that when experiments using cognitive enrichments are successful (in the sense that the animal demonstrates preference or motivation for the enrichment scenario), there is a tendency to attribute this to the stimulation of a positive affective state, thus risking over-attribution.

While Birch's examples involve the occurrence of false positives and false negatives in cognitive research, I argue that these types of errors may also occur in research that makes no assumptions about the mental states of animals, but rather attempts to establish correlations between environmental provisions and welfare outcomes. For example, in a study assessing the effects of natural light and the provision of straw bales on leg strength and activity levels in broiler chickens, researchers reported that the provision of straw bales had no effect on activity levels; they follow this conclusion with a possible explanation related to the number of straw bales provided (Bailie et al 2013). While in this case, the authors' explanation for why no correlation was observed cautions against the acceptance of this as evidence that there is no correlation between these two variables, many other examples offer no explanation, leaving the findings open to interpretation. It is at this point that social and ethical values come to play an important role that may determine whether or not the research findings result in modifications at the farm level that improve welfare in meaningful ways. Because many of these improvements, however minimal, entail a cost to producers (sometimes passed on to consumers), especially considering the scale of animal agriculture operations, I argue that there is a tendency to place a

heavier burden of proof on findings that suggest the need for changes to housing or other provisions. For example, Lusk et al (2018) conducted an economic study modeling the cost to US chicken producers and consumers if the whole meat chicken industry were to convert to slower growing breeds (which arguably have a potential for better welfare outcomes), finding that the annual cost to consumers would be close to 630 million US Dollars, while annual costs to producers would be an estimated 3.5 billion US Dollars. Revising the roles that values play in this process may help correct this potential bias, and better assess inductive risk.

3.5 The role of values in science: an overview

3.5.1 Types of values

The term value, in the sense used in value-theory, corresponds to subjective features of experience, such as emotion or preference (McMullin 1982). On the other hand, value can also be used to describe an inherent property of something that is considered positive or desirable, for example, the productivity of a dairy cow in terms of how much milk she produces in relation to the resources she consumes (McMullin 1982). This second definition of value is presumably objective, since, in this example, judging whether a cow is productive (defined in the restricted sense proposed above) does not depend on anyone's particular beliefs or preferences regarding cows or productivity, but on a ratio of inputs to outputs. These two definitions are examples of a common distinction made between two types of values: subjective and objective. In the context of discussions of values in science, definition and categorization of values goes beyond the objective/subjective divide, and is influenced by what are held to be the goals and guiding principles of science (Douglas 2009). In the 1950s, the distinction between epistemic and non-epistemic values was introduced, and played a central role in supporting the value-free ideal of science (Douglas 2009). The epistemic/non-epistemic distinction has been critiqued, largely

based on the argument that non-epistemic values influence and shape epistemic values, thus blurring the lines between the two kinds, and defeating the purpose of maintaining these values separate for the sake of scientific integrity (Douglas 2009). Regardless, this nomenclature is still used, and is therefore described below, together with other common terms and categories for values in science.

3.5.2 Epistemic, cognitive, non-epistemic, and social/ethical values

Epistemic values are defined as those that promote the defining characteristics of science, and therefore mirror the values of a good scientific theory, and support the attainment of knowledge (McMullin 1982). In more technical terms, a value is said to be purely epistemic if its role in theory choice is likely to improve the epistemic status of a theory, in other words, make that theory more likely to generate or support knowledge in a particular field or on a specific topic (McMullin 1982).

A more recent term that refers to the kind of values described above is “cognitive values”. Lacey (2003) defines cognitive values as: “(...) characteristics (criteria) of “good” (rationally acceptable, desirably held) beliefs, and “good” (soundly accepted) theories. While on the face of it this definition seems very similar to older definitions of epistemic values, Douglas claims that the term “cognitive values” offers a more accurate definition, as it more precisely refers to characteristics of the scientific process that help scientists analyze the evidential and inferential aspects of theories, rather than just support the robustness of a theory according to scientific principles (2009). Many of the values identified as cognitive are in fact consistent with foundational principles of science (e.g., simplicity, explanatory power, scope, consistency, predictive precision), but deploying them as values entails that they play a role in guiding a process, (scientific cognition), rather than embodying a goal (knowledge or truth-seeking)

(Douglas 2009). This, according to Douglas is another important distinction between epistemic and cognitive values; epistemic values act more like criteria for theories as tools to attain the goal of science (reliable, generalizable knowledge about the world), whereas cognitive values do not always play a direct role in attaining this goal (2009). For example, fruitfulness is a cognitive value insofar as it generally improves the productivity of science, but it is not a criterion for determining whether a scientific theory is “good” based on the principles of science (Douglas 2009).

Social and ethical values are often categorized simply as “non-epistemic”, but each kind has distinct roles and implications that may be at times be contradictory (Douglas, 2009). Ethical values are those used to make judgments about whether an action is right or wrong; social values are those that are considered positive values within a specific social context (Douglas 2009). As Douglas points out, social values can sometimes be opposed to ethical values in relation to particular issues. Animal welfare provides some good examples of this. For instance, it can be argued that productivity is a value of American society; a process or individual is deemed to be productive in the measure that it or they produce something that is valued by society, but even more so if that productivity is abundant and efficient (which could be perceived as separate but related values). In the context of animal agriculture, productivity is particularly valued, but can be at odds with ethical values regarding our obligations towards animals.

3.5.3 Roles of values in science

Douglas proposes that rather than focusing on semantic differences, a more useful approach to discussing the role of values in science is to examine the influence they exert over the scientific process, and thus determine their acceptability (Douglas 2009, 96). She divides the roles of values into direct and indirect, where direct roles are those where values are used to

determine decisions and can be stand-alone reasons to motivate choices, and indirect roles are those in which values can help to guide a decision, but can never be the sole or primary justification for making a choice (Douglas 2009, 96-97).

Elliot (2017, 166) provides a list of instances or stages within the scientific process during which direct values could and do play legitimate roles. This includes: a) establishing the research agenda; b) development of research questions and selection of methods; c) determining the aims of scientific inquiry in particular contexts (e.g. wild animal welfare research vs farm animal welfare research); d) response to and communication of uncertainty: including how much evidence to demand in order to draw conclusions, and how boldly to communicate said conclusions; and d) description, framing, and communication of results. A specific aspect of scientific communication where social and ethical values may play a legitimate direct role involves choices regarding terminology in science communication Elliot (2017,120). An example from animal welfare research was the trend to avoid using terms that referred to affective states in animals, for example suffering, joy, and boredom. This intentional choice was grounded, at least in part, in concerns of appearing unscientific by referring to states that we cannot conclusively say exist in animals, based on the available evidence. More recently, researchers are moving away from this tendency, and are more willing to use these terms openly in scientific publications.

3.6 Values in animal welfare science

Discussions of the role of values and/or ethics in animal welfare science most often address how ethical frameworks, moral principles, or values have influenced the development or direction of animal welfare science, but there is less discussion on values as embedded in the process and methods of animal welfare science, which, per the considerable body of literature on the roles of values in science, is a relevant topic. Furthermore, it has been argued that animal welfare scientists, who often play the dual role of scientist and advisor, carry a particular kind of moral burden, as their decisions entail significant consequences for animals and society (Birch 2017). Whereas the heritage of the value-free ideal still supports the notion that social/ethical values should not have a role in any stage of the scientific process, it can be argued that because animal welfare science developed in response to ethical concerns, that animal welfare scientists should, at all times, keep the ethical application of their work in sight, if that work is to be relevant (Fraser 2008).

The social values underpinning the systems within which humans and animals interact must also be considered. Two important social values of the industrial worldview, which includes modern intensive animal agriculture, are progress and productivity (Fraser 2008, 54). These values of industrialism are embedded not only in the practices of intensive animal production, but also permeate the sciences that inform its continued progress, including animal welfare science. Aspects of animal welfare that have been linked to the values of progress and productivity include health, growth, and survival; these values translate quite seamlessly to commonly used indicators in commercial assessment of animal welfare, for example daily weight gain and mortality. A common argument in favor of intensive, industrialized animal production systems is that controlling environments and preventing disease and predation results in better health and higher productivity (Fraser 2008, 58). A related value, efficiency is achieved

in part by controlling as many aspects of a system as possible. I argue that in the context of farm animal welfare research, these values are consistently present in the background, and may influence research in ways that can be legitimate, through indirect roles, but can also overflow into direct roles if there is little awareness of how much influence they exert.

An important and potentially productive way in which values could play a more active role in animal welfare science is by using them, as Douglas (2017) suggests, to manage the inductive risk that is inherent to socially mandated sciences. Two specific ways in which Douglas suggests that social and ethical values may be used in legitimate roles include the external assessment of evidence, and, as part of that, the open, interdisciplinary debate of disagreements regarding the interpretation of evidence, in ways that are accessible to the public and may help them understand the issues and act upon this knowledge. Douglas' (2017) proposal for the external assessment of evidence includes the engagement of experts in the assessment of evidence, not only as experts, but as members of the community, with the responsibilities that entails, including more careful consideration and adequate communication of inductive risk.

With regard to interdisciplinary debates to clarify interpretations and uses of evidence, this is something essential for a field as interdisciplinary as animal welfare science. Indeed, to effect significant change for animal welfare science, it may be necessary to move beyond interdisciplinarity and toward a transdisciplinary community, which entails not only the collaborative production of knowledge, but direct engagement with the co-production and use of that knowledge outside of academia (Toomey et al 2015).

3.7 Conclusion

Douglas (2017) argues that because science cannot non-arbitrarily set standards for sufficient evidence, we must necessarily consider the context in which research is being conducted and its intended uses. For animal welfare, this includes the ethical motivations related to the humane treatment of animals, as well as practical social concerns regarding the effect of changing animal production practices. Examples of these concerns are for example price, availability, land use, and other environmental considerations. Douglas (2017) also suggests that there should be a universal standard for what is acceptable evidence in socially relevant science (2017). Sources of disagreement with regard to the sufficiency of evidence include lack of understanding across disciplinary boundaries (especially important in an interdisciplinary field such as animal welfare science), distrust of unfamiliar methods, disagreement about which kinds of errors matter, and more generally, inductive risk considerations (Douglas 2017).

Chapter 4: Standards, science, and ethics

4.1 Introduction

The previous chapters have discussed the ethical roots of animal welfare science, the development of animal welfare science qua science, its theoretical underpinnings, and the definitions, structure and organization of what is considered evidence in the practice of animal welfare science. The limitations of evidence, and the various roles of social/ethical and cognitive values in the scientific process have also been examined. This chapter addresses how animal welfare concepts, frameworks, and research are used to develop and justify animal welfare standards. This is preceded by a more general discussion of standards, which touches on definitions, categories, development, and implementation, including social and ethical implications.

Standards are, by definition, normative; their function is to measure things within a given category or kind against a model that has been established as the ideal. Here, “ideal” is used not in the sense of that which is the best possible manifestation of things within the category being standardized, but simply a model that has been selected as best for a given purpose. That purpose may be practical, and carry no moral weight. For example, tomato plantations using harvesting machines require beds, plants, and the tomatoes themselves to be of a certain size in order for the machine to be able to harvest efficiently (Busch 2000). This says nothing about whether or not the tomatoes are objectively good; in fact, homegrown, non-standard tomatoes may be considered better in terms of taste, freshness, and other quality attributes. Similarly, uniformity is valued in broiler chickens, not only because it is an indicator of efficient productivity, but because it allows the processing of chickens on standardized equipment during slaughter and processing. Thus, the standard shape, size, and ripeness of tomatoes, as well as the specific size

and weight of chickens in these systems is determined in terms of the technical requirements for mechanical harvest.

It could also be argued that at least some of these attributes are determined by market requirements, for instance, the tomatoes must be within a certain price range, be able to withstand transport without being damaged, and have a sufficiently long shelf life. While some of these may be considered quality attributes, they still do not determine that this particular tomato is objectively the best, all things considered. In this example, the implementation of standards is motivated by social values not directly related to the product being standardized, but to values embedded in the process of bringing it from the field to consumers, such as efficiency, profitability, and reliability. Standards are also not necessarily normative in a moral sense; to use the example above, the farmer who runs an efficient tomato operation by standardizing his tomato production may be considered a “good” farmer in the sense of efficiency and profits, but his motivations for implementing the standards are not moral, nor is he considered morally better for doing so. However, standards are as diverse as they are ubiquitous, and so are their motivations and implications, which in some cases do involve the moral sphere.

4.2 Classification of standards

Busch (2011, 23) states that different kind of standards lead to different individual and social behaviors; they also require different kinds of justification. Standards may be used to ascertain that something is the best, that it fits an exemplary measure or weight, that it is of superb quality, or that it simply meets the threshold for the defining characteristics of its kind (Busch 2011, 25). Busch’s classification of standards includes four kinds: Olympic standards, filters, ranks, and divisions (2011, 42). As the name suggests, Olympic standards are those that single out one or a small group of the best things or beings of a kind, within a specified time and

space (Busch 2011, 43); for example, at the World Dairy Expo, only one dairy cow is selected as “Supreme Champion Cow”, for the period of one year following the competition.

Filters, on the other hand, are minimum standards in the sense that they establish a threshold for belonging to a particular kind, or meeting the minimum conditions for a kind of practice (Busch 2011, 43). The European Union’s 1999 directive on housing for laying hens (CEC 1999/74/EC) is an example of a filter standard, as it establishes the minimum acceptable conditions according to scientific research and advice, rather than the best possible conditions for laying hens. Prior to the directive, hens were given on average less than 450 cm² (69.75 in²) of cage space per hen; in 1989, a research publication reported that brown hens occupied 475 cm² (73.62 in²) just to stand still, and that in order to turn around, they needed 1,272 cm² (197.16 in²) (Appleby, 2003). Rather than recommend the maximum amount of space needed as reported by the research, the European Commission Scientific Veterinary Committee determined that conventional (battery) cages should provisionally increase space allowance to 550 cm² per hen, and would be phased out by January 1, 2012, in favor of enriched cages (furnished with a nest, litter, perch, and claw shortening device) providing at least 750 cm² per hen (Scientific Veterinary Committee 1996). The justification for this determination was that it was difficult to draw conclusions on space requirements because they presumably vary according to resources provided and group size (Scientific Veterinary Committee 1996).

Ranks are a third kind of standard, and as the name suggests, they involve sorting all things or beings within a kind into a hierarchical order, in which it is implied that those sorted into higher ranks are better in some relevant way than those sorted into lower ranks (Busch 2011, 45). The United States Department of Agriculture (USDA) standards for grading shell eggs are an instance of this; the USDA provides detailed guidelines describing the tests and requirements

for shell eggs to be graded AA, A, or B, and this grade will influence the market access and price of these eggs (USDA 2000).

Finally, division standards are similar to ranks in that they sort things into different categories based on observable features, but in this case the categories are unranked (Busch 2011, 46-47). Beef labeled “Certified Angus Beef” has passed a certification process involving a division standard. In order to receive this label, beef must come from cattle of predominantly Angus genetics, and meet 10 specifications related to its appearance and quality (Certified Angus Beef 2016). In this case, while it could be argued that this standard is actually a rank because Angus beef is a better product than non-Angus beef, it is also true that the belief that this is a “better” product is based on the preference of a particular group of consumers, and therefore it is not necessarily objectively or ethically better. This last point relates to what Busch (2011, 51) refers to as the fuzzy boundaries between types of standards; while it is often easy to identify a standard by the categories discussed above, in practice, the different kinds are often linked together.

Another way to categorize standards according to their scope and objectives is to consider whether they are public or private; according to Hobbs (2010, 137; 142), public standards carry with them a legal obligation for compliance, while private standards are ostensibly voluntary, but can also become mandatory if the market adopts them as a requirement. Busch (2011, 26-27) agrees that the public/private distinction does not carry much weight, for not only can private standards become de facto mandatory, but it can also be the case that public standards can encourage or discourage actions by providing incentives or sanctions, but do not ultimately exclude participants, often because there are complex political relations at stake. Some of the functions of private standards include product differentiation, supply chain management,

and protection or improvement of a company's reputation (Hobbs 2010,142). Some of these functions, such as product differentiation, require a labeling system, as well as a strategy for the public promotion of that system, in order to be effective.

In the United States, there are no public (government or legal) standards for farmed animal welfare, and very few regulatory instruments (laws or regulations at the federal or state levels) that oversee the welfare of farmed animals. Some species, like domestic poultry, are specifically excluded from all legal regulation as regards their welfare. Thus, in addition to the roles described above, it can be argued that in the United States, private standards also fulfil the role of promoting production practices in the interest of improving animal welfare. This goal is communicated in many of the mission statements of private standards organizations. Some examples are listed below:

“Humane Farm Animal Care (HFAC) is the leading non-profit certification organization dedicated to improving the lives of farm animals in food production from birth through slaughter. The goal of the program is to improve the lives of farm animals by driving consumer demand for kinder and more responsible farm animal practices.” (Certified Humane, 2018).

“We are a non-profit organization of farmers, scientists, retailers, manufacturers, and animal advocates, all collaborating with a common purpose: to improve farm animal welfare.” (Global Animal Partnership, 2018).

Both of the above are examples of private standards, but can also be described as ranking standards, since part of their goal is to provide a stamp of approval that establishes superiority, in terms of animal welfare practices, in comparison to similar non-certified products. In addition to private standards run by independent certifiers, industry groups in the United States, including United Egg Producers (UEP) and The National Chicken Council (NCC) have also developed standards or guidelines that can be voluntarily adopted by producers. The UEP certification

program, UEP Certified, offers third-party audit certification for egg producers using both cage-based and cage-free systems. UEP Certified provides the following description of the program on their website:

“Egg farmers are committed to providing the best care possible for their hens. To demonstrate this commitment, United Egg Producers (UEP) developed guidelines for optimal hen well-being – guidelines that are backed by decades of research and recommendations from an independent Scientific Advisory Committee. UEP Certified established guidelines for conventional cage housing in 2002 and for cage-free in 2006.” (UEP Certified, 2018).

Despite the fact that the standards created by these different organizations would be similarly classified per the categories previously described (they are all private standards with filter functions), the subtle differences in the way these organizations describe their mission and that of their standards provides some insight into how the goals of organizations are applied to standards, and ultimately to the whole ecosystem that they exist in. Furthermore, it is important to keep in mind that these private standards are applied in the context of what are already highly standardized industries. Therefore, industries participating in private animal welfare certification programs often can run into conflicts at different levels, such as value conflicts between their organization and the certifier, and practical conflicts stemming with the incompatibility of their internal standards, those of the certifier, and those of their clients, among others.

4.3 Standards and their ecosystems

Regardless of the type of standard applied, Busch (2000) argues that standards establish a moral economy, not only by defining what is acceptable or unacceptable in terms of tangible and intangible attributes, but also by excluding and disciplining people and things that fall outside of the established norm. Furthermore, the reach of standards goes beyond that which is being standardized; by establishing a standard for a product, we are also indirectly standardizing the

environment in which it is produced and sold, including infrastructure, workers, distributors, and consumers. According to Busch (2000), even if what is being standardized is generally considered a “thing”, the process also results in the standardization of environments, workers, and standards themselves. Each of these are discussed below, using the example of broiler production in the United States.

With regard to “things”, standardization was largely a byproduct, if not a requirement, of industrialization, although it had been a defining element for the military long before the industrial era (Busch 2010). During the rise of the industrial era, the transformation of markets was one of the most influential factors in the creation and implementation of standards. During the early 1900s, national standards organizations were founded, and the exponential increase in mass production of consumer goods fostered the accelerated proliferation of standards (Busch 2010). The broiler industry was a relative latecomer to industrialization and standardization. Although the rise of the broiler industry as we know it today is often attributed to the increasing demand for affordable protein during and after World War II, this industry had been already growing steadily, especially in the southern states, where it came to replace the failing cotton industry (Gisolfi 2017). However, the rationing of red meat by the US government in 1942 provided a final push toward rapid industrialization and growth; the army became the largest buyer of chickens, and demanded standardization, inspection, and increased production (Gisolfi 2017).

In the broiler industry, the need to standardize eventually included the genetic makeup of the birds; as production became more integrated, poultry feed distributors acquired ownership over all other parts of the process, and the farmers became little more than hired hands. Integrators signed “feed conversion contracts” with farmers (now known as “growers”), whereby

growers were paid for pounds of chicken produced per pound of feed, and included a “tournament system” that rewarded growers for greater feed conversion rates (Gisolfi 2017). Of course, it was in the best interest of integrators to retain as much control as possible, and one way to that was to make the birds as uniform as possible.

The genetic makeup of modern broilers was the result of a contest held in 1947 by the Great Atlantic & Pacific Tea Company (A&P), the country’s largest poultry retailer, in partnership with USDA (Wiehoff 2013). The contest- named “The Chicken of Tomorrow”- resulted in the modern broiler genetics that dominate the market today (Wiehoff 2013). The winning breed or hybrid was from Arbor Acres, a family operation that was eventually bought and made global by Nelson Rockefeller (Wiehoff, 2013). Arbor Acres was later acquired by Aviagen, one of the largest global poultry genetic companies (Wiehoff, 2013). In 2017, Aviagen merged with one of its largest competitors, Hubbard; as a result, it is estimated that Aviagen now has a share of 75-90% of the chicken parent stock globally (Competition and Markets Authority 2018).

While economic interests were the main driver of the extensive standardization of the broiler industry, from a practical perspective, chickens need to be extremely uniform in terms of size and conformation in order to maintain the efficiency of the systems designed to process them. All parts of the system, from hatcheries, to feeders, transport systems, and slaughter lines are designed with this uniform bird in mind. Failures in uniformity not only compromise the efficiency of the system, but may actually result in compromised welfare for the animals. For example, during the stunning process, which commonly involves an automated line that takes the birds through an electrified water bath, smaller birds may fail to be adequately stunned because their heads are not completely immersed. Similarly, birds who are too large may have their

wings enter the water bath before their heads do, suffering painful electric shocks that do not result in loss of consciousness (Humane Slaughter Association 2016). Thus, within this highly standardized system, even slight deviations from the standard can result in serious suffering.

As Busch (2000) points out, the standardization of commodities and the systems within which they are produced necessarily entail the standardization of workers, whose jobs are often repetitive and monotonous, and amount to little else than machine tending. In the poultry industry, jobs that previously required some level of specialized knowledge regarding the biology and welfare of animals are now limited to following standard operating procedure manuals. Within the industry, slaughter plant work is among the riskiest, worst paid, and not surprisingly, presents some of the highest turnover rates. A 2012 report on working conditions of poultry slaughterhouse workers states that this sector employs close to 225,000 workers, a majority of whom are immigrants, and who are paid on average forty percent less than workers in other manufacturing sectors (Alexander 2012).

According to Busch (2000), the standardization of workers has two dimensions: uniformity, and discipline. In the poultry industry, uniformity is established through systematic de-skilling of workers; all are able to perform a series of non-specialized tasks, thus becoming highly replaceable. Discipline is tied to de-skilling and replaceability; workers who put themselves at risk, either because of carelessness or exhaustion, are considered liabilities and replaced. But also, the fast pace of the work creates stringent, often impossible, disciplinary expectations. In 2017, the US Government Accountability Office (GAO) released a report on workplace safety and health of workers in meat and poultry processing plants (Better Outreach, Collaboration, and Information Needed to Help Protect Workers at Meat and Poultry Plants). The report highlights, among other concerning insights, that conditions for workers are not only

dangerous because of the nature of the work, but may compromise basic workers' rights, such as provision of necessary breaks (US GAO 2017). On poultry slaughter lines, chickens are processed at an average speed of 140 birds per minute, and the industry has requested permission to increase line speeds to 175 birds per minute. In these conditions, workers cannot walk away from the line without putting themselves and the animals at risk. A common policy is that if a worker walks away from the line without having someone replace them, this is considered an automatic "quit" (Alexander 2012).

The issue of de-skilling workers in standardized environments is related to the final form of standardization discussed here: the standardization of standards themselves. According to Busch (2000), one of the ideas behind the development of standards is to act as a substitute for human judgment, in other words, they aim to create conditions in which the human element is removed from the production and objective assessment of goods and processes. The vertically integrated broiler industry in the United States is a prime example of this. By removing variability to the greatest extent possible, from the genetic makeup and anatomical characteristics of chickens, to the feed, the harvesting process, and finally, slaughter, processes have been made uniform to the extent that humans are almost an accessory. Uniformity is key to the efficiency that characterizes the industry, and that has led to its exponential growth. The core value of efficiency is one of the keys to standardizing the standards themselves; it serves as both a starting point and an end-goal for all standards within the system.

4.4 Private farm animal welfare standards in the United States: motivations and development

In the United States, private, non-regulatory, or voluntary farm animal welfare standards, whether developed by industry or by private certifiers, have very different origin stories from the industry standards described in the previous section. However, because these voluntary animal

welfare standards apply to industrial animal production, they are necessarily part of the same ecosystem of industry standards, and as such, they must be considered with that context as a backdrop. Note that my use of “industry standards” here refers to standards for production practices in the industry, and not to animal welfare standards established by industry groups. For the purposes of this work, I will consider both industry and private standards for farm animal within the same category (voluntary or non-regulatory animal welfare standards, henceforth “animal welfare standards”), described in this section.

While industry standards emerged largely in response to a need for increased efficiency and production volume consistent with the values of industrialization, private animal welfare standards in the United States appear more recently, and have a more diverse set of motivations. First, it can be argued that the development of these standards follows the growing global interest in animal welfare. This growing interest is patent in many spheres, including higher education and academic research. According to Mench (2008), in the wake of the Brambell Report, the number of course offerings focused on animal welfare at both the undergraduate and graduate level increased significantly; similarly, there has been a steady increase in academic research on animal welfare, fueled by the availability of funding from both industry and government.

Second, the lack of legislation on farm animal welfare may also be a motivation for the emergence of animal welfare standards. Animal welfare legislation has seen rather little development in the United States, especially compared to Europe. At the time of publication of the Brambell Report, there were two laws pertaining to livestock in the United States, the Twenty-Eight Hour Law, which was passed in 1873 to protect livestock during transport to slaughter, and the Humane Methods of Slaughter Act (HMSA), originally passed in 1958, which mandated that livestock be stunned prior to slaughter (Mench 2008). Today, these are still the

only two laws applicable to livestock, and both specifically exclude poultry, which constitute the vast majority of farmed land animals.

A third possible motivation for the emergence of animal welfare standards is consumer demand. With the exception of animal welfare standards developed by industry groups, private standards are seen as catering to niche markets (Mench 2008). This may be true in the sense that private certification almost always results in a price premium due to the requirements of the certification standards, which in many cases increase the cost of production. However, the number of consumers concerned about animal welfare continues to grow, even if this does not always translate to increased purchases of certified or higher welfare products. A consumer survey commissioned by the National Chicken Council in 2017 reported that forty-three percent of surveyed consumers were concerned about how chickens were bred to optimize meat production, while forty-one percent and forty percent reported concerns about how chickens are housed, and how they are raised, respectively (Graber and Keller 2017).

4.5 The role of animal welfare concepts in the development of standards

As laid out in previous sections of this chapter, the final form of a standard, which may simply seem like a laundry list of requirements, embodies histories and motivations that may not be evident, but that impinge on its goals and potential within the broader landscape of industry practices, market dynamics, and consumer preferences. If efficiency and profit are at the core of standards for industry practices, the desire to improve animal welfare is arguably at the core of animal welfare standards. However, the ubiquitous disagreements on how to define animal welfare, which of its elements to prioritize, and ways to assess it are also patent in the development of private animal welfare standards. Fortunately, the practical constraints of

developing a standard that can be applied to production conditions somewhat limit the range of concepts and measures that can be included.

Mellor (2015) proposes three categories of animal welfare standards, which roughly align with the three functional areas of animal welfare previously discussed: baseline reference standards, natural living standards, and preference standards. Baseline reference standards, which include most current standards, are those requiring small, incremental improvements in order to meet the basic physical and physiological needs of animals (Mellor 2015, 18). A second group of standards are those considering the promotion of natural living as the basis of good welfare (Mellor 2015, 18). Under these standards, we should strive to provide for animals conditions such that they may experience the full range of welfare states that are potentially available to them; these conditions are informed by knowledge about the natural, wild, or ancestral environments of these animals (Mellor 2015, 18). The third category of standards draws from preferences, or “what animals want”, as Mellor puts it. This approach involves using scientific approaches to identify resources and conditions that animals would choose themselves, purportedly because they are conducive to positive welfare states, and using that information to implement changes and eventually standardize them (Mellor 2015, 19).

Finally, we can think of animal welfare standards as relying on animal-based measures or environment-based measures (or inputs). For example, Mench (2003, 498-500) describes animal-based performance standards (i.e. what the animal should be able to do, such as stand, turn around, groom), as well as engineering standards (describing the requirements of the physical environment for a particular species, for example, the amount of floor space, or the provision of enrichments). Mench (2003, 500) also notes that environment-based standards are often easier to implement and enforce across large numbers of facilities, while performance standards can be

challenging to measure and entail greater investment in terms of time and training of professional auditors.

4.6 Ethical and scientific justifications for animal welfare standards

Philosophically, we can consider justifications for animal welfare standards in terms of how well various ethical theories account for the need to create and implement standards. For example, Daigle (2014, 640) considers whether the use of automatic milking systems (for which undoubtedly there are standards) would be justified under a utilitarian ethic, since these systems allow the animal to choose when to be milked, therefore respecting their interests. But, as Daigle contends, even if the interests of the animals are being considered, many other competing interests (such as those of the farmer) are at stake, and perhaps not all interests are being considered under the equality principle which utilitarian theory requires (2014, 640).

Furthermore, since one of the main purposes of animal welfare standards is to regulate our use of animals in the context of animal production, other ethical issues may come into play, such as those pertaining to food safety regulation and public health. When there is doubt regarding whether and how different interests should be weighed, or whether or not an animal or group of animals is of moral concern in a way that warrants regulations and standards, ethics can only get us so far in resolving the issue, hence the need for scientific evidence that is acceptable to the different parties involved regardless of their ethical stance on the matter. However, as previously discussed, scientific assessment of welfare also presents limitations in terms of the sufficiency of evidence, and the management of inductive risk in the face of uncertain conclusions. Plausibly, scientific justifications have more traction because of the authority of science, however, in some cases, ethical concerns and justifications prevail. In an analysis of the EU standards for housing of laying hens, Savory (2004, 153) points out that their implementation

was influenced more by public opinion than science; cases like this are an important reminder that ultimately, animal welfare is a social cause, and thus animal welfare a socially mandated science.

4.7 Conclusions

As discussed, the process of defining and applying animal welfare measures requires navigating differences of opinion among scientists, ethicists, industry, consumers, activists, and other stakeholders. Even after long negotiations, the resulting changes can be minimal despite requiring great investment, discouraging those who supported their implementation, and vexing those who must absorb the cost despite their personal opinions on the matter. In that sense, some animal welfare advocates, such as Harrison, argue that accepting small changes can delay and diminish the prospect of achieving meaningful compromise (Sandoe et al 2003, 475). Mellor (2015, 18) agrees that if we continue to rely on minimal baseline reference standards, we are at risk of regarding small improvements as being more important than if we had started with a higher standard. On the other hand, producers, retailers and consumers who are faced with absorbing the cost of implementing new standards may seek ways to avoid doing so; animal welfare standards are still largely a local strategy in an increasingly global market, and similar products produced at lower standards (and lower cost) may become a viable alternative to those who do not wish to bear the cost. For example, Grethe (2007) has discussed the issue of whether or not the implementation of high animal welfare standards in developed countries will lead to the proliferation of “low animal welfare havens” in developing countries.

Furthermore, it is worth considering whether the ways we are collecting and analyzing information on animal welfare standards are useful in promoting further changes and improvement. Since improvement depends in part on this information and feedback, faulty or

incomplete interpretations could lead us in directions that will ultimately not result in the improvement of *de facto* welfare (the actual state of the animal), but only in better readings of welfare indicators or perceived welfare.

The formulation, implementation, and improvement of animal welfare standards is strongly driven by both public concern for animal welfare and the continued development of animal welfare science. Animal welfare measures that allow us not only to quantify basic welfare states, but also to aggregate and analyze data in integrative ways could lead to a better understanding of the overall welfare and quality of life of an animal; but, using standards to determine whether an animal has “a life worth living” may still prove difficult, to say the least, and may demand more than just better science. The continued advancement of research methods and technologies that allow us to better understand animals from their perspective can do much to improve the precision and accuracy of standards and welfare assessment, but the decision to implement these standards ultimately lies in the hands of those who hold social and economic power over the lives of farmed animals. Furthermore, every step in the process of justification, development, implementation, and assessment of a standard involves tradeoffs, which are sometimes genuinely justified by science or ethics, but other times are simply the effect of practical or political constraints. While animal ethics and animal welfare science have long coexisted, and positions and arguments from one are often invoked by the other, there is a need for a deeper analysis of where gaps between ethical and scientific arguments for animal welfare exist. This is especially true in the context of developing standards and policies, where both ethical reasoning and scientific evidence are used to justify conditions and practices for the use of animals.

Chapter 5: Practical perspectives on the development and implementation of animal welfare standards

5.1 Introduction

When the Brambell committee set about the task of creating the groundwork for the science of animal welfare, one of the expected outcomes was that this emerging field would lead to a better understanding of how the welfare of animals is influenced by different production methods (Fraser 2008, 65). This increased knowledge would, in turn, result not only in better welfare for the animals, but also in improved effectiveness of the newer, industrial models of animal production (Fraser 2008, 65). As argued in the previous chapter, one of the roles of standards is to translate and integrate the motivations, objectives, and values of animal welfare science and animal production into tools that set targets and provide a roadmap for achieving those targets. Thus, animal welfare standards are the junction at which the values embedded in animal welfare science, ranging from the social and ethical values that drove its emergence, to the scientific (cognitive) values of its contributing disciplines, converge with the values of producers, industry, consumers, and other downstream stakeholders.

In order to examine the use of evidence in animal welfare standards, it is important to recall the ultimate goals of animal welfare science, or at least what they are perceived to be by stakeholders. Animal welfare scientists, producers, and animal welfare standards professionals may have different perceptions of what those goals are, and as such, there are discrepancies as to what kind of evidence is most relevant, as well as how it should be used and analyzed in order to best achieve those goals. While standards can play the role of integrating and synthesizing information from all of these sources into a coherent, practical instruments, it is important to keep in mind that standards themselves may be value-laden. Divergent values may manifest as differences in the goals of standards that apply to the same kind of practice, such as animal

production. Relevantly, the process for developing private animal welfare standards, at least in the United States, is not standardized. As such, the organizations who create, implement, and market these standards (to both producers and consumers) have the freedom to decide what their overarching goals and priorities are, how information will be collected, analyzed and used, and ultimately, how the standards will be applied and enforced .

Because one of the objectives of this project is to make practical recommendations regarding the use of scientific evidence in balance with social concerns in animal welfare standards, it was considered important to provide detailed examples of private animal welfare standards, as well as practical insights into the process of developing standards, provided by individuals with relevant practical experience in the development and application of these instruments. For this purpose, four people with extensive experience on the topic of voluntary farm animal welfare standards were interviewed using an open-ended, semi-structured approach. The two private US animal welfare standards described are the United Egg Producers Guidelines for Commercial Laying Hens, and the Global Animal Partnership Standard for Chickens Raised for Meat. These examples serve as a backdrop for discussion of the insights obtained from the interviews. This chapter provides an analysis and discussion of these practical insights in the context of the theoretical considerations addressed in previous chapters.

5.2 Methods

The standards selected as examples are considered representative of private animal welfare standards in the US. Another criterion for selection was transparency, in other words the availability of publicly accessible information regarding their development, as well as the instruments themselves. Interview participants were selected based on their knowledge and practical experience in the development of farm animal welfare standards in the United States.

Efforts were made to select interviewees who represented different interests or stakeholder groups, including scientists, industry, and animal advocates. Four people agreed to be interviewed: Mr. Chad Gregory, President and CEO of United Egg Producers (UEP), Dr. Joy Mench, Professor Emerita of Animal Science at the University of California Davis and Chair of the UEP Scientific Committee, Ms. Anne Malleau, Executive Director of Global Animal Partnership (GAP), and Ms. Adele Douglass, founder and CEO of Humane Farm Animal Care (Certified Humane Program). For practical purposes, in the section of this chapter that discusses the interviews, quotes and insights are attributed using the interviewees' initials rather than their full names. More information about each of the interviewees and the organizations they represent is provided in section 5.4.1.

A study proposal for the interviews was prepared and submitted to Michigan State University's Human Research Protection Program Institutional Review Board. Once the study was approved, interviewees were provided with a Minimal Risk Informed Consent document. Interviews followed an open-ended, semi-structured format; an interview guide was prepared in advance to ensure that all relevant topics were addressed. Topics outlined in the interview guide included: a) Personal experiences relevant to this research; b) Perceptions of animal welfare science; c) Standards, science, and the public. Interviewees did not have access to the interview guide. One of the interviews (Mr. Chad Gregory) was conducted in person; the other three were conducted by phone. The duration of the interviews was between 40 and 60 minutes.

5.3 Examples of private animal welfare standards in the US

5.3.1 United Egg Producers Guidelines for Commercial Laying Hens

In 1999, United Egg Producers (UEP), a cooperative of egg farmers, representing ownership of close to 95% of US laying flocks, commissioned an independent Scientific Advisory Committee for Animal Welfare¹³; the committee was asked to review the scientific literature on specific topics relevant to the well-being of laying hens, to identify areas where further research was needed, and to develop recommendations based on their findings (UEP 2017). The committee presented the recommendations in 2000, and UEP launched the “UEP Certified” program in 2002 (UEP 2017). Since at the time the Committee was convened 95% of egg production in the U.S. came from caged hens, the initial guidelines were developed with these animals in mind, however, in 2006, UEP asked the Committee to develop recommendations for cage-free hens as well (UEP 2017). In addition to the provisions related to housing, the UEP guidelines also contain detailed requirements for beak trimming, molting, handling, catching and transport, euthanasia and depopulation, biosecurity and animal health, and reporting animal cruelty and neglect.

The UEP recommended space allowance for caged laying hens is 67- 86 in² (432- 554 cm²) of usable space per bird¹⁴, with the slope of the cage floor not exceeding 8 degrees, such that birds should be able to stand comfortably upright (UEP 2017). Additionally, for birds older than 18 weeks, each bird should have at least 1 inch (2.54 cm) of linear feeding trough space,

¹³ The UEP Scientific Advisory Committee for Animal Welfare is composed of ten members, eight of whom are academic researchers, one industry representative and one representative of a professional body (American Veterinary Medical Association). All are experts in different disciplines relevant to animal welfare. The list of committee members and their qualifications is publicly available on the UEP Certified website (<https://uepcertified.com/scientific-advisory-committee/>)

¹⁴ The UEP guidelines include breed-based differentiations with regard to space allocation. For all housing equipment installed after August 1, 2008, the minimum space allowance is 67 in² (432 cm²) for White Leghorn hens, and 76 in² (490 cm²) for Brown Egg Layers (UEP, 2017).

and there should be one drinking nipple or cup per every 12 birds (UEP 2017). Furthermore, the UEP guidelines establish provisions for lighting intensity and air quality (UEP 2017).

The UEP guidelines for cage-free laying hens allow three kinds of housing systems: multitier aviaries (stacked systems consisting of tiered or A-frame structures integrating perches, feeding lines, drinkers, nests, and sometimes platforms and/or ramps), partially slatted systems (single tier systems in which part of the floor is covered with a slatted surface, and part is a litter-based surface), and floor systems (barns with floors completely covered in litter) (UEP 2017). For systems with partial litter coverage, at least 15% of usable floor area should be covered with litter (UEP 2017). The recommended space allowance for cage-free laying hens is 1-1.5 ft² of usable floor space per hen (929- 1392 cm²) (UEP 2017). Additionally, each bird should have at least 1.5 inches (3.81 cm) of linear feeding trough space, and birds should not be more than 26 feet from a feeder or drinker (UEP 2017). Cage-free systems must also provide nests and perches (not provided or required in battery cage systems). UEP guidelines require 9 ft² (8361 cm²) of nest space per every 100 hens, and 6 inches (15.24 cm) of linear perch space elevated at least 16 inches above floor level for at least 20% of the total number of hens in the flock (UEP 2017). Furthermore, the UEP guidelines establish provisions for lighting intensity and air quality (UEP 2017). As mentioned, the guidelines also establish minimum requirements for cage-free egg production, but have yet to make recommendations for enriched colony housing systems (UEP 2017).

Compliance with the UEP guidelines is achieved through a process of certification and periodic audits (UEP 2017). For initial certification, producers must submit an application and pay the required fees; once the application is accepted, an auditor visits the production facilities, and upon successfully passing the audit, the producer receives a certificate that is valid for one

year (UEP 2017). UEP works with independent auditors from the USDA Agricultural Marketing Service, or Validus, a private animal welfare and food safety auditing firm (UEP 2017).

5.3.2 Global Animal Partnership Standard for Chickens Raised for Meat

Global Animal Partnership (GAP) is a non-profit charitable organization founded in the US in 2008 (Duncan et al 2012). Its main purpose is to offer a private animal welfare certification system, unique within the range of similar certifications by virtue of its tiered approach. GAP's trademark 5-Step[™] Animal Welfare Rating Standards Program offers producers the possibility of becoming certified at six different levels (Step 5+ certification is an addition to Steps 1-5) (Duncan et al 2012; GAP 2018). Each of the five steps is identified by its defining characteristics: Step 1 centers on no crates, no cages, and no crowding; Step 2 on enriched environments (referring to behavioral enrichments); Step 3 on outdoor access (seasonal or temporary); Step 4 on pasture-centered systems (year-round access to pasture and/or the outdoors), Step 5 on animal-centered systems requiring no physical alterations (such as tail-docking, teeth-clipping, or beak-trimming), and Step 5+ on animal centered-systems with whole life on farm (i.e. born, raised, slaughtered and processed on one certified farm) (GAP 2018).

The reasoning behind the tiered approach is to allow a broad range of producers to participate in the program, with the option to remain at the step at which they are certified, or to upgrade to the next step through meeting additional requirements (Duncan et al 2012). Since in most cases, animal products from systems with a higher step certification are more expensive, this system also gives consumers the choice of products with a range of different prices, while also providing them with the necessary information (through labeling) about the attributes of different production systems (Duncan et al 2012).

GAP's process for creating standards includes research and development (conducted in-house), external review by experts and producers, an open comment period, revision, field testing, and finally, approval by GAP's board (GAP 2018). GAP does not have a fixed scientific advisory committee, but rather relies on a roster of experts selected according to their expertise; experts include not only academics, but also producers from different stages of production and different geographies (if applicable) (Anne Malleau, personal communication 2018).¹⁵ The GAP board of directors, who provide final approval of the standards, is composed of nine members; the current board is constituted by three producers, four representatives from non-profit animal welfare organizations, and two food industry representatives (GAP 2018). Board member information is publicly available on GAP's website (GAP 2018).

The standard for meat chickens (broilers) was launched by GAP in 2008, along with standards for pigs and beef cattle. Since its launch, the standard has been revised two times. Its current version, v3.1 was issued in April 2018, replacing an earlier iteration of version 3 (v3.0) issued in July 2017 (GAP 2018). The current version is expected to be revised after four certification cycles from date of issue (each certification cycle is 15 months). The standard includes very detailed and specific requirements organized into ten areas: 1) animal source (genetics) and health, 2) animal care and management, 3) feed and water, 4) housing, 5) outdoor conditions, 6) loading, 7) transport, 8) rodent, wild bird and predator control, 9) protocols, record-keeping and training documents, and 10) slaughter (GAP 2018). While the standard is much too detailed and lengthy to describe here in its entirety, two excerpts are provided below for the purposes of illustrating the structure, language, and specificity of the instrument (GAP 2018).

¹⁵ GAP does not disclose the names of their scientific and expert advisors, either on their website or in the standards documents (Anne Malleau, personal communication, 2018).

“1.2 Growth Rates

Producers are not required to weigh their chickens– compliance is currently determined by GAP from the breeder’s published data. See Appendix I for a list of breeds/strains, their average growth rate determined from the breeder’s published data, and the maximum step level they can achieve. Any operation with questions around acceptable breeds/strains should contact GAP directly prior to scheduling their audit.

1.2.1 The maximum average growth rate must not exceed 68 grams (0.150 lbs.) per day (Steps 1-3).

1.2.2 The maximum average growth rate must not exceed 50 grams (0.110 lbs.) per day (Step 4).

1.2.3 The maximum average growth rate must not exceed 45 grams (0.099 lbs.) per day (Step 5).

1.2.4 The maximum average growth rate must not exceed 35 grams (0.077 lbs.) per day (Step 5+).”

“4.6 Stocking Density

Stocking density must be calculated using the final target weight of all chickens per flock per area at the time of catching plus a maximum of 4% variation to calculate the number of chickens to be placed within a given barn/house. Expected livability must not be included in the calculation.

4.6.1 Chickens must have enough space to express natural behavior, including standing, turning around, and preening, without touching another chicken. (Steps 1-5+)

4.6.2 For birds placed from 1 January 2018 to 30 June 2020, stocking density must not exceed 6.5 lbs./ft² (32 kg/m²) For birds placed from 1 July 2020 onwards, stocking density must not exceed 6 lbs./ft² (29 kg/m²) (Steps 1-2)

[4.6.2 a]: GAP recognizes that Step 1 and 2 operations need time to adjust bird numbers, egg sets and potentially build new housing in order to meet the stocking density required from 1 July 2020 onwards. Operations should use the time allotted to make progress towards meeting the lower stocking density across all their Step 1 and Step 2 housing.

4.6.3 Stocking density must not exceed 6 lbs./ft² (29 kg/m²) (Step 3)

4.6.4 Stocking density must not exceed 5.5 lbs./ft² (27 kg/m²) (Steps 4-5+)”

The process for becoming certified includes a pre-audit preparation process, submitting an application in which the producer must specify which step level they wish to become certified in, scheduling an audit (with audit cycles being 15 months) (GAP 2018). Once an application is submitted, the next step is to schedule an audit. Audits are conducted by third party providers, and results then submitted to GAP for certification (Anne Malleau, personal communication 2018).

5.4 Interviews

5.4.1 Background of interviewees and organizations

As previously stated, one of the criteria used for the selection of participants was their affiliation with stakeholder groups that represent different perspectives within the sphere of farmed animal welfare standards in the United States. Background information on the interviewees, the organizations they represent, and their roles within those organizations is provided here as a backdrop for the insights gained from the interviews.

The first interview conducted was with Mr. Chad Gregory, President and CEO of UEP. Mr. Gregory has served as President of UEP since 1999, and was preceded in this role by his father, Gene Gregory. Thus, his involvement with UEP, and with the US egg industry in general is long-standing; prior to assuming his current role, he was UEP's Director of Membership, a role that allowed him to be in close connection with producers and their operations. UEP was established in 1968 as a federated Capper-Volstead Agriculture Cooperative comprised of five regional marketing cooperatives (UEP 2018). The Capper-Volstead Act (P.L. 67-146), also known as the Cooperative Marketing Associations Act (7 U.S.C. 291, 292) was adopted by the United States Congress on February 18, 1922, and it exempts agricultural associations from some of the constraints of antitrust laws, thus allowing members to collectively market their products (USDA Rural Business and Cooperative Development Service 1995). In contrast to other animal industry groups that do not require their members to be producers (for example, the National Chicken Council), UEP is a producer-only organization, and, according to Mr. Gregory, includes approximately 90% of all egg producers in the United States. The activities of the organization cover a broad range of topics, including production, biosecurity, labor, environment, food safety, and legislation (UEP 2018). Specific activities or programs relevant to animal welfare reported by UEP on their public website include: a) establishing the Egg

Nutrition Center, a credible research, science and education division of the American Egg Board (AEB); b) leading in standards for hen well-being with the launch of UEP Certified (2002) and UEP Certified Cage-Free (2006); c) Supporting various research initiatives on food safety and hen well-being, such as the Coalition for Sustainable Egg Supply; and d) facilitating effective dialogue about transition in hen housing.

The second interview was with Dr. Joy Mench, whose main roles include academic research and scientific advising for industry groups, including UEP. Her academic roles at University of California Davis included joint appointments in the Department of Avian Sciences and the Department of Animal Science, and Director of the Center for Animal Welfare at the UC Davis (she is currently Professor Emerita) (UC Davis 2016). During her extensive research career, she has authored over 150 papers on animal behavior and welfare, the majority of which focused on poultry. Her research on poultry includes studies on handling, transport, husbandry practices, as well as multi-disciplinary projects on the sustainability of alternative housing systems for laying hens and broilers. As part of her research on this last topic, she was one of the principal researchers for the Coalition for Sustainable Egg Supply, a multi-stakeholder commercial scale project that compared various laying hen housing systems and their potential impacts on food safety, the environment, hen health and well-being, worker health and safety and food affordability (Coalition for Sustainable Egg Supply).

The third interviewee, Ms. Anne Malleau, currently serves as Executive Director for Global Animal Partnership (GAP). Ms. Malleau has a professional background in animal science, applied ethology, and business management. Prior to her current role with GAP, she worked with Whole Foods Market in several capacities, including Global Animal Production and Welfare Coordinator (Whole Foods 2012). Ms. Malleau was one of the main drivers behind the

creation of the GAP 5-Step (™) Animal Welfare Rating Standards, a US-based farm animal welfare certification program that is unique in its tiered approach to animal welfare rating and certification (Duncan et al 2012). In partnership with Whole Foods Market, GAP piloted its 5-Step (™) program in 2008, with standards for production and transport for broilers, pigs, and beef cattle (Duncan et al 2012). The program has since expanded to include standards for other species, including turkeys, lambs, goats, bison, and laying hens (GAP 2018). As of December 2011, GAP reported that 1,740 farms and ranches, producing close to 140 million animals, were participating in their 5-Step (™) program (Duncan et al 2012). According to Malleau, in 2017, the number of animals impacted by the program was close to 290 million.

The fourth and final interview conducted was with Ms. Adele Douglass of Humane Farm Animal Care (HFAC), the organization responsible for the Certified Humane certification program. Ms. Douglass, current CEO, founded HFAC in 2003. She has a background in legislation, having worked for former U.S. Representative Bill Green (R-NY). Her experience in farm animal welfare prior to founding HFAC includes the creation of the Free Farmed Program for Farm Animal Services, and working as director of the American Humane Association's Washington, D.C., office for 13 years (HFAC 2018). Outside of her work with HFAC, she serves on several industry animal welfare committees including the Food Marketing Institute, National Council of Chain Restaurants, and Burger King. The Certified Humane program includes standards for beef cattle, broilers, laying hens, dairy cows, goats, pigs, sheep, turkeys, and bison (Certified Humane 2018). The program certifies operations in the United States, Australia, Brazil, Canada, Chile, Hong Kong, and Peru (Certified Humane 2018). The overarching requirements of the standards for all species include: meeting the standard from birth through slaughter, no cages, crates, or tie stalls, freedom to "do what comes naturally" (for example,

wing flapping and dustbathing in chickens, and ability to root in pigs), and quality diets without animal by-products, antibiotics or growth hormones (Certified Humane 2018).

5.4.2 Personal experience and views on farm animal welfare standards

This section summarizes and discusses aspects of the interviews focused on the participants' personal experiences and views of animal welfare standards. One of the themes that was central in all of the narratives of personal experiences working with standards was the central role that experts played in motivating the development of animal welfare standards within each of these organizations. For example, when UEP was faced with the task of updating their guidelines in 1997, their approach was to bring together a panel of experts who could advise them in this process (CG). A similar path was followed by AD when she decided to start the Certified Humane program; she stated that one of the first things she did was reach out to high profile animal welfare scientists; Humane Farm Animal Care currently has a committee of forty international experts (AD). AM mentioned that an inherent strength of the GAP team is that they all have graduate level degrees in animal welfare, with most being classically trained in animal science and welfare. She believes that for her organization, having this expertise in-house is a valuable way to establish credibility.

While the industry itself has a substantial amount of expertise internally, CG stated that perhaps the reason they sought advisors in academia is the perception that research coming from or funded by industry is less trustworthy. Thus, for UEP, for instance, the scientific advisory committee is not only a point of pride, but a safeguard of sorts against the view that industry research is not reliable. This view was echoed by JM, who stated that in her opinion, UEP's scientific committee is a good example of how experts can work with industry in an independent and transparent way; for instance, she mentioned that an important feature of the

UEP committee is that the chair selects and appoints all of the members, independently from any considerations or recommendations from UEP's leadership.

A second theme that emerged in this part of the interviews was the view that in order to be useful, animal welfare science must have practical applications. A simple interpretation of this, and one that most of the interviewees seemed to point to, is that practicality has to do with the identification and solution of problems. At a deeper level, it's plausible to think that in commercial production settings such as those these standards apply to, something is primarily identified as a problem if it primarily affects productivity (or relatedly, profitability). Although from a scientific perspective it is generally understood that productivity alone is not an adequate measure of welfare, there are often strong correlations between the improvement of welfare and the improvement of productivity. Thus, it is valid to ask whether producers are more likely to identify as problems to be solved by animal welfare scientists those that affect productivity (a problem that could be defined as "productivity infiltrates welfare").

JM's perception of her role as an expert seems to agree with this; she stated that she has always seen her research as problem oriented, sometimes to determine if something is really a problem, and other times working on solutions to problems that have already been identified. AM mentioned that the aspect of practicality was something that motivated her to pursue a degree in business administration after completing her Master's degree in applied ethology; she saw this as an opportunity to learn how to apply her expertise, rather than continuing to deepen (and possibly narrow) it. Through practice and application, AM stated, there are things that she would not have learned otherwise, such as understanding that engagement with diverse stakeholders is multi-tiered; there is some level of feedback that you can get in meetings or over the phone, but robust knowledge of the systems she works with can only be gained by going out

to the farms, and engaging with the farmers and their networks. AD shared a related experience, saying that as a person who grew up in New York City, she had a vision of how farm animals should be treated, but it was not close to reality. For AD, the journey of learning about farm animal welfare standards in practice started with visiting the RSPCA and learning about their Freedom Foods program, which eventually led to her starting Certified Humane in the United States. From these various perspectives, it can be concluded that in the context of collaborations between scientists, producers, and standards organizations, the value or attribute of “practicality” is understood to imply problem-solving, the development of broad and applicable knowledge, and approaches grounded in experience.

5.4.3 Perceptions on the role of standards and animal welfare certifications

From the perspective of industry, CG sees standards as an imperative for moving the industry forward, in the sense that they are required by clients (not necessarily by the final consumer, but by retail and food business clients). JM agrees with the notion that standards, even when voluntary, are becoming a requirement in order for certain animal industries to keep their clients. She believes that many of these standards are not completely voluntary, but "quasi voluntary", because once certification is demanded by a retail customer as a condition for business, producers have little choice but to comply. Regardless, AD mentioned that the fact that they are voluntary is important; this puts some power back into producers' hands, and they can make the choice rather than having it directly imposed upon them by political mechanisms. JM also sees value in voluntary standards beyond how they may affect the producer-client dynamic. For one, in the context of the United States, where regulatory standards are limited, the role of non-regulatory animal welfare standards is essential, as they are, at the moment, the only relatively transparent way of tracking animal welfare improvements. Meanwhile, AD believes

that despite putting pressure on producers, the implementation of voluntary standards, if they accomplish their goal of making the animals more productive and healthier, producers will see that the standards make a difference; when animals are not stressed production tends to be better.

CG and AM both talked about specific risks of the voluntary standards system. CG mentioned that for UEP specifically, in their condition of being a producer group, there are legal liabilities related to the perception that if their members (who are the majority of egg producers in the United States) all agree to be certified by UEP, then they are in breach of antitrust laws, and can be the object of a lawsuit. Another specific risk mentioned by AM is that retailers often have very real concerns about their bottom line related to how much demand there will be for a certified product once both producer and retailer have made the necessary investments to offer it to consumers.

An important part of conversations with interviewees involved knowing more about the specific processes their organizations followed to develop and implement a standard. For UEP, the process starts with recommendations made by the scientific committee; the recommendations are generally made in response to questions brought to the committee, who abstain from offering advice on what they are not consulted on (CG). These recommendations are then passed on to an animal welfare committee constituted by producers for consideration. If the producer animal welfare committee approves, they then submit the recommendations to the UEP board for final approval (CG). Once a recommendation or standard is approved, it is immediately implemented; there is no trial period or piloting of the changes, as producer stakeholders are trusted to be able to determine the feasibility of measures or guidelines during the vetting process (CG). According to CG, until now, every recommendation received from the scientific committee has been adopted; some recommendations have had to be phased in or grandfathered in due to the nature

of the changes. For example, when they increased space allowance for hens in cages, the change had to be implemented gradually in accordance with production cycles, and in order to avoid a market imbalance (the changes were phased in as flocks were replaced).

AM says that something to keep in mind about the process of making standards is that all programs are a negotiation between what is scientifically recommended and what is practically possible. She says that it's relatively easy to get a producer to implement a reasonable change, but whether or not the supply chain will adopt the change is a different matter. Incremental programs provide a way to phase more complex changes in gradually. The GAP process, AM mentions, starts by creating a draft in house, then putting together a working group of no more than 10 producers representing levels and stages of production stages of production, as well as different Step certification levels. They also try to engage producers from different geographies, as depending on the species production can be very different depending on what part of the country you are in, and they are interested in knowing how things are done and what is practical on farm. At this point, AM mentions, what they look for are honest answers rather than consensus at this point, and they want to hear from farmers as farmers and not as representatives of their companies. When companies do come to the discussion table, it can be very ineffective because there is a different power dynamic- there is a high level of competition (AM). After this consultation process, the draft goes out for scientific consultation, to scientific advisors who are hand-picked based on their expertise (AM). The full draft is always sent to the experts, although they may be asked to comment only on specific sections (AM). Within the current process, it has not been GAP's practice to share the names of advisors publicly (for example, on the website), but they are in the process of putting together a scientific committee, and they plan to make that information available when it is finalized (AM). The final stage of the process for GAP

standards is making the draft standard available for public comment, and field testing the standard. At this stage, certifiers are also asked for comments, and they play an important role in channeling comments from producers who may not feel comfortable sharing something directly with GAP.

Like UEP, Certified Humane starts the process with recommendations from the scientific committee (AD). Depending on the issue, the committee may decide to break into sub-committees, and at that point call in producers to contribute from their perspective (AD). Together, the committee and the producer representatives will create a draft, which will be sent out to all producers who are already in the program; they will have between six to eight weeks to provide comments and feedback; people outside of the program are not invited to comment (AD). Once comments are received, a second draft goes to the standards committee (this is different from the scientific committee), which is constituted by three scientists, three producers, three external standards reviewers, and a representative from the retail sector (AD). The final draft then goes to the board, who can approve the standard or not, but cannot make changes (AD).

Another topic brought up by three of the interviewees was auditing. AM mentioned that the importance of a standard being reliably auditable cannot be overstated; standards must be easy to understand and apply to anyone who works with it. She finds it interesting that many comparisons of standards look closely at what is being audited, but not at how, and that some standards are notably better than others in that sense. In order to make sure that GAP standards are auditable, they conduct mock audits of their draft standards, and review the results of those mock audits together with the feedback received from producers and other stakeholders. CG mentioned that one thing UEP is very satisfied with is their auditing process; one aspect of that is

that their audits are pass/fail, and that they are audited by third party certifiers. AD also mentioned auditing as one of the strengths of the Certified Humane program, emphasizing that their inspectors are required to have academic preparation and experience directly related to the species they are auditing. They do not outsource their audits to third parties.

When asked whether they believed that voluntary animal welfare standards in the United States needed to be improved in any sense, the lack of uniformity was a common theme. AM referred to lack of uniformity in auditing models; some require that every farm that is part of an operation be audited, while others only a sample within a group under the same management. The group models, she said, tend to be less expensive, but are not as reliable. CG finds it interesting that although many standards committees share committee members their standards are different, making it clear that the committees are not responsible for this lack of uniformity. JM mentioned lack of uniformity in content, process, robustness and scope. She referred to a comparative analysis conducted by the Food Marketing Institute, which looked at producer organization animal welfare guidelines, pointing out inconsistencies with regard to key elements of a comprehensive policy.

One of the concerns stemming from this variability is that it creates a problem for producers; they often have to comply with several standards, which creates more work (and often represents additional expense) (JM). Complying with several standards or having a number of certifications attached to your product can also add to consumer confusion (JM). From the perspective of standards organizations, AM believes that this lack of uniformity creates an unfair playing field, as different standards are often compared as “apples to apples”, when in fact they are not.

Another issue is clarity or readability of the standard. In many cases, AM mentions, the standard is not clear, leading to inaccurate interpretation. Two strategies to mitigate this are calibration, and having a well-developed policy manual to guide decisions in cases where the standard lacks clarity (AM). JM sees lack of clarity as a liability in terms of how standards are interpreted and audited by different third-party auditors. Differences in auditing approaches for the same standard can sometimes be due to training deficiencies, but lack of clarity or specificity in the standard is also a common cause.

With regard to differences between standards, AD also brought up that the motivations and values behind different organizations may also be a reason for the discrepancies we see between one standard and another. For example, she believes that some programs are money-driven, while others are extremely idealistic, and even though to a lay person they may sound wonderful, they do not necessarily represent what the animals need. In the end, she says, farmers know the difference and choose the standard that suits them best. This, of course, raises the question of what values and motivations drive farmers in their selection of a given standard. In relation to this, CG mentioned that for producers, issues like a few extra inches of space can make a big difference from an economic perspective, and this may ultimately be the reason why they choose one standard over another.

Regarding what factors drive the development of standards, AM stated that sometimes the process is advocacy driven, but that advocates can also sometimes hinder the process. In GAP's case, she believes the program is currently more driven by industry and retailers. In this sense, she considers it part of her job to balance what the science says, what is possible, and what the advocates want.

5.4.4 Perceptions of animal welfare science

When asked what they believed the role of animal welfare science to be within the broader landscape of animal welfare standards, all interviewees generally agreed that the goal was to provide the necessary information to improve animal welfare. For AM, it's about trying to understand the impact of environment and management on an animal, while also understanding an animal's motivation so we can develop systems that support that. While health and productivity matter, this is more about the animal herself. AD mentioned some specific aspects that can be informed by scientific findings, including how to best raise or manage animals, and how to meet their needs. CG believes the science to be key to any standard, without it, he said, we would not be able to have standards. One reason for this, he said is that without scientific backing, industry is vulnerable. He also mentioned that his engagement with experts and scientific research has changed the way he thinks about the issues. He recalled that 15 years ago there was a strong focus on things that could be objectively measured, whereas now it is recognized that outcome-based measures are better assessment tools, more suited to the complexity of assessing welfare in production settings. Critical inputs, thought, will still be considered; he sees this as a transitional stage, where the focus is shifting toward outcomes (which are not part of current UEP guidelines), but critical inputs are still included.

With regard to challenges of animal welfare science, AM identified two areas: a) scalability: since research is generally conducted on a smaller scale, applying those results at a larger scale often produces different results; b) repeatability: “you can get something to work experimentally that you can't get to work in the field.” Another issue is when there is not enough evidence to support a change. With regard to making amendments in the face of insufficient evidence, AM said “we sometimes want to make an improvement, and all we get is a tradeoff- you not only have to think about what you're trying to do right now, but three steps ahead.

Incremental improvement sounds great on paper, but in practice is much more difficult because there are many practical issues that impact incremental approaches and what they look like.”

JM believes that animal welfare science has produced conclusive evidence in many areas. Some are more difficult than others; an example of a difficult area is pain alleviation (JM). Even when robust evidence is produced, there may be limitations in its application (so in those cases it does not fulfill the goal of improving the welfare of animals) (JM). There can also be complications in both producing the evidence and applying the knowledge; some limitations involve balancing other goals of animal production, for example economic concerns, health concerns, and public health issues (JM).

The problem of “waiting on the evidence”, or worse, having emerging research provide evidence that a proposed change is not necessary or even detrimental is also a challenge. AM mentioned that having a tiered program helps to mitigate this in some sense, because producers in the program know that they are expected to continue to make changes as they advance through the Steps, but she also acknowledged that some producers get tired of this, and feel like they are being constantly harassed. There are the economic issues attached to this, and that creates a negative environment for dialogue. She says one solution to this is improving the ways they communicate their "final destination". For example, their new standard for broilers requires 1% window space, but in addition to setting that threshold, they communicate to producers all of the approaches that are acceptable in order to meet that threshold, as well as a reasonable timeframe for implementation.

While Certified Humane does not have a tiered program, AD says that they (Certified Humane) are also able to navigate the difficulties presented by inconclusive science; in their case they attempt to find the most ethical way to do things. For example, while there is not enough

evidence that the use of mammalian sub products in food carries a risk, they have made the decision to require that they not be used in feed, based on expert opinions. AD claims that an important part of making decisions in this way is to be transparent; in their case, they always disclose the names of their expert advisors.

Regardless of limitations, JM believes that scientific evidence is important and influential in the production of AW standards, especially in groups with a strong presence of scientists, or where the opinions of scientists weigh heavily. In these cases (where scientists have it is particularly important that scientist committees be independent (understood as, independent of the commercial or political interests of the group they are working for).

Other important roles of science mentioned by interviewees include navigating differences of opinion (or value conflicts) between different stakeholders. For example, CG believes that science plays an important role in avoiding debates with NGOs about what is good for the welfare of the animals. AD noted that some producers are more resistant to change, and may require some sort of “proof”, and science can play that role. She recalls the case of a farmer she visited who claimed to hold strong Christian values, and yet believed that his method of euthanizing animals by putting them in a 55-gallon drum and leaving them to die was efficient and not problematic. In this case, AD said, trying to debate this based on values would have been unproductive, but referring to scientific reasons for why this was wrong was more likely to change the farmer’s mind.

5.4.5 Science, standards, and public trust

The final section of the interview involved questions about the role of the public in influencing or supporting standards, and whether they consider scientific evidence in forming an opinion of standards. A useful clarification made by JM is that “the public” can have several

meanings, and it helps to be specific. For instance, she believes that at the retail level, the public may have little influence, because individual consumers generally have little knowledge of the details of animal welfare science, standards, and certifications. Food businesses, such as restaurants, are a different case, because their relationship with consumers is more direct; for example McDonald's, cares deeply about their public image, because of their heavy public presence (JM). In other words, the influence of the public tends to be proportional to the public presence of particular kinds of businesses. However, another important consideration is that consumers can be strongly influenced by the media and activist organizations (for example, by means of videos depicting situations of poor welfare or animal abuse) (JM). But even this impact is often short-lived- consumers quickly forget and/or return to their regular habits (JM).

AD also believes that on some levels the public plays an important role in pushing for better animal welfare standards. For Certified Humane, public support is important, and one of their strategies to engage the public is to be transparent by making as much information as possible publicly available (AD). She says that in some cases, this transparency has backfired, as it has been used by animal rights activists such as PETA to attack them and their certified producers (AD). She believes that these groups threaten transparency.

Despite the roles that the public may play in the advancement of animal welfare standards, JM believes that it is unlikely that any change to the way standards are created, implemented, or communicated to the public will have an effect on how the public views them or understands them. It is, however, important for standards to be transparent. One of the reasons for this is the acknowledgment that science is not value free, and we have to be transparent about our choices, and the processes by which we (scientists) use them to direct animal welfare (JM). Transparency is also important because standards informed by science are subject to change as

new findings are made (JM). Being transparent about the process helps create an understanding that the standards are not the final word, but rather the best choice at a particular point in time given what we know (JM).

Finally, from an industry perspective, CG noted that while he believes the public does play a role in the adoption of animal welfare standards, it's often an inconsistent one, and one that can hurt industry. One reason for this is that their attitudes do not necessarily match their actions; so, they might vote on legislation that has a significant impact on producers, but they are not always willing to uphold those commitments when making purchases (CG). Also, when making decisions about animal welfare, consumers often fail to see the bigger picture, such as environmental impacts, economics, and public health implications (CG). According to CG, there are many studies that find that consumers trust farmers, but not necessarily industrial farmers. Because of this gap in trust, consumers may not trust the standards enforced by organizations that represent industrial farmers (CG). On the other hand, some consumers don't think about labeling or certifications at all (CG). For the industry, the guidelines and certification are more of a way of protecting themselves than they are a response to consumers (CG).

5.5 Discussion and conclusions

One of the main objectives of this chapter was to determine, through the examination of standards themselves, as well as consultation with experts in the field of animal welfare standards, whether or not standards accomplish the goal of integrating viewpoints from different stakeholders in a practical way that ultimately matters to the welfare of animals. Insights into the process of standard creation provided by interviewees were particularly useful in answering this question. In all cases examined, and regardless of which stakeholders were ultimately invited to the table, scientific perspectives took priority. However, for each of the standards examined, the

importance of science was explained in different terms, plausibly because of the different nature of the organizations leading the standard process.

In GAP's case, science provides credibility, and a solid foundation for the justification of the requirements detailed in their standards. However, in her interview, AM makes clear that science and the evidence it provides is necessary, but not sufficient to justify standards. She emphasizes the importance of including producer viewpoints from the outset, and further testing standards in the field before approval. UEP's view of the importance of science, as expressed by CG, is also that it provides a solid foundation for their standards, but overall, it provides them with credibility, an important issue for an industry that is increasingly scrutinized and often perceived to lack transparency. In both cases, two of the central cognitive values of science, impartiality and autonomy, were also emphasized as reasons for prioritizing science in the standard-making process.

Another theme that emerges from the discussions described in this chapter is the relationship between standards and the public. If we accept that animal welfare science is a socially mandated science, and that standards are a way to apply that science in ways that fulfil that mandate, then it should be the case that the public plays an important role in shaping standards and fostering their adoption. With regard to this, JM makes an interesting remark, which is that what is defined as "the public" can in fact have different meanings, and whether or not they have influence depends on who we include in that sphere. Additionally, because in animal production supply chains tend to be long, the interest and influence of the public may be related to how close they are to the primary source. Furthermore, because of the physical and epistemic distancing of consumers from their sources of food, they may be subject to influences from other stakeholders, for instance activists or the media. Additionally, CG points out an

important issue in consumer behavior with regard to purchasing of animal products and animal welfare, which is the disconnect that often exists between attitudes and behaviors. For producers, frustrations about this disconnect are exacerbated by a growing public narrative about increasing concern for animal welfare, which is not always consistent with the purchasing behaviors of consumers, leading producers to make large investments that are not always met with consumers' willingness to pay higher prices.

Lastly, this chapter helps identify areas which may be improved or further explored in the process of translating animal welfare science into practical applications such as standards. One topic is the notable lack of uniformity, specifically in private animal welfare standards in the United States. This lack of uniformity is not surprising given the paucity of legal frameworks providing minimum guidelines for animal production. This leaves private standards with the responsibility and the freedom to draw their own lines in the sand, under the influence of forces like market dynamics and consumer demand. A second topic, and one that returns to this work's central issue of the relevance of evidence, is related to a discrepancy in the rate at which robust evidence is produced, and the need to continue adjusting standards in response to external influences. Interviewees pointed out specific issues such as sufficient evidence not being ready or available when needed, or new, contradictory evidence emerging after a consensus has been reached on a particular standard. The fact that in spite of these challenges private standards continue to make progress and move from niche markets to the mainstream supports the notion that scientific evidence, while considered foundational to these processes, is necessary but perhaps not sufficient for their success.

Chapter 6: Conclusions and recommendations

6.1 Conclusions

The overarching goal of this project is to address the question of whether or not evidence produced by animal welfare science is sufficient to justify the development and implementation of animal welfare standards. From an ethical/philosophical perspective, it is clear that evidence alone, however robust, is not sufficient; as discussed in this work, a range of values, both social and epistemic, permeate the practice of animal welfare science and its applications. From a policy perspective, the question might seem unnecessary, given the prevalence and growth in numbers of standards, both public and private, in the face of what often seems like a scientific stalemate between animal industries, scientists, activists, and other stakeholders with divergent ethical views regarding the use of animals. However, beyond these central questions lies a deeper curiosity about the circumstances and conditions that provide evidence, or science generally for that matter, with normative traction. Within the broader context of this work, evidence can be understood as a proxy for the cognitive values of science that arguably bear some of this normative weight.

The valuing of science and its embedded cognitive values as a dynamic and evolving social process that sets the standard for the systematic production of socially relevant generalizable knowledge is arguably part of the answer. There is an implicit trust in science to systematically provide tried and tested information to guide decisions about a broad range of things that matter to us, including our relationships with animals. However, as illustrated by the discussions presented in this work, part of this normative traction can also be attributed to the intricate relationships between science and ethics in animal welfare science. The analysis of the different roles of values in science provides useful frameworks and insights that further clarify

how these relationships might influence the moral relevance of evidence. Philosophers of science have defined distinct stages of the scientific process, as well as acceptable roles for social and ethical values and criteria for when these values might play direct roles (as primary drivers of decisions in the scientific process) or indirect roles (in support of other decision-making criteria) (Douglas 2009, 96-97).

Examples of direct roles of values that can ultimately influence the ethical relevance of include the selection of methods. Dawkins (1980) argues that scientific questions about the welfare of animals can only be consistently and meaningfully answered if there is agreement on the methods used to answer them; the process of agreement presumably involves a value judgment that is not solely based on the scientific or technical adequacy of available methods, but on research question that reflect the value-laden worldviews of the researchers and their institutional contexts. This arguably produces evidence that is already imbued with values; however, this should not compromise its integrity when it results from a scientific process in which values play legitimate and justifiable roles. Another conclusion of the analysis of evidence and its normative dimensions is that even when it bears legitimate moral weight, it is often not sufficient to justify animal welfare policies, standards, or practices.

Even with improvement in the ways animal welfare science produces and synthesizes evidence, the inherent issues of inductive risk and underdetermination persist. Thus, subjective judgments must often be made regarding acceptable risks, harms, and benefits of scientific applications. It is important to note that subjective judgments are not always morally based. There is, for instance, the subjective weighing of evidence based on epistemic/cognitive values, or the prioritization of some kinds or bodies of evidence over others, in line with factors like prevailing theories or schools of scientific or disciplinary thought. Animal welfare standards like

those described and analyzed in this work constitute instances in which subjective judgments (both factual and moral) are made about the relevance of particular kinds of evidence to the determination of socially acceptable practices in the instrumental use of animals. Like policies, standards, in the process of systematizing and normalizing what is socially and morally acceptable, also grant moral traction to the evidence used to justify them.

A final point on the relationships between science and ethics in animal welfare science is an observation on the parallels between the limitations of ethical frameworks and the conceptual frameworks of animal welfare science. In both cases, despite the existence of multiple frameworks with diverse proposals for the conceptualization, organization, and explanation of knowledge, these often fail to center the most pressing questions and issues in animal welfare science (for instance, subjective experience). One view of this is that, despite progress in terms of the volume and robustness of scholarship, there is relatively little novelty with regard to the kinds of topics being researched and discussed. However, a more positive outlook may point to the recent growing interest in research on positive welfare states and expressions, such as projects researching play behavior in broiler chickens (e.g. Baxter et al 2019 and Vasdal et al 2019).

An examination of the history and development of both animal welfare science and animal ethics seems to point to a shared notion that these fields have grown in a more or less linear, cumulative, and progressive fashion. For instance, historical analyses of animal ethics commonly refer to the expansion of our moral universe, which has resulted in increasing moral consideration of animals. However, alternative views on how paradigms come to be and are eventually replaced by others would tell us that the nature of the process is not linear and cumulative, but rather cyclical. Thus, in accordance with this view, paradigms and worldviews

do not simply evolve, but rather experience periods of crisis and revolution during which they are replaced with new, often very different paradigms (Kuhn 1996). Thus, the task may go beyond updating conceptual frameworks or reorganizing existing ones to center issues like affect and subjective experience, but rather remaining open to the possibility that they no longer serve the needs of rapidly changing social landscapes and emerging issues.

6.2 Practical recommendations and future research directions

Interdisciplinarity is arguably one of the defining characteristics of animal welfare science, and a topic which holds promise for fruitful collaboration between philosophers and scientists. Interdisciplinary collaboration presents significant challenges including institutional, disciplinary, and epistemic barriers (Tuana 2013). Effective interdisciplinary research entails a high level of epistemic integration; philosophers can play an important role in supporting this kind of work by assuming a facilitative role, which helps bring to the surface hidden assumptions that may break down communication (O'Rourke and Crowley 2013). In applied fields such as animal welfare, interdisciplinary collaboration need not be constrained to the academic setting. Insights into the standard-making process provided by this work also illustrate the dynamics of interdisciplinary collaboration beyond academia, for instance in the establishment of scientific advisory committees. Criteria for the selection of disciplines and committee members, as well as analyses of the power dynamics within these committees deserve further examination, and if better understood, could improve the functioning of these committees.

Potential outcomes of improved interdisciplinary collaboration between scientists and philosophers in animal welfare science include a better understanding of valid roles for values, subjective measures, and value judgments in animal welfare science. Examples of specific outcomes are guidelines for balancing objective and subjective criteria in welfare assessment,

and criteria for the selection of subjective criteria that can be systematically and reliably integrated into welfare assessment. Ideally, tools like these could improve scientists' confidence in navigating the legitimate roles of ethical and social values in the framing, conduct, and interpretation of research. Similarly, participation in this kind of interdisciplinary research would change the role of ethicists and philosophers, from distant theoretical analysts of the ethical dimensions of science or post-facto judges of scientific interventions, to engaged practitioners of scientifically and socially relevant philosophy (Fehr and Plaisance 2010).

Finally, it is important to consider that the knowledge produced through these collaborations is only as effective as the social processes and institutions used to apply and implement it. The recent proliferation of private and public (legal or regulatory) animal welfare standards has been positive in the sense of responding to a growing public interest in the welfare and protection of animals, but it is important to keep in mind that they are also shaped and influenced by many other forces, including market dynamics, global supply chains, and the increasing integration and consolidation of animal industries. As discussed in this work, while most standards rely on scientific evidence to justify the practices and thresholds that they enforce, the processes by which this is done, including setting criteria for acceptable or better yet robust evidence, are not uniform. Furthermore, these processes are not always transparent. While some private standards bodies offer detailed public information on their policies and processes, including selection of their scientific committees and advisory boards, this is not always the case. Not surprisingly, although private standards draw on the same body of evidence, there are discrepancies in the requirements that they set for similar species and production systems. This is indicative of important differences in processes and interpretation of the evidence. Given the importance of animal welfare standards as regulatory instruments, expressions of social values,

and depositories of public trust, their processes and the values that inform them must also be transparent in order to ensure that they are in fact delivering on their promise to improve animal welfare.

APPENDICES

Appendix I: Interview guide

Interview guide for study STUDY00000186

Study Title: The role of scientific evidence in animal welfare science, standards, and policy: a critical analysis

PURPOSE OF RESEARCH

The purpose of this research study is to examine how scientific evidence is used to support and justify the creation of animal welfare standards. It attempts to trace the path of this evidence, from its origins in scientific research, to its use by public and private standards bodies, non-profit organizations, and private businesses, to create animal welfare standards and policies. As intended result of this research is to make recommendations regarding best practices in the use of evidence for the creation and implementation of animal welfare standards and policy.

PURPOSE OF THIS DOCUMENT

The purpose of this document is to guide the researcher in conducting an open-ended, unstructured interview with a study participant who has previously consented to the interview. The interview may be conducted in person, over the phone, or using video-conferencing technology. As the format of the interview is unstructured, it will be conducted as an informal conversation; this document will not be given to subject participants prior to or during the interview, and will only serve as an aid for the researcher in covering all topics of interest as outlined below. As stated on the informed consent document, subjects are free to decline to answer any question that they are not comfortable answering.

THEMES AND GUIDING QUESTIONS

Personal experience relevant to the research

1. Please tell me about your role and experience in creating and/or implementing animal welfare standards.
2. If you are actively involved in the development and/or implementation of animal welfare standards, please tell me about the process you and your organization follow in doing so.

General perceptions on the role of animal welfare standards and certifications

1. In your opinion, what is the main role (or roles) of animal welfare standards in the context of animal production?
2. Do you believe that third party animal welfare certifiers are trustworthy in terms of delivering on the commitments of their standards or certifications?

3. If you believe that there is need for improvement of current third-party standards and certifications, what are some specific aspects that can be improved?
4. In your opinion, should third party certification become more widely adopted?
5. In your opinion, what factors most strongly influence the development of animal welfare standards?

Perceptions on animal welfare science

1. What do you believe is the overarching goal of animal welfare science?
2. Do you believe that animal welfare science is able to produce conclusive evidence to improve or modify animal production practices?
3. Do you believe that animal welfare standards in the United States are generally grounded in science and/or scientific evidence?
4. How much of an influence do scientific findings and evidence have on the development of animal welfare standards and policy?

Standards, science, and public trust

1. It is often argued that the public (either as consumers, or as part of activist organizations) play a role in pushing for higher standards and changes in animal production practices. Do you agree or disagree with this, and why?
2. Do you believe that the public is generally aware of the science behind the development of animal welfare standards?
3. Do you believe that the public generally trusts science and evidence to guide animal welfare standards and production practices?
4. Do you believe that the public generally trusts third party certifications (i.e. labeling that indicates that animal products were produced using certain practices or standards that make them superior to non-certified products)?

Appendix II: Informed consent template

Research Participant Information and Consent Form

You are being asked to participate in a research study. Researchers are required to provide a consent form to inform you about the research study, to convey that participation is voluntary, to explain risks and benefits of participation, and to empower you to make an informed decision. You should feel free to ask the researchers any questions you may have.

Study Title: The role of scientific evidence in animal welfare science, standards, and policy: a critical analysis

1. PURPOSE OF RESEARCH

The purpose of this research study is to examine how scientific evidence is used to support and justify the creation of animal welfare standards. It attempts to trace the path of this evidence, from its origins in scientific research, to its use by public and private standards bodies, non-profit organizations, and private businesses to create animal welfare standards and policies. The intended result of this research is to make recommendations regarding best practices in the use of evidence for the creation and implementation of animal welfare standards and policy.

2. WHAT YOU WILL DO

This research will use unstructured interviews; in other words, the investigator will ask questions on the general topic of the project, but the questions are not formulated in advance, and the interview will be much like an informal conversation. The interviewee is welcome to skip or decline to answer any question that they would prefer not to answer, and may also choose to end the interview at any point during the conversation.

The interview may be conducted in person, over the phone, or using video conferencing technologies; the interviewee and investigator will jointly agree on the preferred method for the interview, as well as the best time to conduct it. The interview is expected to take between 30 and 60 minutes. The investigator will take hand-written notes during the interview. If the interviewee agrees, an audio recording device may be used. There will be no video recording or photography.

3. POTENTIAL BENEFITS

You will not benefit personally from being in this study. However, we hope that, in the future, other people might benefit from this study because it is expected to contribute to general knowledge in a field that may be of interest to you.

4. POTENTIAL RISKS

There are no psychological, emotional, physical, or legal, risks associated with this research. Although there is no such thing as a “risk free” study, your participation in this research presents no foreseeable risks.

5. PRIVACY AND CONFIDENTIALITY

Data collected during interviews, as well as any identifying information, including signed consent forms will be kept for five years. All digital data will be stored on a password protected laptop computer belonging to the principal investigator. Handwritten notes will be transcribed shortly after the interview and then shredded. Any audio recordings will be transferred to the password protected computer, and then deleted from the original recording device. The information will only be accessible to the principal investigator and co-investigator.

6. YOUR RIGHTS TO PARTICIPATE, SAY NO, OR WITHDRAW

You have the right to say no to participate in the research. You can stop at any time after it has already started. There will be no consequences if you stop and you will not be criticized. You will not lose any benefits that you normally receive.

7. COSTS AND COMPENSATION FOR BEING IN THE STUDY

There are no anticipated costs associated with this study, and no compensation, monetary or otherwise, is offered. Communication costs, if any, will be covered directly by the investigator.

8. ALTERNATIVE OPTIONS

Not applicable.

9. CONFLICT OF INTEREST

No conflicts of interest declared.

10. CONTACT INFORMATION

If you have concerns or questions about this study, such as scientific issues, how to do any part of it, or to report an injury, please contact study coordinator Dr. Paul B. Thompson at:

480 Wilson Rd Room 330
East Lansing MI 48824 US
Ph. 517-432-0316
Email: thomp649@msu.edu

If you have questions or concerns about your role and rights as a research participant, would like to obtain information or offer input, or would like to register a complaint about this study, you

may contact, anonymously if you wish, the Michigan State University's Human Research Protection Program at 517-355-2180, Fax 517-432-4503, or e-mail irb@msu.edu or regular mail at 4000 Collins Rd, Suite 136, Lansing, MI 48910.

11. DOCUMENTATION OF INFORMED CONSENT.

Your signature below means that you voluntarily agree to participate in this research study.

Signature

Date

You will be given a copy of this form to keep.

12. IDENTIFICATION AND RECORDING OF THE INTERVIEW

§ I agree to allow my identity to be disclosed in reports and presentations.

Yes No Initials _____

§ I agree to allow audio-taping of the interview.

Yes No Initials _____

Note: if allowed, audio will be recorded using a recording app on a smartphone. The audio files will then be exported to a password-protected computer, and permanently deleted from the recording device.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Achinstein, Peter. 2003. *The Book of Evidence*. First issued as an Oxford Univ. Press paperback. Oxford Studies in Philosophy of Science. Oxford: Oxford University Press.
- Alexander, Charlotte S. 2012. "Explaining Peripheral Labor: A Poultry Industry Case Study." *Berkeley Journal of Employment and Labor Law* 33, no. 2.
- Altmann, Jeanne. 1974. "Observational Study of Behavior: Sampling Methods." *Behaviour* 49, no. 3: 227–266.
- Appleby, Michael C. 1999. *What Should We Do About Animal Welfare?* Oxford; Malden, MA, USA: Blackwell Science.
- . 2003. "The European Union Ban on Conventional Cages for Laying Hens: History and Prospects." *Journal of Applied Animal Welfare Science* 6, no. 2:103–21.
- Appleby, Michael C., Joy A. Mench, I. Anna. S. Olsson, and Barry O. Hughes, eds. 2011. *Animal Welfare*. 2nd ed. Wallingford, Oxfordshire, UK ; Cambridge, MA, USA: CABI.
- Bailie, C. L., M. E. E. Ball, and N. E. O'Connell. 2013. "Influence of the Provision of Natural Light and Straw Bales on Activity Levels and Leg Health in Commercial Broiler Chickens." *Animal* 7, no. 4: 618–26.
- Baxter, M., C. L. Bailie, and N. E. O'Connell. 2019. "Play behaviour, fear responses and activity levels in commercial broiler chickens provided with preferred environmental enrichments." *Animal* 13, no. 1: 171-179.
- Bekoff, Marc, and Lori Gruen. 1993. "Animal Welfare and Individual Characteristics: A Conversation against Speciesism." *Ethics & Behavior* 3, no. 2: 163–175.
- Bentham, J. 2001. *An introduction to the principles of morals and legislation*. Holmes Beach, Fla: Gaunt.
- Bentley, Janine. 2017. "U.S. Per Capita Availability of Red Meat, Poultry, and Fish Lowest Since 1983." *Amber Waves*, February 6, 2017. <https://www.ers.usda.gov/amber-waves/2017/januaryfebruary/us-per-capita-availability-of-red-meat-poultry-and-fish-lowest-since-1983/>.
- Birch, Jonathan. 2017. "Animal sentience and the precautionary principle." *Animal Sentience: An Interdisciplinary Journal on Animal Feeling* 2, no. 16: 1.
- Blokhuis, H. J. (ed). 2009. "Welfare Quality Assessment Protocol for Poultry." Lelystad, The Netherlands: Welfare Quality Consortium.

- Brambell, F.W.R (chairman). 1965. Report of the Technical Committee to enquire into the Welfare of Animals kept under Intensive Livestock Husbandry Systems. London: Her Majesty's Stationery Office (HMSO).
- Broom, Donald M. 1986. "Indicators of Poor Welfare." *British Veterinary Journal* 142, no. 6: 524–526.
- . 2011. "A History of Animal Welfare Science." *Acta Biotheoretica* 59, no. 2: 121–37.
- . 2014. *Sentience and Animal Welfare*. Wallingford, Oxfordshire: CABI.
- Broom, Donald M., and Andrew Ferguson Fraser. 2007. *Domestic Animal Behaviour and Welfare*. 4th ed. Wallingford, UK ; Cambridge, MA: CABI.
- Broom, Donald M., and K. G. Johnson. 1993. *Stress and Animal Welfare*. 1st ed. Chapman & Hall Animal Behaviour Series. London ; New York: Chapman & Hall.
- Buller, Henry, and Emma Roe. 2014. "Modifying and Commodifying Farm Animal Welfare: The Economisation of Layer Chickens." *Journal of Rural Studies* 33 (January): 141–49.
- Busch, Lawrence. 2000. "The Moral Economy of Grades and Standards." *Journal of Rural Studies* 16, no. 3: 273–83.
- . 2010. "Standards, Law, and Governance." *Journal of Rural Social Sciences* 25, no. 3: 56–78.
- . 2011. *Standards: Recipes for reality*. Cambridge, MA: MIT Press.
- Certified Angus Beef. "Certified Angus Beef. ® Brand." Accessed May 30, 2016. <http://www.certifiedangusbeef.com/>.
- Certified Humane. "Certified Humane Raised and Handled." Accessed January 20, 2018. <https://certifiedhumane.org/>
- . "Our Standards." Accessed April 9, 2018. <https://certifiedhumane.org/how-we-work/our-standards/>.
- Coalition for Sustainable Egg Supply. "Coalition for Sustainable Egg Supply." Accessed March 29, 2018. <http://www2.sustainableeggcoalition.org/about>.
- Cockram, Michael S., and Barry O. Hughes. 2011. "Health and Disease." In *Animal Welfare*, edited by Michael C. Appleby, Joy Mench, I. Anna S. Olsson, and Barry O. Hughes, Second edition. Oxfordshire, UK: CABI.
- Competition and Markets Authority. 2018. "Anticipated Acquisition by Aviagen Group Holding Inc. of Hubbard Holding SAS Decision on Relevant Merger Situation and

Substantial Lessening of Competition.” ME/6727-17. U.K. Government Assets Publishing Service,

Daigle, Courtney Lynd. 2014. “Incorporating the Philosophy of Technology into Animal Welfare Assessment.” *Journal of Agricultural and Environmental Ethics* 27, no. 4 (August): 633–47.

Dairy Cattle Show. “World Dairy Expo.” Accessed May 30, 2016.
<https://worlddairyexpo.com/pages/Dairy-Cattle-Show.php>.

Dawkins, Marian Stamp. 1980. *Animal Suffering: The Science of Animal Welfare*. London; New York: Chapman and Hall.

———. 1989. “Time Budgets in Red Junglefowl as a Baseline for the Assessment of Welfare in Domestic Fowl.” *Applied Animal Behaviour Science* 24, no. 1: 77–80.

———. 2003. “Behaviour as a Tool in the Assessment of Animal Welfare.” *Zoology* 106, no. 4: 383–87.

———. 2008. “The Science of Animal Suffering.” *Ethology* 114, no. 10: 937–45.

———. 2012. *Why Animals Matter: Animal Consciousness, Animal Welfare and Human Well-Being*. Oxford: Oxford University Press.

De Waal, F. B. M., Stephen Macedo, Josiah Ober, and Robert Wright. 2006. *Primates and Philosophers: How Morality Evolved*. The University Center for Human Values Series. Princeton, N.J: Princeton University Press.

Donovan, Josephine, and Carol J. Adams, eds. 2007. *The Feminist Care Tradition in Animal Ethics: A Reader*. New York: Columbia University Press.

Douglas, Heather E. 2009. *Science, Policy, and the Value-Free Ideal*. Pittsburgh, PA: University of Pittsburgh Press.

———. 2017. “Why Inductive Risk Requires Values in Science.” In *Current Controversies in Values and Science*, edited by Kevin C. Elliott and Daniel Steel. New York: Routledge, Taylor & Francis Group.

Duncan, I.J.H. 2005. “Science-Based Assessment of Animal Welfare: Farm Animals.” *Rev. Sci. Tech. Off. Int. Epiz* 24, no. 2: 483–92.

Duncan, I.J.H., and D.G.M. Wood-Gush. 1971. “Frustration and Aggression in the Domestic Fowl.” *Animal Behaviour* 19, no. 3: 500–504.

- . 1972. “Thwarting of Feeding Behaviour in the Domestic Fowl.” *Animal Behaviour* 20, no. 3: 444–51.
- Duncan, I.J.H., M. Park, and A.E. Malleau. “Global Animal Partnership’s 5-Step™ Animal Welfare Rating Standards: A Welfare-Labeling Scheme That Allows for Continuous Improvement.” *Animal Welfare* 21, no. 1 (2012): 113–16.
- Elliott, Kevin Christopher. 2017. *A Tapestry of Values: An Introduction to Values in Science*. New York, NY: Oxford University Press.
- Espeland, Wendy Nelson, and Mitchell L. Stevens. 1998. “Commensuration as a Social Process.” *Annual Review of Sociology* 24, no. 1: 313–43.
- Farm Animal Welfare Council (FAWC). “Farm Animal Welfare in Britain: Past, Present, and Future.” 2009. Accessed April 28, 2018.
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/319292/Farm_Animal_Welfare_in_Great_Britain_-_Past__Present_and_Future.pdf.
- Fehr, Carla, and Kathryn S. Plaisance. 2010. “Socially Relevant Philosophy of Science: An Introduction.” *Synthese* 177, no. 3: 301–16.
- Fraser, David. 2008. *Understanding Animal Welfare: The Science in Its Cultural Context*. UFAW Animal Welfare Series. Oxford ; Ames, Iowa: Wiley-Blackwell.
- Fraser, David., D. Weary, E. Pajor, and B. Milligan. 1997. “A Scientific Conception of Animal Welfare That Reflects Ethical Concerns.” *Animal Welfare*, 6 (January): 187-205.
- Gisolfi, Monica R. 2017. *The Takeover: Chicken Farming and the Roots of American Agribusiness*. Environmental History and the American South. Athens: University of Georgia Press.
- Global Animal Partnership. “Improving the Lives of Farm Animals Step by Step.” Accessed April 3, 2018. <https://globalanimalpartnership.org/>.
- Goldenberg, Maya J. 2005. “Evidence-Based Ethics? On Evidence-Based Practice and the ‘Empirical Turn’ from Normative Bioethics.” *BMC Medical Ethics* 6, no. 1:11-20.
- . 2014. “Evidence-Based Medicine.” Edited by William C. Cockerham, Robert Dingwall, and Stella R. Quah. *The Wiley Blackwell Encyclopedia of Health, Illness, Behavior, and Society*. John Wiley & Sons Ltd.
- Goodpaster, Kenneth E. 1978. “On Being Morally Considerable.” *The Journal of Philosophy* 75 no. 6: 308–25.

- Graber, Roy, and Jennifer Keller. 2017. "Infographic: Chicken Consumers' Top Concerns." WattAgNet, July 25. <http://www.wattagnet.com/articles/31412-infographic-chicken-consumers-top-concerns>.
- Grethe, Harald. 2007. "High Animal Welfare Standards in the EU and International trade—How to Prevent Potential 'low Animal Welfare Havens'?" *Food Policy* 32, no. 3: 315–333.
- Griffin, Donald Redfield. 1976. *The question of animal awareness: Evolutionary continuity of mental experience*. New York: Rockefeller University Press.
- Gunnarsson, Stefan. 2006. "The Conceptualisation of Health and Disease in Veterinary Medicine." *Acta Veterinaria Scandinavica* 48, no. 1 (November): 20-71.
- Hempel, Carl. 1966. *Philosophy of Natural Science*. New Jersey: Prentice Hall.
- Hinde, Robert A. 1970. *Animal Behaviour: A Synthesis of Ethology and Comparative Psychology*. 2nd ed. New York: McGraw Hill.
- Hobbs, Jill E. 2010. "Public and Private Standards for Food Safety and Quality: International Trade Implications." *eJADE: Electronic Journal of Agricultural and Development Economics* 11, no. 1: 136-152 .
- Humane Farm Animal Care. "Who Is behind Certified Humane?" Accessed April 19, 2018. http://certifiedhumane.org/wp-content/uploads/2015/11/Who-Is-Behind-Certified-Humane_11-1-2015.pdf
- Humane Slaughter Association. 2016. HSA online guide for the electrical waterbath stunning of poultry. Accessed December 18, 2018. <https://www.hsa.org.uk/downloads/publications/hsaonlineguidewaterbathpoultryapril2016.pdf>
- Jabareen, Yosef. 2009. "Building a Conceptual Framework: Philosophy, Definitions, and Procedure." *International Journal of Qualitative Methods* 8, no. 4: 49–62.
- Keeling, Linda J., Jeff Rushen, and Ian J.H. Duncan. 2011. "Understanding Animal Welfare." In *Animal Welfare*, edited by M. C. Appleby, J. A. Mench, I.A.S. Olsson, and B. O. Hughes, 13–16. Wallingford, Oxfordshire ; Boston, MA: CABL.
- Keyserlingk, Marina A. G. von, and Daniel M. Weary. 2017. "A 100-Year Review: Animal Welfare in the Journal of Dairy Science—The First 100 Years." *Journal of Dairy Science* 100, no. 12: 10432–44.
- Kincaid, Harold, ed. 2012. *The Oxford Handbook of Philosophy of Social Science*. Oxford Handbooks. Oxford; New York: Oxford University Press.

- Korsgaard, Christine M. 2011. "Interacting with Animals: A Kantian Account." In *The Oxford Handbook of Animal Ethics*, edited by Tom L. Beauchamp and R. G. Frey. USA: Oxford University Press.
- Korte, S. Mechiel, Berend Olivier, and Jaap M. Koolhaas. 2007. "A New Animal Welfare Concept Based on Allostasis." *Physiology & Behavior* 92, no. 3: 422–28.
- Kuhn, Thomas S. 1996. *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.
- Lacey, Hugh. 2005. *Is Science Value Free? Values and Scientific Understanding*. Paperback ed. Philosophical Issues in Science. London: Routledge.
- Lacey, Hugh, and John Staddon. 2003. "The Behavioral Scientist Qua Scientist Makes Value Judgments." *Behavior and Philosophy* 31: 209–23.
- Lund, Vonne. 2006. "Natural Living—a Precondition for Animal Welfare in Organic Farming." *Livestock Science* 100, nos. 2–3: 71–83.
- Lusk, Jayson L, Nathanael M. Thompson, and Shawna L. Weimer. 2018. "The Cost and Market Impacts of Slow Growth Broilers." Unpublished manuscript, last modified October 31, 2018. Accessed December 17, 2018.
<https://static1.squarespace.com/static/502c267524aca01df475f9ec/t/5bdaf60e562fa73fc1584bef/1541076494784/slow+growth+costs+paper+3.pdf>.
- MacIntyre, A. C. 1959. "Hume on 'Is' and 'Ought.'" *The Philosophical Review* 68 (4): 451–68. <https://doi.org/10.2307/2182491>
- McFarland, D. J., and R. M. Sibly. 1975. "The Behavioural Final Common Path." *Phil. Trans. R. Soc. Lond. B* 270, no. 907: 265–293.
- McKenna, Erin, and Andrew Light, eds. 2004. *Animal Pragmatism: Rethinking Human-Nonhuman Relationships*. Bloomington, IN: Indiana University Press.
- McMullin, Ernan. 1982. "Values in Science." In *Proceedings of the Biennial Meeting of the Philosophy of Science Association*. East Lansing, MI.
- Mellor, D.J. 2015. "Positive Animal Welfare States and Reference Standards for Welfare Assessment." *New Zealand Veterinary Journal* 63, no. 1 (January): 17–23.
- . 2016a. "Moving beyond the 'Five Freedoms' by Updating the 'Five Provisions' and Introducing Aligned 'Animal Welfare Aims.'" *Animals* 6, no. 10: 59–66.
- . 2016b. "Updating Animal Welfare Thinking: Moving beyond the 'Five Freedoms' towards 'A Life Worth Living.'" *Animals* 6, no. 3: 21–41.

- Mellor, D.J., and C.S.W. Reid. 1994. "Concepts of Animal Well-Being and Predicting the Impact of Procedures on Experimental Animals." In *Improving the Well-Being of Animals in the Research Environment*. Proceedings of the conference held at the Marriott Hotel, Sydney, October 1993 / R.M. Baker, G. Jenkin and D.J. Mellor (eds.).
- Mellor, D.J., and N.J. Beausoleil. 2015. "Extending the 'Five Domains' Model for Animal Welfare Assessment to Incorporate Positive Welfare States." *Animal Welfare* 24, no. 3: 241–53.
- Mench, Joy A. 1998. "Thirty Years After Brambell: Whither Animal Welfare Science?" *Journal of Applied Animal Welfare Science* 1, no. 2: 91–102.
- . 2003. "Assessing Animal Welfare at the Farm and Group Level: A United States Perspective." *Animal Welfare (UFAW)* 12, no. 4: 493–503.
- . 2008. "Farm Animal Welfare in the U.S.A.: Farming Practices, Research, Education, Regulation, and Assurance Programs." *Applied Animal Behaviour Science*, Farm Animal Welfare since the Brambell Report, 113, no. 4: 298–312.
- Miller, Neal E. 1959. "Liberalization of basic SR concepts: Extensions to conflict behavior, motivation and social learning." *Psychology: A study of a science, Study 1*: 196–292.
- Mintel Group Ltd. 2017. "Poultry, US: November 2017." London.
- National Chicken Council. 2019. "Broiler Chicken Industry Key Facts 2019." Accessed April 23, 2019. <https://www.nationalchickencouncil.org/about-the-industry/statistics/broiler-chicken-industry-key-facts/>
- National Chicken Council and U.S. Poultry and Egg Association. 2019. "2018 Poultry and Egg Economic Impact Study." Prepared by John Dunham & Associates, Inc. Brooklyn, New York. Accessed March 29, 2019. <http://www.uspoultry.org/files/2018-Poultry-Impact-Methodology.pdf>
- Nordenfelt, Lennart. 2006. *Animal and Human Health and Welfare: A Comparative Philosophical Analysis*. Oxfordshire, UK: CABI.
- Nussbaum, Martha C. 2004. "Beyond Compassion and Humanity: Justice for Non-human Animals." In *Animal Rights: Current Debates and New Directions*. Sunstein, Cass R., and Martha Craven Nussbaum, eds. Oxford ; New York: Oxford University Press.
- Olanrewaju, H.A., J.P. Thaxton, W.A. Dozier III, J. Purswell, W.B. Roush, and S.L. Branton. 2006. "A Review of Lighting Programs for Broiler Production." *International Journal of Poultry Science* 5, no. 4: 301–8.
- Olsson, I. Anna S, Charlotte M Nevison, Emily G Patterson-Kane, Chris M Sherwin, Heleen A Van de Weerd, and Hanno Würbel. 2003. "Understanding Behaviour: The Relevance of

- Ethological Approaches in Laboratory Animal Science.” *Applied Animal Behaviour Science* 81, no. 3: 245–64.
- O’Rourke, Michael, and Stephen J. Crowley. 2013. “Philosophical Intervention and Cross-Disciplinary Science: The Story of the Toolbox Project.” *Synthese* 190 (11): 1937–54.
- Palmer, Clare. 2010. *Animal Ethics in Context*. New York: Columbia University Press.
- Paul, Elizabeth S., Emma J. Harding, and Michael Mendl. 2005. “Measuring Emotional Processes in Animals: The Utility of a Cognitive Approach.” *Neuroscience & Biobehavioral Reviews* 29, no. 3: 469–91.
- Petherick, J. Carol, Elaine Seawright, and Dave Waddington. 1993. “Influence of Motivational State on Choice of Food or a Dustbathing/Foraging Substrate by Domestic Hens.” *Behavioural Processes* 28, no. 3: 209–20.
- Proctor, Helen S., Gemma Carder, and Amelia R. Cornish. 2013. “Searching for Animal Sentience: A Systematic Review of the Scientific Literature.” *Animals* 3, no. 3: 882–906.
- Regan, Tom. 2004. *The Case for Animal Rights*. Berkeley: University of California Press.
- Rollin, Bernard E. 2006. *Animal Rights & Human Morality*. 3rd ed. Amherst, N.Y: Prometheus Books.
- Rushen, Jeffrey. 2003. “Changing Concepts of Farm Animal Welfare: Bridging the Gap between Applied and Basic Research.” *Applied Animal Behaviour Science* 81, no. 3: 199–214.
- Rutgers, Bart, and Robert Heeger. 1999. “Inherent Worth and Respect for Animal Integrity.” In *Recognizing the Intrinsic Value of Animals*, edited by Marcel Dol, Martje Fentener van Vlissingen, Soemini Kasanmoentalib, Thijs Visser, and Hub Zwart, 41–51. Animals in Philosophy and Science. The Netherlands: Van Gorcum.
- Sandøe, P, S B Christiansen, and M C Appleby. 2003. “Farm Animal Welfare: The Interaction of Ethical Questions and Animal Welfare Science.” *Animal Welfare* 12, no. 4 (November): 469–78.
- Savory, C. J. 2004. “Laying Hen Welfare Standards: A Classic Case of ‘Power to the People’.” *Animal Welfare* 13, Supplement 1: 153–158.
- Scientific Veterinary Committee. 1996. Report on the Welfare of Laying Hens. Brussels, Belgium: Commission of the European Communities Directorate-General for Agriculture VI/B/II.2
- Shepherd, E. M., and B. D. Fairchild. 2010. “Footpad Dermatitis in Poultry.” *Poultry Science* 89, no. 10: 2043–51.

- Singer, Peter. 2009. *Animal Liberation: The Definitive Classic of the Animal Movement*. Updated ed., 1st Ecco Ed., 1st Harper Perennial ed. New York: Ecco Book/Harper Perennial.
- Sumner, L. W. 1988. "Animal Welfare and Animal Rights." *Journal of Medicine and Philosophy* 13, no. 2:159–75.
- Thompson, Paul B. 2003. "Crossing Species Boundaries Is Even More Controversial than You Think." *American Journal of Bioethics* 3, no. 3: 14–15.
- . 2004. "Getting Pragmatic About Animal Welfare." In *Animal Pragmatism: Rethinking Human-Nonhuman Relationships*. McKenna, Erin, and Andrew Light, eds. Bloomington, IN: Indiana University Press.
- Tinbergen, Niko. 1989. *The Study of Instinct*. Oxford [England] : New York: Clarendon Press ; Oxford University Press.
- Tonsor, Glynn T., and Nicole J. Olynk. 2011. "Impacts of Animal Well-Being and Welfare Media on Meat Demand." *Journal of Agricultural Economics* 62, no. 1:59–72.
- Toomey, Anne H, Nils Markusson, Emily Adams, and Beth Brockett. 2015. "Inter- and Trans-Disciplinary Research: A Critical Perspective." *Global Sustainable Development Report*.
- Tuana, Nancy. 2013. "Embedding Philosophers in the Practices of Science: Bringing Humanities to the Sciences." *Synthese* 190, no. 11: 1955–73.
- Turnbull, Margaret Greta. 2017. "Underdetermination in Science: What It Is and Why We Should Care." *Philosophy Compass* 13, no. 2.
- U.S. Government Accountability Office. 2017. "Better Outreach, Collaboration, and Information Needed to Help Protect Workers at Meat and Poultry Plants." GAO 18-12. Workplace Safety and Health.
- United Egg Producers Certified. 2018. Accessed March 29, 2019. <https://uepcertified.com/>
- United Egg Producers. 2017. "Animal Husbandry Guidelines for U.S. Laying Flocks- 2017 Edition," Accessed December 17, 2018. <https://uepcertified.com/wp-content/uploads/2015/08/UEP-Animal-Welfare-Guidelines-20141.pdf>
- . n.d. "History of United Egg Producers." *United Egg Producers* (blog). Accessed April 29, 2018. <https://unitedegg.com/about/history/>.

- United States Department of Agriculture (USDA), Agricultural Marketing Service. 2000. United States Standards, Grades, and Weight Classes for Shell Eggs. AMS 56. Effective July 20, 2000.
- United States Department of Agriculture Economic Research Service. 2017. "Ag and Food Sectors and the Economy." Ag and Food Statistics: Charting the Essentials. Accessed October 18, 2017. <https://www.ers.usda.gov/data-products/ag-and-food-statistics-charting-the-essentials/ag-and-food-sectors-and-the-economy/>
- University of California Davis. 2016. "Joy A. Mench." UC Davis Department of Animal Science. Accessed October 18, 2018. <http://animalscience.ucdavis.edu/faculty/emeriti/mench.html>.
- USDA Rural Business and Cooperative Development Service. 1995. "Understanding Capper-Volstead." Cooperative Information Report 35. United States Department of Agriculture.
- Vasdal, Guro, Judit Vas, Ruth C. Newberry, and Randi Oppermann Moe. 2019. "Effects of environmental enrichment on activity and lameness in commercial broiler production." *Journal of Applied Animal Welfare Science* 22, no. 2: 197-205.
- Vicino, Greg A., and Lance J. Miller. 2015. "From Prevention of Cruelty to Optimizing Welfare: Opportunities to Thrive." 34th International Ethological Conference, Queensland, Australia. <http://behaviour-2015.m.asnevents.com.au/schedule/session/6656/abstract/24659>.
- Webster, John. 2005a. The assessment and implementation of animal welfare: theory into practice. *Rev. Sci. Tech. Off. Int. Epiz.* 24, no. 2: 723–734.
- . 2005b. *Animal Welfare: Limping towards Eden: A Practical Approach to Redressing the Problem of Our Dominion over the Animals*. UFAW Animal Welfare Series. Oxford, UK ; Ames, Iowa: Blackwell Publishers.
- Weed, Douglas L. 2005. "Weight of Evidence: A Review of Concept and Methods." *Risk Analysis* 25, no. 6: 1545–1557.
- Weeks, C.A, T.D. Danbury, H.C. Davies, P Hunt, and S.C Kestin. 2000. "The Behaviour of Broiler Chickens and Its Modification by Lameness." *Applied Animal Behaviour Science* 67, nos. 1-2: 111–25.
- Wemelsfelder, Françoise. 1997. "The Scientific Validity of Subjective Concepts in Models of Animal Welfare." *Applied Animal Behaviour Science* 53, nos. 1-2: 75–88.

Whole Foods Market. 2012. “Anne Malleau.” June 11, 2012. Accessed October 18, 2018.
<https://www.wholefoodsmarket.com/person/anne-malleau>.

Wiehoff, Dale. 2013. “How the Chicken of Tomorrow Became the Chicken of the World.”
Institute for Agriculture and Trade Policy. March 26, 2013.
<https://www.iatp.org/blog/201303/how-the-chicken-of-tomorrow-became-the-chicken-of-the-world>.

World Organization for Animal Health (OIE). 2017. “Terrestrial Animal Health Code.”
Accessed October 18, 2018. <http://www.oie.int/en/international-standard-setting/terrestrial-code/access-online/>