Engaging in effectiveness: highlighting the role of challenge in well-being and welfare

Becca Franks

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2012

© 2011
Becca Franks
All rights reserved

ABSTRACT

Engaging in effectiveness:
highlighting the role of challenge in well-being and welfare

Becca Franks

What makes an animal's life worth living? Animal welfare scientists have been investigating this question in captive animals for nearly half a century. It has also attracted the attention of academics in other fields because this line of inquiry may improve how we not only manage animals in our care but also think about our own well-being. Concurrently, theories of human well-being and behavior are beginning to play a greater role in animal welfare science. Thus, though the overlap is still limited, the fields of animal welfare and human well-being are converging. To facilitate this integration, I propose *engaging in effectiveness* as common ground from which to generate hypotheses regarding well-being/welfare patterns in human and other species. By engaging in effectiveness, I mean devoting one's resources to 1) obtaining desired results—*value effectiveness*, 2) establishing what is real—*truth effectiveness*, and 3) managing what happens—*control effectiveness*. In a series of experiments, I tested the ability of the engaging in effectiveness model to account for human and rat behavior. The first set of studies (in humans only) confirmed that self-reported effectiveness was strongly correlated to well-being and expectations of future effectiveness/success. The second set of studies found that the frequency of effective engagement was positively correlated to effectiveness (in humans) and negatively correlated to signs of poor welfare (in rats). The third set of studies (in humans and rats) explored the opposing roles that challenges may play in welfare. By providing

opportunities to be effective, challenges may enhance welfare. Conversely, by their potential to cause ineffectiveness/failure on any one of the three domains (value, truth, or control), challenges may decrease welfare. In the final set of studies (in rats only), by manipulating engagement opportunities in the homecage, preliminary validity for a novel measure of welfare was demonstrated. These four sets of studies support the engaging in effectiveness model, highlight the role of challenge in welfare/well-being, and suggest new avenues of research in humans and other animals.

TABLE OF CONTENTS

AWKNOLEDGEMENTS

I am in awe of the wisdom, support, encouragement, and guidance I have received throughout the years. In particular, of E. Tory Higgins as a mentor beyond the call of duty and almost beyond belief; he is simply inspirational. Without the impressive scientific judgment, practicality, and insight of Frances A. Champagne, this research would have never happened. I am also immeasurably grateful to Niall Bolger for his perceptive consideration of this research and his dedication to my development as a statistics-enthusiast and scientist in general. Over the years, I have been fortunate to benefit from many thought-provoking conversations with and the keen observations of James Curley. Finally, I wish to thank Lisa Son for donating her time and expertise to helping me improve this research.

Beyond my committee, the first person who deserves more thanks than I can give is Katie Manley, a dedicated and smart budding-scientist.

Many others have contributed to making this work possible and I wish I could thank you more. Faculty mentors who have guided me along the way include: Sarah Woolley, Peter Balsam, Dave Krantz, Walter Mischele, Ram Hassin, Herb Terrace, Heidi Lyn, Patricia Cole, Diana Reiss, Don Moore, Robert Shapiro, and Hank Jedema. Fellow students who have guided, but have also battled and triumphed with me are: Rahia Mashood, Steen Senhert, Michele Wan, Cate Jensen, Kira Boesch, James Cornwell, Dave Meile, Baruch Eitiam, and all the Higgins and Champagne

lab members past and present. Research support from Kathryn Gudsnuk and Vicky Du was indispensible. Thank you all.

Last but not least, whatever is useful and good about me and through me, this work, would not be there without Lisa Haber-Thomson, Jen Giddings, my unreasonably supportive parents, Beth and Peter Franks, and my wonderful husband, Luca Benedetti.

One fall afternoon a friend, an ornithologist, was counting migrating birds near Demarcation Bay on the north coast of Alaska, at a place called Pingokralik. On several tundra ponds he was also following the progress of three or four families of red-throated and arctic loons. Loons are unable to walk on land, and they require plenty of open water for taking off. Early in September, when the red-throated loon chicks were barely half their parents' size, the coast was buffeted by snow squalls. Within a few days the tundra ponds were frozen over. My friend emerged from his tent one morning to find a red-throated loon and its chick paddling about energetically in an effort to maintain a small patch of open water. The other parent, which had spent the night at sea, flew by every half hour or so with food in its beak, but it could no more land than the other bird could take off.

The next day it warmed up enough so the pondbound adult could take off and the other bird could land with food for the chick. The loons—there were other families in similar straits nearby—persevered in this manner, even as the human observer was driven off to a more permanent shelter. He did not know the fate of the loon chicks (the adults may well have abandoned them). What he remembers seeing were the adults flying back and forth strongly from the sea, dark spots fading in the snow squalls. Resolute, even in the face of poor timing. Successful animals.

-Barry Lopez, Arctic Dreams

INTRODUCTION

Animal welfare science is a relatively young field, but it addresses almost impossibly large questions: How do you define and measure 'good' welfare? What conditions promote healthy and productive animal life? And in the extreme, what makes life worth living? These weighty questions have been gaining attention from not only animal welfare scientists, who study how these questions apply to captive animals, but also biologists and psychologists in general, veterinarians, legislators (responding to public demand), farmers, and philosophers (Fraser, 1999). *Acta Biotheoretica*, a mathematically and philosophically based biology journal, devoted a recent issue (June 2011) to topics in animal welfare. In the introduction, the editors write that as a concept, animal welfare was originally defined "in terms of regaining balance" and poor welfare as the "impossibly to do so" (Hagen, Van den Bos, & Buning, 2011). Since that original definition, welfare researchers have applied a diversity of definitions, and though no consensus has emerged, there are several themes.

First, because of the impossibility of measuring subjective states in animals many scientists advocate basing animal welfare on physiological measurements of health, disease, and stress reactivity (Barnett & Hemsworth, 1990). Elaborating on this biologically grounded position, other welfare researchers promote the ability to 'cope with one's environment' (Broom, 1991) and allostasis, i.e. stability through change (Korte, Olivier, & Koolhaas, 2007), as more flexible, yet still scientifically grounded definitions of welfare. In contrast, a large contingent argues that motivational and behavioral needs are the critical and measurable determinants of an organism's welfare (Dawkins, 1990; Held & Spinka, 2011; Hughes & Duncan, 1988; Mason, 1991). Most recently, tests of an individual's behavioral response to ambiguous environmental

demand have opened the door for other researchers to favor a definitions of welfare centered on affective and cognitive states (Boissy et al., 2007; Mendl & Paul, 2004).

Human well-being is, in some sense, a species-specific, highly specialized version of animal welfare. Though a thorough review of the human well-being literature is impractical here—the definitions of well-being are at least as heterogeneous as those of welfare—it is worth noting that human research inspired many of the theoretical and measurement breakthroughs in animal welfare science (e.g. Dawkins, 1990; Mendl & Paul, 2004). Furthermore, advances in welfare research have strengthened the connection, making comparative work more feasible and compelling (Appleby & Sandoe, 2002). Despite this overlap, however, the two fields tend to rely on different theoretical frameworks about what is and what causes good welfare. For example, the positive psychology movement emphasizes self-determination (Ryan & Deci, 2000), happiness (Diener, 2000), and wisdom and creativity (Seligman & Csikszentmihalyi, 2000) as markers of well-being (i.e. good welfare) in humans. Central aspects of these markers are likely to be irrelevant to the lives of other species, especially evolutionarily distant relations in which, nonetheless, welfare-like patterns are documented: e.g. fish (Galhardo, Correia, & Oliveira, 2008), octopus (Moltschaniwskyj et al., 2007), and bees (Bateson, Desire, Gartside, & Wright, 2011).

From this diversity, are ubiquitous patterns identifiable and can they be accounted for with a single theoretical platform? In an attempt to knit together the two literatures and multiple perspectives while generating testable, species-general hypotheses—particularly ones related to the role challenges play in good welfare—I propose a framework in which good welfare is linked to what happens over time when an organism *engages in effectiveness*. The use of these terms requires some explanation.

To engage means to devote resources—for example, mental, behavioral, hormonal, etc.—to an activity. Engagement on its own is not indicative of good welfare. For example, a person ruminating about past failures or a bird stereotypically plucking its feathers are each engaged but do not have good welfare. In order to identify good welfare, it is necessary to know the object of engagement.

Effectiveness (Higgins, 2011) means being successful in three domains: value (having desired results), truth (establishing what is real or true), and control (managing what happens). Though being *in*effective reduces welfare, the absence of effectiveness is not necessarily equivalent to poor welfare. For example, consider a person struggling through a mathematical proof or a guppy fleeing a predator. In both cases, the individuals may not be currently effective nor is there any assurance that ultimately they will be: the person may never solve the problem or the guppy may swim too slowly. Nevertheless, while they are engaged in pursuing effectiveness, i.e. working towards the solution or eventual escape, they can have good welfare.

Framing welfare as *engaging* in effectiveness underscores that it requires a dynamic process and cannot be achieved through a steady state, even if the individual has desired outcomes, knows what to expect, and is in control. Instead, the critical element is the action of pursuing these states. The original ('regaining balance' quoted in Hagen et al., 2011), coping (Broom, 1991), and allostasis (Korte et al., 2007) definitions and their respective programs of research support placing the emphasis on movement and process. Naming *effectiveness* as the object of engagement resonates with the biological (Barnett & Hemsworth, 1990) and especially motivational (Dawkins, 1990) approaches, but leverages recent theoretical work that provides a detailed, principled account of what it is that animals (including humans) ultimately want (Higgins, 2011). Accordingly, this theory of effectiveness (Higgins, 2011) affirms that all

animals need biological health and desirable outcomes but that they also need to establish what is real and manage what happens. In other words, good welfare cannot come solely from good results; truth and control effectiveness are also essential.

The observation that welfare/well-being involves truth and control effectiveness has broad support within the human literature (e.g. Diener & Seligman, 2002) and similar patterns have been indentified in nonhuman animals as well. The advantages of control effectiveness were demonstrated in the 1960s when now classic research on learned helplessness in laboratory rats revealed the protective health benefits of having control over one's environment (e.g. Lefcourt, 1966). Earlier, Harry Harlow had found that monkeys voluntarily interact with cognitive puzzles in the absence of external rewards (Harlow, 1950)—a first indication that truth effectiveness may be important in the life of a nonhuman animal. More recently, evidence of the importance of truth effectiveness has accumulated in additional species. For example, goats were found to interact with watering devices that required learning despite having access to a device that did not require learning (Langbein, Siebert, & Nurnberg, 2009). Furthermore, several species including primates (Novak, Kinsey, Jorgensen, & Hazen, 1998) and pigs (Puppe, Ernst, Schon, & Manteuffel, 2007) have shown signs of enhanced welfare when given solvable cognitive challenges (truth successes). These patterns have prompted animal welfare researchers to advocate considering truth-like (cognition and predictability, for example) and control-like domains in enrichment-program design (Meehan & Mench, 2007; Sambrook & Buchanan-Smith, 1997).

Poor welfare, according to the engaging with effectiveness framework, is associated with what happens over time when engagement in effectiveness fails. Human examples of how this can occur are easily recognizable: 1) disengagement, e.g. a depressed individual experiencing no

pleasure or pain; 2) engaging in ineffectiveness, e.g. an anxious individual focusing on negative, uncertain, uncontrollable outcomes; or 3) excessive engagement in any one of the three domains: a) value, e.g. addiction, b) truth, e.g. neuroticism, c) control, e.g. obsessive-compulsive disorder.

For most of these malfunctions, nonhuman animal corollaries are identifiable. When living in poor housing conditions, rats show signs of diminished optimism, i.e. disengaging from success (Harding, Paul, & Mendl, 2004), and enhanced pessimism, i.e. engaging with failure (Burman, Parker, Paul, & Mendl, 2009) especially regarding ambiguous outcomes. This research has inspired similar work conducted with dogs (Burman et al., 2011), sheep (Sanger, Doyle, Hinch, & Lee, 2011), mice (Flecknell, Leach, & Bateson, 2011), starlings (Brilot, Asher, & Bateson, 2010), and young chickens (Salmeto et al., 2011). Though I am unaware of research demonstrating a nonhuman animal displaying signs of excessive engagement in the truth domain, animals may show signs of over-emphasizing the other two domains. When kept in poor environments, animals are more likely to develop self-administered drug habits (e.g. Bezard et al., 2003), a possible example of excessive engagement in the value domain, and engage in stereotypical behaviors, a possible example of excessive engagement in the control domain (e.g. Mason & Mendl, 1997).

Together, these patterns suggest that engaging in effectiveness is a potentially useful theoretical model. It classifies environments that lead to good welfare as ones that afford opportunities to engage in species-specific, ecologically relevant effectiveness. Additionally, it suggests that poor welfare comes from living in environments that decrease engagement, diminish effectiveness, and/or possibly overemphasize one type of effectiveness.

From this framework, I derive several hypotheses. 1) Individuals with good welfare show increased signs of engaging in effectiveness. 2) The frequency of effective engagement

(i.e. devoting resources to success) is associated with higher indications of good welfare and lower indications of poor welfare. 3) Challenges can improve welfare by providing opportunities to engage in effectiveness, but they can also be harmful to welfare if they: a) lead to undesirable outcomes or are too frequent (value failure), c) are unpredictable (truth failure), and/or d) are unmanageable (control failure). 4) Environments are beneficial to welfare insofar as they provide opportunities to engage in effectiveness.

To investigate these hypotheses, a comparative approach is instrumental. Human studies enable quick access to (self-reported) subjective states, which streamlines the validation and refinement of theoretical associations. Nonhuman studies, on the other hand, typically facilitate environmental control and manipulation, thus elucidating causal-pathways. Moreover, demonstrating these patterns in more than one species goes some way to indicating the breadth of their generalizability—across situations and species. As such, the association between effective engagement and welfare was examined with humans and its generalizability was tested using laboratory rats (Hypotheses 1 and 2). Similarly, the relationship between challenges and welfare was explored with humans and extended to rats (Hypothesis 3). Finally, the effect of environmental manipulations of welfare was tested with rats alone (Hypothesis 4).

STUDY SECTION 1: EFFECTIVENESS IN HUMANS

Individuals with good welfare show increased signs of engaging in effectiveness.

Insofar as the engaging in effectiveness framework is the intersection between well-being and welfare, humans with high well-being should also report high effectiveness. Furthermore, because engaging in effectiveness means focusing on successful outcomes, highly effective individuals should have greater expectations of success when contemplating real life challenges. Because good welfare is an active, dynamic process and not a passive acceptance of positive outcomes, the model also predicts that highly effective individuals will not expect their problems to be easy (requiring no effort), but will simply have more confidence in eventual success. Together with prior research showing that effectiveness captures a phenomenon distinguishable from optimism (Franks, 2010), this prediction demonstrates an additional dissimilarity. Optimistic people are identified by their expectation that good things will happen to them and that bad things will not (Scheier et al., 1994). Highly effective people, on the other hand, do not expect their challenges to be easy or absent, but rather, are uniquely confident in their ability to be enact a successful outcome.

To assess well-being in humans "Satisfaction with Life Scale" was applied. It is a concise self-reported measure (Diener, Emmons, Larsen, & Griffin, 1985) and a standard in psychology and social science research; according to "Web of Science," the original article has been sited over 2,500 times. It asks for endorsements of the following statements (see Appendix A for complete questionnaire): 1) "In most ways my life is close to my ideal." 2) "The conditions of my life are excellent." 3) "I am satisfied with my life." 4) "So far I have gotten

the important things I want in life." And, 5) "If I could live my life over, I would change almost nothing."

To assess engaging in effectiveness in humans, several new questions (12-14 depending on the version) designed to probe an individual's feelings of success on the value, truth, and control effectiveness domains were employed. For the complete list see Appendix A, but some sample items include: "I have what I need to get what I want" (value); "I always give up when I become confused" (truth); and "Organizing is one of my strengths" (control). Though the effectiveness measure and the well-being measure are substantially different in their theoretical approach, wording, and content, a few of the effectiveness items could be considered to overlap with ones on the well-being measure. Those items are "I have what I need to get what I want" (number 5), "There are a lot of things going wrong in my life" (number 7), and "I am in a bad situation" (number 12). As such, two analyses were conducted: one testing the association between well-being and the effectiveness questions as a whole and one testing it without the 'contaminated' items (excluding numbers 5, 7, and 12).

Study 1.1: Effectiveness and Well-being

1.1 Methods

Individuals participated in a short (less than 15 minutes) on-line experiment through Amazon's Mechanical Turks (MTurks) recruitment program.[1] MTurks is a fast, cheap, and reliable method of on-line data collection that provides a more diverse sample than typical college-based

[1] Note: All human-run studies were conducted with MTurks, which reports having a subject pool of over 500,000 people in 190 countries. Participation in this study was restricted to US residents who had not already participated in a study related to this program of research.

recruitment or other on-line services (Buhrmester, Kwang, & Gosling, 2011). After linking to a webpage generated in and hosted by Qualtrics (a web-based survey design software), participants responded to questions about their well-being (Diener, see Appendix A), effectiveness (see Appendix A), and demographic information. This study ran on two occasions, but because the results and demographic information were nearly identical across the two studies, they were combined into one.

1.1 Demographics

Fifty-one females and 51 males participated (N=102). The median age was 28 years (range 18-77). Thirty-nine percent had at least a bachelor's degree. They lived across the United States in primarily suburban neighborhoods (59%). The median reported income was $20,000 - $49,000 per year. The majority were white, non-Hispanic (79%).

1.1 Results

The 12 effectiveness questions showed good internal reliability (α: 0.75). The composite measure exhibited no significant association with any of the demographic information except a positive correlation to age ($r = 0.30$, $p = 0.002$; 95% confidence interval [CI] = 0.12, 0.47; all subsequent CIs refer to 95% coverage). Effectiveness was positively correlated to well-being ($r = 0.70$, $p < 0.0001$; CI = 0.58, 0.79). Even after removing the 'contaminated' items (numbers 5, 7, and 12), the positive correlation between effectiveness and well-being remained high ($r = 0.61$, $p < 0.0001$; CI = 0.47, 0.72).

Study 1.2a & Study 1.2b: Effectiveness and Expectations of Success

1.2a & 1.2b Methods

Individuals participated in a short (less than 15 minutes) on-line experiment through Amazon's MTurks recruitment program. After linking to a webpage generated in and hosted by Qualtrics (a web-based survey design software), they responded to questions about their effectiveness (see Appendix A) and demographic information. Before the demographic questions, they responded to a filler questionnaire and wrote about a current challenge.

In Study 1.2a, they wrote about a problem they were trying to solve. They were prompted to "think about a problem you are currently experiencing. In the space below, describe the context and the details of the problem and how you plan to solve it." In Study 1.2b they wrote about a personal goal they were trying to attain. They were prompted to "think of an important, personal goal that you have not yet attained. In the space below, please write about your goal." See Appendix B for complete wording.

After writing about their challenge, they answered two questions about the difficulty of and their expected success in resolving the challenge. They were asked to report "How likely is it that you will figure out this problem [reach this goal]" on a scale from *extremely unlikely* (-3) to *extremely likely* (+3) and "How easy/difficult will it be to figure out this problem [reach this goal]" on a scale from *extremely difficult* (-3) to *extremely easy* (+3) (Appendix B).

In study 1.2b, the order of the writing activity and effectiveness questions was counter-balanced and there was no evidence of order effects.

1.2a & 1.2b Demographics

The demographics were nearly identical in the two studies and are reported together. There were 52 participants in Study 1.2a and 100 participants in Study 1.2b. The sample was 53% (80) female. The median age was 27 years (range 18-61). Fifty-one percent had a bachelor's degree or higher. They lived across the United States in primarily suburban neighborhoods (59%). The median reported income was $20,000 - $49,000 per year. The majority were white, non-Hispanic (75%).

1.2a Results

The 12 effectiveness questions again showed good internal reliability (α: 0.84). The composite measure exhibited no significant association with any of the demographic information except gender, with females reporting effectiveness scores 0.36 standard deviations higher than males ($t(50) = 2.19$, $p = 0.03$; CI 0.03, 0.69).

On average, people thought their problem would be somewhat difficult to solve. On a scale from −3 (*extremely difficult*) to +3 (*extremely easy*), the mean was −0.79 (CI −1.2, −0.4), which was significantly less than zero ($t(51) = 3.83$, $p = 0.0004$). Nevertheless, participants generally expected that they would be successful in resolving their problem. On a scale from −3 (*extremely unlikely*) to +3 (*extremely likely*), the mean was 1.9 (CI 1.6, 2.3), which was

significantly higher than zero ($t(51) = 10.96$, $p < 0.0001$). There was no evidence of an association between effectiveness and perceived difficulty ($p > 0.27$). There was, however, a positive correlation between effectiveness and expected success ($r = 0.53$, $p = 0.0001$; CI 0.30, 0.70), which remained strong even when controlling for perceived difficulty (partial-$r = 0.51$, $p = 0.0001$; CI 0.28, 0.69). See Figure 1.

1.2b Results

The 12 effectiveness questions again showed good internal reliability (α: 0.84). The composite measure exhibited no significant association with any of the demographic information.

On average, people thought their goal would be quite difficult to attain. On a scale from −3 (*extremely difficult*) to +3 (*extremely easy*), the mean was −1.60 (CI −1.79, −1.41), which was significantly less than zero ($t(99) = 16.60$, $p < 0.0001$). Nevertheless, participants generally expected that they would be successful in attaining their goal. On a scale from -3 (*extremely unlikely*) to +3 (*extremely easy*), the mean was 1.52 (CI 1.25, 1.79), which was significantly higher than zero ($t(99) = 11.37$, $p < 0.0001$). Again, there was no evidence of an association between effectiveness and perceived difficulty ($p > 0.7$), but there was a positive association between effectiveness and expected success ($r = 0.37$, $p = 0.0002$; CI 0.19, 0.53), which remained significant even after controlling for perceived difficulty (partial-$r = 0.37$, $p = 0.0001$; CI 0.18, 0.53). See Figure 2.

Summary of Study Section 1

As predicted, human well-being was strongly related to self-reported effectiveness despite relatively little item overlap. The "Satisfaction with Life Scale" probes an individual's contentment with life overall. The effectiveness questions target an individual's experiences working through the effectiveness domains of value, truth, and control. Even so, three effectiveness items could be considered 'contaminated' in that they bear a resemblance to the well-being items. After eliminating these items the relationship to well-being remained strong. Also in line with the engaging in effectiveness predictions, highly effective individuals had greater expectations of success, but did not expect their challenges to be more or less difficult than individuals with low effectiveness. Highly effective individuals' success expectations remained high even when controlling for challenge difficulty.

STUDY SECTION 2: FREQUENCY OF EFFECTIVE ENGAGEMENT

The frequency effective engagement (i.e. devoting resources to success) is associated with higher indications of good welfare and lower indications of poor welfare.

To assess this prediction in humans, participants were asked to write about a time when they had an experience of effective engagement—times "when everything works perfectly—when you are so engaged and activated by what you are doing that you lose track of time; ideas, thoughts, actions, just flow." Because the experience of effective engagement can result from a success in any one or a combination of the three domains (value, truth, and control), the engaging in effectiveness model predicts that people would spontaneously report the successful resolution of truth and control dilemmas as well as value success. In other words, though a portion of the reported experiences were expected to be straightforwardly positive (vacations, time with friends and family, etc.), some were expected to include challenges (figuring out a problem at school, managing a difficulty at work, etc.). Additionally, if welfare is linked to what happens over time when an individual engages in effectiveness, regardless of the source of the engagement (challenge or otherwise), the *frequency* of effective engagement should be positively associated with welfare. Accordingly, participants were asked how often they had these experiences and their answers were related to their self-reported effectiveness and well-being.

In ecologically relevant and species-specific ways, these patterns should be similar in nonhuman animals. Rats are nocturnal, omnivorous foragers and as such, two important ways of being effectiveness are obtaining the safety conferred by darkness and highly caloric food items.

Rats whose behavior (engagement) leads to higher frequencies of these outcomes should show lower signs of poor welfare.

Welfare assessment, like its definition, is an ongoing area of research (Broom, 2011), but one convenient proxy measure may be fecal boli elimination in novel environments. Boli production is classically associated with individual differences in rodent stress and emotionality (Archer, 1973; Hall, 1934) and it has been found to be related to anxiety-like behaviors (e.g. Ennaceur, Michalikova, & Chazot, 2006) and living in stressful housing conditions (e.g. Calvo-Torrent, Brain, & Martinez, 1999). Thus, though it may be better classified as a measure of stress and not poor welfare *per se*—which may be a distinct phenomenon, depending on one's prefferred definition of welfare (Broom, 1991; Dawkins, 1990; Korte, 2001; Korte, Olivier, & Koolhaas, 2007; Mendl & Paul 2004)—boli production provides an easy assessment of an at least closely related construct. From the engaging in effectiveness model, therefore, it follows that in a novel environment, rats who are more successful in maintaining darkness and obtaining treats (effective engagement) will produce less boli, i.e. display lower indications of stress.

To assess this prediction, rats were placed in novel environment, an eight-arm radial maze under bright illumination (See Figure 3). Two of the arms were designed to be *success-arms*: reaching the end of one activated a treat-dispenser (*treat-arm*), reaching the end of the other temporarily turned off the light, which affecting the entire maze (*darkness-arm*). There were also two arms designed to be *failure-arms*: one that turned the light back on (*light-arm*) and one that activated a treat-dispenser, but did not release a treat (*nontreat-arm*). Finding an increase in the time spent in the treat- and darkness-arms but a decrease in the time spent in the light- and nontreat-arms would confirm that the treats and darkness are generally desirable, effective outcomes for rats. If true, the success-arms would provide several ways of being

effective: value effectiveness from the darkness/treats outcome, control effectiveness from the rat's control over these outcomes, and truth effectiveness from learning the contingencies. Hence, according the engaging in effectiveness model signs of poor welfare should be negatively related to darkness time and treat activations.

Study 2.1a & Study 2.1b: Effective Engagement and Well-being

2.1a & 2.1b Methods

Individuals participated in a short (less than 15 minutes) on-line experiment through Amazon's MTurks recruitment program. After linking to a webpage generated in and hosted by Qualtrics (a web-based survey design software), they responded to questions about their effectiveness, well-being (Study 2.1b, Diener), and demographics. In study 2.1b two new items were added to the effectiveness questions (see Appendix A).

Before answering the demographic questions, participants wrote about a recent engaging experience and answered a question about how frequently they have these types of experiences. They were prompted to think about one of those times "when everything works perfectly—when you are so engaged and activated by what you are doing that you lose track of time; ideas, thoughts, actions, just flow." The question about the frequency of these occurrences ranged from "*never*" (0) to "*all the time*" (4). See Appendix B for complete wording.

Two coders marked the writing samples. Coders were instructed to "find participants whose writing sample indicates that their engaging experience came from some type of challenge—a

demand (not necessarily negative), work (schoolwork included), a problem or trouble, something difficult to accomplish, a bad situation, etc.—and not from something that was simply and mostly good." The inter-rater reliability was good (agreement = 82%, kappa = 0.64, z=7.30, $p <$ 0.0001) and samples were only included in the analysis if both coders agreed it contained a challenge.

A writing sample coded as not containing a challenge was "I was on a beach in Bermuda. The sight of the pink sand and turquoise water filled me with a sense of euphoria. I was so engaged that I nearly missed my airport transfer." A writing sample coded as containing a challenge was "In Iraq during a combat patrol. There was smoke, dust, sand, bullets flying, loud sound, did not think of anything just went through motions seemed like automatic response. No time to think just react." See Appendix C for additional examples.

2.1a & 2.1b Demographics

The demographics were nearly identical in the two studies and are reported together. There were 53 participants in Study 2.1a and 75 participants in Study 2.1b. The sample was 54% (69) female. The median age was 29 years (range 18-88). Forty-five percent had a bachelor's degree or higher. They lived across the United States in primarily suburban neighborhoods (56%). The median reported income was $20,000 - $49,000 per year. The majority were white, non-Hispanic (72%).

2.1a Results

The 12 effectiveness questions showed good internal reliability (α: 0.78). The composite measure exhibited no significant association with any of the demographic information.

Forty-two percent of the engagement experiences came from a challenge. Recounting a challenge or not did not influence self-reported effectiveness ($p > 0.8$).

Participants reported an average frequency of engagement experiences of 1.96, which was in the middle of the 0-4 scale and corresponded to "*off and on*." Overall, frequency was positively correlated with effectiveness ($r = 0.27$, $p < 0.05$; CI 0.00, 0.50) and this relationship was not modulated by whether the participant wrote about a challenge or not (interaction term, $p > 0.3$).

2.1b Results

The 14 effectiveness questions showed good internal reliability (α: 0.88). The composite measure exhibited no significant association with any of the demographic information except gender, with females reporting effectiveness scores 0.49 standard deviations higher than males ($t(73) = 3.68$, $p < 0.001$; CI 0.23, 0.76). Effectiveness was highly correlated to well-being ($r = 0.63$, $p < 0.0001$; CI = 0.48, 0.75).

Thirty-nine percent of the engagement experiences came from a challenge. Recounting a challenge did not influence self-reported effectiveness or well-being (both p's > 0.5).

Participants reported an average frequency of engagement experiences of 1.87, which was in the middle of the 0-4 scale and corresponded with "*off and on*." Overall, frequency had a marginal

positive correlation with effectiveness ($r = 0.20$, $p < 0.08$; CI = −0.02, 0.41) and was positively correlated to well-being ($r = 0.42$, $p = 0.0002$; CI = 0.21, 0.59). Neither relationship was modulated by whether the participant wrote about a challenge or not (interaction term p's > 0.4). Well-being was independently predicted by both effectiveness (partial-$r = 0.62$, $p < 0.0001$, CI = 0.44, 0.73) and frequency of engagement (partial-$r = 0.38$, $p < 0.001$ CI = 0.17, 0.56).

Study 2.2: Effective Engagement and Stress

2.2. Subjects and housing

The subjects in this study were 60 Long-Evans female rats bred and housed in our facility at the Department of Psychology at Columbia University in accordance with IACUC regulations. When testing began, they were just under 2 months of age (see Figure 4 for a diagram of rat studies). From weaning, they were group-housed (four rats per cage) in large enriched cages kept at constant temperature and humidity with a 12D:12L light schedule. Lights turned on at approximately 9:00 and off at 21:00. In addition to periodic food enrichment (3-4 times per week of various cereals, fruits, vegetables, nuts, etc.), rat chow and water were available ad lib. Each cage contained a large opaque plastic insert that provided shelter and environmental complexity.

2.2 Methods

Habituations and tests were conducted throughout the day during the light cycle, between the hours of 10:00 and 19:00. The rats were tested in a radial arm maze built by ScientificDesign. The maze was located in a small room across a hallway from the colony room and contained

eight arms projecting from a central hub (see Figure 3 and Appendix D). In this phase of the experiment half of the arms were blocked from entry (see Figure 3). Each of the four open arms contained contingencies that a computer with AnyMaze software automatically activated when the animal reached the end. AnyMaze tracked the rat's movement in real-time via a video camera mounted above the maze. Two of the open arms (success-arms) were adjacent and each contained an opportunity to be effective: reaching the end of one turned off the overhead light for 30 seconds (dark-arm) and reaching the end of the other released a food reward (treat-arm). At the end of the two arms directly opposite to the success-arms were two failure-arms: one arm turned on the overhead light (light-arm) and the other activated the food dispenser mechanism without actually dispensing a treat (nontreat-arm).

A week prior to testing, the rats habituated to the maze for four minutes with the light off and in the presence of their cage mates. Tests lasted ten minutes and were repeated four times over the course of two weeks. By experimenter error, one group of rats (i.e. N=4) was not tested on the fourth day. At the end of each individual's test, boli were counted and the maze was washed down with 70% ethanol solution. The AnyMaze software automatically recorded the amount of time that the light was off (*darkness time*), the number of treat activations, and the amount of time the animals spent in each section of the maze—of particular interest, the end of dark-arm, end of treat-arm, end of nontreat-arm, end of light-arm. The data from days 2-4 were analyzed, giving the animal one day to habituate to the apparatus on its own.

2.2 Statistical Models

Multilevel (mixed) models were used in Stata v11.2 to account for and examine the repeated observations of each animal (Gelman & Hill, 2006). All such models included at least a random intercept (which accounts/tests for individual differences in absolute response level) and when appropriate, a random linear slope (which tests for individual differences in response patterns). For count data, a generalized multilevel model with a log-link (Poisson model) was used. Additionally, the multilevel nature of the data allowed for the investigation of within- vs. between-individual effects (van de Pol & Wright, 2009). Experimental day was coded such that the intercept of the model was the predicted level of behavior on the first day examined, i.e. the second day in the maze was coded as 0, the third as 1, the fourth as 2.

2.2 Results

Time spent at the end of each of the success-arms increased over the three days and time spent at the end of the failure-arms decreased (See Table 1, all $p < 0.007$). These trajectories were equally true for all animals: using an independent random effects covariance structure revealed no evidence for individual slope trajectories on any of the end-times. In terms of average level of behavior, however, individual animals tended to be consistent and reliably different from one another (See Table 1, all $p < 0.006$, See Figure 5). For example, rats that spent the most amount of time in the treat-arm on day three were the same rats that spent the most amount of time in the treat-arm on day four.

Over the three days, the amount of darkness and number of successful treat activations increased and individuals were consistent in their relative tendency to produce these outcomes (Table 1, all

$p < 0.001$). In other words, for example, the rats that achieved the most darkness on day two were the same rats that achieved the most darkness on day four.

The number of boli decreased over the three days and was consistent within individual (See Table 1, all $p < 0.001$). Even after controlling for this decrease, the number of boli produced was negatively related to darkness time and negatively related to treat number (N=60, average number of observations per individual n = 2.9; Poisson model darkness coefficient = −0.002, z = 2.20, $p < 0.03$, CI = −0.004, −0.0002; Poisson model treat coefficient = −0.06, z=3.56, $p <$ 0.0005, CI = −0.1, −0.03; see Figures 6 and 7).

Teasing apart between- vs. within-individual effects revealed that animals who created more darkness produced less boli (between-individual effect: Poisson model coefficient = −0.007, z = 1.99, $p < 0.05$, CI = −0.013, −0.00) and at the same time, that more darkness on any given day was associated with less boli (within-individual effect: Poisson model coefficient = −0.002, z = 1.77, $p < 0.08$, , CI = −0.004, −0.0002). There was only very slight evidence of a difference in the magnitude of the darkness effects (Poisson model coefficient = −0.005, z = 1.42, p = 0.16, CI = −0.01, 0.002). The animals that activated the most number of treat rewards overall were the same animals that produced the least number of boli (between-individual effect: Poisson model coefficient = −0.13, z = 4.27, $p < 0.001$, CI = −0.19, −0.07) and the number of treat activations on any given day tended to relate to fewer boli (within-individual effect: Poisson model coefficient = −0.03, z = 1.64, p = 0.10, CI = −0.07, 0.01). The between-individual treat effect was, however, significantly greater than the within-individual treat effect (Poisson model

coefficient = −0.10, $z = 2.73$, $p < 0.01$, CI = −0.17, −0.03). These models were run controlling for day of experiment; the results were similar when not controlling for day.

Summary of Study Section 2

For both humans and rats, the data confirm an association between frequency of effective engagement and welfare. The more often people had effective engagement experiences, the higher their well-being and self-reported effectiveness. The more effective rats were at obtaining treats and darkness, the less signs of stress they displayed. At this stage, these studies do not demonstrate that the frequency of effective engagement causes changes in welfare. It is possible that instead, individuals with poor welfare are less effective and/or more disengaged. Indeed, the within- vs. between-individual analyses in the rat study (Study 2.2) indicates that both causal pathways may be true. The within-individual effect suggests that day-to-day variations in effectiveness (treat and darkness) related to stress. At the same time, the between-individual effect suggests that individuals who tend to be effective (on treat activations or darkness) also tend to be less stressed individuals. Both patterns are compatible with the larger model engaging in effectiveness model. Nevertheless, experimental research is required to determine the presence of bi-directional influence. Finally, the human studies suggest that one important source of effective engagement comes from overcoming or working through a challenge; approximately 40% of the effective engagement experiences contained a challenge.

STUDY SECTION 3: THE IMPORTANT, DUAL ROLE OF CHALLENGES

Challenges can improve welfare by providing opportunities to engage in effectiveness, but they can also be harmful to welfare if they: a) lead to undesirable outcomes or are too frequent (value failure), c) are unpredictable (truth failure), and/or d) are unmanageable (control failure).

Frequency of successful engagement and good welfare were found to be related. Similarly, though not as straightforwardly, the frequency of challenges should be related to welfare. Challenges here refer to any disruption that imposes a demand on an individual. High frequency or magnitude of challenge may hinder an individual from obtaining successful outcomes (value failure). In these cases, challenges are bad for welfare. Absence of challenges, however, is also detrimental to welfare insofar as it takes away opportunities to engage in being effective (potentially leading to disengagement). Taken together, the model predicts that people with high effectiveness will report having close to 'the right' number of challenges. Self-reported effectiveness should decrease as challenges become too frequent *and* as challenges become too infrequent.

Furthermore, effectiveness should be negatively related to depressive symptomatology (one form of poor welfare) and the connection between depressive symptomatology and challenge frequency should also show a similar curvilinear pattern. This pattern would be particularly surprising for depressive symptomatology insofar as an *absence* of challenge is not normally thought of as a being associated with higher symptoms of depression. According to the engagement in effectiveness model, however, eliminating challenges may increase depressive symptomatology by reducing opportunities to engage in effectiveness. To assess depressive

symptomatology, the Center for Epidemiologic Studies Depression Scale (CES-D) was used. It is a freely available, well-validated continuous scale instrument for quantifying an individual's experience with the common symptoms of depression in the past week.

From the engagement in effectiveness framework, two additional ways in which challenges can impair effectiveness is through their unpredictability (truth failure) and uncontrollability (control failure). As such, when rats are exposed to unpredictable challenges over which they have no control, they should show signs of diminished welfare, i.e. more boli in novel environments. Several unplanned negative events were used as a 'natural experiment' of this prediction. To show that challenges are not uniformly distressing, but can lead to increased engagement in effectiveness, a planned experiment was also conducted: Half the animals were presented with a moderate challenge in their homecage and their engagement with success was measured in the maze. Insofar as manageable challenges can increase welfare, these challenges should increase an animal's engagement with success in the maze: increase darkness time and increase treat activations. Nevertheless, as the timing of these challenges was unpredictable (truth failure) and the outcome potentially negative (potential value failure), the model also indicates that these challenges could simultaneously increase stress, i.e. increased boli production in the maze.

Study 3.1a & Study 3.1b: Challenge Frequency and Depression

3.1a & 3.1b Methods

Individuals participated in a short (less than 15 minutes) on-line experiment through Amazon's MTurks recruitment program. After linking to a webpage generated in and hosted by Qualtrics

(a web-based survey design software), they responded to questions about their effectiveness (see Appendix A), depressive symptomatology (Study 3.1b, CES-D, see Appendix A), and demographics.

Before answering the demographic questions, participants answered questions about the frequency of challenging experiences in their life. They were prompted "All of us experience challenges—times during which our abilities, talents, creativity, and resources are put to the test" and then asked "Right now, how would you describe the number of challenges in your life?" The scale ranged from "*too few*" (−3) to "*just right*" (0) to "*too many*" (+3) (see Appendix B). In Study 3.1a, two new items were included in the effectiveness questionnaire (see Appendix A).

3.1a & 3.1b Demographics

The demographics were nearly identical in the two studies and are reported together. There were 76 participants in Study 3.1a and 51 participants in Study 3.1b. The sample was 52% (65) female. The median age was 29 years (range 18-60). Forty-two percent had a bachelor's degree or higher. They lived across the United States in primarily suburban neighborhoods (56%). The median reported income was $20,000 - $49,000 per year. The majority were white, non-Hispanic (74%).

3.1a Results

The 14 effectiveness questions showed good internal reliability (α: 0.86). The composite measure exhibited no significant association with any of the demographic information except education, to which it was positively correlated ($r = 0.25$, $p < 0.05$; CI 0.03, 0.45).

A slight majority (51%) of the participants reported having more than just the right number of challenges in their life (+1 to +3 on the challenges scale). The plurality (37%) of participants reported that the number of challenges in their life was exactly "*just right*," and the remaining (12%) reported having less than just the right number (−1 to −3 on the challenges scale). Having increasingly too many challenges relative to "*just* right" (from 0 to +3, N=67) was negatively related to effectiveness ($r = -0.36$, $p = 0.003$, CI: −0.56, −0.13). Having increasingly too few challenges relative to "*just right*" (from 0 to −3, N=37) was marginally negatively related to effectiveness ($r = -0.26$, $p = 0.11$, CI: −0.54, 0.07). In a multiple-regression, the quadratic term (the square of challenges) was more predictive of effectiveness than the linear term (the untransformed challenge variable) (effectiveness quadratic: $t = 3.69$, df = 73, $p < 0.001$; effectiveness linear: $t = 1.26$, df = 73, $p > 0.2$), reflecting a ∪-shaped curve of too few challenges and, especially, too many challenges being associated with lower effectiveness.

3.1b Results

The 12 effectiveness questions showed good internal reliability (Cronbach's alpha: 0.75). The composite measure exhibited no significant association with any of the demographic information, but was strongly and negatively correlated to depressive symptomatology ($r = -0.73$, $p < 0.0001$, CI: −0.84, −0.57).

A large percentage (47%) of the participants reported having more than just the right number (+1 to +3 on the challenges scale). The plurality (33%) of participants reported that the number of challenges in their life was exactly '*just right*,' and the remaining (20%) reported having less

than just the right number (−1 to −3 on the challenges scale). Having increasingly too many challenges relative to "*just right*" (from 0 to +3, N=41) was positively related to depressive symptomatology and negatively related to effectiveness (depression: $r = 0.73$, $p < 0.001$, CI: 0.55, 0.85; effectiveness: $r = -0.54$, $p < 0.001$, CI: −0.72, −0.27). Having increasingly too few challenges relative to "*just right*" (from 0 to −3, N=27) was also positively related to depressive symptomatology and negatively related to effectiveness (depression: $r = 0.53$, $p = 0.005$, CI: 0.18, 0.75; effectiveness: $r = -0.40$, $p = 0.04$, CI: −0.68, −0.02). In a multiple-regression, the quadratic term (the square of challenges) was more predictive of depressive symptomatology and of effectiveness than the linear term (the untransformed challenge variable; depression quadratic: $t(48) = 4.86$, $p < 0.001$; depression linear: $t(48) = 1.57$, $p = 0.12$; effectiveness quadratic: $t(48) = 3.54$, $p = 0.001$; effectiveness linear: $t(48) = 1.3$, $p > 0.2$; See Figures 8 and 9). As in Study 3.1a, this pattern of results reflect a U-shaped curve of too few challenges and, especially, too many challenges being associated with higher depressive symptomatology and lower effectiveness.

Study 3.2: Unpredictable, Uncontrollable Challenges and Stress

3.2 Subjects and housing

The female rats in this study were the same as reported in Study 2.2 (See Figure 4) and were kept under the same housing conditions.

3.2 Methods

Over the course of the testing period reported above (Study 2.2), a number of unplanned disturbances were recorded. These negative events included, for example, flooded and/or dirty

home-cages, experimenter error leading to a rat remaining in a small holding cage for an extended period (20 minutes), and computer malfunction. Of the 176 tests, 22 were classified as being preceded by a disturbance of some kind and were analyzed as 'natural experiments' of unpredictable, uncontrollable challenges.

3.2 Statistical models

Again using multilevel models (random intercept only and controlling for experiment day) in Stata v11.2, the effect of these disturbances on the rats' behaviors was assessed.

3.2 Results

The disturbances had no effect on the number of treats or darkness time (p's > 0.7). However, as predicted, they did increase the number of boli produced (controlling for experiment day, Poisson model coefficient; 0.70, z=4.33, $p < 0.001$, CI = 0.38, 1.01).

Study 3.3: Manageable Challenges, Engaging in Effectiveness, and Stress

3.3 Subjects and housing

The female rats in this study were the same as reported in Study 2.2 and 3.2 (See Figure 4) and were kept in the same housing conditions unless otherwise noted.

3.3 Methods

Two months after the testing period reported above (Study 2.2 and 3.2), half of the rats (N=30) were given noxious novel objects twice a week in their homecage for three weeks. The novel

object was a metallic teabag anchored to the front of the cage and filled with a paper towel soaked in either bleach or fantastic® household cleaner. After placing the novel object in the cage, the rats' behavior was scanned for 15 minutes for signs of burying responses. Burying is one way rats engage in defense against predators (see De Boer & Koolhaas, 2003; Matuszewich et al., 2007) and would thereby confirm that rats responded to the objects as a threatening challenge. All treatment cages were observed attempting to bury the novel object. The other cages received no novel objects.

One week after the novel object treatments, all animals were tested twice in the automated maze. In this phase, the same four doors were open and the contingencies remained the same, but testing only lasted for four minutes (*vs*. 10 minutes the previous phase). During the test weeks, no novel objects were placed in the cages.

3.3. Statistical models

Using multilevel models (random intercept, controlling for experiment day and previous mean level of behavior) in Stata v11.2, the effect of these novel objects on the rats' behaviors was examined.

3.3. Results

The treatment increased the amount of darkness the rats achieved and marginally increased the number of treats they activated (N=60, occasions=2; effect on darkness: 16.11 seconds, z=2.36, p=0.02, CI=2.76, 29.48; effect on treats [Poisson model]: 0.12, z=1.50, p=0.13, CI= -0.04, 0.27).

The treatment also increased boli production (N=60, occasions=2, effect on boli [Poisson model]: 1.26, $p = 0.005$, $z = 2.80$, CI = 0.38, 2.14).

Summary of Study Section 3

In both humans and rats, challenges related to welfare. Critically, the human data demonstrated that the relationship was non-linear. The significance of the quadratic term in the multiple-regression is indicative of a strong curvilinear association between effectiveness and challenges, in this case a ∪-shaped curve. Too many or too few challenges were associated with decreased effectiveness and increased depressive symptomatology. In particular, that too little challenge was associated with higher depressive symptoms is an under-recognized pattern, though it is clearly predicted from the engaging in effectiveness model. Similarly, the rat data indicated a complex relationship between challenge and welfare. The natural experiments supported the prediction that challenges are likely to lead to signs of poor welfare when they are undesirable, unpredictable *and* uncontrollable (value, truth, and control failures). When a challenge was at least manageable, however, the experimental manipulation in the homecage demonstrated that challenges can increase engagement in effectiveness even while increasing signs of stress.

Thus, challenges appear to relate to welfare in several ways. People with high self-reported effectiveness expect to be more successful in overcoming their own challenges than people with low self-reported effectiveness (Studies 1.2a & 1.2b). In a separate set of studies, a substantial portion of participants spontaneously recounted challenges as an ideal example of a time when they were engaged in being effective (Studies 2.1a & 2.1b). Having too few challenges was associated with lower effectiveness and higher depressive symptomatology in

humans (Studies 3.1a & 3.1b). In rats and humans, however, challenges were also potentially harmful to welfare, especially when they were unpredictable, uncontrollable (Study 3.2) and/or when they were too frequent (Studies 3.1a & 3.1b). As such, adding challenges to an environment poses a potential trade-off as was found in Study 3.3: it simultaneously increased engagement in effectiveness and signs of stress. More research is required to explore the nature of this complex relationship, yet the studies presented here demonstrate that challenges play an important, albeit dual role in determining an individual's welfare.

STUDY SECTION 4: ENVIRONMENTAL MANIPULATIONS

Environments are beneficial to welfare insofar as they provide opportunities to engage in effectiveness.

If welfare comes from engaging in effectiveness over time, it follows that removing all forms of engagement, including challenges, should decrease welfare. Two months after the homecage challenges (reported in Study 3.3 above), when all animals had returned to baseline levels of behavior, two-thirds of our rats were isolated (*isolated rats* vs. *original-housed rats*) to remove as many forms of engagement as possible: no social companions, shelter, treats, or challenges. At this point in the study program, however, welfare measurement became problematical. Aside from its somewhat questionable relationship to poor welfare (see Korte, 2001; Korte, Olivier, & Koolhaas, 2007), boli production became an untenable measure of welfare because it had dropped to near zero levels in the maze. Boli counts are a convenient measure of stress, but only in novel environments (Archer, 1973; Hall, 1934). By the ninth maze test, it was no longer novel, rendering boli counts meaningless. Indeed, even by the fourth day all animals were producing few boli in the maze (see Study 2.2). An alternative measure of welfare was therefore required.

Building on Mendl's research (Mendl, Burman, Parker, & Paul, 2009) and the human studies reported here (the association between effectiveness and success expectancies Studies 1.2a & 1.2b), an animal's response to ambiguity can serve as an indicator of poor welfare. In an additional final minute of the maze test, therefore, rats were allowed to explore the *ambiguous-arms*, those arms located between the success-arms (dark- and treat-arms) and the failure-arms

(light- and nontreat-arm; see Figure 3). Given the opportunity to explore the ambiguous-arms, rats with poor welfare should visit fewer arms (disengage from effectiveness) and, consistent with an increased expectation of failure (engage with ineffectiveness), spend more time avoiding them. Because the ambiguous arms could have contained greater rewards than those found in the unambiguous arms, worse punishments, nothing, or anything in between, behavior towards them is likely to be more sensitive to changes in welfare—i.e. more reflective of tendencies to engage/disengage in effectiveness/ineffectiveness—than behavior towards the success arms. If there were any behavioral differences between the isolated rats and original-housed rats in the success-arms, however, it should be similar to the pattern predicted for the ambiguous-arms: poor welfare manifesting as disengaging from effectiveness (fewer treats and less darkness).

Finally, to begin to test the relative worth of challenges versus more traditional forms of enrichment, half of the isolated rats received challenges (*challenge-treatment*) and the other half food-enrichment (*enrichment-treatment*). This manipulation compliments previous research demonstrating that cognitive challenges can be enriching (e.g. Langbein et al., 2009), but liberalizes it to include non-cognitive challenges. After one week of treatment, rats were again tested in the maze. One indication that non-cognitive challenges can increase welfare just as well as traditional enrichment would be to find no behavioral differences between the enrichment treatment group and the challenge treatment group.

Study 4.1: Effects of Barren Housing

4.1 Subjects and housing

The female rats in this study were the same as those reported in Study 2.2, 3.2, and 3.3, but some rats were re-assigned to new housing conditions (see Figure 4). The control animals remained under the original housing conditions (original-housed rats, N=20, five cages of four animals each). The isolated rats (N=40) were re-housed into standard laboratory cages (smaller than their homecages) with ad lib food and water but no enrichment (no social companions, treats, shelter, or novel objects) and in isolation. The conditions were assigned so that control and treatment animals from the previous experiment (Study 3.3) were evenly distributed into the isolated and original housing.

4.1 Methods

Testing in this phase began two months after the previous testing period (Study 3.3). Just prior to re-housing, when all animals were still in their original groups and homecages, they were re-tested in the automated maze. The maze conditions were the same as the previous test (4 arms, same contingencies, four minutes) except the failure contingencies were now more negative. A loud burst of white noise occurred when the animal went down either failure-arm, but the noise repeated if the animal remained in the light-arm (see Appendix D).

Directly after this re-testing, the rats were assigned to either original or isolated housing and were kept under these conditions for two weeks. At this point, all animals were tested again in the automated maze. During this round of maze testing, the four previously closed doors (ambiguous-arms) opened in the fourth minute of the test and the animal's behavior was recorded for one minute. Two of these ambiguous-arms were located between the dark- and

light-arm (*dark/light-ambiguous*). The other two ambiguous arms were located between the treat- and nontreat-arm (*treat/nontreat-ambiguous*). Both the dark/light-ambiguous arms caused the light to go off. Both the treat/nontreat ambiguous arms dispensed treats.

4.1 Statistical models

Controlling for the amount of darkness and treats in the pre-isolation test, Generalized Linear Models (GLMs) were used to test the effect of isolation: regular regression for amount of darkness, Poisson regression model for treat number, logistic regression for the probability than an ambiguous arm was visited, and regular regression for length of time spent in various maze locations. The last minute of the test was analyzed separately from the first four minutes—segregating behavior pre- and post-access to the ambiguous-arms.

4.1 Results

In the interim since the previous round of testing (two months since Study 3.3), the effect of the challenge appeared to have worn off: The challenge rats were now behaviorally indistinguishable from their non-challenge counterparts (challenge effect on darkness, treats, boli: all $p > 0.4$; darkness effect: 0.22, CI = −18.10, 18.54; treat effect: 0.09, CI = −0.16, 0.34; boli effect: −0.58, CI = −2.35, 1.18).

After the housing manipulation, in the first four minutes of the test, the isolated rats did not differ from the original-housed rats in time spent in the dark-arm or treat-arm (p's > 0.7), though there was slight evidence that they ate 1.35 fewer treats on average ($z = 1.64$, $p = 0.10$, CI = −0.36, 0.03). In the last minute of the test, the isolated rats did not differ from the original-housed rats

in the time they spent in either of the success-arms (treat-arm effect: $p = 0.19$, CI = −1.94, 9.48; darkness-arm effect: $p > 0.8$) nor did they differ in the in the amount of overall darkness or treats they achieved (p's > 0.4).

Examining behavior towards the ambiguous arms, only three animals (two original-housed rats and one isolated rats) visited all four ambiguous arms, but 20 animals visited three ambiguous arms. The probability that an animal visited three or more ambiguous arms was less likely for isolated rats (OR = 0.35, $z = 1.85$, p=0.06, CI 0.11, 1.06). Of all the ambiguous arms, this pattern was especially true for the ambiguous arm nearest the treat-arm (OR = 0.28, $z = 2.12$, p=0.03, CI 0.08, 0.91). Compared to original-housed rats, isolated rats spent 6.70 fewer seconds in the ambiguous arms ($t(56) = 2.81$, $p = 0.007$; CI 11.5, 1.9). Of all the ambiguous arms, this pattern was especially true for the ambiguous arm nearest the light-on-repeating-noise-arm (seconds: 2.90, $t(56) = 2.54$, $p = 0.01$; CI 5.2, 0.6).

Study 4.2: Effects of "Treats vs. Challenges"

4.2 Subjects and housing

The subjects in this study were the same as reported in Study 2.2, 3.2, 3.3, and 4.1 (see Figure 4). The original-housed rats and isolated rats remained in the same housing conditions except that half of the isolated rats received food enrichment every day (enrichment-treatment, N=20) and the other half were twice given challenging novel objects, this time, large plastic structures (challenge-treatment, N=20).

4.2 Methods

After a week of treatment, all animals including the original-housed rats (N= 20) were tested in the automated maze under the same conditions as the previous test: four minutes with the same four clear success/failure contingencies, a final fifth minute with access to the four ambiguous arms as well.

4.2 Statistical models

GLMs (applying logistic or Poisson models where appropriate) were used. The results reported below do not include covariates, but the pattern of results remains the same when controlling for the behavior in the previous test: dark duration, treat activation, or ambiguous arms visited.

4.2 Results

After treatment, there was no evidence of a difference between the behavior of the isolated rats (now enrichment- or challenge-treatment) and original-housed rats (effect of isolated rats compared to original-housed rats, treat: $p > 0.4$; darkness effect: 9.21, $p > 0.14$, CI = −3.33, 21.74; visit 3 or more ambiguous-arms: $p > 0.8$; duration in ambiguous-arms: $p > 0.3$). In other words, with treatment (either enrichment or challenge), the behavior of the isolated animals (still in isolation) became indistinguishable from that of the controls. Moreover, comparing enrichment-rats and control-rats, there was no strong evidence of a differential effect of treatment type: darkness, treats, arms visited and length of visit (effect of ET compared to CT: darkness effect: 8.52, $p = 0.26$, CI −6.49, 23.54; treats effect: 0.20, $p = 0.11$, -0.04, 0.45; visit three or more ambiguous-arms: $p > 0.8$; duration in ambiguous-arms: $p > 0.6$).

Summary of Study Section 4

With these studies, the engagement in effectiveness framework predicted how challenges shape the welfare of rats. The first study used a novel method of welfare measurement to show that barren, isolated housing conditions caused decreased engagement with success and increased concern with failure and that while slight hints of these patterns could be found in the success-arms, it was the ambiguous-arms that brought them into clear relief. Moreover, the results from the success-arms demonstrate that having poor welfare is not simply a matter of being risk adverse. The isolated rats did not spend more time in the low-risk, success arms. If anything, they showed a tendency to disengage from even sure-fire successful behaviors, achieving slightly fewer treat activations in the first four minutes. The final study took a first step in comparing the effect of enrichment vs. non-cognitive challenges. Both treatments returned the isolated animals to normal levels of behavior and that the effect of challenge treatment was indistinguishable from the enrichment treatment. This result contributes to the accumulating evidence that certain types of challenges can function like standard forms of enrichment.

DISCUSSION

This research began with a general engaging in effectiveness perspective, yet it revealed a specific and understudied actor in good welfare: challenge. Challenge expectancies were a defining characteristic of highly effective people. In considering their own challenges, effective people had greater expectations of success than less effective people, even though they did not expect their challenges to be easier (Studies 1.2a & 1.2b). Challenge was a common source of engaging experiences and, moreover, regardless of whether the source of the engaging experience was a challenge or not, the frequency of effective engagement was related to greater effectiveness and well-being (Studies 2.1a & 2.1b). Challenge at 'just the right' frequency was associated with higher effectiveness and lower depressive symptomatology in humans and the reverse was found for both 'too many' challenges (value failure) and 'too few' challenges (disengagement; Studies 3.1a & 3.1b). In rats, unpredictable (truth failure), uncontrollable (control failure) challenges induced stress (Study 3.2). Manageable challenges, on the other hand, increased signs of engagement in darkness pursuit and treat activation—two ways of being effective—while also increasing stress (Study 3.3). Finally Study 4.2 provided some confirmation that poor welfare stemming from a lack of engagement opportunities (Study 4.1) was rescued equivalently well by challenge treatment as standard enrichment treatment. Together, these studies bolster the proposal that appropriate challenges are central to good welfare (Chamove & Moodie, 1990; Meehan & Mench, 2007) and demonstrate the value of the engaging in effectiveness model in anticipating what characteristics make a challenge "appropriate." Moreover, this model was useful because though alternative perspectives of

welfare and well-being are compatible with aspects of this research, none can account for its entire trajectory.

Both biology and affect/cognition played critical roles in the measurement of welfare in our female rats. The stress physiology that leads to an increased boli production in novel environments provided an indicator of poor welfare in Studies 2.2, 3.2, and 3.3. When this measurement became unsuitable, the affective/cognitive perspective (Harding et al., 2004) inspired a new measurement of welfare (Studies 4.1 and 4.2). A strictly biological perspective of welfare, however, has little to say about the role challenges play in good welfare and is insufficient for understanding human well-being. Similarly, far from suggesting the benefits of challenges, a strictly affective/cognitive perspective is primarily concerned with the harmful nature of challenge, for by definition, challenges can reduce positive affect. The cognitive/affective model also runs into generalizability problems with species that may have rudimentary, at best, cognitions and affect (Mendl, Paul, & Chittka, 2011).

The coping model (Broom, 1991), supported by our finding that unmanageable challenges induced stress (Study 3.2), also focuses on the negative aspects of challenges, though it is more detailed in its account. According to the coping model, challenges are negative because they may interfere with an animal's ability to deal with its environment and because they may be precisely what demanded coping in the first place. Again, however, this model is silent about why challenges are good for welfare.

Models that emphasize allostasis (Korte, Olivier, & Koolhaas, 2007) and behavioral and motivational (Dawkins 1990; Hughes & Duncan, 1988) do suggest, on the other hand, a potential benefit of challenges. Decades ago, primatologists posed the question of whether "alarming events" are good for captive animals reasoning that if behavior in captivity should mimic

behavior in nature (for the sake of good welfare), challenges would be necessary to produce the full spectrum of behavioral responses (Chamove & Moodie, 1990). Though they found evidence in support of the affirmative (Moodie & Chamove, 1990), few studies since then have directly explored this intriguing idea. In our studies, giving female rats the opportunity to express a species-specific behavior, i.e. burying a threat, produced signs of enhanced welfare (Studies 3.3 and 4.2), crediting the use of behavioral/motivational need as a blueprint for good welfare. However, our research also demonstrated that challenges were not uniformly beneficial to welfare. In humans and rats, challenges sometimes increased distress (value, truth, and control failures: Studies 3.1a & 3.1b, 3.2, 3.3).

Though a behavioral/motivational perspective on welfare is congruent with this pattern of results, it does not predict them *a priori*. Without a predictive framework of what an animal wants, it is impossible to foresee what types of challenges would benefit what sorts of individuals and when. Indeed, economic models of rational behavior and evolutionary models of optimal foraging suggest that animals should choose avoid the costs of challenges, making it seem as though they would prefer or want a challenge-free life. An engagement in effectiveness model, on the other hand, outlines an answer to why challenge can be good. Because animals benefit from engaging in effectiveness, a challenge will benefit an animal's welfare to the extent that it affords engagement in effectiveness opportunities but will be harmful to welfare if it causes engagement with effectiveness failure: disengagement, engagement with ineffectiveness, value failure, truth failure, or control failure.

Finally, human models of welfare—i.e. well-being research—provide many insights and potential directions for future research. In the current program, well-being research helped refine, confirm, and extend the theoretical approach (Studies 1.1, 1.2a & 1.2b, 2.1a & 2.1b, and

3.1a & 3.1b). Nevertheless, many themes in the well-being literature may be ill suited for animal welfare research in general. Cross-fertilization between the two fields requires a model that is species-neutral, yet capable of accounting for sophisticated, species-specific patterns. These studies have indicated that an engaging in effectiveness perspective may provide that necessary common ground and thereby suggest new directions for future research.

Though the engaging in effectiveness model does seem to be uniquely capable of accounting for this body of results, much work remains in establishing its relevance to welfare and well-being. First, experimental work is required with human participants. Though modulating human well-being in the lab may prove to be difficult, longitudinal and prospective studies may provide a powerful intermediate for testing causal pathways. For example, a recent study found that self-reported effort in a breathing-task accounted for increases in well-being over time; though success expectancy and adherence were also related, the determining factor was self-reported effort (Gaitan-Sierra & Hyland 2011). These results fit well with the engaging in effectiveness perspective on challenge: Challenge—or level of perceived challenge, i.e. self-reported effort—can improve welfare by increasing environmental demand and thus generating an experience of effective engagement.

Even without manipulation, the challenge-welfare relationship could be further explored with diary studies by tracking the welfare of participants as they encounter and surmount (or are stymied by) challenges. This type of observational-exploratory research could also address the relationship between engagement in effectiveness and health. Are highly effective people biologically and behaviorally healthier than less effective people? Does having engaging experiences improve health, and if so, how and why? By extension, does engagement with effectiveness confer physiological benefits to laboratory rats? The controlled environments and

shorter lifespan of many nonhuman captive animals could facilitate exploration of these questions of causality.

In addition to biological outcomes, research should continue exploring the behavioral outcomes of good welfare. The final study (4.2) indicated that behavior in the maze returned to baseline following enrichment. It is possible, however, that this result is an artifact of the experimental design in that it was the rats' second exposure to the ambiguous arms. Previous experience with these arms may have rendered them useless indicators of welfare status. This possibility would be somewhat discredited if it could be shown that even after multiple exposures, behavior in the ambiguous arms remained sensitive to welfare changes. Additional research along these lines is needed.

What does the engagement with effectiveness model have to say about individual-environment fit? It is possible that individuals (humans and nonhumans) place differential emphasis on each of the three effectiveness domains. If that is the case, it may turn out that a 'truth-emphasizer,' for example, gains more welfare benefits from opportunities to engage in truth effectiveness compared to other individuals *and* other opportunities of being effective. Indeed, some of the work on cognitive challenges indicates the presence of these types of individual differences (Langbein et al., 2009). Testing the interaction between individual and environment with an engagement in effectiveness framework may help clarify when and why enrichment improves welfare at the individual level.

More broadly, this research highlights the need for research exploring relationship between challenges and good welfare. The results of Studies 3.3, 4.1 and 4.2 should be replicated in other rat cohorts and additional species. Aside from replication, there are many unanswered questions regarding the role that challenges play in good welfare. How do

challenges that are predictable but hard to manage impact welfare? If animals receive frequent, but manageable and predictable challenges early in development, will they become more effective at dealing with different types of challenges later in life? Will they become more effective overall? Many possibilities exist and as yet, relatively little is known.

The engagement in effectiveness perspective of challenges casts new light on the seemingly paradoxical result that animals (including humans) sometimes value stimuli and outcomes associated with higher costs. For example, in direct contrast to a cost-benefit model of value, researchers found that starlings preferred colors associated with longer versus shorter flights (Kacelnik & Marsh, 2002). The engagement in effectiveness model predicts this result and suggests a variety of extensions: After the birds learned that the efforts were going to be rewarded (truth and value effectiveness), the longer flights (control effectiveness) afforded them a greater experience of engaging in effectiveness and were thus more valuable. Expanding on this interpretation, the engaging in effectiveness framework would predict that when an equivalent amount of effort produces the same reward (control and value effectiveness), tasks are more valuable when they require more versus less learning (truth effectiveness). This hypothesis could be tested in rats by presenting them with a learning challenge in which an equivalent number of lever presses produced a reward, but in one case, the pattern of lever presses was more difficult to learn than in the other. An individual rat's preference for learning challenges could then be assessed by giving it a choice between chambers associated with each type of learning. Modifying this experimental design to vary by each domain (value, truth, and control effectiveness), individual differences in domain-emphasis may be detectable and, applied across species, this paradigm could reveal species level differences in domain preference as well.

Similarly, the engagement in effectiveness perspective could augment the theory that risky, even dangerous play may be necessary in the development of young children. Recently psychologists have suggested that risky play is a way that children explore challenges in their environment and thereby boost their self-confidence and coping-skills (Sandseter & Kennair 2011). According to the engaging in effectiveness model, challenging play provides diverse ways of engaging in effectiveness and thus enhances welfare. Studies with laboratory rats could experimentally examine this hypothesis by giving one group of rats challenging play opportunities (e.g., tall structures with ropes, ramps, and bridges) and another group of rats, non- or less-challenging play opportunities (e.g. running wheels, toys, etc). Later in life, these rats could be tested for differences in their ability to engage in species-specific ways of being effective (e.g., maintaining darkness or obtaining treats). According to engaging in effectiveness model, the challenging play would confer *general* benefits to adult rats beyond the direct skills they may have acquired as a result of their interaction with the play devices.

Undoubtedly there are several applications of challenge-welfare research, but reintroduction programs may be an especially germane venue. Scientists have noted that animal welfare science is an underutilized sister-field to conservation biology (Teixeira, De Azevedo, Mendl, Cipreste, & Young, 2007) and challenge-welfare research in particular suggests several intriguing proposal. For example, behavioral variability is of central importance to reintroduction programs as it is likely to be crucial to survival in the wild (Watters & Meehan, 2007). An engagement with effectiveness framework suggests ways in which conservation biologists could increase the behavioral variability of the animals in their care. Captive environments that included species-specific, predictable, and manageable challenges would maximize the opportunities to engage in effectiveness and increase the diversity of

environmental demands to which the captive animal must devote its resource. In theory, these environments should produce a greater range of behavior than environments that did not contain such challenges. Aside from decreased behavioral diversity, captive animals have also been found to be deficient in specific behaviors, including anti-predator and vigilant behaviors (Rantanen, Buner, Riordan, Sotherton, & Macdonald, 2010; Zidon, Saltz, Shore, & Motro, 2009). Critically, vigilance behavior was found to be a key predictor of post-release survival (e.g. Kreger, Hatfield, Estevez, Gee, & Clugston, 2006). Study 3.3 demonstrated that challenges can increase vigilance in the form of darkness (safety) pursuit. It would be interesting to discover if this increased vigilance in the lab would translate to greater fitness in the wild and whether challenge enrichment in general contributed to post-release success.

In the 1990s researchers found some evidence that challenges enhanced welfare, though the results were mixed (Moodie & Chamove, 1990). Like that research, the current program opens as many questions as it answers. Nevertheless, an engagement in effectiveness framework may prove to be an instrumental guide in planning and interpreting future research in animal welfare science.

REFERENCES

Appleby, M. C., & Sandoe, P. (2002). Philosophical debate on the nature of well-being: Implications for animal welfare. *Animal Welfare, 11*(3), 283-294.

Archer, J. (1973). TESTS FOR EMOTIONALITY IN RATS AND MICE - REVIEW. Animal Behaviour, 21(MAY), 205-235. doi: 10.1016/s0003-3472(73)80065-x

Barnett, J. L., & Hemsworth, P. H. (1990). THE VALIDITY OF PHYSIOLOGICAL AND BEHAVIORAL MEASURES OF ANIMAL-WELFARE. *Applied Animal Behaviour Science, 25*(1-2), 177-187.

Bateson, M., Desire, S., Gartside, S. E., & Wright, G. A. (2011). Agitated Honeybees Exhibit Pessimistic Cognitive Biases. *Current Biology, 21*(12), 1070-1073. doi: 10.1016/j.cub.2011.05.017

Bezard, E., Dovero, S., Belin, D., Duconger, S., Jackson-Lewis, V., Przedborski, S., . . . Jaber, M. (2003). Enriched environment confers resistance to 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine and cocaine: Involvement of dopamine transporter and trophic factors. *Journal of Neuroscience, 23*(35), 10999-11007.

Boissy, A., Manteuffel, G., Jensen, M. B., Moe, R. O., Spruijt, B., Keeling, L. J., . . . Aubert, A. (2007). Assessment of positive emotions in animals to improve their welfare. *Physiology & Behavior, 92*(3), 375-397. doi: 10.1016/j.physbeh.2007.02.003

Brilot, B. O., Asher, L., & Bateson, M. (2010). Stereotyping starlings are more 'pessimistic'. *Animal Cognition, 13*(5), 721-731. doi: 10.1007/s10071-010-0323-z

Broom, D. M. (1991). ANIMAL-WELFARE - CONCEPTS AND MEASUREMENT. *Journal of Animal Science, 69*(10), 4167-4175.

Broom, D. M. (2011). A History of Animal Welfare Science. Acta Biotheoretica, 59(2), 121-137. doi: 10.1007/s10441-011-9123-3

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science, 6*(1), 3-5. doi: 10.1177/1745691610393980

Burman, O., McGowan, R., Mendl, M., Norling, Y., Paul, E., Rehn, T., & Keeling, L. (2011). Using judgement bias to measure positive affective state in dogs. *Applied Animal Behaviour Science, 132*(3-4), 160-168. doi: 10.1016/j.applanim.2011.04.001

Burman, O. H. P., Parker, R. M. A., Paul, E. S., & Mendl, M. T. (2009). Anxiety-induced cognitive bias in non-human animals. *Physiology & Behavior, 98*(3), 345-350. doi: 10.1016/j.physbeh.2009.06.012

Calvo-Torrent, A., Brain, P. F., & Martinez, M. (1999). Effect of predatory stress on sucrose intake and behavior on the plus-maze in male mice. *Physiology & Behavior, 67*(2), 189-196. doi: 10.1016/s0031-9384(99)00051-7

Chamove, A. S., & Moodie, E. M. (1990). ARE ALARMING EVENTS GOOD FOR CAPTIVE MONKEYS. *Applied Animal Behaviour Science, 27*(1-2), 169-176. doi: 10.1016/0168-1591(90)90016-7

Dawkins, M. S. (1990). FROM AN ANIMALS POINT OF VIEW - MOTIVATION, FITNESS, AND ANIMAL-WELFARE. [Review]. *Behavioral and Brain Sciences, 13*(1), 1-&.

De Boer, S. F., & Koolhaas, J. A. (2003). Defensive burying in rodents: ethology, neurobiology and psychopharmacology. [Review]. *European Journal of Pharmacology, 463*(1-3), 145-161. doi: 10.1016/s0014-2999(03)01278-0

Diener, E. (2000). Subjective well-being - The science of happiness and a proposal for a national index. *American Psychologist, 55*(1), 34-43.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). THE SATISFACTION WITH LIFE SCALE. *Journal of Personality Assessment, 49*(1), 71-75. doi: 10.1207/s15327752jpa4901_13

Diener, E., & Seligman, M. E. P. (2002). Very happy people. [Article]. *Psychological Science, 13*(1), 81-84. doi: 10.1111/1467-9280.00415

Ennaceur, A., Michalikova, S., & Chazot, P. L. (2006). Models of anxiety: Responses of rats to novelty in an open space and an enclosed space. *Behavioural Brain Research, 171*(1), 26-49. doi: 10.1016/j.bbr.2006.03.016

Flecknell, P., Leach, M., & Bateson, M. (2011). Affective state and quality of life in mice. *Pain, 152*(5), 963-964. doi: 10.1016/j.pain.2011.01.030\

Franks, Becca (2010). Feeling effective and self-assessed health. Unpublished Masters of Philosophy Requirement, Columbia University

Fraser, D. (1999). Animal ethics and animal welfare science: bridging the two cultures. *Applied Animal Behaviour Science, 65*(3), 171-189. doi: 10.1016/s0168-1591(99)00090-8

Galhardo, L., Correia, J., & Oliveira, R. F. (2008). The effect of substrate availability on behavioural and physiological indicators of welfare in the African cichlid (Oreochromis mossambicus). *Animal Welfare, 17*(3), 239-254.

Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.

Gaitan-Sierra, Carolina; Hyland, Michael E. (2011) Nonspecific mechanisms that enhance well-being in health-promoting behaviors. Health Psychology, Sep 26, 2011, No Pagination Specified. doi: 10.1037/a0025582

Hagen, K., Van den Bos, R., & Buning, T. d. C. (2011). Editorial: Concepts of Animal Welfare. *Acta Biotheoretica, 59*(2), 93-103. doi: 10.1007/s10441-011-9134-0

Hall, C. S. (1934). Emotional behavior in the rat I Defecation and urination as measures of individual differences in emotionality. Journal of Comparative Psychology, 18(3), 385-403. doi: 10.1037/h0071444

Harding, E. J., Paul, E. S., & Mendl, M. (2004). Animal behavior - Cognitive bias and affective state. *Nature, 427*(6972), 312-312. doi: 10.1038/427312a

Harlow, H. F. (1950). LEARNING AND SATIATION OF RESPONSE IN INTRINSICALLY MOTIVATED COMPLEX PUZZLE PERFORMANCE BY MONKEYS. *Journal of Comparative and Physiological Psychology, 43*(4), 289-294. doi: 10.1037/h0058114

Held, S. D. E. H. S. D. E., & Spinka, M. (2011). Animal play and animal welfare. *Animal Behaviour, 81*(5), 891-899. doi: 10.1016/j.anbehav.2011.01.007

Higgins, E. T. 2011. *Beyond Pleasure and Pain: How motivation works*. New York, NY: Oxford University Press.

Hughes, B. O., & Duncan, I. J. H. (1988). THE NOTION OF ETHOLOGICAL NEED, MODELS OF MOTIVATION AND ANIMAL-WELFARE. *Animal Behaviour, 36*, 1696-1707.

Kacelnik, A. & Marsh, B. 2002. Cost can increase preference in starlings. Animal Behaviour 63, 245-250.

Korte, S. M. (2001). Corticosteroids in relation to fear, anxiety and psychopathology. Neuroscience and Biobehavioral Reviews, 25(2), 117-142.

Korte, S. M., Olivier, B., & Koolhaas, J. M. (2007). A new animal welfare concept based on allostasis. *Physiology & Behavior, 92*(3), 422-428. doi: 10.1016/j.physbeh.2006.10.018

Kreger, M. D., Hatfield, J. S., Estevez, I., Gee, G. F., & Clugston, D. A. (2006). Behavioral profiles of the captive juvenile whooping crane as an indicator of post-release survival. *Zoo Biology, 25*(1), 11-24. doi: 10.1002/zoo.20075

Langbein, J., Siebert, K., & Nurnberg, G. (2009). On the use of an automated learning device by group-housed dwarf goats: Do goats seek cognitive challenges? [Article]. *Applied Animal Behaviour Science, 120*(3-4), 150-158. doi: 10.1016/j.applanim.2009.07.006

Lopez, B. (1986). Artic Dreams. New York City, NY: Vintage Books, Random House.

Lefcourt, H. M. (1966). INTERNAL VERSUS EXTERNAL CONTROL OF REINFORCEMENT - A REVIEW. *Psychological Bulletin, 65*(4), 206-&. doi: 10.1037/h0023116

Mason, G., & Mendl, M. (1997). Do the stereotypies of pigs, chickens and mink reflect adaptive species differences in the control of foraging? *Applied Animal Behaviour Science, 53*(1-2), 45-58.

Mason, G. J. (1991). STEREOTYPIES - A CRITICAL-REVIEW. *Animal Behaviour, 41*, 1015-1037. doi: 10.1016/s0003-3472(05)80640-2

Matuszewich, L., Karney, J. J., Carter, S. R., Janasik, S. P., O'Brien, J. L., & Friedman, R. D. (2007). The delayed effects of chronic unpredictable stress on anxiety measures. *Physiology & Behavior, 90*(4), 674-681. doi: 10.1016/j.physbeh.2006.12.006

Meehan, C. L., & Mench, J. A. (2007). The challenge of challenge: Can problem solving opportunities enhance animal welfare? *Applied Animal Behaviour Science, 102*(3-4), 246-261. doi: 10.1016/j.applanim.2006.05.031

Mendl, M., Burman, O. H. P., Parker, R. M. A., & Paul, E. S. (2009). Cognitive bias as an indicator of animal emotion and welfare: Emerging evidence and underlying mechanisms. *Applied Animal Behaviour Science, 118*(3-4), 161-181. doi: 10.1016/j.applanim.2009.02.023

Mendl, M., & Paul, E. S. (2004). Consciousness, emotion and animal welfare: insights from cognitive science. *Animal Welfare, 13*, S17-S25.

Mendl, M., Paul, E. S., & Chittka, L. (2011). Animal Behaviour: Emotion in Invertebrates? *Current Biology, 21*(12), R463-R465. doi: 10.1016/j.cub.2011.05.028

Moltschaniwskyj, N. A., Hall, K., Marian, J., Nishiguchi, M., Sakai, M., Shulman, D. J., . . . Warnke, K. (2007). Ethical and welfare considerations when using cephalopods as experimental animals. *Reviews in Fish Biology and Fisheries, 17*(2-3), 455-476. doi: 10.1007/s11160-007-9056-8

Moodie, E. M., & Chamove, A. S. (1990). BRIEF THREATENING EVENTS BENEFICIAL FOR CAPTIVE TAMARINS. *Zoo Biology, 9*(4), 275-286. doi: 10.1002/zoo.1430090403

Novak, M. A., Kinsey, J. H., Jorgensen, M. J., & Hazen, T. J. (1998). Effects of puzzle feeders on pathological behavior in individually housed rhesus monkeys. *American Journal of*

Primatology, 46(3), 213-227. doi: 10.1002/(sici)1098-2345(1998)46:3<213::aid-ajp3>3.0.co;2-l

Puppe, B., Ernst, K., Schon, P. C., & Manteuffel, G. (2007). Cognitive enrichment affects behavioural reactivity in domestic pigs. *Applied Animal Behaviour Science, 105*(1-3), 75-86. doi: 10.1016/j.applanim.2006.05.016

Rantanen, E. M., Buner, F., Riordan, P., Sotherton, N., & Macdonald, D. W. (2010). Vigilance, time budgets and predation risk in reintroduced captive-bred grey partridges Perdix perdix. *Applied Animal Behaviour Science, 127*(1-2), 43-50. doi: 10.1016/j.applanim.2010.08.002

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78.

Salmeto, A. L., Hymel, K. A., Carpenter, E. C., Brilot, B. O., Bateson, M., & Sufka, K. J. (2011). Cognitive bias in the chick anxiety-depression model. *Brain Research, 1373*, 124-130. doi: 10.1016/j.brainres.2010.12.007

Sambrook, T. D., & BuchananSmith, H. M. (1997). Control and complexity in novel object enrichment. *Animal Welfare, 6*(3), 207-216.

Sandseter Ellen Beate Hansen Leif Edward Ottesen Kennair (2011). Children's Risky Play from an Evolutionary Perspective: The Anti-Phobic Effects of Thrilling Experiences. Evolutionary Psychology 9(2): 257-284

Sanger, M. E., Doyle, R. E., Hinch, G. N., & Lee, C. (2011). Sheep exhibit a positive judgement bias and stress-induced hyperthermia following shearing. *Applied Animal Behaviour Science, 131*(3-4), 94-103. doi: 10.1016/j.applanim.2011.02.001

Scheier, M.F., Carver C. S., and Bridges, M.W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A re-evaluation of the Life Orientation Test. Journal of Personality and Social Psychology, 67, 1063-1078.

Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology - An introduction. [Editorial Material]. *American Psychologist, 55*(1), 5-14.

Teixeira, C. P., De Azevedo, C. S., Mendl, M., Cipreste, C. F., & Young, R. J. (2007). Revisiting translocation and reintroduction programmes: the importance of considering stress. *Animal Behaviour, 73*, 1-13. doi: 10.1016/j.anbehav.2006.06.002

van de Pol, M. V., & Wright, J. (2009). A simple method for distinguishing within- versus between-subject effects using mixed models. [Editorial Material]. *Animal Behaviour, 77*(3), 753-758. doi: 10.1016/j.anbehav.2008.11.006

Watters, J. V., & Meehan, C. L. (2007). Different strokes: Can managing behavioral types increase post-release success? [Article]. *Applied Animal Behaviour Science, 102*(3-4), 364-379. doi: 10.1016/j.applanim.2006.05.036

Zidon, R., Saltz, D., Shore, L. S., & Motro, U. (2009). Behavioral Changes, Stress, and Survival Following Reintroduction of Persian Fallow Deer from Two Breeding Facilities. *Conservation Biology, 23*(4), 1026-1035. doi: 10.1111/j.1523-1739.2008.01163.x

FIGURES AND TABLES

Figure 1: Effectiveness and success expectancies: solving a problem

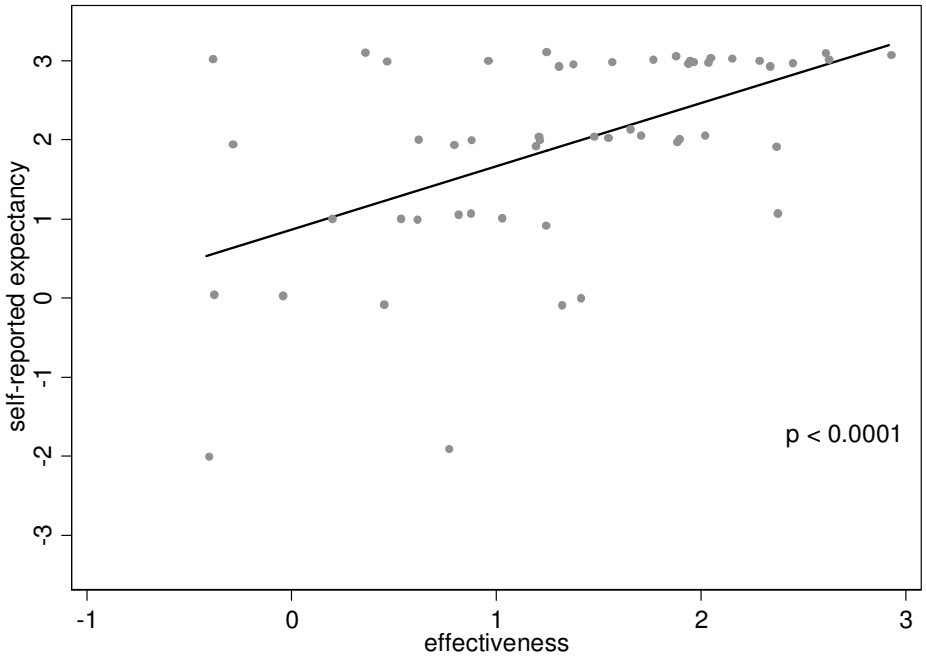


Figure 2: Effectiveness and success expectancies: achieving a personal goal

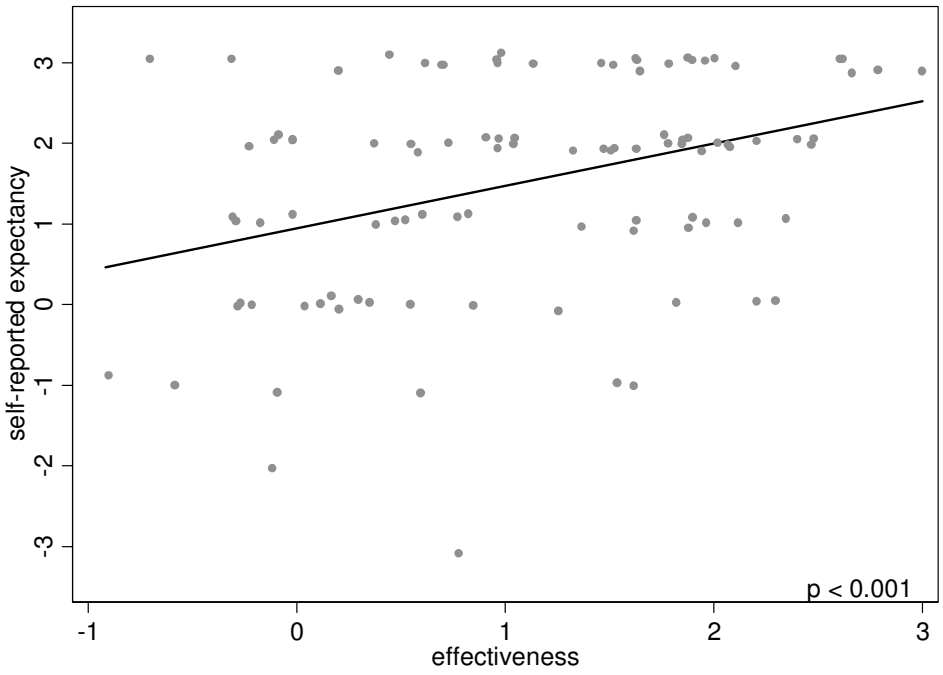
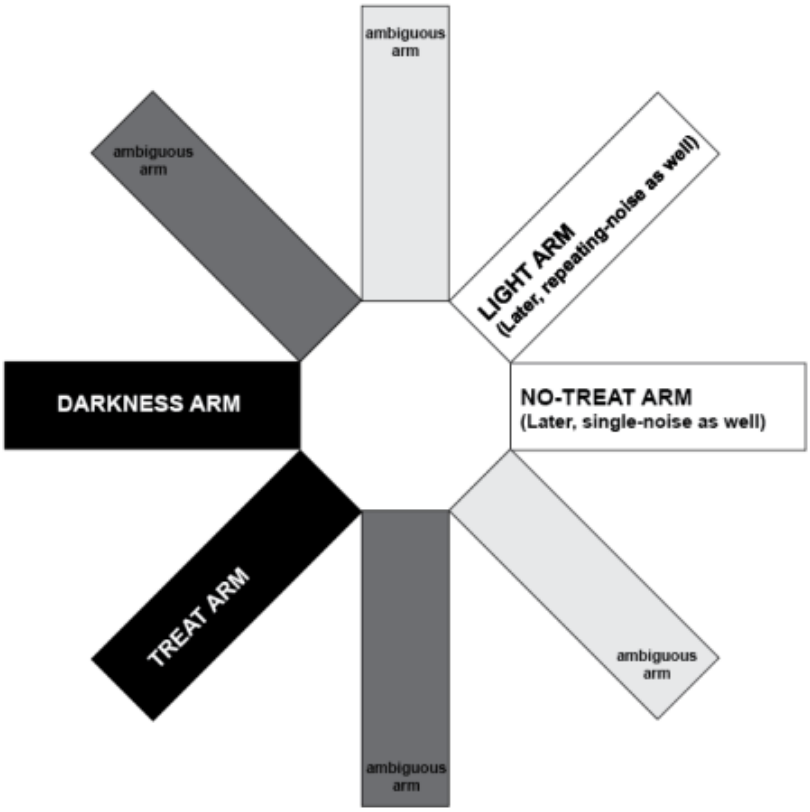
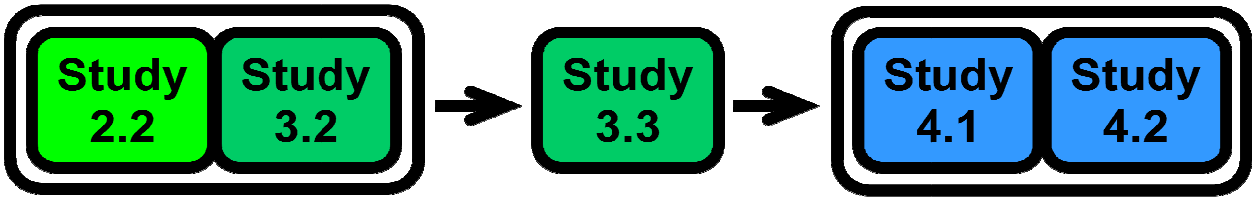


Figure 3: Radial Arm-Maze Overhead View



*Note: the ambiguous arms were only accessible to the rats in Studies 4.1 and 4.2.

Figure 4: Rat Studies Diagram



Studies 2.2. and 3.2 occurred in the same setting. Study 3.3 occurred 2 months after the previous studies. Studies 4.1 and 4.2 occurred 2 months after the previous study.

Figure 5: Individual responses to each arm over time

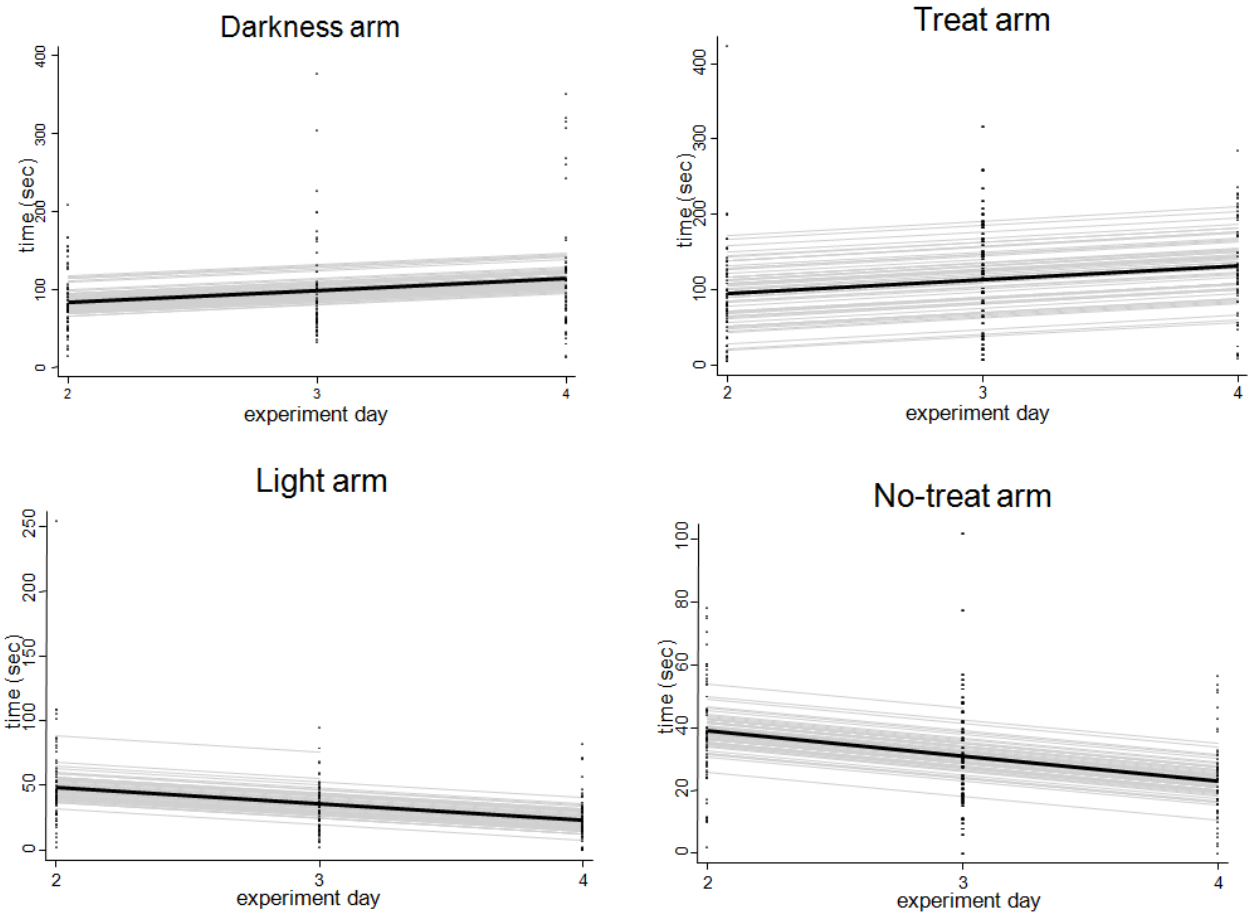


Figure 6: Stress and darkness

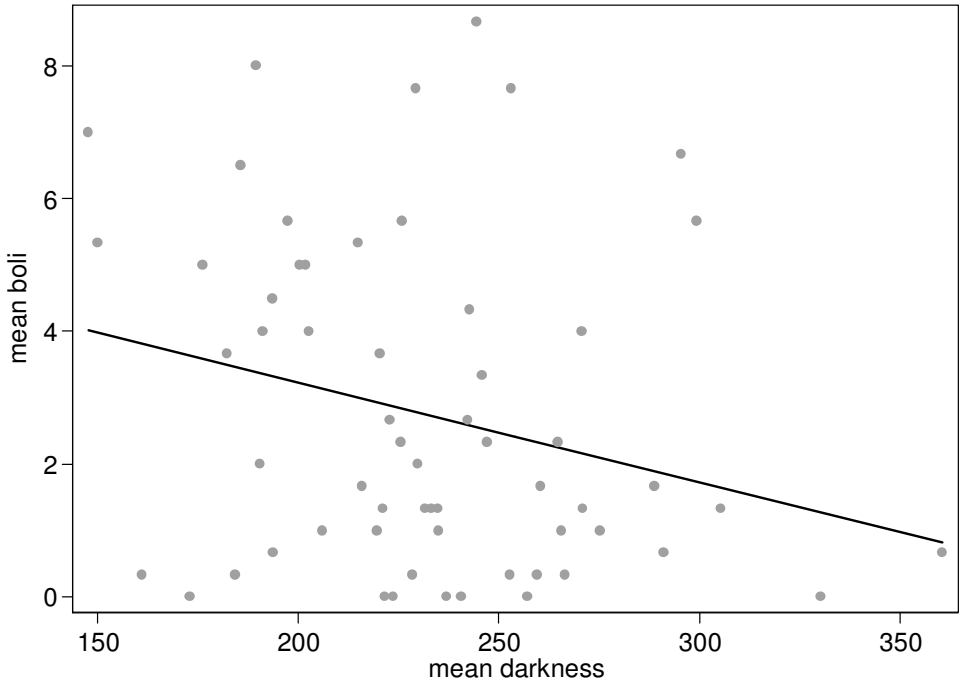


Figure 7: Stress and treats

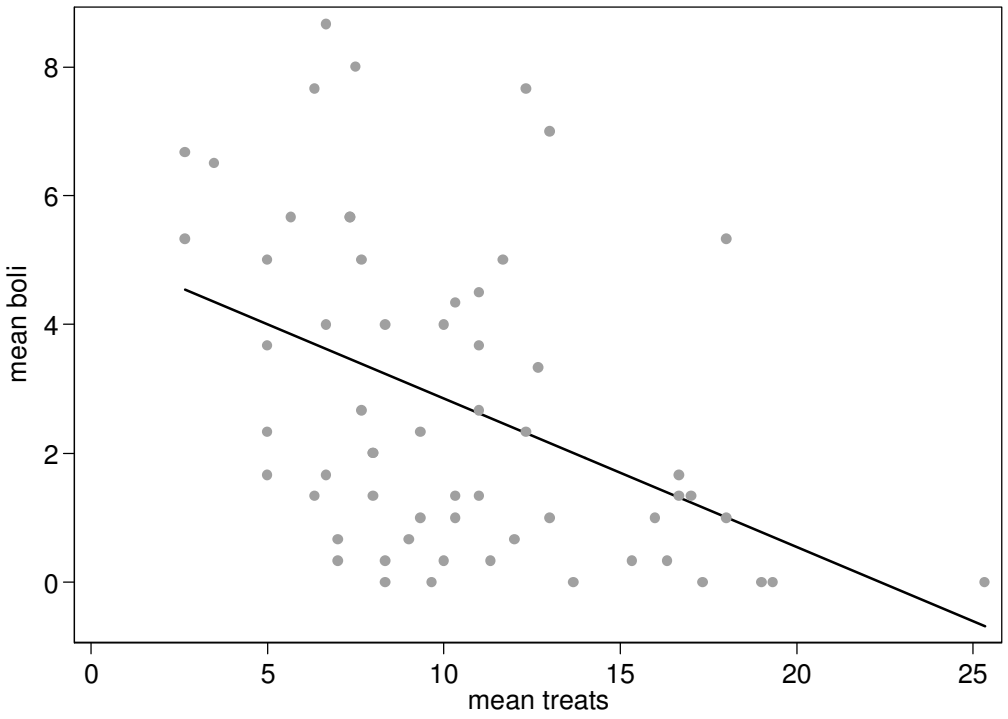


Figure 8: The number of challenges and effectiveness

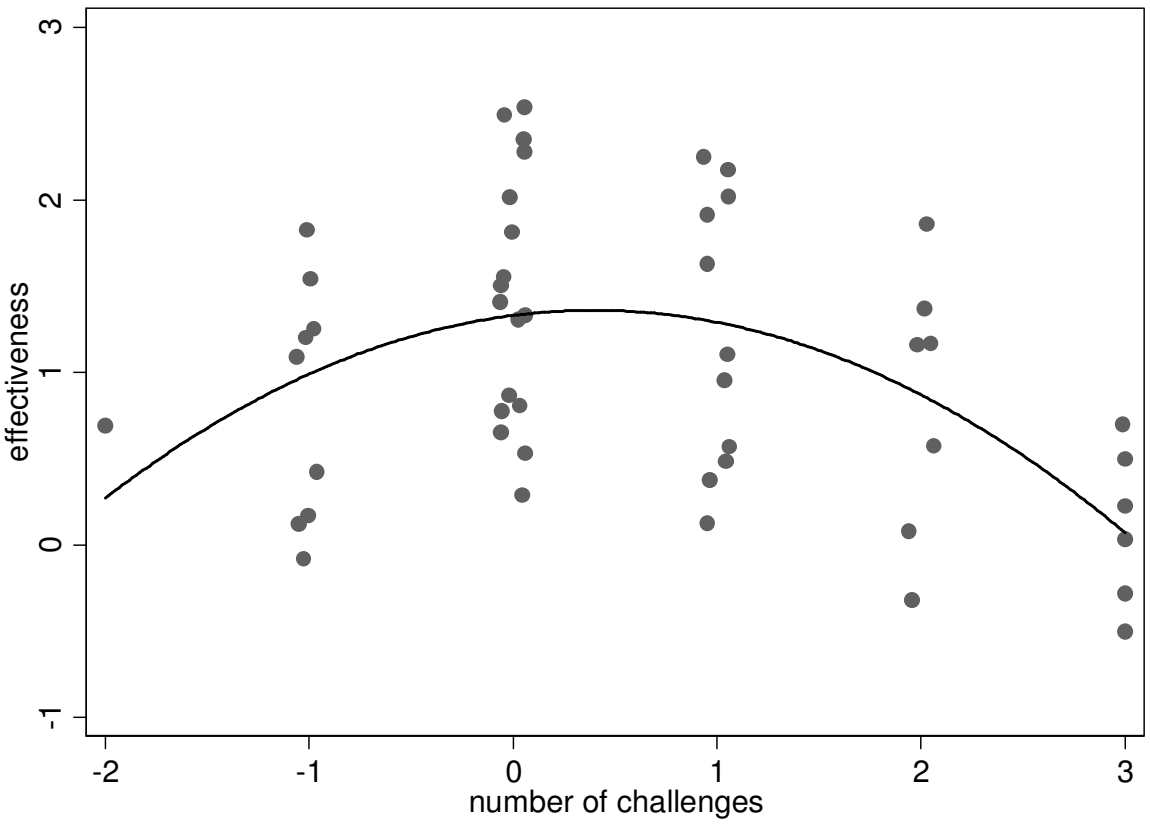


Figure 9: The number of challenges and depression

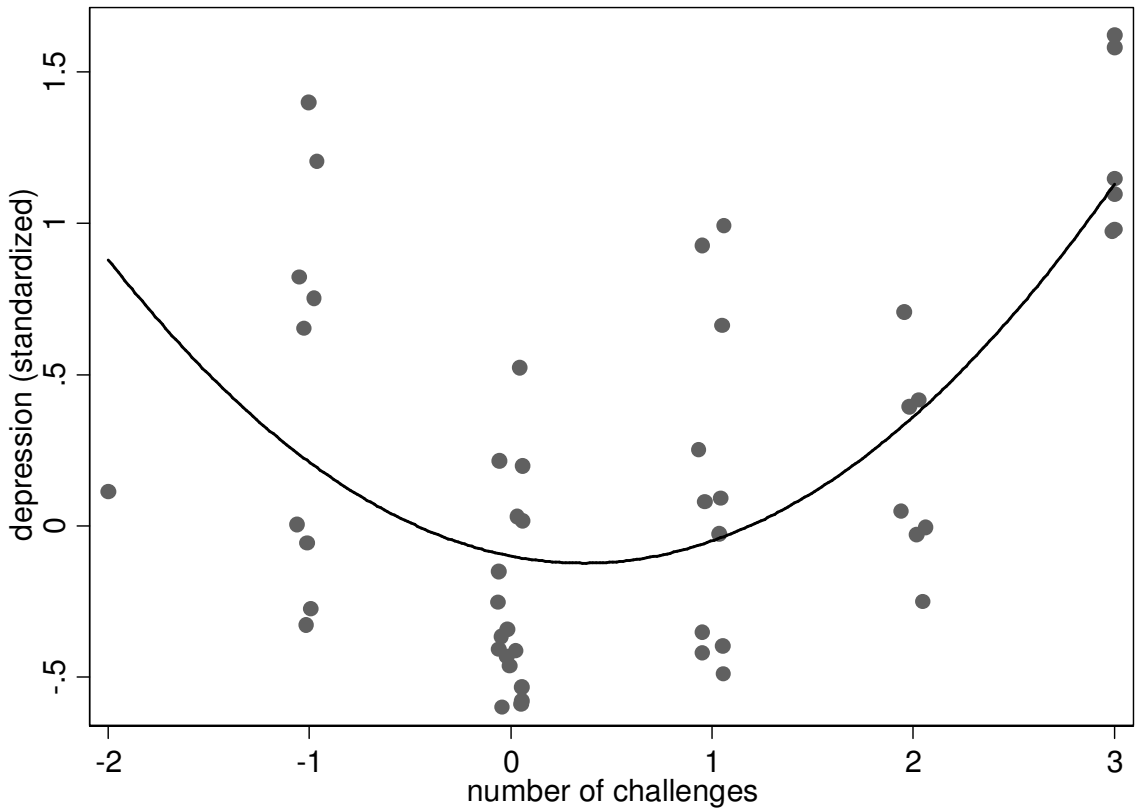


Table 1: Multilevel models of rat behavior in the automated radial arm maze

Table 1.

Outcome	Intercept (SD)	Intercept SD: 95% CI	Experimental Day Slope (CI)
Time in success (darkness) arm	84 (23)	12, 43; p=0.006	14.79 (4.31, 26.27); z=2.77, p=0.006
Time in success (treat) arm	94 (43)	33, 56; p<0.0001	18.73 (10.13, 27.34); z=4.27, p<0.0001
Time in failure (light) arm	48 (13)	8, 20; p<0.0001	-12.20 (-16.3, -8.09); z=5.82, p< 0.0004
Time in failure (no-treat) arm	39 (7)	5, 12; p<0.0001	-7.66 (-10.24, -5.09); z=5.82, p< 0.0001
darkness time	216 (27)	17, 43; p=0.001	16.99 (6.83, 27.14); z=3.29, p=0.001
treat activations	2 (0.4)	0.3, 0.5; p<0.0001	0.18 (0.12, 0.24); z=6.26, p<0.0001
fecal boli	0.8 (1)	0.8, 1.3; p<0.0001	-0.25 (-0.37, -0.14); z=4.32, p<0.0001

*Note: All individual N=60, all mean observations n=2.9

APPENDIX A: QUESTIONNAIRES

Effectiveness Questions:

1. I don't put a lot of effort into life.
2. I often feel too tired to engage in my daily activities.
3. Usually, I am very good at figuring things out.
4. I see little beauty in the world.
5. I have what I need to get what I want.
6. Organizing is one of my strengths.
7. There are a lot of things going wrong in my life.
8. I am easily distracted from my pursuits.
9. The people around me have faith in my abilities.
10. I always give up when I become confused.
11. I think that life is meaningful.
12. I am in a bad situation.

Effectiveness Questions:

1. I don't put a lot of effort into life.
2. I often feel too tired to engage in my daily activities.
3. I am very good at figuring things out.
4. Every day I see something beautiful.
5. I have what I need to get what I want.

6. Organizing is one of my strengths.
7. There are a lot of things going wrong in my life.
8. I am easily distracted from my goals.
9. The people around me have faith in my abilities.
10. I always give up when I become confused.
11. I think that life is meaningful.
12. I am in a bad situation.
13. There is a good balance between order in chaos in my life.
14. My life is boring.

Well-being: Satisfaction with life (Diener et al, 1985)

In most ways my life is close to my ideal.

The conditions of my life are excellent.

I am satisfied with my life.

So far I have gotten the important things I want in life.

If I could live my life over, I would change almost nothing.

Depression: CES-D

	During the Past Week			
	Rarely or none of the time (less than 1 day)	Some or a little of the time (1-2 days)	Occasionally or a moderate amount of time (3-4 days)	Most or all of the time (5-7 days)
1. I was bothered by things that usually don't bother me.	☐	☐	☐	☐
2. I did not feel like eating; my appetite was poor.	☐	☐	☐	☐
3. I felt that I could not shake off the blues even with help from my family or friends.	☐	☐	☐	☐
4. I felt I was just as good as other people.	☐	☐	☐	☐
5. I had trouble keeping my mind on what I was doing.	☐	☐	☐	☐
6. I felt depressed.	☐	☐	☐	☐
7. I felt that everything I did was an effort.	☐	☐	☐	☐
8. I felt hopeful about the future.	☐	☐	☐	☐
9. I thought my life had been a failure.	☐	☐	☐	☐
10. I felt fearful.	☐	☐	☐	☐
11. My sleep was restless.	☐	☐	☐	☐
12. I was happy.	☐	☐	☐	☐
13. I talked less than usual.	☐	☐	☐	☐
14. I felt lonely.	☐	☐	☐	☐
15. People were unfriendly.	☐	☐	☐	☐
16. I enjoyed life.	☐	☐	☐	☐
17. I had crying spells.	☐	☐	☐	☐
18. I felt sad.	☐	☐	☐	☐
19. I felt that people dislike me.	☐	☐	☐	☐
20. I could not get "going."	☐	☐	☐	☐

APPENDIX B: SURVEY WORDING

Study 1.2a: Effectiveness and Success Expectancies: figuring out a problem:

Think about a problem you are currently experiencing. In the space, below describe the context and details of the problem and how you plan to solve it:

How likely is it that you will figure out this problem?

Extremely unlikely	Unlikely	Somewhat unlikely	Fifty/fifty	Somewhat likely	Likely	Extremely likely
○	○	○	○	○	○	○

How easy/difficult will it be to figure out this problem?

Extremely easy	Easy	Somewhat easy	Neutral	Somewhat difficult	Difficult	Extremely difficult
○	○	○	○	○	○	○

Study 1.2b: Effectiveness and Success Expectancies, personal goal

Now, take some time to think of a important, personal goal that you have not yet attained.

Once you have thought of a goal, press the arrow to continue.

Picture the details of this goal...

- **Where will you be?**
- **What are your surroundings?**
- **What people and objects are involved?**
- **Are there specific sounds or smells associated with it?**

Now think about the things that you have that will help you reach your goal. These could be . . .

- **Personal attributes like intelligence, kindness, experience, determination, sense of humor, etc.**
- **Relationships like family, friends, colleagues, etc.**
- **Groups to which you belong like community organizations, racial identity, or religious institutions, etc.**
- **Or important values like faith, art/music, or nature, etc.**

In the space below, please write about your goal. What are the important resources that you have? How can they help you figure things out or provide you with a sense of discovery and insight about the situation? Finally, how will your resources and understanding help you be proactive, confident, organized and effective at managing what happens?

[Remember the details of your goal. Where will you be? What are your surroundings? What people and objects are involved? Are there specific sounds or smells associated with it? And think about the resources you have: personal attributes like intelligence, kindness, experience, determination, sense of humor…relationships like family or friends… groups like community organizations or your racial group…important values like faith, art/music, or nature].

How likely is it that you will reach this goal?

Extremely unlikely	Unlikely	Somewhat unlikely	Fift-fifty	Somewhat likely	Likely	Extremely likely
○	○	○	○	○	○	○

How easy/difficult will it be to reach this goal?

Extremely easy	Easy	Somewhat easy	Neutral	Somewhat difficult	Difficult	Extremely difficult
○	○	○	○	○	○	○

Studies 2.1a & 2.1b: Engaging Experiences, Well-being, and Effectiveness

There are times in life when everything works perfectly-when you are so engaged and activated by what you are doing that you lose track of time; ideas, thoughts, actions just 'flow.'

In the space below, please write about a time when you has such an experience. What was the context (sights, sounds, smells, etc.)? What were you doing? How did you feel?

How often do you have these types of experiences?

Never	Rarely	Off and on	Frequently	All the time
○	○	○	○	○

Studies 3.1a & 3.1b Challenge Frequency, Effectiveness, and Depression:

All of us experience challenges-- times during which our abilities, talents, creativity, and resources are put to the test.

Right now, how would you describe the number of challenges in your life?

Too many			Just right			Too few
○	○	○	○	○	○	○

APPENDIX C: WRITING EXAMPLES

2.1a & 2.1b: Examples of challenging experiences reported in engagement writing

I work in an emergency hospital and we had a "crash"- a patient who came in not breathing, non responsive. Our team jumped into action and we barely needed to communicate out loud. People ran and got the items we needed and hooked up monitors as needed. The whole time just flowed and it was about 20 minutes later that we had to pronounce but we didn't realize it.

*****

Several months ago, I was feeling overwhelmed by work. I had several major projects due for work and I wasn't sure I was going to be able to get them all done in time. I ended up staying up late one night to work on them. I turned on some upbeat music which really motivated me. I worked all night and by morning I had not only completed the projects but I had done an excellent job. I felt fantastic. I love when I am able to focus and complete something.

*****

In Iraq during a combat patrol. There was smoke, dust, sand, bullets flying, loud sound, Did not think of anything just went through motions seemed like automatic response. No time to think just react.

*****

In law school on my constitutional law final exam, a three hour exam, I opened the exam booklet and saw a question an essay question about the equal protection clause. Equal protection jurisprudence is my favorite topic. During that three hours I became so focused and became so involved in the question that I completely lost track of the exam time. I had brought a cup of coffee into the exam with me, but completely forgot that it was there and let it get cold. I don't remember sights, sounds,

smells, etc. I was just really focused, like the rest of the world had ceased to exist. It was just me and the question.

2.1a & 2.1b: Examples of non-challenging experiences reported in engagement writing

The last time I felt this way, We went camping up in the mountains to get away from Everything... It was nature at it's purest... Birds, beautiful scenery, trees, water flowing, etc. It was such Beauty in such a simple way, like we didn't have a care or worry in the world around us....

*****

I was on a beach in Bermuda. The sight of the pink sand and turquoise water filled me with a sense of euphoria. I was so engaged that I nearly missed my airport transfer.

*****

When I took my kids to the park. We were just having a great time laughing and playing. The smell of fresh cut grass was in the air and the sound of kids playing. There was a lot of smiling going on! It felt really great.

*****

When I got together with 8 girlfriends I knew from long ago, and we hadn't seen each other for many, many years, some of us 38 years. I was very engaged with talking with them that for the few days I was with them, I totally lost track of time. Yes, I knew it was morning or I knew it was dark, but time, specific time, was of no importance at all. I felt very good and extremely happy. We were having a great time whether we were on the beach, in a restaurant, in the hotel rooms, at the beach house, on the ferry, etc. I have not felt so happy and laughed that hard and long for a long, long, time.

APPENDIX D: RADIAL ARM MAZE

Side View