Psychological Value Theory:

The Psychological Value of Human Lives and Economic Goods

Dale J. Cohen, Amanda R. Cromley, Katelyn E. Freda, and Madeline White

The University of North Carolina Wilmington

Word Count: 18,800

Author Note

Dale J. Cohen, Amanda R. Cromley, Katelyn E. Freda, and Madeline White, Department of Psychology, University of North Carolina Wilmington. We thank Philip Quinlan for his invaluable advice and encouraging words during the long review process. We also thank the members of the Cohen Cognition and Perception Laboratory for their assistance and support. The current work received no funding. There are no real or potential conflicts of interest that may impact the current research. Portions of this research were presented at the 57th and 59th Annual Meeting of the Psychonomic Society (2017) and the North Carolina Cognition Group (2018).

Correspondence concerning this article should be addressed to Dale J. Cohen, Department of Psychology, University of North Carolina Wilmington, 601 South College Road, Wilmington, NC 28403-5612. E-mail: cohend@uncw.edu

Abstract

Here, we present a strong test of the hypothesis that sacrificial moral dilemmas are solved using the same value-based decision mechanism that operates on decisions concerning economic goods. To test this hypothesis, we developed Psychological Value Theory. Psychological Value Theory is an expansion and generalization of Cohen and Ahn's (2016) Theory of Subjective Utilitarianism. Psychological Value Theory defines a new theoretical construct termed *Psychological Value*, measures Psychological Value using a traditional psychophysics paradigm, and predicts preferential choice from those measurements using a value-based computational model. We evaluate the validity of Psychological Value Theory across six experiments. In Experiment 1, we use Psychological Value Theory to estimate the perceived Psychological Value of human lives and economic goods. The data reveal that perceived Psychological Value of lives is highly influenced by individual differences of people but minimally influenced by the number of people in a group. In Experiments 2-5 we demonstrate that when used as input in a value-based computational model, perceived Psychological Values of human lives accurately predict participants' RT and response choices to sacrificial moral dilemmas. In Experiment 6, we replicate these findings for decisions involving economic goods. We cross-validate our results with multiple datasets using multiple methods. We conclude that the same value-based processes underlying economic decisions also underlie choices involving human lives.

Keywords: Utility Theory; Preferential Choice; Moral Judgment; Value; Subjective Utilitarian; Social Cognition

Psychological Value Theory:

The Psychological Value of Human Lives and Economic Goods

Decisions involving economic goods (e.g., which detergent to buy) are often considered fundamentally different than decisions involving human lives (e.g., whose life to save when you can't save everyone). We assume a common theoretical construct, that we term Psychological Value, drives both of these decisions. Here, we define Psychological Value, directly measure it, use these measurements to *a priori* predict preferential choices in complex social and economic decision-making tasks. We call this *Psychological Value Theory*.

Sacrificial moral dilemmas and the Value of Human Lives

Sacrificial moral dilemmas are traditionally designed so that the value of human lives is the primary motivation for taking action. To accomplish this, sacrificial moral dilemmas have a formulaic structure: a larger group of people is in danger by default, and a smaller group of people (or a single person) is safe. The observer can choose to "do nothing" or "act." If the observer chooses to "act," the large group will be saved, but as a consequence, the smaller group of people will die. The moral justification for taking action in dilemmas with this structure is provided by the Axiom of Monotonicity. When applied to human lives, the Axiom of Monotonicity states that *more lives have a higher value than fewer lives*—all else being equal (e.g., Fishburn, 1970). A consequentialist ethical system can use the higher value of the larger group as a justification for taking action (e.g., Conway & Gawronski, 2013; Greene, Nystrom, Engell, Darley, & Cohen, 2004). For this reason, it has become commonplace for researchers to classify the "act" response as the "characteristically Utilitarian" option (e.g., Bialek & De Neys, 2017; Conway & Gawronski, 2013; Greene, 2007; Greene et al., 2004; Koenigs et al., 2007; Suter & Hertwig, 2011; Trémolière, De Neys, & Bonnefon, 2012).

Importantly, the Axiom of Monotonicity only applies when one is comparing different quantities of the same type (e.g., 5 workmen vs 1 workman). In many sacrificial moral dilemmas, however, the characters vary greatly. For example, in an oft used dilemma, a mother suffocates her baby to stop an army from killing townspeople (termed *the crying baby dilemma*; e.g., Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Koenigs et al., 2007; Paxton, Ungar & Greene, 2011). Despite the intuition that a mother values "her baby" more than "townspeople," researchers determine the characteristically Utilitarian option by simply calculating the number of lives saved relative to lives lost. Greene et al. (2004) makes this explicit when he states, "In each of these difficult dilemmas, an action that normally would be judged immoral (e.g., smothering a baby) is favored by strong utilitarian considerations (e.g., saving many lives)" (p. 391).

When the Axiom of Monotonicity drives the classification of the "act" response as "characteristically Utilitarian" despite the differences in the characters (e.g., the crying baby dilemma), researchers are implicitly assuming that observers accept the *Axiom of Equal Valued Lives* (or are applying an ethical system that holds that belief). The Axiom of Equal Valued Lives states that the values of all human lives are equal regardless of any feature of the human being (e.g., race, ethnicity, religion, social standing, demographics, etc.). If one does not accept the Axiom of Equal Valued Lives, a complicated calculus is required to compute the value of a group of people based on the value of each individual in the group. Although economists may calculate different values of statistical lives (VSLs) for different groups (e.g., Aldy & Viscusi, 2007), psychologists rarely do. There is some evidence, however, to suggest that the individual differences of characters in scenarios influence observers' responses (e.g., Swann, Gómez, Dovidio, Hart, & Jetten, 2010).

The importance of assessing the Axiom of Equal Valued Lives in relation to the Axiom of Monotonicity is apparent when one considers the "do nothing" response to sacrificial moral dilemmas. In contrast to the "act" response, the "do nothing" response is often hypothesized to arise from non-consequentialist reasoning (e.g., killing is "wrong," regardless of the consequences), and is therefore classified as the "characteristically Deontological" option (e.g., Greene et al., 2001; Greene et al., 2004; Greene, Morelli, Lowenburg, Nystrom & Cohen, 2008; Shenhav, & Greene, 2014). Importantly, although some researchers classify the "do nothing" response as "characteristically Deontological" for traditional dilemmas (e.g., Białek & De Neys, 2017), others only do so for specific subtypes of dilemmas (e.g., those that involve inflicting direct, foreseeable, serious harm, termed *personal* dilemmas; Greene et al., 2001; Greene et al., 2004; or only "high conflict" personal dilemmas, e.g., Koenigs et al., 2007). Nevertheless, if individual differences in characters have more influence than quantity on the valuation of human lives, then it is possible for the smaller group (e.g., the crying baby) to be valued more than the larger group (e.g., the townspeople). In this case, one can reasonably conclude that the "do nothing" option is the result of consequentialist reasoning (e.g., the baby is valued more than the townspeople), rather than non-consequentialist reasoning (smothering a baby is wrong). This interpretation holds for all subtypes of scenarios (both impersonal and personal). For this reason, it is important to understand how people value human lives and the influence of these values on observers' judgements to sacrificial moral dilemmas.

In sum, the relative influence of quantity and individual differences on the valuation of human lives remains unknown, as does a precise understanding of the influence of the value of human lives on observers' responses to sacrificial moral dilemmas. Here, we measure the influence of quantity and individual differences on the valuation of human lives —and use those

measured values to drive a value-based sequential sampling model to a priori predict participants' response choice and RTs in sacrificial moral dilemmas.

Psychological Value Theory

Measuring how people value human lives is a theoretically difficult problem. Value is typically inferred *from* choices using the principals of Utility Theory. Utility Theory proposes that *when a task is value-based*, one can mathematically equate choice and value (e.g., Edwards, 1954; Fishburn, 1968/1981). Choice and value do not equate, however, when observers consider factors other than value when making a decision. Thus, to apply Utility Theory, one must *a priori* assume that the task is value-based. Although it is generally assumed that choices involving economic goods and services are value-based (e.g., Barberis, 2013; Smith, 1989; Stigler, 1950), it is also generally assumed that choices involving human lives are not (e.g., Conway & Gawronski, 2013; Greene et al., 2001; Greene et al., 2004; Greene, et al., 2008; Shenhav, & Greene, 2010/2014). Therefore, the value of human lives is understood to be outside the scope of Utility Theory.

To overcome this constraint, we developed *Psychological Value Theory*. Psychological Value Theory is an extension and generalization of our value-based model of moral judgment (Cohen & Ahn, 2016). Psychological Value Theory reverses the standard approach of Utility Theory. Rather than infer value from choices, Psychological Value Theory predicts choices from values. Therefore, Psychological Value Theory does not require the *a priori* assumption that observers respond to a specific task based on the value of the options.

To reverse the standard approach of Utility Theory, one must first define the theoretical construct of value independent of choice[1]. Doing so enables the model to make novel predictions about the construct itself and its' relation to choice. Psychological Value Theory

terms the primary theoretical construct *Psychological Value*, Ψv. Psychological Value is the perception of the importance, worth, or usefulness of an item to the observer. In this definition, we emphasize that Psychological Value is a *perception*, rather than a sensation or conception (see Appendix C; for opposing views, see e.g., Gold & Shadlen, 2007; Padoa-Schioppa, 2013)[2]. As such, we propose that Psychological Value inherits the qualities of other perceptions, thus providing testable predictions. We discuss several such predictions below and in Appendix C.

Perceptions are typically influenced by multiple attributes of a stimulus (e.g., Ashby & Townsend, 1986; Cohen & Lecci, 2001; Dyer & Sarin, 1979; Roberts, & Goodwin, 2002). These attributes, however, may collapse onto a monotone, unidimensional construct (e.g., Signal Detection Theory, Green & Swets, 1966; General Recognition Theory, Ashby & Townsend, 1986; Multi-Attribute Utility Theory, e.g., Dyer & Sarin, 1979; Roberts & Goodwin, 2002). Accepting these common assumptions, we formalize the Psychological Value of an Item (ψ_v) as

$$\psi_V = \sum_{i=1}^{N} C_i P_i \quad (1)$$

where P_i are perceptual features that contribute to Psychological Value, such as emotional connection, monetary value, religious beliefs, etc., and C_i are different weights for each P_i (Cohen & Lecci, 2001). Here, C_i can vary by situation. For example, the contribution of the flavor of an apple to one's Psychological Value of that apple may be great when the observer is nutritionally satisfied but may carry little weight when the observer is nutritionally deprived.

Like other perceptions, Psychological Value is subject to Ashby and Lee's (1993) Axiom of Perceptual Variability. The Axiom of Perceptual Variability states that, "There is trial-by-trial variability in the perceptual information obtained from every object or event. … [thus] … the perception will change even if the stimulus does not" (p. 370). Accepting the Axiom of Perceptual Variability, we represent Ψ_v of an item as a distribution of values along a continuum,

$f(\Psi_v)$. By representing Ψ_v as a distribution, $f(\Psi_v)$, we acknowledge that there is a stochastic component to Ψ_v. Modeling this stochastic component is fundamentally important when making predictions because although an observer will respond inconsistently on a trial-by-trial basis, the overall probability of their behavior (i.e., collapsed over many trials) will be highly predictable (described below). This is a defining property of Psychological Value Theory.

The second step in reversing the standard approach of Utility Theory is to measure Psychological Value. If Psychological Value is a perception, it should be available for empirical estimation using procedures typically used to measure other perceptions. One procedure often used to measure perception is magnitude estimation (e.g., Cohen & Lecci, 2001; Marks & Algom, 1998; Stevens, 1975). When Stevens (1957) developed magnitude estimation, his peers doubted the plausibility of successfully measuring the seemingly intangible relationship between the physical world and the subjective perception of it (Stevens, 1957). To this Stevens wrote, "Many authors have screamed that this is nonsensical, meaningless, and impossible, but those who follow these methods go ahead and do it anyhow. These direct assessments of sensation seem not so impossible after they have been made" (Stevens, 1957, pp. 163). Since that time, magnitude estimation has proved an effective measure of perception in several modalities under a large variety of conditions (for reviews, see Gescheider, 1988; Marks & Algom, 1998; Stevens, 1986).

Cohen and Lecci (2001) have demonstrated that magnitude estimation procedures can be used to directly measure perceptual variability (in addition to measures of central tendency). Therefore, the magnitude estimation procedure should provide a mechanism for measuring perceptual variability associated with Psychological Value, $f(\Psi_v)$. Signal Detection Theory (SDT) proposes that an observers' sensitivity to the difference between two perceptual stimuli is

described by the overlap of the perceptual distributions of those two stimuli (e.g., ROC curve). If $f(\Psi_v)$ describes the perceptual qualia of the value of the item, and one can measure $f(\Psi_v)$ directly, the overlap of the distributions associated with two items, $f(\Psi v_1 \cap \Psi v_2)$, will describe an observer's sensitivity to the higher valued option (HVO) in a value-based task. This is a second defining property of Psychological Value Theory.

Different decision models (e.g., ballistic, heuristic, quantum, accumulation, etc.) rest on different assumptions and produce different predictions from the same input data. Therefore, to unambiguously make precise point predictions, the final step in reversing the standard approach of Utility Theory is to model a single, specific decision mechanism. Psychological Value Theory models the decision process as a *value-based sequential sampling procedure* (VSSP). VSSPs model the relation between the latent value construct and participants' RTs and response choices in a preferential choice task (e.g., Decision Field Theory; Attentional Drift Diffusion; for a review, see Busemeyer, Gluth, Rieskamp, & Turner, 2019). Preferential choices are those which do not have any objectively correct answer (e.g., choosing which of two cars to purchase). Similar to SDT, VSSPs assume that the two choices in a two alternative forced choice task are represented by overlapping latent value distributions (see Figure 1). Information accumulates over time and eventually the information crosses one of two pre-specified decision thresholds at which point the participant responds. The rate of that accumulation—termed *drift rate*—is a function of the degree of overlap of the latent distributions (e.g., Link & Heath, 1975; Ratcliff, 1978; Ratcliff & Rouder, 1998). A large overlap indicates that the two options have relatively similar values and results in a slow accumulation of information (long RTs) and a high error rate. A small overlap indicates that the two options have relatively dissimilar values and results in a fast accumulation of information (short RTs) and a low error rate.

Traditionally, researchers do not measure value directly and instead use VSSPs to estimate the latent value construct (i.e., drift rate) from participants' RT and response choice data collected in a preferential choice task (e.g., Busemeyer et al., 2019; Krajbich, 2019; Krajbich, Armel, & Rangel, 2010; Rangel & Clithero, 2014; see Table 1). These VSSPs have been critiqued as being under-constrained, and thus able to identify parameter values that fit virtually all forms of data (e.g., Jones, & Dzhafarov, 2014; but see Busemeyer, & Johnson, 2004). Recently, researchers have increased the constraints of VSSPs by integrating participants' mean ratings of value surrogates into their estimation of drift rate (e.g., Krajbich et al., 2010; Krajbich, Hare, Bartling, Morishima, & Fehr, 2015; Krajbich, Lu, Camerer, & Rangel, 2012; Krajbich & Rangel, 2011; Milosavljevic, Malmaud, Huth, Koch, & Rangel; 2010; Tavares, Perona, & Rangel, 2017). The ratings of value surrogates provide measures of central tendency but provide no information about the distributional shape or distributional overlap (i.e., drift rate). As such, researchers must assume the shape of the latent value distributions and use the VSSP to estimate their overlap (i.e., the drift rate) from the RT and response data. Psychological Value Theory advances this work by measuring the degree of overlap of the perceived Psychological Value distributions directly, $f(\Psi v_1 \cap \Psi v_2)$, thus further constraining the VSSPs (see Table 1). Because $f(\Psi v_1 \cap \Psi v_2)$ is a measure of drift rate, it can drive the VSSP to predict RT and response choice rather than the other way around. This is a third defining property of Psychological Value Theory.

We have developed a relatively simple, Robust Random Walk procedure (the RRW) to model the decision process of Psychological Value Theory[3]. The RRW is optimized to incorporate the assumptions of Psychological Value Theory, and therefore, is different from traditional VSSPs in two ways. First, rather than estimating drift rate from participants' RTs and

response choices, the RRW takes a direct, non-parametric measure of drift rate as an input and uses that measurement to predict RT and response choice. By taking drift rate as an input rather than estimating it from the response data, the model is far more constrained than traditional VSSPs (thus negating the critique that VSSPs cannot be falsified, e.g., Jones & Dzhafarov, 2014). Second, the step size of each iteration of the random walk is independent of the shapes of the distributions associated with each choice (i.e., non-parametric). The use of a non-parametric measure of drift rate has the advantage of being robust across transformations of the raw perceived Psychological Values data. The result is that the RRW provides relatively robust estimates that make fewer parametric assumptions than traditional VSSPs. The RRW makes strong point predictions simultaneously for the observers RT responses to the higher valued option (HVO), RT responses to lower valued option (LVO), and the probability of choosing the HVO [p(HVO)] for the entire set of item choices. The RRW is briefly described in Appendix A.

Like all VSSPs, the RRW makes relatively straightforward predictions: the speed and accuracy of participants' responses in a two-choice decision task is a function of (a) the overlap of the perceived Psychological Value distributions of the options, $f(\Psi v_1 \cap \Psi v_2)$, and (b) response biases (e.g., the placement of the start point, and the distance the boundaries are from the start point; e.g., Ratcliff & Rouder, 1998). $f(\Psi v_1 \cap \Psi v_2)$—a measure of the similarity of the perceived Psychological Values of the options—is the primary parameter influencing the speed and accuracy of an observer's response (see Figure 1). When the perceived Psychological Values of the options are dissimilar (i.e., $f(\Psi v_1 \cap \Psi v_2)$ is small), responses are fast and accurate. When the perceived Psychological Values of the options are highly similar (i.e., $f(\Psi v_1 \cap \Psi v_2)$ is large), responses are slow and error prone. There are two types of response biases: *speed response biases* and *choice response biases*. *Speed response biases* refer to the separation distance of the

response boundaries. The closer the boundaries are, the faster the observer will respond and the more errors the observer will make (see Figure 2). *Choice response bias* is instantiated in the start point of the process. An unbiased observer's start point is positioned equidistant from each response boundary. In contrast, an observer with a bias toward a particular response has a start point positioned closer to the boundary of the preferred response. This will shift an observers' revealed preference toward the boundary closer to the start point (see Figure 2).

Speed and choice response biases provide a mechanism for Psychological Value Theory to capture the influence of features other than Psychological Value itself. For example, speed response biases capture features such as impulsivity, time pressure, etc. Choice response biases capture one's tendency to prefer a particular option over another. The choice response bias is particularly relevant to the understanding of how people make sacrificial moral judgments. Recall that sacrificial moral dilemmas are typically used to assess whether observers make judgments about human lives based on consequentialist (e.g., Utilitarian) or non-consequentialist (e.g., Deontological) reasoning. Whereas the inherent value-based nature of the RRW captures a consequentialist influence, the choice response bias may capture the non-consequentialist influence. Specifically, the degree to which an observer is reluctant to actively kill another human being (for example) will manifest as a choice response bias toward the "do nothing" boundary. At the extreme, an observer who believes that one should never kill regardless of the consequences (e.g., a Deontologist) will be so biased toward the "do nothing" boundary that they will never respond "act."

The Current Research

Here, we present a strong test of the hypothesis that sacrificial moral dilemmas are value-based tasks whereby observers attempt to identify and save the option with the highest perceived

Psychological Value. In Experiment 1, we use principals of Psychological Value to measure the perceived Psychological Value of a variety of individuals who differ by social status. For each social status, we vary the number of lives being valued. We mirror these manipulations with economic goods. In Experiments 2-5, we assess whether the RRW can predict preferential choice behavior in sacrificial moral dilemmas using the estimates of perceived Psychological Values of human lives collected in Experiment 1 (see Figure 3). In Experiments 6, we assess whether the RRW can predict preferential choice behavior in economic dilemmas using the estimates of perceived Psychological Value of economic goods collected in Experiment 1. These manipulations provide the data necessary to assess the validity of the Axioms of Monotonicity and Equal Valued Lives.

The RRW will successfully predict preferential choice behavior in Experiments 2-6 only if (1) the Psychological Values collected in Experiment 1 described the perceptual qualia of the values of the items, (2) sacrificial moral dilemmas and economic dilemmas are value-based tasks whereby observers attempt to identify and save the option with the highest Psychological Value, and (3) the decision processes modeled in the RRW is valid. If, however, any of the premises are invalid, the entire system will fail and the RRW will not predict response choice or RT. If successful, this will be the first-time participants' choice behavior in a sacrificial moral judgment task has been successfully modeled using direct estimates of Psychological Value to drive a VSSP.

Experiment 1

In Experiment 1, we estimate the perceived Psychological Value of humans, non-human animals, and objects. VSSP researchers often use rating scales as value surrogates. These value surrogates, however, are highly specific questions linked directly to the choice task they are

predicting such as pleasantness ratings of faces, houses, and paintings (Lebreton, Jorge, Michel, Thirion & Pessiglione, 2009), willingness to pay for snack foods, nonfood consumables (Chib, Rangel, Shimojo, & O'Doherty, 2009), movies (Grueschow, Polania, Hare, & Ruff, 2015), and desirability ratings of snack foods (e.g., Colas & Lu, 2017; Gwinn, Leber, & Krajbich, 2019; Krajbich et al., 2010; Lim, O'Doherty & Rangel, 2011). For example, an individual's ratings of snack food preference are used as value surrogates for a preferential choice task that asks the individual to choose his or her preferred snack food. Such rating scales are limited because the ratings are not generalizable to other stimuli classes. For example, one cannot predict preferential choices about which of two cars an observer will purchase by asking about snack food preferences. Whereas such rating scales may provide information about feature preference, they likely do not capture the theoretical, stimulus independent construct of value. We discuss why currency is a sub-optimal measure of Psychological Value in Appendix B.

In contrast to the traditional methodology described above, we use magnitude estimation to measure Psychological Value for the reasons discussed in the Introduction. Magnitude estimation tasks are applicable to multiple classes of stimuli without alteration, and therefore, they can be used to measure the value of objects, food, lives, etc. on the same scale. Thus, magnitude estimation tasks may be more appropriate to measure the theoretical, stimulus independent construct of Psychological Value.

Participants were asked to estimate the Psychological Value of a subset of variations of 123 core items (humans, non-human animals, and objects) using a magnitude estimation task. Unlike Cohen and Ahn (2016), we do not assess the perceived Psychological Value of family or friends. This provides the data required for a strong test of the Axiom of Equal Valued Lives because some may argue that family and friends are a special segment of society (e.g., Kurzban,

DeScioli & Fein, 2012). Second, we systematically vary the number of people in a group. Cohen and Ahn (2016) found that the perceived Psychological Value of one anonymous person was virtually identical to that of five anonymous people. Here, we extend Cohen and Ahn (2016) by systematically assessing the perceived Psychological Value of different sized groups of people. This provides the data required for a strong test of the Axiom of Monotonicity. We suspect that quantity of lives will have a relatively small influence on Psychological Value (e.g., Cohen & Ahn, 2016; Slovic, 2007), and social status will have a relatively large influence on Psychological Value (e.g., Bohnet & Frey, 1999; Millar, Starmans, Fugelsang & Friedman, 2016; Swann et al., 2010).

Method

Participants. Four hundred fourteen naïve participants volunteered and received course credit for their participation. We did not collect demographic data. At UNCW, the student body is 62% female, 83% white, 6% African American, 6% one or more other races, and 4% unknown. About 7% of students are Hispanic. The average student age is 22. Sample size was determined by (a) setting a minimum number of participants (350), (b) estimating the time necessary to collect that number of participants, (c) posting all available experimental slots for the time estimated in "b," and (d) running all participants who signed up. This resulted in the collection of more than the minimum number of participants because of a higher-than-expected sign-up rate.

Apparatus and Stimuli. The magnitude estimation task was completed on a 24-in. LED color monitor Mac with a 72-Hz refresh rate controlled by a Macintosh Mini running an OS X. The resolution of the monitor is 1920 × 1200 pixels. An Apple keyboard was used during the experiment.

Briefly, the participant's task was to provide a personal value of a *probe* stimulus in relation to a previously learned *standard* stimulus. The standard stimulus (a chimpanzee) and its assigned value (1,000) was consistent for all participants and all trials. We chose a chimpanzee because Cohen and Ahn (2016) demonstrated that it did not fall on the extreme high or low end of the items assessed.

There were 486 potential probes created from specific variations of a set of 123 *core items* that consisted of 53 humans, 21 non-human animals, and 49 objects. Depending on the core item, the personalization and/or quantity of the item could be varied to produce a maximum of 6 probes for a single core item. The personalization of probes could be identified as *personal* or *impersonal*. The core item was preceded by the word "your" for personal probes ("your book") but not for impersonal probes ("a book"). Sixty-nine of the core items had probes for both personal and impersonal conditions and were termed *common items*. Fifty-four of the core items were not appropriate for both personalization conditions and, therefore, were presented as either a personal probe or an impersonal probe. These items were termed *unique items*. There was only 1 unique item presented as a personal probe: "you." The remaining 53 unique items were impersonal probes such as "an astronaut."

Probes for 100 of the core items could also vary in quantity; these items were termed *variable ratio items*. Each variable ratio item had three values from three separate quantity conditions: single, small, and large. In the single condition, the quantity was always 1. In the small condition, the assigned quantity ranged from 3 to 9. In the large condition, the assigned quantity ranged from 47 to 53. For example, we had three quantity variations of *book*: "a book," "4 books," or "48 books." For each variable quantity condition (small and large), the quantity assigned to a specific item remained constant (e.g., 4 books but never 5 books). For core items

that were both common and variable ratio, the same quantities were used in the impersonal variations and personal variations. Therefore, the core item *book* was presented as "a book," "your book," "4 books," "your 4 books," "48 books," and "your 48 books."

Twenty-three core items were considered *non-variable.* Non-variable items were always presented in the single quantity condition. Most of the non-variable items were common items, with the exception of one unique item: "you."

For each trial, the standard and the probe were presented together in a dialog box. The standard was centered and presented in dark grey above the probe. The probe was centered and presented in black. Both the standard and the probe were presented as text in 16-point Helvetica font. There was a text box below the probe in which the participant was to enter their estimate of the personal value of the probe. The participants used the number pad on the right side of the keyboard to enter their response which was formatted in real time with commas to reflect the numeric value. For example, as the participant entered their response, a comma would appear between the 3rd and 4th whole number unit (e.g., an estimation of 1000 would appear as 1,000 in the response textbox). This aided the participant in reading the numbers they entered. Participants could enter any real numbers including negative values and/or decimals.

Procedure. Participants completed the experiment in a dark private testing room that contained a desk, a chair, and computer. White noise was playing at low volume from speakers in the ceiling to mask any ambient sounds.

The procedure was similar to the procedure used by Cohen and Ahn (2016). Each trial, participants were presented a standard and a probe and were to estimate the "personal value" of the probe in a relation to that of the standard. Consistent with Cohen and Ahn (2016), we used the term "personal value" to communicate the construct to naïve participants. They were

instructed not to think of "personal value" as monetary value. Personal value was operationalized as follows:

> *…You can define "personal value" in any way you find appropriate. Personal value is not necessarily the same as monetary value. For example, we may ask the personal value of your first report card. Here, the monetary value may differ dramatically from the personal value…*

Participants were to assign a numerical value to each probe in relation to the assigned numerical value of the standard (chimpanzee = 1,000). This was a magnitude estimation task, and therefore, for example, if the participant thought the probe had a personal value 7 times that of the standard, they were to enter 7000. If the probe was half as valuable, they were to input 500, etc.

We presented each participant with four practice trials followed by 192 experimental trials. During the experiment, each participant was presented with one probe for each unique item (54 trials) and two probes for each common item (138 trials). Common items were presented twice: once as an impersonal probe and once as a personal probe. Therefore, a non-variable common item (e.g. "a mother") would have the same probes presented for every participant—one trial for "a mother" and another trial for "your mother." Variable ratio common items (e.g. "a book") were also presented once as a personal probe and once as an impersonal probe, but the quantity variation of each probe was randomly chosen for each participant. Therefore, a single participant may be presented with the probes "4 books" and "your book" during an experiment but never "4 books" and "a book" or "your 4 books" and "your book." We find that when a participant is presented different quantities of the same probe, they will calculate a value of the second probe presented based on the quantity and value of the first probe.

We believe these calculated values are artificial and do not accurately represent Psychological Value[4]. For this reason, we never present different quantities of the same probe to the same participant. For unique variable ratio items, the quantity variation of each probe was also randomly chosen for each participant. The order of the probes for experimental trials was randomized.

Each trial consisted of a 500 ms fixation point followed by the dialog box containing the standard, a probe and a response textbox. The participants entered their response as a numeric value in the textbox and pressed an "OK" button to advance to the next trial. There was one self-timed break that occurred after the 99th experimental trial. After completing the magnitude estimation task, participants completed a 12–item survey. Information from the survey was not used in analysis for this paper and, therefore, will not be further discussed.

All procedures were approved by the UNCW IRB, protocol# 17-0102.

Results

Forty-seven participants' data was removed from the analysis. Seven participants were removed because of technical or human errors that occurred while running the experiment. The remaining participants were removed using two criteria that were intended to ensure that participants expended sufficient effort in the task. The first criterion required that participants' median RT for their trials exceeded 2750 ms. This criterion was set because it is very difficult to read and understand the probe and input an appropriate response on a number pad in less than 3 seconds. The second criterion required that participants' median *log* response (i.e., log(Magnitude Estimate)) be greater than or equal to 1. This criterion was set because participants who respond with little effort will input few values (1 or 2 numbers and then hit enter) without regard to the probe itself. The first criterion removed 38 participants, whereas the

second criterion removed 2 participants. We analyzed the data of the remaining 367 participants. From these participants, we removed individual trials based on RT outliers that suggest a lack of attention on that trial (i.e., RT > 45s or RT < 1000ms). These criteria removed less than 0.5% of the individual trials.

We used a non-parametric bootstrap procedure on the raw data to estimate $f(\Psi v_i \cap \Psi v_j)$ for every pair of items (see Cohen & Ahn, 2016). Recall that $f(\Psi v_i \cap \Psi v_j)$ is theoretically an estimate of perceptual similarity along the Psychological Value continuum for any pair of items. Therefore, the non-parametric estimate of $f(\Psi v_i \cap \Psi v_j)$ should be equivalent to the non-parametric sensitivity measures in SDT such as the ROC. If so, $f(\Psi v_1 \cap \Psi v_2)$ will describe an observer's sensitivity to the higher valued option (HVO) in a value-based task.

Each of the 486 possible variations of the core items were treated as individual items for the bootstrap analysis. For each pair of items (117,855 total pairs), we designated one item as $Item_1$ and the other other as $Item_2$. We then randomly selected a value from each items' distribution and identified the item with the larger value. In the bootstrap procedure, we repeated this process, with replacement, $(N_A+N_B)/2$ times (where N_i is the total number of values collected for $Item_i$) and calculated the proportion of draws that $Item_1$'s value was greater than $Item_2$'s value. For each pair of items, we replicated this procedure 10,000 times and calculated the mean percentage (meanP) in which $Item_1$ was greater than $Item_2$. We then transformed meanP into Overlap,

$$Overlap = 1-(abs(meanP-0.5)/0.5) \quad (2)$$

Overlap ranges from 0-1, where 0 indicates no overlap and 1 indicates complete overlap. We also noted the item that had the higher meanP.

We identified 28 probes to use as stimuli in our sacrificial moral dilemma tasks in Experiments 2-4 (see Table 2, Figure 4). These probes were the single quantity variation or impersonal, unique, variable-ratio core-items (see Figure 5). Twenty-four probes were humans, two were non-human animals and two were objects. The probes median values ranged from 0 to 30,000, demonstrating that there was substantial variation in the $f(\Psi v)$ of an individual human life. The four non-human probes were chosen because they fell on different areas of the personal values continuum. "a cockroach" and "a rabid possum" fell at the low end of the range, "a life-saving antidote" fell at the high end, and "a smartphone" fell towards the middle. Finally, Overlap values of the 28 chosen probes ranged from 0.019 to 0.999 with a median value of 0.50.

For Experiments 5 and 6, we expanded on the set of probes. In Experiment 5, we added the groups of the humans to the individuals used in Experiments 2-4. We also dropped non-human probes as well as "An Adult with a Deadly Contagious Disease" because of the number of characters relative to the other probes. This set contained 69 probes: 23 (humans varying on social status) x 3 (number of individuals in a group). Overlap values of this set of 69 probes ranged from 0.05 to 0.999 with a median value of 0.62.

For Experiment 6, we identified 24 *economic goods* from the set of impersonal, unique, variable-ratio core-items. For each of the 24 economic goods, we included the single variation and all group sizes. The final set contained 72 probes: 24 (impersonal economic goods) x 3 (number of economic goods in a group). Overlap values of this set of 72 probes ranged from 0.068 to 0.999 with a median value of 0.61.

Because Overlap is a non-directional measure, Overlap does not indicate which item has the greater Psychological Value. A separate variable, termed direction (*d*), provides that

information. Therefore, we created the Directional Overlap (O_D) measure. O_D combines Overlap and direction in the following formula:

$$O_D = \text{abs (Overlap - 1)} * d, \quad (3)$$

where *d* equals 1 if the Group A's distribution is greater than the Group B's distribution and -1 if the opposite is true. Directional Overlap ranges from -1 to 1 whereby -1 indicates that there is no overlap and Group A has the lesser values, 0 indicates complete overlap between Groups A and B, and 1 indicates that there is no overlap and Group A has the higher values. We calculated Directional Overlap (O_D) of each group (Group A) relative to every other group (Group B).

Directional Overlap provides a precise estimate of the influence of individual item type and quantity on the perceived Psychological Value of human lives and economic goods. Directional Overlap is, in fact, a directional measure of *effect size* because it is a measure of the overlap of two distributions. In Figure 6, we summarize the directional overlaps of the humans and the economic goods used in the present experiments. To show how participants perceived Psychological Value of each group relative to every other group, for each Group A, we averaged O_D across all Group B's. In Figure 6, Group A is on the x-axis (ordered by average O_D), O_D is on the y-axis, and group size (single, small, and large) of Group A is identified by the shade of gray of the point. As demonstrated in Figure 6, the perceived Psychological Value of human lives is highly influenced by the social status of individuals in the group (e.g., convict vs policeman) but is minimally influenced by the quantity of individuals in the group (e.g., single vs large). The perceived Psychological Value of economic goods follows this same pattern but with slightly more influence of quantity. Below, we calculate the influence of quantity on perceived Psychological Value.

We first calculated the influence of quantity on perceived Psychological Value accepting the Axiom of Equal Valued Lives. To do so, we ran two linear regressions: one for human lives and one for economic goods that compared each person/economic good to one another regardless of social status. For the criterion variable, we calculated the Directional Overlap, O_D, comparing the single, small, and large quantities of each person/economic good with every other person/economic good, regardless of social class (e.g., a nun vs 4 judges; a nun vs 48 soldiers, etc). We did not run this analysis on person/economic goods that had identical quantities because, without a "larger group" the direction variable is ambiguous. This resulted in 2,033 O_Ds for human lives (i.e., 69 groups [i.e., 23 individuals varying on social status x 3 comparisons per individual type] taken 2 at a time equal 2346 combinations – 313 groups with identical quantities), and 2,206 O_Ds for economic goods. An $O_D > 0$ indicates that the smaller quantity group had the greater perceived Psychological Value, an $O_D = 0$ indicates that there is no influence of quantity on perceived Psychological Value, and an $O_D < 0$ indicates that the larger quantity group had the greater perceived Psychological Value. We then calculated a linear regression predicting these O_Ds as a function of *group size difference* (i.e., abs(number of people/ economic goods in group 1 – number of people/economic goods in group 2)). The Axiom of Monotonicity accepting the Axiom of Equal Valued Lives predicts a negative relation between O_D and quantity. The steeper the negative slope, the stronger the influence of quantity.

There was a small, but significant influence of quantity of human lives on O_D, $F(1, 2031) = 15.07$, $p < 0.001$, $r^2 = .01$. Both the intercept (intercept = 0.035, t=2.0, p = 0.04) and the slope (slope = -0.002, t=3.9, $p < 0.001$) were significant (see Figure 7, top row)[5]. There was a small but significant influence of quantity of economic goods on O_D regardless of item type, $F(1, 2204) = 22.7$, $p < .001$, $r^2 = .01$. There was a significant slope (slope = -0.002, t=4.8, $p < 0.001$),

but no intercept effect (intercept = -0.01, t = -0.67, p = .51) (see Figure 7, top row). This pattern of results suggests that the Psychological Value of economic goods increases with quantity, but this increase is minimal (slope = -0.002). The data reveal a minimal influence of quantity on the Psychological Value of human lives and economic goods, with quantity accounting for only about 1% of the variance in O_D. To put these results in perspective, with slope effect of -0.002, a group size difference of 20 items is required for Psychological Value Theory to predict a 2% increase in the likelihood of choosing the group with the greater number of people/economic goods.

We ran a similar analysis that measured the influence of quantity on perceived Psychological Value rejecting the Axiom of Equal Valued Lives. Here, we assess the validity of the Axiom of Monotonicity only on items/people *of the same type*. To do so, for each person/economic good of the *same type* (e.g., a nun), we calculated the Directional Overlap, O_D, comparing the single, small, and large quantities of that type (e.g., a nun vs 4 nuns; a nun vs 48 nuns; and 4 nuns vs 48 nuns). This resulted in 69 O_Ds for human lives (i.e., 23 individuals varying on social status x 3 comparisons per individual type), and 72 O_Ds for economic goods. Although there are fewer O_Ds in this analysis, it retains significant power. Specifically, over 240 independent observations were used to calculate *each* O_D (totaling over 16,500 observations for human lives analysis and over 17,000 observations for economic goods analysis). There was no significant influence of quantity of human lives on O_D, $F(1,67) = 1.54$, $p = .22$, $r^2 = .02$. Nevertheless, there was a small, negative intercept effect (intercept = -0.04, t = 2.3, p = .03). This pattern of results indicates that there is a very small increase in Psychological Value for groups containing more than one person of the same social status, but that effect does not increase with an increased number of lives (see Figure 7, bottom row). There was a small but

significant influence of quantity of economic goods on O_D, $F(1,70) = 11.03$, $p < .01$, $r^2 = .14$, with no intercept effect (intercept = -0.03, $t = 1.0$, $p > .05$) (see Figure 7). This pattern of results suggests that the Psychological Value of economic goods increases with quantity, but this increase is minimal (slope = -0.003)[6]. As such, these results confirm that increasing the quantity of individuals of a specific social status (e.g., a nun vs 48 nuns) or specific economic good (e.g., a motorcycle vs 50 motorcycles) had minimal influence on estimated perceived Psychological Value.

In sum, contrary to the assumptions of the Axioms of Monotonicity and Equal Valued Lives, quantity has minimal influence on our estimates of perceived Psychological Value (whether or not one accepts the Axiom of Equal Valued Lives), and social status of an individual has a relatively large influence on our estimate of perceived Psychological Value. If the assumptions of Psychological Value Theory are valid, quantity will exert a minimal influence on decisions in a sacrificial moral dilemma and social status will exert a large influence on decisions in a sacrificial moral dilemma.

Experiment 2

The data from Experiment 1 reveal subtle differences in the perceived Psychological Value of individuals that differ only on social status suggesting violations of the Axiom of Equal Valued Lives. Psychological Value Theory predicts that participants' RTs and the probability of response choices in Experiment 2 should be well modeled by a VSSP (here, we use our RRW) driven by the estimates of perceived Psychological Value collected in Experiment 1. Because we measure both the similarity in the perceived Psychological Value of the options, $f(\Psi_{V1} \cap \Psi_{V2})$, and preferential choices, this is a very strong (i.e., highly constrained) prediction (i.e., the vast majority of potential data outcomes will falsify this prediction).

Method

Participants. One hundred sixty naïve participants volunteered and received partial course credit for their participation. Sample size was determined using the same procedure as Experiment 1 (minimum estimated 100; see Cohen & Ahn, 2016). This resulted in the collection of more than the minimum number of participants because of a higher-than-expected sign-up rate.

Apparatus and Stimuli. The task was completed on a 24-in. LED color monitor Mac with a 72-Hz refresh rate controlled by a Macintosh Mini running an OS X. The resolution of the monitor is 1920 × 1200 pixels. An Apple keyboard was used during the experiment.

Participants were presented with 10 sacrificial moral dilemma scenarios that were created for Cohen and Ahn (2016; see Cohen and Ahn (2016) for specific criteria used to construct the scenarios and examples of specific scenarios). Each scenario presented a situation in which an item (Item1) was going to be killed or destroyed if the reader did nothing. However, the reader could act to save Item1 but kill or destroy a different item (Item2). For every trial of every participant, "Item1" and "Item2" were randomly chosen from the list of 28 items identified in the Experiment 1. If the item was living, the verb "killed" was used. If the item was not living (e.g., a smartphone) the word "destroyed" was used.

Procedure. Participants completed the experiment in a dark private testing room that contained a desk, a chair, and computer. White noise was playing at low volume from speakers in the ceiling to mask any ambient sounds. The procedure was similar to that of Cohen and Ahn (2016).

We presented participants with 10 different scenarios four times each for a total of 40 scenarios. The order of the 40 scenarios was randomized between participants. For every

participant, every scenario, and every trial, a random pair of items was chosen from the list of 28 probes and inserted into the scenario replacing "Item1" and "Item2." Each scenario ended with a question asking participants if they would save "Item1" causing "Item2" to be killed (or destroyed). Participants responded with either "yes" or "no" by pressing one of two response keys.

To get a RT estimate that did not include reading time, we followed the masking developed by Cohen and Ahn (2016). Each trial consisted of a 500 ms fixation point, followed by a timed masked scenario, followed by the unmasked scenario. The unmasked scenario remained on the screen until the participant entered their response using the "d" and "k" keys on the keyboard. We counterbalanced "yes" and "no" responses between the "d" and "k" keys. There were four practice trials prior to the experimental trials to familiarize participants with the procedure. The practice trials consisted of the same trial format but were innocuous questions about food preferences rather than sacrificial moral dilemmas.

All procedures were approved by the UNCW IRB, protocol# 16-0210.

Results

RT and p(HVO) Analysis

The RT and p(HVO) analysis is similar to the analysis of Cohen and Ahn (2016). Seven participants were removed because they either did not complete the experiment or did not comply with cell phone use instructions. We removed six participants because their average RTs were outliers (i.e., RTs < 2000ms or RTs > 15s) and 12 participants that responded at or below chance (indicating that they did not put effort into the experiment or confused the meaning of the response keys[7]). We then removed less than 1% of the individual trials that were RT outliers (i.e., RT > 45s or RT < 1000ms). We analyzed the data of the remaining 135 participants.

To get a clean RT measure, we first removed the influence of learning by fitting a learning function to the data, log(RT) = a * exp (b * (trial – 1)) + c. There was a significant influence of all parameters (a = 0.86, b = –0.08, c = 8.13), all t's > 8.0, p's < .001. We used the average residuals of this function (RT_{res}) collapsed over participant as our RT outcome variable.

Psychological Value Theory predicts (a) a positive relation between RT and Overlap and (b) a negative relation between the probability of choosing the higher valued option, termed *p(HVO),* and Overlap. We rounded Overlap to the nearest 0.1 ($O_{0.1}$) and used this as our predictor variable[8]. To assess the relation between RT and Overlap, we calculated a linear regression, $RT_{res} = a + (b * O_{0.1})$. We found a significant linear relation, $F(1,9) = 77.11$, $p < 0.001$, $r^2 = 0.91$ (see Table 3 and Figure 8). To assess the relation between p(HVO) and Overlap, we calculated an exponential decay function: $p(HVO) = 0.5 * (1 - O_{0.1}{}^{b}) + 0.5$. There was a significant effect of Overlap, t =13.46 $p < 0.001$, $r^2 = 0.97$ (see Table 3 and Figure 8).

The pattern relating p(HVO) and Overlap was also evident when probes are considered individually (see Figure 9). We plotted the probability of choosing an individual probe (e.g., an orphan) relative to every other probe (termed *comparison probes* here). In these plots, we ordered the x-axis by directional overlap (O_D) whereby the comparison probe farthest to the left is the most likely to be chosen over the individual probe, and the comparison probe farthest to the right is the least likely to be chose over the individual probe. Psychological Value Theory predicts that the points to the left would be low (more likely to choose the comparison probe) followed by a gradual rise as the O_D transitions from a positive to a negative value followed by points high on the y-axis (less likely to choose the comparison probe). The predicted pattern of response choice is evident for the individual probes shown in Figure 9 even though the average number of trials per comparison probe was small (n = 13.6). We chose to highlight these four

individual probes (i.e., an orphan, an adult, a congressman, and a terrorist) because they range from very low Psychological Value (i.e., a terrorist) to very high Psychological Value (i.e., an orphan).

Robust Random Walk Analysis

Here, we assess whether Psychological Value Theory can be used to accurately model participants' choice RT and response choice using the RRW. The RRW *simultaneously* fits p(HVO), the RT to save the Higher Valued Option, and the RT to save the Lower Valued Option. Importantly, the RRW uses Overlap—as measured in Experiment 1—as a direct estimate of drift rate. By fitting the model to multiple dependent variables simultaneously and using Overlap as a direct estimate of drift rate, the RRW is a highly constrained model that is easily falsifiable.

We compared the performance of two models. First, we ran a simple model that assumes no choice response bias (i.e., start point = 0). We then compared the fit of this model to that of a more complex model that assumed a choice response bias. Both the simple and the choice response models shared the following three free parameters: *boundary*; the d_A parameter of the IAB; and T_{er} (the non-decision time). In addition, the RRW was ran with the d_B parameter fixed at 0.2, the noise σ parameter fixed at 1.0, and all other parameters fixed at 0.

The choice response model added a *start point effect* parameter to the base set of parameters. Specifically, to assess whether the participants had a bias toward either the non-action or action boundary, the data was divided into two trial types: scenarios in which the lower valued person was going to be killed by default (LVO) and scenarios in which the higher valued person was going to be killed by default (HVO). Within the context of the RRW, a bias toward the "do not act" boundary translates into a positive bias in the LVO trials and an equal but

negative bias in the HVO trials. To model this, we effect code the LVO condition as 1 and the HVO condition as -1. This effect coding is multiplied by the *start point* parameter which then varies the start point as a symmetric positive/negative effect around a zero start point for the two trial types.

Models can only be meaningfully compared using the BIC when they have the same number of datapoints. Therefore, to compare the simple model to the choice response model, we first prepared the dataset as described for the choice response bias model (i.e., with the data separated by HVO and LVO). When we fit the simple model to this dataset, we set the start point parameter = 0. When we fit the choice response model to this dataset, we let the start point be a free parameter. If the simple and choice response models had essentially equivalent r^2 and BIC values, we favored the simple model (no choice response bias) because it was less complex. Otherwise, we concluded the model with the lower BIC was the best fit model. If the simple model was the best fit model, we reran the simple model without splitting the data by HVO and LVO and report the fit and parameter values from that model.

To fit both models, we used a smart grid search optimization routine that we developed. This routine was implemented in R. For each parameter, the optimization program randomly selects valid values within a set of bounds. It runs the RRW (or any other model) and saves the fit statistic. Here, we use BIC as our fit statistic. For each set of parameters, the RRW calculates the BIC of the simultaneous fit to the RT and p(HVO) data. Because we wish to equally weigh p(HVO) and RT fits, we averaged the RT and p(HVO) residual sum of squares and used that average to calculate the BIC[9]. The smart grid search optimization repeats this N number of times (here, N = 2500). The smart grid search then reduces the bounds of each parameter based on the spread of the parameters in the top 10 fits. This procedure repeats until the fit of the model fails

to improve. This optimization routine accurately identified parameters on simulated data when tested extensively.

When fitting the simple and choice response models, we collapsed across participants to get stable estimates of RT and p(HVO) per Overlap x Boundary Crossed x Bias Effect Code combination. However, we had very few total incorrect responses for some low overlap (high accuracy) trials. To remove unstable estimates, we excluded conditions that had fewer than 40 trials.

Because sequential sampling models are stochastic, one can run the model with the same parameters and get slightly different fits. To quantify the variability of the fits with a single set of parameters, we (1) ran the best fit model 10 times, (2) calculated the average predicted RT and p(HVO) for each $f(\Psi_{V1} \cap \Psi_{V2})$, and the r^2 and BIC of that average predicted fit. The data and the model fits are presented in Figure 8. The black lines in Figure 8 plots the average fit, and the light gray lines plots the 10 individual runs of the best fit model.

The simple model fit the data extremely well with an overall $r^2 = .88$ with a BIC = -147. The choice response bias model modestly improved the fit over the simple model, $r^2 = .89$ with a BIC = -150. The lower BIC of this model suggests that participants had a small choice response bias (see Figure 8).

Table 4 presents the parameter values for this model. There are three features of interest in the table. First, and most important, is that the RRW fit the data very well using the empirical values of $f(\Psi_{V1} \cap \Psi_{V2})$ collected in Experiment 1 in place of estimating drift rate. This result demonstrates that Psychological Value Theory is a useful theory and method for predicting RT and response choice. Second, the d_B parameter is positive. A positive d_B indicates a recency of information bias. That is, participants are weighting more recent evidence greater than distant

evidence in the VSSP. This produces the relatively long "error" trials whereby participants chose to save the smaller valued item over the larger valued item. Third, the start point bias parameter equaled 0.10. This indicates that the participants shifted their start point bias 10% toward the "do not act" boundary. This bias reveals a general but small reluctance to act when having to sacrifice one person to save another.

Discussion

The precise estimates of the perceived Psychological Values of human lives collected in Experiment 1 accurately predicted the RT and response choice data in Experiment 2. These results confirm the validity of the estimates of the perceived Psychological Values of human lives estimated in Experiment 1, the violation of the Axiom of Equal Valued Lives that they demonstrated, and the validity of Psychological Value Theory. Having successfully modeled participants' responses in sacrificial moral dilemmas, we now ask, "Do the present results replicate when the contextual information of the sacrificial moral dilemmas is removed?"

Experiment 3

Recall that sacrificial moral dilemmas have a formulaic structure: person(s) A will be killed by default if nothing is done. However, an actor can behave in some way to save person(s) A, but as a result, person(s) B will be sacrificed. This structure has been applied across a variety of contexts including Trolley tracks, in a self-driving car, medical scenarios, etc. We created a scenario that simplifies sacrificial dilemmas to their formulaic structure. In Experiment 3, we assess whether participants respond to this simplified scenario with the same pattern of results as they do the scenarios used in Experiment 2. If so, we can conclude that the extraneous information often used to distinguish between contexts in sacrificial dilemmas scenarios is irrelevant to participant's choices, and the patterns of results collected using different scenarios

in various research labs likely generalize. Such a finding would suggest that the results of the Trolley dilemma, for example, will generalize to self-driving cars or medical decisions. In addition, if the results generalize, we will use this simplified scenario in Experiment 4, in which we apply Psychological Value Theory to both group and individual data.

Method

Participants. One hundred forty-eight naïve participants volunteered and received course credit for their participation. Sample size was determined by the process described in Experiment 1 (minimum estimated 100).

Apparatus and Stimuli. The stimuli and apparatus were identical to Experiment 2 with two exceptions. First, in an effort to reduce the number of participants who confused the mapping between the keyboard and the response, we created a dialogue box that displays the keyboard mapping in large and salient lettering. This was presented directly before and directly after the practice trials. Second, we embedded the *Items* in the following single, repeating scenario:

Through circumstances out of your control,

Item1

is about to be killed, but

Item2

will not be affected. You have the opportunity to save Item1.

However, if you save Item1, Item2 will be killed.

Would you save

Item1

causing

Item2

to be killed?

Similar to Experiment 2, for every trial for every participant, "Item1" and "Item2" were randomly chosen from the list of 28 items identified in Experiment 1. For the four non-living items, the word "killed" was replaced with "destroyed."

Procedure. The procedure was identical to Experiment 2.

Results and Discussion

RT and p(HVO) Analysis

We analyzed our data similar to Experiment 2. We removed 8 participants because their average RTs were outliers (i.e., RTs < 2000ms or RTs > 15s) and 7 participants that responded at or below chance (indicating that they did not put effort in the experiment or confused the meaning of their response keys[10]). We then removed less than 0.5% of the individual trials that were RT outliers (i.e., RT > 45s or RT < 1000ms). We analyzed the data of the remaining 133 participants.

There was a significant influence of all parameters of the learning function (a = 0.84, b = –0.06, c = 8.06), all t's > 6.0, p's < .001. We used the average residuals of this function (RT_{res}) as our RT outcome variable. There was a significant linear relation between RT and Overlap, $F(1,9) = 72.9$, $p < 0.001$, $r^2 = 0.90$. There was a significant exponential decay function between p(HVO) and Overlap, t =23.70, $p < 0.001$, $r^2 = 0.99$ (see Table 3 and Figure 10). Finally, we compared the data from Experiment 2 to that of Experiment 3. In all cases, the parameters fitted to the data in Experiment 3 were not significantly different from those of Experiment 2 (all z's < 1.3).

Robust Random Walk Analysis

We modeled the RRW in Experiment 3 in the same way as Experiment 2. We also added a third model: the *Fixed Choice Response Model*. The *Fixed Choice Response Model* is the choice response model with all the parameters *fixed to the values estimated in Experiment 2* except the non-decision time (T_{er}). We freed the non-decision time parameter because the non-decision time is (theoretically) unrelated to the decision process modeled by Psychological Value Theory. If the Fixed Choice Response Model is the best fit model, we can conclude that the parameter values identified in Experiment 2 are meaningful beyond the specific experimental dataset which allows Psychological Value Theory to make specific point predictions. Furthermore, the superiority of the Fixed Choice Response Model would demonstrate that the simplified scenario did not change the participants' decision process.

The simple model fit the data extremely well with an overall $r^2 = .92$ and a BIC = -151. The choice response model, with free parameters, provided an equally good fit as the simple model, $r^2 = .92$ with a BIC = -149. Although this is an excellent fit, the BIC and r^2 of the two models are virtually identical. The Fixed Choice Response Model (with 40 runs) fit the data with an overall $r^2 = .92$ with a BIC = -157 (see Figure 11). Although the r^2 is equivalent to the other models, the BIC is quite a bit lower because this model has only one free parameter whereas the simple model and the choice response model have three and four free parameters, respectively. We therefore conclude that the Fixed Choice Response Model is the best fit model. Table 4 presents the parameter values for the model.

In sum, the RRW fit the RT and response choice in the sacrificial moral dilemma task using the estimates of perceived Psychological Value collected in Experiment 1 as a direct measure of drift rate and the parameters of the RRW estimated in Experiment 2. The data from

Experiment 3 demonstrate that participants' pattern of response to the simplified scenario is identical to that of the elaborate scenarios used in Experiment 2. This result suggests that the data collected using different scenarios in various research labs likely generalize. Therefore, we use the simplified scenario in Experiment 4 to assess whether the perceived Psychological Values collected in Experiment 1 accurately predict the preferential choice behavior of individuals.

Experiment 4

Psychological Value Theory is a model of the cognitive and perceptual processes involved in preferential choices. To date, the predictions of Psychological Value Theory have only been tested and validated using group data (Cohen & Ahn, 2016). In Experiment 4, we assess whether these predictions hold for individuals. Because we assess patterns of behavior for individuals, we can also assess whether perceived Psychological Value is stable across individuals.

Experimentally, it is difficult to collect enough responses to a series of sacrificial moral dilemma from an individual participant to plot a function. The simplified scenario used in Experiment 3 demonstrated that participants' responses to the simplified scenario are virtually identical to those of the more detailed scenarios. Because the simplified scenario is significantly shorter than the more detailed scenarios and does not require variation, we can collect about 120 trials from a single participant in about one hour. Responses to 120 sacrificial moral dilemma scenarios generate about 10 responses per bin when we fit our RT and p(HVO) functions. Ten responses per bin is not optimal because it results in noisy RT measures and low-resolution p(HVO) measures. Nevertheless, it may be used to determine whether Psychological Value

Theory predicts individuals' responses from the population values using the RT and p(HVO) analysis. Unfortunately, 120 trials is too few to run the RRW on individual data.

If Psychological Value is highly variable across individuals, then Ψv will predict preference only within the individual it was measured. Therefore, Ψv of a population will be a poor predictor of preferential choice of an individual. If, however, Psychological Value is stable across individuals, then Ψv of a population will be a good predictor of preferential choice of an individual.

Method

Participants. Ninety naïve participants volunteered and received course credit for their participation. Sample size was determined by the process described in Experiment 1 (minimum estimated 60 because trial numbers were increased).

Apparatus and Stimuli. The stimuli and apparatus were identical to Experiment 2.

Procedure. The procedure was identical to Experiment 2 with the exception that every participant was presented 120 experimental trials.

Results

RT and p(HVO) analysis

We analyzed our data similar to Experiment 2. We removed six participants because their average RTs were outliers (i.e., RTs < 2000ms or RTs > 15s) and 1 participant that responded at or below chance. We then removed 1.3% of the individual trials that were RT outliers (i.e., RT > 45s or RT < 1000ms). We analyzed the data of the remaining 83 participants.

To get a clean RT measure, we first removed the influence of learning by fitting a learning function to each participant's data. We do not report the learning fits here because they are not of substance.

When analyzed as a group, there was a significant linear relation between RT and Overlap, F (1,9) = 105.9, $p < 0.001$, $r^2 = 0.93$. Similarly, there was a significant relation between p(HVO) and Overlap, $t = 13.53$, $p < 0.001$, $r^2 = 0.97$ (see Table 3 and Figure 10).

We also calculated the linear regression, $RT_{res} = a + (b * O_{0.1})$, and exponential decay function, $p(HVO) = 0.5 * (1 - O_{0.1}{}^{b}) + 0.5$, on each participant's data. Figure 12 (top left) presents the best fit linear relation between RT and Overlap for each participant. Figure 12 (top right) presents the best fit exponential decay relation between p(HVO) and Overlap for each participant. The shade of gray of the line is determined by the r^2 of the fit, whereby an $r^2=0$ was plotted in white, with the shade of gray transitioning linearly to black when $r^2=1$. Figure 12 (bottom left and right) presents a histogram of the r^2s for each of these fits.

We calculated t-tests to determine whether the parameter values and r^2 values differed significantly from 0 (see Table 3 under Individual Analysis). The average intercept of the linear relation between RT and Overlap was significantly less than zero, $t(82) = 13.8$, $p < 0.001$. The average slope of the linear relation between RT and Overlap was significantly greater than zero, $t(82) = 13.7$, $p < 0.001$. Before calculating the analysis on r^2 values, we transformed them to reflect the direction of the slope. That is, if the slope associated with a particular r^2 was negative, we multiplied it by -1. We termed this transformed measure, r^2_{signed}. This transformation would result in an average r^2_{signed} of zero if the slopes of the best fit lines were distributed randomly around zero. Furthermore, negative slopes (the opposite of our prediction) would count heavily against the average effect size. The average r^2_{signed} of the linear relation between RT and Overlap was significantly greater than zero, $t(82) = 14.0$, $p < 0.001$. The average beta of the exponential decay relation between p(HVO) and Overlap was significantly greater than zero, $t(82) = 19.2$, *p*

< 0.001. The average r^2 of the exponential decay relation between p(HVO) and Overlap was significantly greater than zero, $t(82) = 44.7$, $p < 0.001$.

Robust Random Walk Analysis

We modeled the RRW on the group data in Experiment 4 in the same way as Experiment 3. The simple model fit the data well with an overall $r^2 = .84$ and a BIC = -131. The choice response model with free parameters provided an equally good fit as the simple model, $r^2 = .85$ with BIC = -132. We then ran the Fixed Choice Response Model (with 40 runs). This model fit with an overall $r^2 = .84$ with a BIC = -139 (see Figure 11). The r^2 is essentially identical and the BIC is quite a bit lower than the other two models. We conclude that the Fixed Choice Response Model is the best fit model. Table 4 presents the parameter values for the model.

Discussion

Experiment 4 revealed that Psychological Value Theory's predictions were extremely accurate when predicting an individual's performance. With only 120 trials, $f(\Psi v_i \cap \Psi v_j)$ accounted for an average of 42% of the variance in RT and 70% of the variance in response choice for each individual. Similar to Experiments 3, the group data in Experiment 4 was well modeled by the RRW using the perceived Psychological Values collected in Experiment 1 and the parameters estimated in the RRW in Experiment 2. These data support the conclusion that Psychological Values are relatively stable across individuals. We will discuss the results in more detail in the General Discussion.

In Experiment 5, we pit the predictions of Psychological Value Theory against a model of preferential choice behavior driven by the Axioms of Monotonicity and Equal Valued Lives.

Experiment 5

Recall that researchers studying sacrificial moral dilemmas often rely on the Axioms of Monotonicity and Equal Valued Lives to classify the response that saves the larger group regardless of individual differences as "characteristically Utilitarian" (e.g., Shenhav & Green, 2010; Trémolière & Bonnefon, 2014). Here, researchers are assuming that quantity alone is driving the consequentialist response. The estimates of perceived Psychological Value of human lives collected in Experiment 1 reveal that perceived Psychological Value of human lives is influenced more by the social status of the individual (e.g., convict vs police officer) than the number of individuals in the group (e.g., single vs large). As such, the smaller group often has a greater Psychological Value than the larger group. For example, the data from Experiment 1 reveal that participants value the life of one college student more than the lives of 50 Congressmen. In this instance, saving the smaller group is the consequentialist response.

In Experiment 5, we pit the predictions of Psychological Value Theory against those of the Axioms of Monotonicity and Equal Valued Lives. We pitted the two groups of people against one another using the simplified moral dilemma. Each group was identified by the number and social status of individuals in the group (e.g., 3 nuns). There were 23 different variations of social status (e.g., judge, nun, etc.) and 15 variations of group size (1; 3-9; 47-53). If Psychological Value Theory is valid, $f(\Psi_{V1} \cap \Psi_{V2})$ should accurately predict participants' RTs and response choices better than the number of lives saved relative to lives lost. If quantity is driving the "act" response despite individual differences of the characters (as is often assumed e.g., the Crying Baby dilemma), the number of lives saved relative to lives lost should have more predictive power than $f(\Psi_{V1} \cap \Psi_{V2})$.

Method

Participants. One hundred and six naïve participants volunteered and received course credit for their participation. Sample size was determined by the process described in Experiment 1 (minimum estimated 60 because trial numbers were increased).

Apparatus and Stimuli. The stimuli and apparatus were identical to Experiment 3 with the following exceptions. For every trial for every participant, "Item1" and "Item2" were randomly chosen from the list of 69 items available probes collected in Experiment 1. The 69 available probes were a cross between 23 social statuses and one of three group sizes: single (1), small (3-9), and large (47-53). If group size was single, the probe would be preceded by "a" or "an" item such as "a nun." Each individual probe was assigned a small and large group size that differed by 44, which was presented to the left of the probe. For example, if "nun" was assigned a small value of 3 (e.g., 3 nuns), it's large value would be 47 (e.g., 47 nuns).

Procedure. The procedure was identical to Experiment 4.

Results

We analyzed our data similar to Experiment 2. Five participants were removed because they either did not complete the experiment or did not comply with instructions not to use cell phones. We removed six participants because their average RTs were outliers (i.e., RTs < 2000ms or RTs > 15s) and 10 participants that responded at or below chance (indicating that they did not put effort in the experiment or confused the meaning of their response keys). We then removed 1% of the individual trials that were RT outliers (i.e., RT > 45s or RT < 1000ms). We analyzed the data of the remaining 90 participants both as a group and individually. To get a clean RT measure, we first removed the influence of learning by fitting a learning function to

each participant's data, log(RT) = a * exp (*b* * (trial – 1)) + *c*. For each participant, we used the average residuals of this function (RT_{res}) as our RT outcome variable.

Quantity Analysis

The Axiom of Monotonicity predicts that participants should prefer to save the larger sized group over the smaller sized group—all things being equal. Researchers, however, often vary the characters in some of their most theoretically influential dilemmas. We therefore assess how well quantity predicts responses assuming both the Axioms of Monotonicity and Equal Valued Lives.

We tested the influence of quantity by coding the probability the participant saved the Higher Quantity Option (i.e., saved more lives). We term this measure *p(HQO)* to distinguish it from Psychological Values' accuracy measure, p(HVO). We then calculated a linear regression whereby the criterion variable was p(HQO) and the predictor variable was *group size difference* (i.e, abs(number of people in Item1 – number of people in Item2)). Figure 13 presents these data. To retain only stable estimates of p(HQO), we removed any group size difference that had fewer than 20 total trials contributing to the calculation of p(HQO). Based on this criterion, we did not remove any group size differences in the current analysis. There was a significant linear relation, $F(1, 21) = 5.9$, $p = 0.02$, $r^2 = 0.22$[11]. Both the intercept (0.55) and the slope (0.003) were significant, t's > 2.4, $p < 0.05$. As depicted in Figure 13, non-linear regressions will not significantly improve the fit of the function.

We ran a secondary analysis restricted to scenarios containing only characters of the same social status (a pure test of the Axiom of Monotonicity). Because this secondary analysis has many fewer observations, we binned the predictor variable (group size difference) by rounding it to the nearest even number. Further, to retain some stability in estimates, we removed

any bin that had fewer than 7 total trials contributing to the calculation of p(HQO). This resulted in 8 bins with an average of 38.5 (SD=32) observations per bin. There was no significant relation between p(HQO) and group size difference, $F(1, 7) = 0.04$, $p = 0.85$, $r^2 = 0.01$. Additionally, the intercept of this analysis (intercept = 0.67) did not significantly differ from the intercept of the analysis that included all the scenarios (intercept = 0.55), $t(28) = 1.77$, $p = 0.09$, indicating that participants choose to save the larger group at the same rate regardless of whether the social status of the two groups are the same or different.

Psychological Value Analysis

Psychological Value Theory predicts (a) a positive relation between RT and Overlap and (b) a negative relation between p(HVO) and Overlap. The quantity analysis demonstrated some relation between group size difference and the likelihood that a participant will save the larger group when one accepts the Axiom of Equal Valued Lives. The small influence of group size difference on participants' preferential choices, however, may reflect the small influence of quantity on perceived Psychological Value. If Psychological Values are the driving mechanism, Psychological Values should better predict the probability that participants will save the larger group than group size difference. To assess this, we divided the scenario item pairs into two sets: 1) those pairs in which Psychological Value Theory predicted a preference to save the larger group (termed *consistent*) and 2) those pairs in which Psychological Value Theory predicted preference to save the smaller group (termed *inconsistent*). In the consistent trials, Psychological Value Theory predicts that $p(HQO) > 0.5$ (chance) and p(HQO) should decrease as Overlap increases. In the inconsistent trials, Psychological Value Theory predicts that $p(HQO) \leq 0.5$ because for the inconsistent trials Psychological Value Theory predicts the

opposite of the Axiom of Monotonicity. P(HQO) should increase as Overlap increases because it will approach chance as predicted by Psychological Value Theory.

We quantified the above predictions in the following linear model:

$$p(HQO) = a + b_1(-1*trialType) + b_2(trialType*Overlap), \quad (4)$$

where trialType = 1 for inconsistent trials, trialType = -1 for consistent trials, and Overlap is the average Overlap collapsed across subjects for each group size difference in the scenario. One group size difference was removed in the current analysis because it had less than 20 total trials contributing to the calculation of p(HQO) (see above). The model was significant, $F(1, 41) = 216.6$, $p < 0.001$, $r^2 = 0.91$. All parameters were significant, $a = 0.63$, $b_1 = 0.99$, and $b_2 = 1.21$, *t*'s > 6, $p < 0.01$ (see Figure 13). To assess whether quantity can account for any additional variance, we added group size difference as a predictor variable. Although the group size difference was marginally significant, slope = 0.001, t=2.2, p = 0.03, it increased the total variance explained by the model by only 1% ($r^2 = 0.92$ vs $r^2 = 0.91$). This supports the conclusion that the influence of group size difference on participants' preferential choices was primarily a result of the small influence of quantity on perceived Psychological Value.

We ran a secondary analysis restricted to scenarios containing only characters of the same social status (a pure test of the Axiom of Monotonicity). Because this secondary analysis has lower power, we binned Overlap by rounding it to the nearest 0.02. Further, we removed any bin that had fewer than 7 total trials contributing to the calculation of p(HQO). Because there were only 2 datapoints in the inconsistent condition, we ran our standard predicted exponential decay analysis on the consistent datapoint (6 bins, Mean observations per bin = 40, SD=33). There was a significant relation between p(HQO) and Overlap, $t = 6.4$ $p = 0.001$, $r^2 = 0.93$[12].

Below, we analyze the data consistent with the overall predictions of Psychological Value Theory.

RT and p(HVO) Analysis. When analyzed as a group, there was a significant linear relation between RT and Overlap, $F(1,8) = 241.0$, $p < 0.001$, $r^2 = 0.97$. There was a significant relation between p(HVO) and Overlap, $t = 14.5$ $p < 0.001$, $r^2 = 0.97$ (see Table 3 and Figure 10).

The same pattern of results was present when analyzed individually (see Table 3 and Figure 14). Individual performance was assessed by calculating the linear regression, $RT_{res} = a + (b^* O_{0.1})$, and exponential decay function, $p(HVO) = 0.5 * (1 - O_{0.1}{}^b) + 0.5$, on each participant's data. All parameter values and r^2s were significantly different from zero (dfs = 89, all ts > 10, ps < 0.001).

Robust Random Walk Analysis. We modeled the RRW on the group data in Experiment 5 in the same way as Experiment 3. The simple model fit the data extremely well with an overall $r^2 = .92$ with a BIC = -194. The choice response bias model with free parameters provided a slightly less good fit than the simple model, $r^2 = .91$ with BIC = -185. We then ran the Fixed Choice Response Model (with 40 runs). This model fit with an overall $r^2 = .91$ with a BIC = -199 (see Figure 11). The BIC is quite a bit lower than the other two models. We therefore conclude that the Fixed Choice Response Model is the best fit model. Table 4 presents the parameter values for the model. The superiority of the Fixed Choice Response Model is important because the perceived Psychological Values used in Experiment 5 were different than the perceived Psychological Values used in Experiment 2. This provides further evidence that quantity does not influence the decision process. There was insufficient data to run a pure monotonicity analysis using the RRW.

Discussion

The results of Experiment 5 reveal that the difference in the number of lives in the large and small groups has minimal influence on choice in sacrificial moral dilemmas (whether or not one accepts the Axiom of Equal Valued Lives). In contrast, perceived Psychological Value predicts choice extremely well in all cases. These results support the conclusion that quantity has minimal influence on perceived Psychological Values, and perceived Psychological Values drive choice in sacrificial moral dilemmas.

Because Psychological Value Theory is a general model of how Psychological Value drives preferential choice, it predicts that the processes driving judgments in sacrificial moral dilemmas should be the same as those driving economic judgments. Below, we assess whether Psychological Value Theory predicts participants preferential choice behavior when making economic sacrificial decisions.

Experiment 6

Experiment 6 was identical to Experiment 5 with the exception that the items were economic goods that were destroyed rather than people that were killed. As such, these dilemmas asked participants if they would save one group of economic goods by letting another group of economic goods be destroyed. Because the items we used in these the dilemmas are considered "goods" (e.g., kayaks, jewelry, pencils, etc.), the judgments are traditionally considered "economic" rather than "moral."

Psychological Value Theory proposes that the same value-based processes drive decisions that involve human lives and those that involve economic goods and services. Therefore, $f(\Psi_{V1} \cap \Psi_{V2})$ should predict RT and the p(HVO) in Experiment 6. If, however, decisions involving human lives recruit uniquely different cognitive and perceptual processes

than those involving economic goods, Psychological Value Theory will not predict participants' RT and p(HVO) in Experiment 6.

Method

Participants. Eighty-five naïve participants volunteered and received course credit for their participation. Sample size was determined by the process described in Experiment 1.

Apparatus and Stimuli. The apparatus and stimuli in Experiment 6 were identical to those of Experiment 5 with two exceptions. First, the word "killed" was replaced with "destroyed" in the scenario. Second, the 72 available probes were a cross between 24 economic goods (e.g., a smartphone, a water bottle, etc.) and one of three group sizes: single (1), small (3-9), and large (47-53). In all other ways, the experiments were identical.

Procedure. The procedure was identical to that used in Experiment 5.

Results

We analyzed and report our data similar to Experiment 5. Two participants were removed because they either did not complete the experiment or did not comply with cell phone use instructions. We removed 15 participants because their average RTs were outliers (i.e., RTs < 2000ms or RTs > 15s) and 3 participants that responded at or below chance (indicating that they did not put effort in the experiment or confused the meaning of their response keys). We then removed 2.5% of the individual trials that were RT outliers (i.e., RT > 45s or RT < 1000ms). We analyzed the data of the remaining 65 participants individually. To get a clean RT measure, we first removed the influence of learning by fitting a learning function to each participant's data, $\log(RT) = a * \exp(b * (trial - 1)) + c$. For each participant, we used the average residuals of this function (RT_{res}) as our RT outcome variable.

Quantity Analysis

The quantity analysis was conducted identical to that of Experiment 5. We first assessed the influence of quantity accepting the Axiom of Equal Valued Lived (including all items). To do so, we calculated a linear regression whereby the criterion variable was p(HQO) and the predictor variable was the group size difference in the scenario (see Figure 13). To retain only stable estimates of p(HQO), we removed any group size difference that had fewer than 20 total trials contributing to the calculation of p(HQO). Based on this criterion, we removed no group size differences in the current analysis. There was a significant linear relation, $F(1, 21) = 4.8$, $p = 0.04$, $r^2 = 0.19$[13]. Both the intercept (0.55) and the slope (0.002) were significant, t's > 2.0, $p < 0.05$. We then assessed the influence of quantity rejecting the Axiom of Equal Valued Lived (restricting the analysis to only identical items), resulting in 7 bins, Mean observations per bin = 27, SD=16. The results revealed no significant relation between p(HQO) and group size difference, $F(1, 6) = 0.36$, $p = 0.57$, $r^2 = 0.06$. The intercept of this analysis (intercept = 0. 96) was significantly greater than the intercept of the analysis that included all the scenarios (intercept = 0. 55), $t(28) = 13.3$, $p < .01$, indicating that participants choose to save the items in the larger group more often when the items are identical vs when the items differ.

Psychological Value Analysis

Psychological Value Theory predicts (a) a positive relation between RT and Overlap and (b) a negative relation between p(HVO) and Overlap. The quantity analysis shows some relation between group size difference and the likelihood that a participant will choose the larger group when all the items were included in the analysis. To assess if this effect was driven by Psychological Values, we conducted the same linear regression (Equation 4) as Experiment 5. The model was significant, $F(1, 43) = 507.2$, $p < 0.001$, $r^2 = 0.96$. All parameters were

significant, a = 0.55, b_1 = 0.78, and b_2 = 0.74, *t*'s > 4, p < 0.01 (see Figure 13). To assess whether quantity can account for any addition variance, we added group size difference as a predictor variable into the previous analysis. The group size difference predictor variable was not significant, t=0.97, p = 0.34. This suggests that the influence of group size difference on participants' preferential choices was solely a result of the small influence of quantity on perceived Psychological Value. We then assessed the influence of quantity restricted to only identical items (7 bins, Mean observations per bin = 20, SD=23). The analysis revealed a significant relation between p(HQO) and Overlap, t =6.8 p = 0.03, r^2 = 0.83[14].

Below, we analyze the data consistent with the overall predictions of Psychological Value Theory.

RT and p(HVO) Analysis. When analyzed as a group, there was a significant linear relation between RT and Overlap, $F(1,8) = 182.8$, $p < 0.001$, $r^2 = 0.80$. There was a significant relation between p(HVO) and Overlap, t =14.7, $p < 0.001$, $r^2 = 0.98$ (see Table 3 and Figure 10).

The same pattern of results was present when analyzed individually (see Table 3 and Figure 14). Individual performance was assessed by calculating the linear regression, $RT_{res} = a + (b * O_{0.1})$, and exponential decay function, $p(HVO) = 0.5 * (1 - O_{0.1}{}^b) + 0.5$, on each participant's data. All parameter values and r^2*s* were significantly different from zero (dfs = 64, all ts > 9, ps < 0.001).

Robust Random Walk Analysis. We modeled the RRW on the group data in Experiment 6 in the same way as Experiment 3. The simple model fit the data very well with an overall r^2 = .90 with a BIC = -164. The choice response model with free parameters improved the fit of the simple model with an overall r^2 = .92 with BIC = -171 (see Figure 11). Specifically,

the start bias effect parameter was estimated at -0.08. The negative start point effect parameter indicates that participants were biased towards the "action" boundary.

We then ran the Fixed Choice Response Model (with 40 runs). This model fit with an overall $r^2 = .79$ with a BIC = -134 (see Figure 11). The BIC is quite a bit higher than the other two models, and the r^2 is quite a bit lower. The Fixed Choice Response Model did not fit the current data well because the start point bias in the Fixed Choice Response Model is positive whereas the start point bias exhibited in the data (and identified by the choice response model) is negative. We therefore conclude that the choice response model is the best fit model. Table 4 presents the parameter values for this model. There was insufficient data to run a pure monotonicity analysis using the RRW.

Discussion

The results of Experiment 6 parallel that of Experiment 5. Specifically, participants' responses were best predicted by similarity in the perceived Psychological Value of the items rather than the difference in the quantity of the items—whether or not the analysis was restricted to identical items. The fact that the RRW fit extremely well driven by the perceived Psychological Values of economic goods estimated in Experiment 1 demonstrates that participants' economic decisions are well described by Psychological Value Theory. This supports the conclusion that the same value-based processes drive decisions involving human lives as those involving economic goods. Furthermore, similar to Experiment 1, quantity has minimal influence on perceived Psychological Values, and perceived Psychological Values drive choice in economic dilemmas.

There were two interesting differences between the results of Experiments 5 and 6. First, in Experiment 6 (economic goods), participants were more likely to save the larger quantity

group when both groups contained identical items versus when they contained different items. This same effect was not present in Experiment 5 (human lives). Because p(HQO) was well predicted by similarity in Psychological Value—but not by group size difference—for identical items, this result can likely be modeled by the RRW (e.g., a start point bias). Unfortunately, we had insufficient data from identical items to explore this hypothesis using the RRW. Future research should address this and other questions relating to how perceived Psychological Value drives economic decisions. Second, participants exhibited a bias toward the "action" boundary in Experiment 6 versus the "do nothing" boundary in Experiment 5. We discuss the implications of this finding in the General Discussion.

General Discussion

Here, we conducted a strong test of the hypothesis that sacrificial moral dilemmas are solved by attempting to identify and save the option with the greatest perceived Psychological Value. To do so, we collected estimates of perceived Psychological Values of a variety of people and economic goods in Experiment 1. In Experiments 2-6, we assessed whether the perceived Psychological Values estimated in Experiment 1 predicted participants' preferential choices in sacrificial moral and economic dilemma tasks using a VSSP (see Table 1). The results were unambiguous: the perceived Psychological Value data collected in Experiment 1 accounted for about 90% of the variance in participants' RT and response choices in Experiments 2-6. Specifically, in Experiment 2, the RRW, driven by the Psychological Values collected in Experiment 1, accurately predicted participants' RTs and response choices ($r^2 = .89$). In Experiments 3-5, the RRW, driven by the Psychological Values collected in Experiment 1 and using the decision relevant parameter values estimated in Experiment 2, again accurately predicted participants' RTs and response choices (average $r^2 = .89$). In Experiment 6, the RRW,

driven by the Psychological Values collected in Experiment 1 revealed that participants exhibited a choice response bias opposite of that exhibited in Experiments 2-5 ($r^2 = .92$). Thus, we cross-validated our results with multiple datasets using multiple methods. Together, Experiments 1-6 are a strong demonstration that sacrificial moral dilemmas are value-based tasks, that perceived Psychological Value is a useful theoretical construct that explains known facts and predicts new ones, and that Psychological Value Theory is a valid model of how Psychological Value drives value-based choice behavior.

The current experiments are the first to use a VSSP to accurately predict participants' RTs and response choices from directly measured $f(\Psi_{V1} \cap \Psi_{V2})$ in a complex social decision task (see Table 1). Previously, researchers have constrained VSSPs by linking conditions to ratings of value surrogates or currency when inferring drift rate from participants' responses (e.g., Krajbich et al., 2010; Krajbich & Rangel, 2011; Milosavljevic et al., 2010). Here, we have constrained the VSSPs further by using the estimates of Psychological Value collected in Experiment 1 as direct measures of drift rate. In addition, we have applied the VSSP to a seemingly complex social decision that is generally not believed to be value-based: sacrificial moral dilemmas. Despite the increased constraints and the difficulty of the problem, our model predicted participant performance with a high degree of precision and accuracy.

Sacrificial Moral Dilemmas as Value-Based Tasks

Traditionally, researchers hypothesize two separate and independent decision mechanisms involved in moral judgment: a slow, rational decision mechanism that considers characteristically Utilitarian ethical consequences of the dilemma and a fast, emotional decision mechanism that reacts to the ethical nature of the action, despite the consequences (e.g., Cushman, Young & Greene, 2010; Greene et al., 2001; Greene et al., 2004; Greene et al., 2008;

Greene & Haidt, 2002). The extant literature discusses these mechanisms as if they are in conflict, which results in one mechanism producing a response (e.g., Amit & Greene, 2012; Ciaramelli, Muccioli, Làdavas, & Di Pellegrino, 2007; Greene et al., 2001; Greene et al., 2004; Greene et al., 2008). Psychological Value Theory, in contrast, captures both consequentialist and non-consequentialist influences simultaneously within a single, stochastic decision process.

By instantiating Psychological Value Theory's decision process in the RRW, we advance the theoretical and predictive value of theory. For example, the RRW quantifies both the consequentialist and non-consequentialist influences. The RRW is—like all VSSPs—primarily consequentialist because choosing the option with the greatest perceived Psychological Value is the primary goal. As such, $f(\Psi_{V1} \cap \Psi_{V2})$ quantifies the consequentialist influence. The RRW incorporates non-consequentialist influences in the response bias parameters. These are non-consequentialist because they are driven by factors other than the primary goal (i.e., choosing the higher valued option). For example, a speed response bias with compressed boundaries may indicate impulsive behavior or time pressure. Similarly, a choice response bias toward the "do nothing" boundary may indicate an aversion to "killing" which researchers traditionally classify as characteristically Deontological.

Table 4 presents the parameters required to fit the RRW in the current experiment. The first thing to notice about Table 4 is that Experiments 3-5 are best fit by fixing the decision-relevant parameters to the values obtained in Experiment 2. This is a powerful demonstration of the robust nature of the process, our ability to replicate the effect across experiments, and Psychological Value Theory's ability to a priori predict responses based on the perceived Psychological Value of the items. Importantly, the RRW fit the data equally well in experiments involving human lives and experiments involving economic goods. This supports

the claim that decisions involving human lives and decisions involving economic goods are explained by the same underlying decision process.

Critically, the free parameters of the RRW are readily interpretable. The non-decision times are captured by T_{er}. Non-decision times are ancillary processes unrelated to perceived Psychological Value and response biases. For example, T_{er} captures the time associated with reading, motor movement, etc. The d_b parameter is a measure of the how the accumulation of value information is weighted over time. A positive d_b indicates a recency effect whereby more recent information is given greater weight than older information. The recency effect is present in all experiments.

The speed response bias is captured by the Boundary parameter. The Boundary parameter is larger for scenarios involving human lives than those involving economic goods. When the Boundary parameter is compared across Experiments 5 and 6, the data reveal that for equally difficult decisions (as defined by $f(\Psi v_1 \cap \Psi v_2)$), participants take more time to make a decision about a human life than about an economic good. Because wider boundaries produce more accurate responses, this indicates that participants are biased toward accuracy (rather than speed) when making decisions about human lives.

A choice response bias is captured by the start point parameter. The RRW revealed that observers shift their start point 10% closer to the "do nothing" boundary when confronted with a dilemma in which human lives are involved. This suggests that observers prefer to mistakenly let a higher valued group die by default than to accidentally kill a higher valued group through their action. In contrast, when observers are confronted with a dilemma involving economic goods, they shift their start point 8% closer to the "act" boundary. This suggests that observers prefer to mistakenly take action to destroy a higher valued economic good than to mistakenly let

a higher valued economic good be destroyed by default. It is tempting to develop an interpretation involving characteristically Deontological reasoning for the dilemmas involving human lives (e.g., a reluctance to kill). However, we cannot readily identify a parsimonious characteristically Deontological explanation for the opposite direction bias found in the dilemmas involving economic goods. Thus, characteristically Deontological reasoning may not be driving the "do nothing" response.

One alternative interpretation of the choice response biases found in our dilemmas is that this bias is related to the observers' need or want to control the outcome. That is, a bias toward the "act" boundary may represent a willingness to take control whereas a bias toward the "do nothing" boundary may represent a willingness to relinquish control. Such an interpretation suggests that people would rather be in control when it comes to the fate of economic goods and relinquish control when it comes to the fate of human lives. This is just conjecture, of course, but it provides an avenue for future research. Thus, by quantifying biases in readily interpretable parameters, the RRW provides the first steps to understand the biases.

Using the RRW to explicitly model the decision processes also increases the constraints and precision of Psychological Value Theory. There are many proposed decision mechanisms for two alternative forced choice tasks (e.g., ballistic, heuristic, quantum, etc.,). Each of these decision mechanisms have different assumptions and will produce different predictions from the same input. By modeling a specific decision mechanism, the RRW increases the information content of Psychological Value Theory which increases our understanding of the cognitive processes involved in solving sacrificial moral dilemmas. Furthermore, the RRW simultaneously predicts the p(HVO), the RT for choosing the HVO, and the RT for choosing the LVO. Because these three dependent variables are interdependent within the system modeled by

the RRW, fitting the data using the RRW is much more constrained than fitting independent functions to p(HVO) and RT (as Cohen & Ahn, 2016 did). Again, more constraints increase the information content of the model (e.g., Glöckner, & Betsch, 2011).

The Psychological Value of Human Lives

There is a common assumption that more lives have greater value (e.g., Dickert, Västfjäll, Kleber, & Slovic, 2012; Greene et al., 2001; Slovic, 2007; Tversky & Kahneman, 1992). For example, Slovic (2007) proposed that the value of saving a human life is a monotonically increasing logarithmic function of the number of lives saved (e.g., Dickert et al., 2012). The number of lives saved has the biggest influence on behavior when the quantity of lives saved is small. As the quantity of lives saved increases, the added saved lives have diminishing influence on observers' choice behavior. Our data, in contrast, reveal that quantity of lives has minimal influence on the perceived Psychological Value of the group. This is true both when social status is held constant and when social status varies. As such, error prone systems (such as human cognitive and perceptual systems) will not reliably perceive the larger group as having a higher Psychological Value than the smaller group. This likely explains why researchers have found that increases in quantity of lives do not motivate participants to act in a variety of experimental situations such as donating to charity (e.g., Dickert, Kleber, Västfjäll, & Slovic, 2016), saving lives in a sacrificial dilemma (e.g., Greene et al., 2008; Koenigs, et al., 2007), helping others (e.g., Fetherstonhaugh, Slovic, Johnson, & Friedrich, 1997), making behavioral economic decisions, (e.g., Hsee, & Rottenstreich, 2004; Tversky & Kahneman, 1981), etc.

Our data also shows that observers perceive subtle differences in Psychological Value between individuals who differ only on social status. From a general viewpoint, this violation of the Axiom of Equal Valued Lives is expected because there is abundant evidence that people

will display a preference for one person or group over another (e.g., Brewer, 1979; Devine, 1989; Hewstone, Rubin, & Willis, 2002; Monteith, & Walters, 1998). Nevertheless, researchers designing sacrificial moral dilemmas often classify the consequentialist option solely on the basis of the number of lives saved (e.g., Greene et al., 2004; Koenigs et al., 2007). We have shown, in contrast, that individual differences can lead to large differences in perceived Psychological Value. When combined with the minimal influence of quantity, individual differences of the characters in sacrificial moral dilemmas can easily result in the smaller group having a greater perceived Psychological Value than the larger group. When this happens, the "do nothing" response is the consequentialist option. Recall, within the context of Psychological Value Theory, the alternative response results from an error prone system (rather than a separate and independent decision mechanism).

We therefore recommend that researchers do not rely on assumptions such as Monotonicity and Equal Valued Lives to identify the relative Psychological Value of humans (e.g., Conway & Gawronski, 2013; Greene et al., 2001; Greene et al., 2004; Greene et al., 2008; Koenigs et al., 2007). Rather, to identify the group of people with the higher perceived Psychological Value, researchers *should measure perceived Psychological Value of each group*. To make precise point predictions, researchers should have a measure of value *similarity*. Psychological Value Theory provides that measure. We discuss Psychological Value Theory in more detail below.

Psychological Value Theory

Psychological Value Theory defines the theoretical construct of Psychological Value independent of choice, measures that construct independent of choice, and *a priori* predicts

preferential choice from those measurements using a VSSP. Psychological Value Theory rests on the following premises:

(1) the theoretical construct of perceived Psychological Value exhibits perceptual variability, is measurable, is perceived consistently across individuals, and is the primary feature that drives preferential choice including sacrificial moral dilemmas.

(2) the overlap of perceived Psychological Value distributions, $f(\Psi v_1 \cap \Psi v_2)$, describes an observer's sensitivity to the higher valued option (HVO) in a value-based task, and

(3) the RRW describes the single process, stochastic decision mechanism involved in value-based choices including sacrificial moral dilemmas.

If any one of these premises were not valid, the Psychological Values collected in Experiment 1 could not have accurately predicted participants' RTs and response choices in Experiments 2-6. Thus, our results provide evidence that the system of premises supporting Psychological Value Theory are valid (e.g., Popper, 1968).

We proposed a new theoretical construct, termed *Psychological Value*, that is central to Psychological Value Theory. Some may ask whether Psychological Value actually *exists*. Whether or not a theoretical construct exists is a question that physicists have long ago reframed (e.g., Carnap, 1966). The relevant question is not whether the theoretical construct is "real." Rather the relevant question is whether the theoretical construct is necessary to explain known facts and predict new ones. To quote Carnap (1966),

> The theoretical terms are convenient symbols. The postulates containing them are adopted because they are useful, not because they are "true". They have no surplus meaning beyond the way in which they function in the system. It is

> meaningless to talk about the "real" electron or the "real" electromagnetic field. (p. 33-34)

That is, theoretical constructs are features of theories that are necessary to explain known facts and predict new ones. So we ask, "Is Psychological Value necessary within the system described by Psychological Value Theory to explain known facts and predict new ones?"

By postulating Psychological Value and its' features, we were able to anticipate how it functions across a variety of conditions. For example, we identified the method for measuring the perceptual event and derived the relative similarity of two perceptual value events *from their individual measurements*. Importantly, the similarity measure we derived, $f(\Psi v_1 \cap \Psi v_2)$, has a true zero that indicates identical perceptual events (and carries similar information as the ROC from SDT). Such precise measurement permits the similarity estimate to be used directly as the driving evidence in a sequential sampling model of decision making (i.e., the RRW). The RRW produces point predictions for every item pair with a minimal number of free parameters. When applied to a complex social decision task (i.e., sacrificial moral dilemmas), the RRW produced highly accurate predictions. Thus, within the system described by Psychological Value Theory, *Psychological Value* (as defined herein) is necessary to explain known facts and predict new ones.

A fundamental assumption of Psychological Value Theory is that Psychological Value is perceived. Within the field of decision neuroscience, there is skepticism that perceptual mechanisms can be applied to the construct of value (e.g., Gold & Shadlen, 2007; Padoa-Schioppa, 2013). For example, Gold and Shadlen (2007) state,

> Choices on both perceptual and value-based tasks often appear to be governed by a random process. For perceptual tasks, this randomness is typically explained by

> considering the evidence as a mixture of signal plus noise. The DV and decision rule are both formulated to minimize the effects of this noise in pursuit of a particular goal. However, for value-based decisions, the randomness is often assumed to be part of the decision process itself. That is, a subjective measure like utility is used to assign the relative desirability of each choice. The decision rule is then probabilistic: a random selection weighed by these relative measures … (p. 560)

The viewpoint is supported by Padoa-Schioppa (2013), who states,

> … the concept of accumulation of evidence over time, which is central to perceptual decisions (Gold & Shadlen, 2007), does not equally apply to economic decisions. Indeed the ‘‘evidence’’ in economic decisions (i.e., offer values) is immediately available, not delivered gradually over time. (p. 1332)

Psychological Value Theory and our data contribute to this debate. We directly measure the variability associated with perceived Psychological Value and use those measurements in a VSSP to a priori predict response choice. By successfully doing so, we present strong evidence that the randomness in value-based decisions results from the evidence (i.e., the perceived Psychological Value) rather than the decision process. This result provides support for the premise that Psychological Value is perceived.

The theoretical construct of perceived Psychological Value is the foundational premise from which the system that comprises Psychological Value Theory cascades. It is therefore reasonable to ask whether one can identify other more elemental theoretical constructs that comprise perceived Psychological Value (e.g., issues of “right” and “wrong,” “mine” vs “yours,” etc., see Figure 3). Equation 1 explicitly specifies that perceived Psychological Value is

multidimensional. Therefore, Psychological Value Theory allows for Psychological Value to be decomposed into multiple features. Because Psychological Value Theory places perceived Psychological Values within the context of a well understood and quantified decision theory, researchers can use this knowledge to study the composition of perceived Psychological Values. Indeed, Cohen and Lecci (2001) present a procedure that can be used as a road map for how one might decompose the features that comprise perceived Psychological Values. We believe, however, that Psychological Value, like color (which is composed of hue, saturation, and intensity), is perceived as a gestalt rather than as a composition of features. Nevertheless, we encourage researchers to use Cohen and Lecci's (2001) procedure (among others) to explore these issues.

In the introduction, we specified several defining properties of Psychological Values (e.g., perceptual variability, monotone, multidimensional, etc.). *Defining properties* are essential for the model to function. In Appendix C, we present several *auxiliary properties* of Psychological Values. Auxiliary properties increase the constraints of the model but are not essential. Therefore, if auxiliary properties are invalidated, the model will still hold, but will have more free parameters.

Conclusion

Here, we present a strong test of the hypothesis that sacrificial moral dilemmas are value-based tasks whereby the observer attempts to identify and save the item with the greatest value. To do so, we introduce Psychological Value Theory which predicts choice from independent measurements of value. We instantiated the decision mechanism proposed by Psychological Value Theory in a new Robust Random Walk (RRW) procedure. Using the principles of Psychological Value Theory, we measured the perceived Psychological Value of human lives

and economic goods in Experiment 1. We then used those measurements to predict participants' reaction times and preferential choices in complex social decisions in Experiments 2-6. The RRW accounted for about 90% of the variance in participants' RTs and response choices. We cross validated our results with multiple datasets using multiple methods. These experiments provide strong evidence that (a) responses to sacrificial moral dilemmas are simple preferential choices based on the perceived Psychological Value of the options, (b) perceived Psychological Value of lives is highly influenced by individual differences of people but minimally influenced by the number of people in a group, and (c) within the system described by Psychological Value Theory, the theoretical construct termed *Psychological Value* is necessary to explain known facts and predict new ones.

Open Practices Statement

The analyses of Experiments 2-6 rely on five R packages written by the first author. These packages can be retrieved from Github at the following URLs:

https://github.com/ccpluncw/ccpl_R_chutils
https://github.com/ccpluncw/ccpl_R_chValues
https://github.com/ccpluncw/ccpl_R_chMorals
https://github.com/ccpluncw/ccpl_R_RRW
https://github.com/ccpluncw/ccpl_R_smartGridSearch

The data and R code for the analysis of Experiment 2-6, as well as the distributional overlaps of perceived Psychological Values from Experiment 1 relevant to the current experiments, can be downloaded at:
https://github.com/ccpluncw/ccpl_data_SVThumanLives2020.git

Please cite this article if the R packages and/or data are used in any way to produce a publication. These packages and data cannot be used without the first author's consent for any for profit endeavor.

Although the experiments were not pre-registered, the procedures and analyses in Experiments 1 replicated the procedures and analyses of Cohen and Ahn (2016) Experiment 1. Furthermore, Experiments 3-6 replicated and cross-validated Experiment 2, thus demonstrating the robust nature of the effects.

Summary of Revisions Across Versions

Summary of Revisions from Version 1 to Version 2

There were major revisions from Version 1 to Version 2. These occurred over the course of about a year. They were driven by two factors: 1) reviewer comments and 2) my deeper understanding and the continued improvement of the Robust Random Walk (RRW). Below is a list of the major changes:

1. I changed the name of the theory from "**Subjective Values Theory**" to "**Psychological Value Theory**." I did so because the term "Subjective Values" is a buzz word in the literature, so the term comes with surplus meaning. The surplus meaning sometimes creates confusion. I now name the theory after the theoretical construct (Psychological Value) from which the rest of the theory cascades.
2. I greatly shortened and focused the introduction (and corresponding General Discussion). I did so at the request of reviewers. This occurred at the expense of some of the important points I wanted to make. Nevertheless, the new introduction is more readable, precise, and focused. As such, it is an improvement.
3. I moved much of the technical information is in the appendices. This, again, improves readability.
4. As a result of reviewer comments and my own adjustments, I have improved the RRW. These changes include the following. First, I now often fix the nSD parameter at 1.0, because freeing that parameter only negligibly reduces the fit statistic. Second, when calculating the BIC and the r^2, I equalize the influence of the p(HVO) and RT data. This ensures that neither the DV will have undue influence on the model fit. Third, I minimize on the BIC statistic, rather than the r^2 statistic. Finally, I have improved the flexibility of the RRW by implementing a generalized mechanism to assess a priori, parameter specific model predictions.

References

Aldy, J. E., & Viscusi, W. K. (2007). Age differences in the value of statistical life: revealed preference evidence. *Review of Environmental Economics and Policy*, *1*(2), 241-260.

Amit, E., & Greene, J. D. (2012). You see, the ends don't justify the means: Visual imagery and moral judgment. *Psychological science*, *23*(8), 861-868.

Ashby, F. G., & Lee, W. W. (1993). Perceptual variability as a fundamental axiom of perceptual science. *Advances in Psychology, 99*, 369–399. https://doi.org/10.1016/s0166-4115(08)62778-8

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological review, 93*(2), 154. https://doi.org/10.1037//0033-295x.93.2.154

Ashby, N. J. S., Jekel, M., Dickert, S., & Glöckner, A. (2016). Finding the right fit: A comparison of process assumptions underlying popular drift-diffusion models. Journal of Experimental Psychology: Learning, Memory, and Cognition, 42(12), 1982–1993. doi:10.1037/xlm0000279

Attneave, F. (1971). Multistability in perception. *Scientific American, 225*(6), 62-71. https://doi.org/10.1038/scientificamerican1271-62

Barberis, N. C. (2013). Thirty years of prospect theory in economics: A review and assessment. *Journal of Economic Perspectives, 27*(1), 173-96. https://doi.org/10.1257/jep.27.1.173

Bhui, R. (2019). Testing Optimal Timing in Value-Linked Decision Making. Computational Brain & Behavior, 2(2), 85–94. doi:10.1007/s42113-019-0025-9

Białek, M., & De Neys, W. (2017). Dual processes and moral conflict: Evidence for deontological reasoners' intuitive utilitarian sensitivity. *Judgment and Decision Making*, *12*(2), 148.

Bohnet, I., & Frey, B. S. (1999). Social distance and other-regarding behavior in dictator games: Comment. *American Economic Review, 89*(1), 335-339. https://doi.org/10.1257/aer.89.1.335

Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological bulletin, 86*(2)*,* 307. https://doi.org/10.1037//0033 2909.86.2.307

Busemeyer, J. R., Gluth, S., Rieskamp, J., & Turner, B. M. (2019). Cognitive and neural bases of multi-attribute, multi-alternative, value-based decisions. *Trends in cognitive sciences*.

Busemeyer, J. R., & Johnson, J. G. (2004). Computational models of decision making. *Blackwell handbook of judgment and decision making*, 133-154.

Carnap, R. (1966). Chapters 26. *Philosophical foundations of physics* (Vol. 966). New York: Basic Books.

Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *Journal of Neuroscience, 29*(39), 12315-12320.https://doi.org/10.1523/jneurosci.2575-09.2009

Cohen, D. J., & Ahn, M. (2016). A subjective utilitarian theory of moral judgment. *Journal of Experimental Psychology: General, 145*(10), 1359. https://doi.org/10.1037/xge0000210

Cohen, D. J., & Lecci, L. (2001). Using magnitude estimation to investigate the perceptual components of signal detection theory. *Psychonomic Bulletin & Review, 8*, 284–293. http://dx.doi.org/10.3758/BF03196163

Colas, J. T., & Lu, J. (2017). Learning where to look for high value improves decision making asymmetrically. *Frontiers in psychology, 8*, 2000. https://doi.org/10.3389/fpsyg.2017.02000

Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: A process dissociation approach. *Journal of Personality and Social Psychology, 104*(2), 216-235. https://doi.org/10.1037/a0031021

Coren, S. (2012). Sensation and perception. *Handbook of Psychology, Second Edition*, *1*. https://doi.org/10.1002/9781118133880.hop201007

Cunningham, M. R., Roberts, A. R., Barbee, A. P., Druen, P. B., & Wu, C. H. (1995). " Their ideas of beauty are, on the whole, the same as ours": Consistency and variability in the cross-cultural perception of female physical attractiveness. *Journal of personality and social psychology, 68*(2), 261. https://doi.org/10.1037//0022-3514.68.2.261

Cushman, F., Young, L., & Greene, J. D. (2010). Our multi-system moral psychology: Towards a consensus view. *The Oxford handbook of moral psychology*, 47-71.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of personality and social psychology, 56*(1), 5-18. https://doi.org/10.1037//0022-3514.56.1.5

Dickert, S., Kleber, J., Västfjäll, D., & Slovic, P. (2016). Mental imagery, impact, and affect: A mediation model for charitable giving. *PloS one*, *11*(2), https://doi.org/10.1371/journal.pone.0148274

Dickert, S., Västfjäll, D., Kleber, J., & Slovic, P. (2012). Valuations of human lives: normative expectations and psychological mechanisms of (ir) rationality. *Synthese, 189*(1), 95-105. https://doi.org/10.1007/s11229-012-0137-4

Dyer, J. S., & Sarin, R. K. (1979). Measurable Multiattribute Value Functions. Operations Research, 27(4), 810–822. doi:10.1287/opre.27.4.810

Edwards, W. (1954). The theory of decision making. *Psychological bulletin, 51*(4), 380. https://doi.org/10.1037/h0053870

Fetherstonhaugh, D., Slovic, P., Johnson, S., & Friedrich, J. (1997). Insensitivity to the value of human life: A study of psychophysical numbing. *Journal of Risk and uncertainty, 14*(3), 283-300. https://doi.org/10.1023/a:1007744326393

Fiedler, S., & Glöckner, A. (2015). Attention and moral behavior. *Current Opinion in Psychology*, *6*, 139-144. doi:10.1016/j.copsyc.2015.08.008

Fink, B., & Penton-Voak, I. (2002). Evolutionary psychology of facial attractiveness. *Current Directions in Psychological Science, 11*(5), 154-158. https://doi.org/10.1111/1467-8721.00190

Fishburn, P. C. (1968). Utility theory. *Management science, 14*(5), 335-378. https://doi.org/10.1287/mnsc.14.5.335

Fishburn, P. C. (1970). *Utility theory for decision making* (No. RAC-R-105). Research analysis corp McLean VA. https://doi.org/10.21236/ad0708563

Fishburn, P. C. (1981). Subjective expected utility: A review of normative theories. *Theory and decision, 13*(2), 139-199. https://doi.org/10.1007/bf00134215

Gescheider, G. A. (1988). Psychophysical scaling. *Annual Review of Psychology, 39*, 169-200. https://doi.org/10.1146/annurev.psych.39.1.169

Girgus, J. J., Rock, I., & Egatz, R. (1977). The effect of knowledge of reversibility on the reversibility of ambiguous figures. *Perception & Psychophysics, 22*(6), 550-556. https://doi.org/10.3758/bf03198762

Glöckner, A., & Betsch, T. (2011). The empirical content of theories in judgment and decision making: Shortcomings and remedies. *Judgment and Decision Making*, *6*(8), 711.

Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work?. *Trends in cognitive sciences*, *6*(12), 517-523.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual review of neuroscience*, *30*.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1). New York: Wiley.

Greene, J. D. (2007). Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains. *Trends in cognitive sciences*, *11*(8), 322-323.

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition, 107*(3), 1144-1154. https://doi.org/10.1016/j.cognition.2007.11.004

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron, 44*(2), 389-400. https://doi.org/10.1016/j.neuron.2004.09.027

Greene, J. D., Sommerville, R., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293*(5537), 2105-2108. https://doi.org/10.1126/science.1062872

Grueschow, M., Polania, R., Hare, T. A., & Ruff, C. C. (2015). Automatic versus choice-dependent value representations in the human brain. *Neuron, 85*(4), 874-885. https://doi.org/10.1016/j.neuron.2014.12.054

Gwinn, R., Leber, A. B., & Krajbich, I. (2019). The spillover effects of attentional learning on value-based choice. *Cognition, 182*, 294-306. https://doi.org/10.1016/j.cognition.2018.10.012

Hawkins, G. E., Forstmann, B. U., Wagenmakers, E.-J., Ratcliff, R., & Brown, S. D. (2015). Revisiting the Evidence for Collapsing Boundaries and Urgency Signals in Perceptual Decision-Making. Journal of Neuroscience, 35(6), 2476–2484. doi:10.1523/jneurosci.2410-14.2015

von Helmholtz, H. (1924). Helmholtz's treatise on physiological optics (Trans. from the 3rd German ed.) (J. P. C. Southall, Ed.). Rochester, NY, US: Optical Society of America. http://dx.doi.org/10.1037/13536-000

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual review of psychology, 53*(1), 575-604. https://doi.org/10.1146/annurev.psych.53.100901.135109

Hsee, C. K., & Rottenstreich, Y. (2004). Music, Pandas, and Muggers: On the Affective Psychology of Value. Journal of Experimental Psychology: General, 133(1), 23–30. doi:10.1037/0096-3445.133.1.23

Jameson, D., & Hurvich, L. M. (1961). Complexities of perceived brightness. *Science, 133*(3447), 174-179. https://doi.org/10.1126/science.133.3447.174

Jones, M., & Dzhafarov, E. N. (2014). Unfalsifiability and mutual translatability of major modeling schemes for choice reaction time. *Psychological review*, *121*(1), 1-32.

Jones, M., Love, B. C., & Maddox, W. T. (2006). Recency effects as a window to generalization: separating decisional and perceptual sequential effects in category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(2), 316. https://doi.org/10.1037/0278-7393.32.3.316

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature*, *446*(7138), 908-911.

Krajbich, I. (2019). Accounting for attention in sequential sampling models of decision making. *Current opinion in psychology, 29*, 6-11. https://doi.org/10.1016/j.copsyc.2018.10.008

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature neuroscience, 13*(10), 1292. https://doi.org/10.1038/nn.2635

Krajbich, I., Bartling, B., Hare, T., & Fehr, E. (2015). Rethinking fast and slow based on a critique of reaction-time reverse inference. *Nature communications, 6*. https://doi.org/10.1038/ncomms8455

Krajbich, I., Hare, T., Bartling, B., Morishima, Y., & Fehr, E. (2015). A common mechanism underlying food choice and social decisions. *PLoS computational biology, 11*(10), e1004371. https://doi.org/10.1371/journal.pcbi.1004371

Krajbich, I., Lu, D., Camerer, C., & Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in psychology, 3*, 193. https://doi.org/10.3389/fpsyg.2012.00193

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences, 108*(33), 13852-13857. https://doi.org/10.1073/pnas.1101328108

Kurzban, R., DeScioli, P., & Fein, D. (2012). Hamilton vs. Kant: Pitting adaptations for altruism against adaptations for moral judgment. *Evolution and Human Behavior*, *33*(4), 323-333. doi:10.1016/j.evolhumbehav.2011.11.002

Lebreton, M., Jorge, S., Michel, V., Thirion, B., & Pessiglione, M. (2009). An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron, 64*(3), 431-439. https://doi.org/10.1016/j.neuron.2009.09.040

Lemer, C., Dehaene, S., Spelke, E., & Cohen, L. (2003). Approximate quantities and exact number words: Dissociable systems. *Neuropsychologia*, *41*(14), 1942-1958.

Leopold, D. A., & Logothetis, N. K. (1999). Multistable phenomena: changing views in perception. *Trends in cognitive sciences, 3*(7), 254-264. https://doi.org/10.1016/s1364-6613(99)01332-7

Lim, S. L., O'Doherty, J. P., & Rangel, A. (2011). The decision value computations in thevmPFC and striatum use a relative value code that is guided by visual attention. *Journal of Neuroscience, 31*(37), 13214-13223. https://doi.org/10.1523/jneurosci.1246-11.2011

Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika, 40*, 77–105. http://dx.doi.org/10.1007/BF02291481

Marks, L. E., & Algom, D. (1998). Psychological scaling. In M. H. Birn- baum (Ed.), *Measurement, judgment, and decision making* (pp. 81-178). New York: Academic Press. https://doi.org/10.1016/b978-012099975-0.50004-x

Mill, J. S. (1998). *Utilitarianism* (R. Crisp, Ed.). New York, NY: Oxford University Press. (Original work published 1861)

Millar, C., Starmans, C., Fugelsang, J., & Friedman, O. (2016). It's personal: The effect of personal value on utilitarian moral judgments. *Judgment & Decision Making, 11*(4). https://doi.org/10.1016/j.cognition.2014.05.018

Mormann, M. M., Malmaud, J., Huth, A., Koch, C., & Rangel, A. (2010). The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. https://doi.org/10.2139/ssrn.1901533

Monteith, M. J., & Walters, G. L. (1998). Egalitarianism, moral obligation, and prejudice-related personal standards. *Personality and Social Psychology Bulletin, 24*(2), 186-199. https://doi.org/10.1177/0146167298242007

Özgen, E. (2004). Language, learning, and color perception. *Current Directions in Psychological Science, 13*(3), 95-98. https://doi.org/10.1111/j.0963-7214.2004.00282.x

Padoa-Schioppa, C. (2013). Neuronal origins of choice variability in economic decisions. *Neuron*, *80*(5), 1322-1336.

Paxton, J. M., Ungar, L., & Greene, J. D. (2012). Reflection and reasoning in moral judgment. *Cognitive science*, *36*(1), 163-177.

Popper, K. R. (1968). *The logic of scientific discovery*. New York: Harper & Row. https://doi.org/10.4324/9780203994627

Rangel, A., & Clithero, J. A. (2014). The computation of stimulus values in simple choice. In *Neuroeconomics* (pp. 125-148). Academic Press. https://doi.org/10.1016/b978-0-12-416008-8.00008-5

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59-108. https://doi.org/10.1037//0033-295x.85.2.59

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two- choice decisions. *Psychological Science, 9*, 347–356. http://dx.doi.org/ 10.1111/1467-9280.00067

Roberts, R., & Goodwin, P. (2002). Weight approximations in multi-attribute decision models. Journal of Multi-Criteria Decision Analysis, 11(6), 291–303. doi:10.1002/mcda.320

Rock, I. E. (1997). *Indirect perception*. The MIT Press.

Rock, I., Wheeler, D., & Tudor, L. (1989). Can we imagine how objects look from other viewpoints?. *Cognitive Psychology, 21*(2), 185-210. https://doi.org/10.1016/0010-0285(89)90007-8

Shenhav, A., & Greene, J. D. (2010). Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron, 67*(4), 667-677. https://doi.org/10.1016/j.neuron.2010.07.020

Shenhav, A., & Greene, J. D. (2014). Integrative moral judgment: dissociating the roles of the amygdala and ventromedial prefrontal cortex. *Journal of Neuroscience, 34*(13), 4741-4749. https://doi.org/10.1523/jneurosci.3390-13.2014

Slovic, P. (2007). "If I look at the mass I will never act": Psychic numbing and genocide. *Judgment and Decision Making, 2*(2), 79–95.

Smith, V. L. (1989). Theory, experiment and economics. *Journal of Economic Perspectives*, *3*(1), 151-169.

Stevens, S. S. (1986). *Psychophysics: Introduction to its perceptual, neural, and social prospects.* Oxford: Transaction.

Stevens, S. S. (1975). *Psychophysics: Introduction to its Perceptual, Neural, and Social Prospects.* Wiley, New York

Stevens, S. S. (1957). On the psychophysical law. *Psychological review, 64*(3), 153. https://doi.org/10.1037/h0046162

Stevens, S. S. (1955). The measurement of loudness. *The Journal of the Acoustical Society of America, 27*(5), 815-829. https://doi.org/10.1121/1.1908048

Stigler, G. J. (1950). The development of utility theory. I. *Journal of Political Economy, 58*(4), 307-327. https://doi.org/10.1086/256962

Suter, R. S., & Hertwig, R. (2011). Time and moral judgment. *Cognition*, *119*(3), 454-458.

Swann Jr, W. B., Gómez, Á., Dovidio, J. F., Hart, S., & Jetten, J. (2010). Dying and killing for one's group: Identity fusion moderates responses to intergroup versions of the trolley problem. *Psychological Science, 21*(8), 1176-1183. https://doi.org/10.1177/0956797610376656

Tavares, G., Perona, P., & Rangel, A. (2017). The attentional drift diffusion model of simple perceptual decision-making. *Frontiers in neuroscience*, *11*, 468.

Trémolière, B., & Bonnefon, J. F. (2014). Efficient kill–save ratios ease up the cognitive demands on counterintuitive moral utilitarianism. *Personality and Social Psychology Bulletin, 40*(7), 923-930. https://doi.org/10.1177/0146167214530436

Trémolière, B., De Neys, W., & Bonnefon, J. F. (2012). Mortality salience and morality: Thinking about death makes people less utilitarian. *Cognition*, *124*(3), 379-384.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty, 5*(4), 297-323. https://doi.org/10.1007/bf00122574

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453-458. https://doi.org/10.1126/science.7455683 United Nations. (1948). Universal declaration of human rights. UN General Assembly.

Table 1

A summary of the measurements, estimates, and assumptions of Utility Theory, Sequential Sampling, and Psychological Value Theory. Utility Theory (with and without attribute ratings) and SSPs estimate values from behavioral responses. Sequential Sampling with preference ratings estimates the distributional overlap of values (i.e., drift rate) from behavioral responses while assuming the shape of and constraining the placement of the value distributions. By measuring both the value distributions and the behavioral responses, Psychological Value Theory is the most constrained model of all those presented.

Theory	**Model**	**Value Distributions**			**Response**		**Response Bias**	
		Shape	Placement	Overlap	Preferential Choice	Reaction Time	Choice Bias	Speed Bias
Utility Functions	V' = *f*(R)	✗	e	✗	✓	✗	✗	✗
Sequential Sampling	V' = *f*(R)	a	e	e	✓	✓	e	e
Sequential Sampling w/Preference Ratings	V' = *f*(R + V_p)	a	✓	e	✓	✓	e	e
Psychological Value Theory	**R = *f*(V)**	✓	✓	✓	✓	✓	e	e

Note: ✓ = Measured; e = Estimated; a = Assumed; ✗ = none; V' = *f*(R) = Value is estimated from measured responses; V' = *f*(R + V_p) = Value is estimated from measured responses and constrained by measured value placements; R = *f*(V) = Measured responses are predicted from measured values.

Table 2

Response Quartiles of Estimated Psychological Values for the Probes Selected for Experiments 2-4

Probe	Psychological Values		
	25%	50%	75%
A Pedophile	0.00	0.00	11.25
A Rapist	0.00	0.00	100.00
A Terrorist	0.00	0.00	100.00
A Cockroach	0.00	2.00	22.50
A Rabid Possum	0.00	10.00	100.00
An Assassin	0.00	10.00	1,000.00
A Thief	0.00	100.00	1,000.00
A Gang Member	1.00	200.00	1,000.00
A Smartphone	100.00	600.00	1,700.00
A Convict	1.00	650.00	1,500.00
A Chimpanzee (the standard)		**1,000.00**	
An Addict	100.00	1,000.00	5,000.00
An Adult with a Deadly Contagious Disease	50.00	1,000.00	5,500.00
A Congressman	500.00	2,000.00	10,000.00
A Nun	1,000.00	3,000.00	15,000.00
A Celebrity	1,000.00	3,000.00	20,000.00
A Billionaire	1,000.00	4,000.00	50,000.00
A Judge	1,000.00	5,000.00	10,000.00
A Homeless Adult	1,200.00	5,000.00	10,000.00
An Olympian	1,500.00	5,000.00	12,500.00
An Astronaut	2,000.00	5,000.00	31,000.00
A College Student	2,750.00	8,500.00	72,500.00
An Adult	2,000.00	10,000.00	60,000.00
A Mentor	2,000.00	10,000.00	75,000.00
A Police Officer	2,000.00	10,000.00	85,000.00
An Elderly Person	3,000.00	10,000.00	100,000.00
An Orphan	4,000.00	10,000.00	100,000.00
A Soldier	6,000.00	15,000.00	1,000,000.00
A Life-Saving Antidote	7,000.00	30,000.00	5,000,000.00

Note. Values for the response quartiles represent the raw responses compared to the standard, chimpanzee = 1000. The probes are listed in ascending order of the estimated Psychological Values in the following order of response quartiles: 50%, 75%, 25%.

Table 3

The parameter estimates and r^2 for the RT and p(HVO) analyses for Experiments 2-6.

	Reaction Times $RT = f(\Psi v_1 \cap \Psi v_2)$			Response Choices $p(HVO) = f(\Psi v_1 \cap \Psi v_2)$	
Group Analysis	Slope	Intercept	r^2	Beta	r^2
Experiment 2	0.37	-0.21	0.91	1.91	0.97
Experiment 3	0.44	-0.23	0.90	1.99	0.99
Experiment 4	0.42	-0.22	0.93	2.07	0.97
Experiment 5	0.28	-0.17	0.97	1.84	0.97
Experiment 6	0.24	-0.15	0.80	2.61	0.98

Individual Analysis	Slope		Intercept		r^2_{signed}		Beta		r^2	
	M	SD	M	SD	M	SD	M	SD	M	SD
Experiment 4	0.42	0.28	-0.22	0.15	0.42	0.27	2.36	1.11	0.70	0.14
Experiment 5	0.29	0.27	-0.18	0.17	0.32	0.30	2.39	1.15	0.68	0.18
Experiment 6	0.27	0.21	-0.18	0.14	0.35	0.29	2.14	1.83	0.75	0.14

Note: All parameter values for all experiments are significant at $p < 0.001$.

Table 4

The parameter estimates and r^2 for the Robust Random Walk for Experiments 2-6.

Experiment	**T_{er}**	**Boundary**	**d_B**	**SPE**	**r^2**
Experiment 2	-1.18	49.65	0.26	0.10	.89
Experiment 3	-1.52	49.65*	0.26*	0.10*	.92
Experiment 4	-1.27	49.65*	0.26*	0.10*	.84
Experiment 5	-1.10	49.65*	0.26*	0.10*	.91
Experiment 6	-0.60	22.36	0.12	-0.08	.92

Note: SPE = Start Point Effect
*indicates a fixed parameter value set from those estimated in Experiment 2

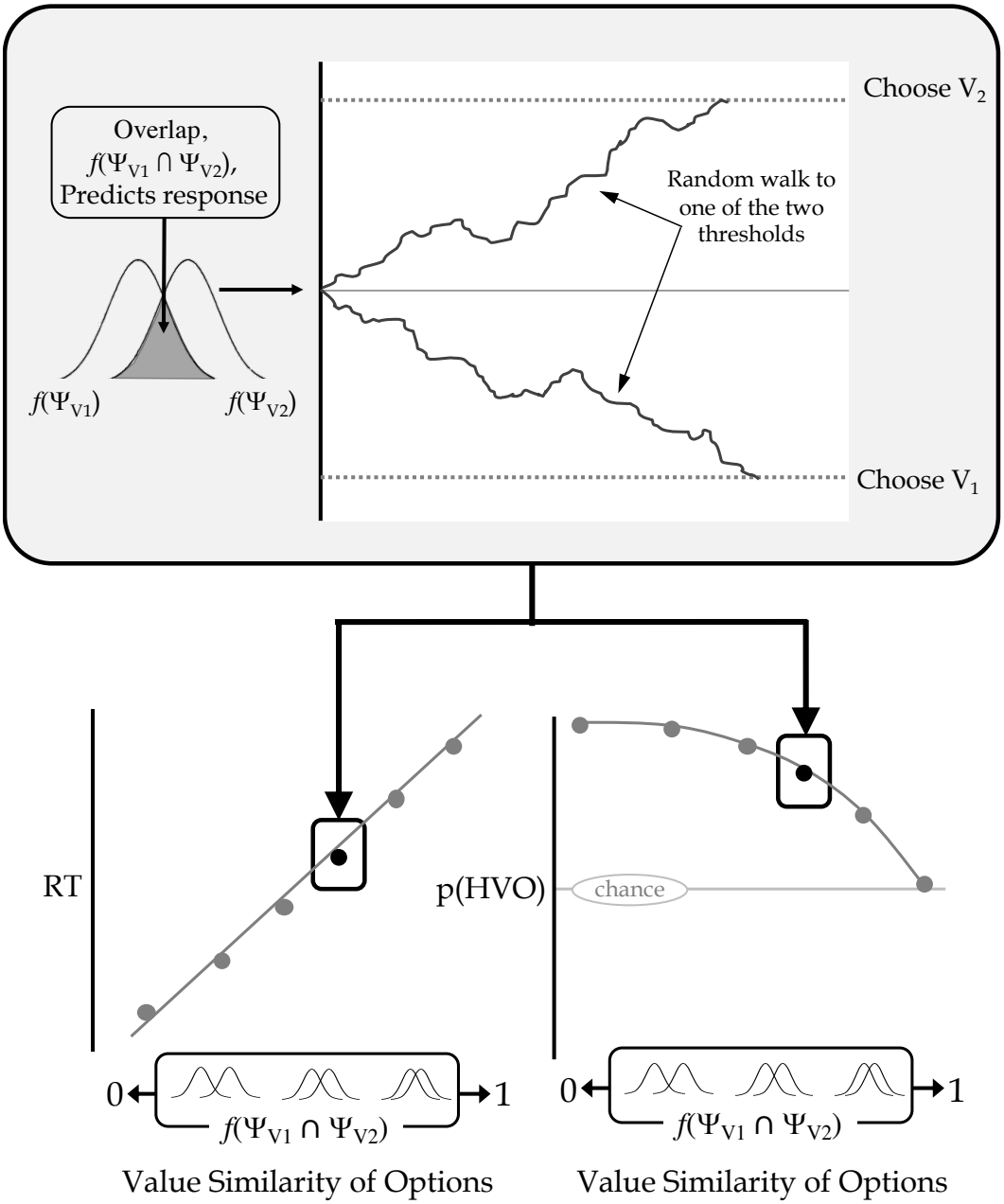

Figure 1: An illustration of a simplified VSSP. At the start of the process, there is little evidence for either choice, though the start point can be biased closer to one or the other boundary. Over time, evidence accumulates and randomly walks toward one or the other boundary. Each boundary represents a response choice. When enough evidence accumulates to cross a boundary, the observer responds consistent with that boundary. Value similarity of the two options, $f(\Psi v_1 \cap \Psi v_2)$, drives the speed that evidence is accumulated. As $f(\Psi v_1 \cap \Psi v_2)$ increases, RT increases and the probability of choosing the higher valued option, p(HVO), decreases.

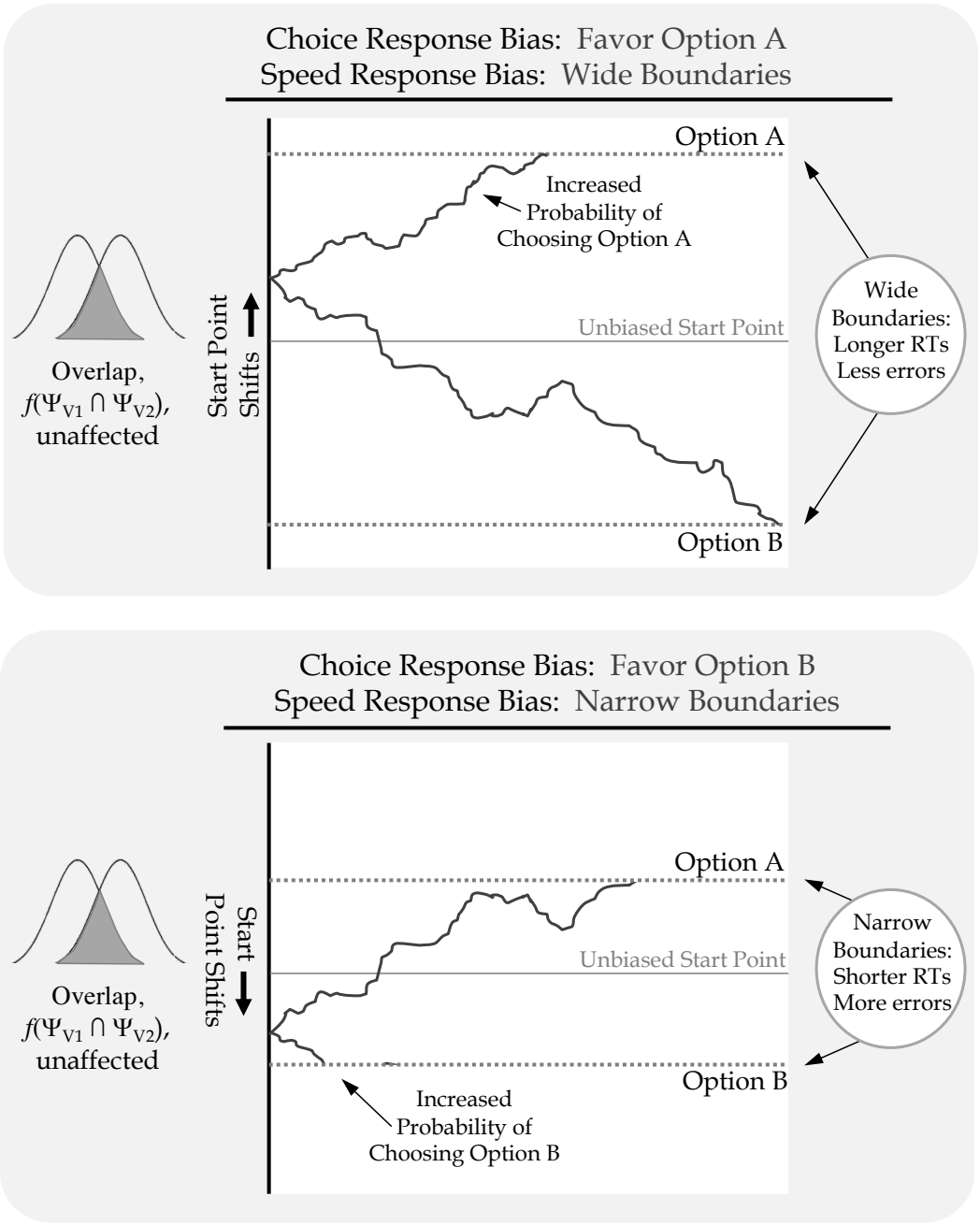

Figure 2: An illustration of both a choice response bias and a speed response bias. The choice response bias increases probability of choosing the option that is closer to the start point. The RTs associated with that response choice will tend to be faster than the RTs associated with the opposing boundary. The speed response bias changes the overall speed and accuracy of the responses. Wide boundaries lead to slow accurate responses and narrow boundaries lead to fast inaccurate responses. Both choice and speed response biases change p(HVO) without a corresponding shift in the relative value of the options, $f(\Psi v_1 \cap \Psi v_2)$.

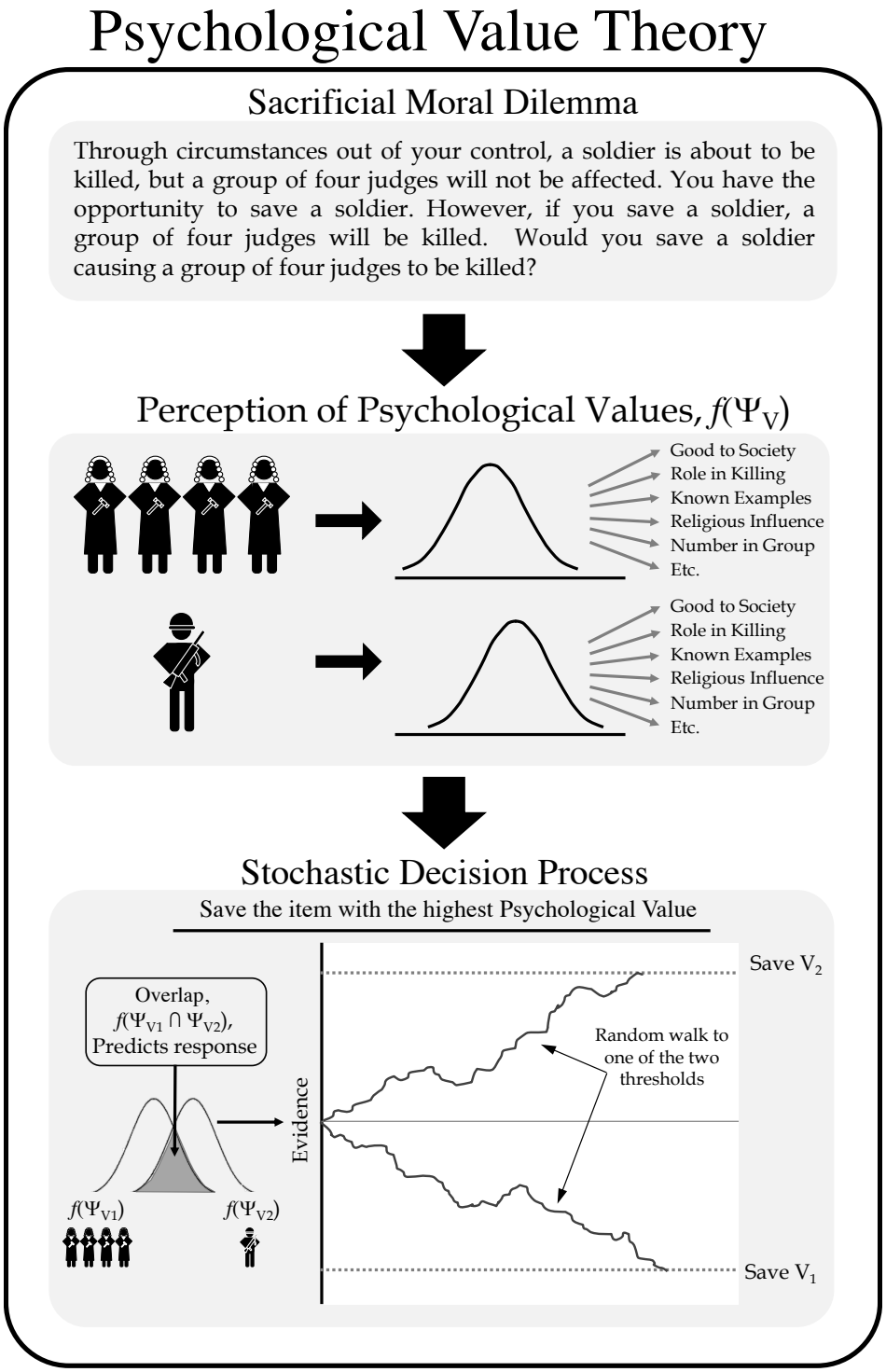

Figure 3: When applied to sacrificial moral dilemmas, Psychological Value Theory assumes that people have a perceived Psychological Value of each item in the dilemma. This perceived Psychological Value may be influenced by many features including Utilitarian, Deontological, Religious, Memory, etc. A single, stochastic process attempts to identify and save the group with the highest perceived Psychological Value.

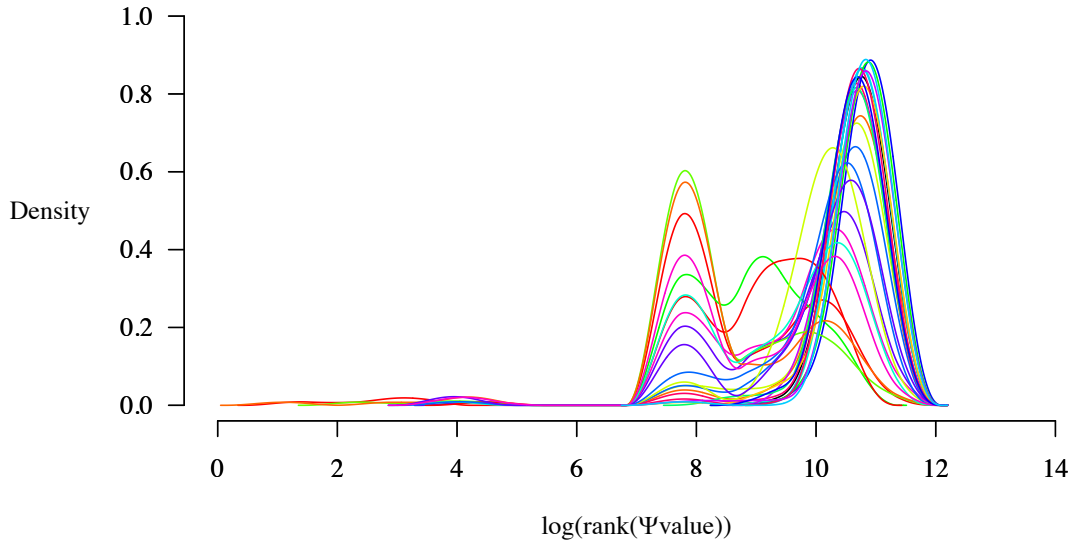

Figure 4. The relative placement of $f(\Psi_v)$ for all 28 probes used in Experiment 1-3. The presented distributions are kernel density estimates of the log transformed ranked values using the density function in R.

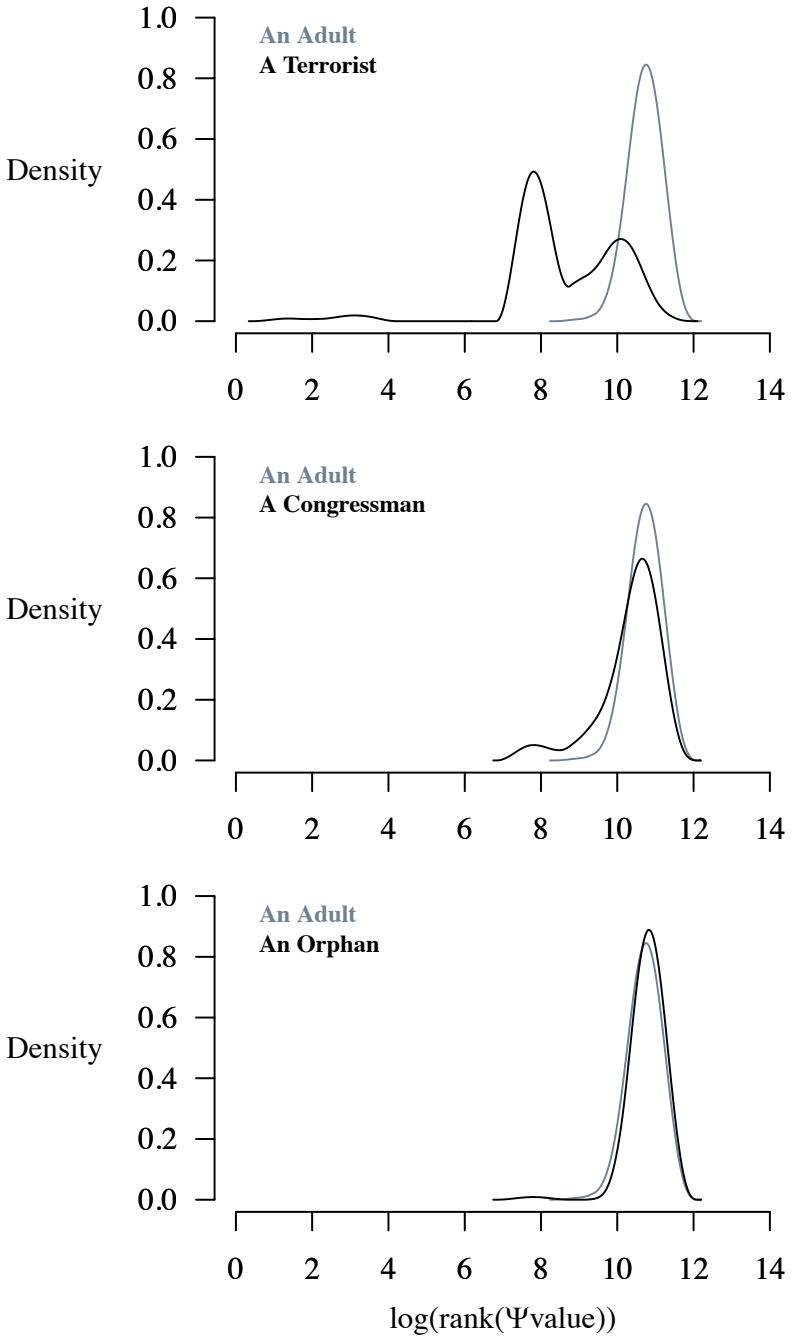

Figure 5. The placement of $f(\psi_v)$ for an adult relative to a terrorist, a congressman, and an orphan from Experiment 1. These individual probes have increasing overlap values. The presented distributions are kernel density estimates of the log transformed ranked values using the density function in R.

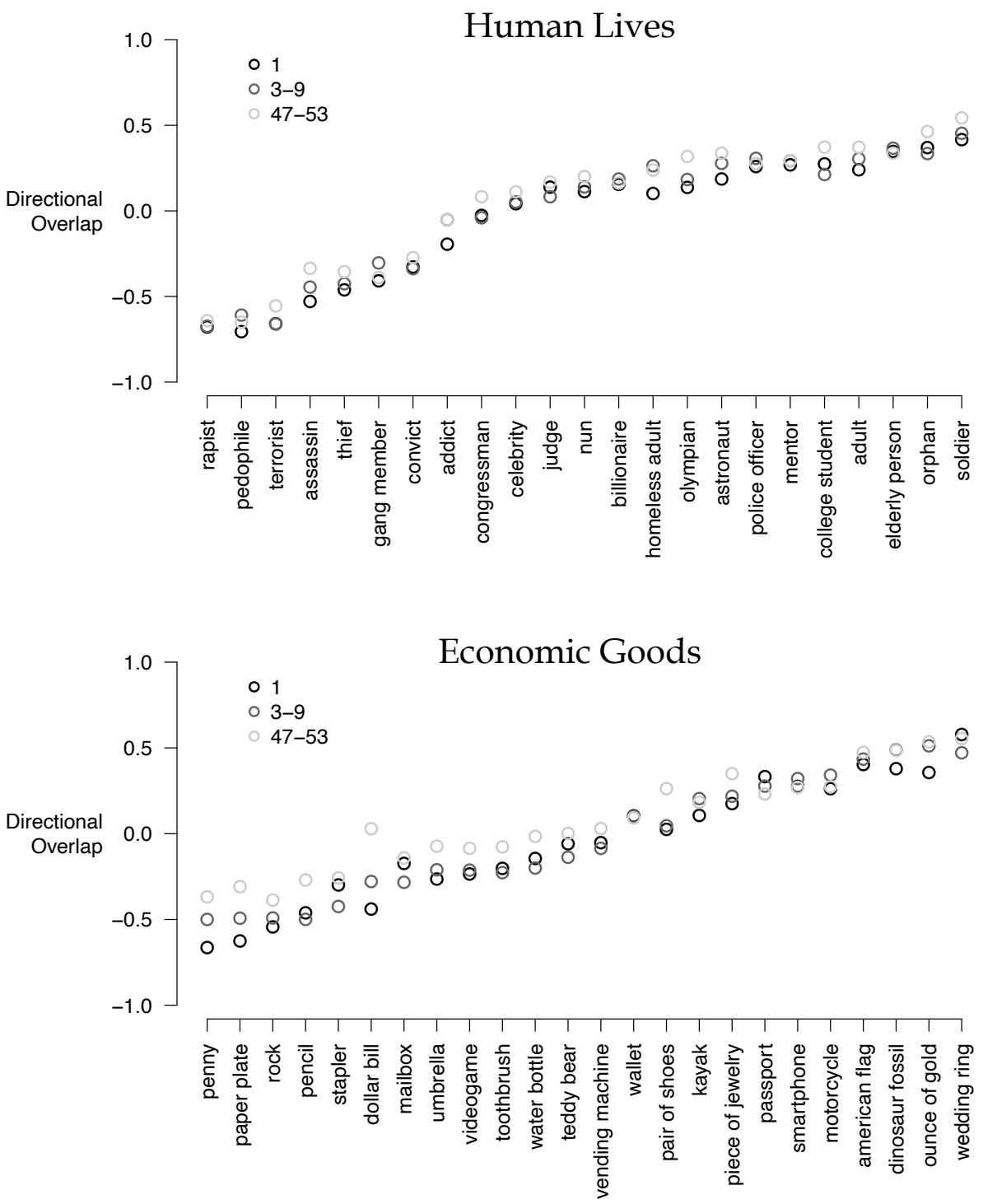

Figure 6: The relative Psychological Values of the probes of different quantities collected in Experiments 1 and used in Experiments 2-6 presented as directional overlap (y-axis; see the text). Higher directional overlap corresponds to higher valued options. The graphs reveal that perceived Psychological Value is highly influenced by individual differences in the person/economic good and minimally influenced by changes in the quantity of people/economic goods.

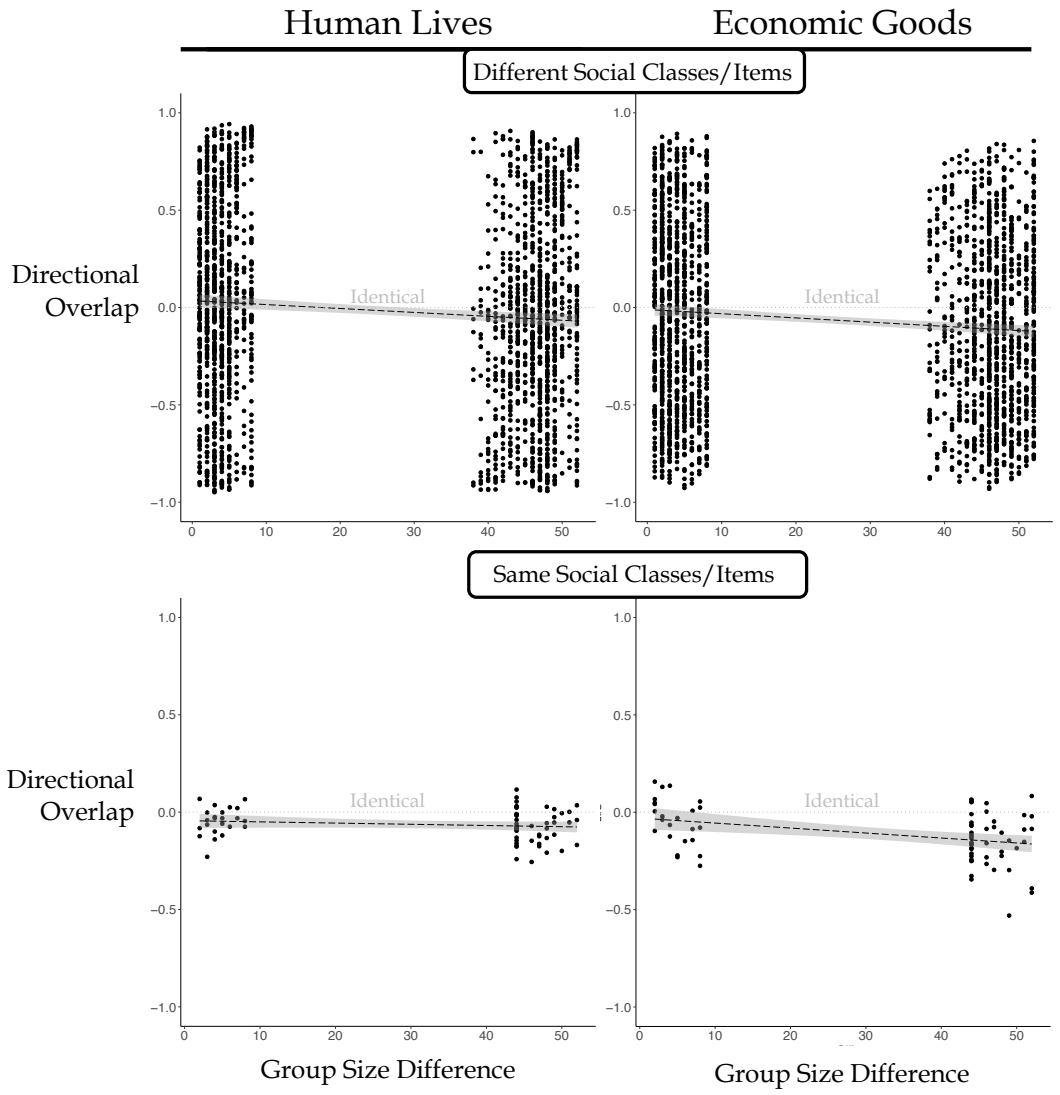

Figure 7. The influence of quantity on the perceived Psychological Value of human lives (left) and economic goods (right). The y-axis displays the value similarity of two groups (Directional Overlap, O_D). The x-axis displays the quantity difference of the two groups being compared. $O_D = 0$ indicates that there is no influence of quantity on perceived Psychological Value; $O_D < 0$ indicates that the larger quantity group has the greater Psychological Value than the smaller group; and $O_D > 0$ indicates that the smaller quantity group has the greater Psychological Value than the larger group. The farther the O_D is from 0, the larger the influence of quantity. The top row shows comparisons of people/economic goods regardless of individual differences between the groups (e.g., a nun vs 4 judges). The bottom row shows comparisons only of the same individual/economic good (e.g., a nun vs 4 nuns). In all cases, the influence of quantity on the perceived Psychological Value is minimal.

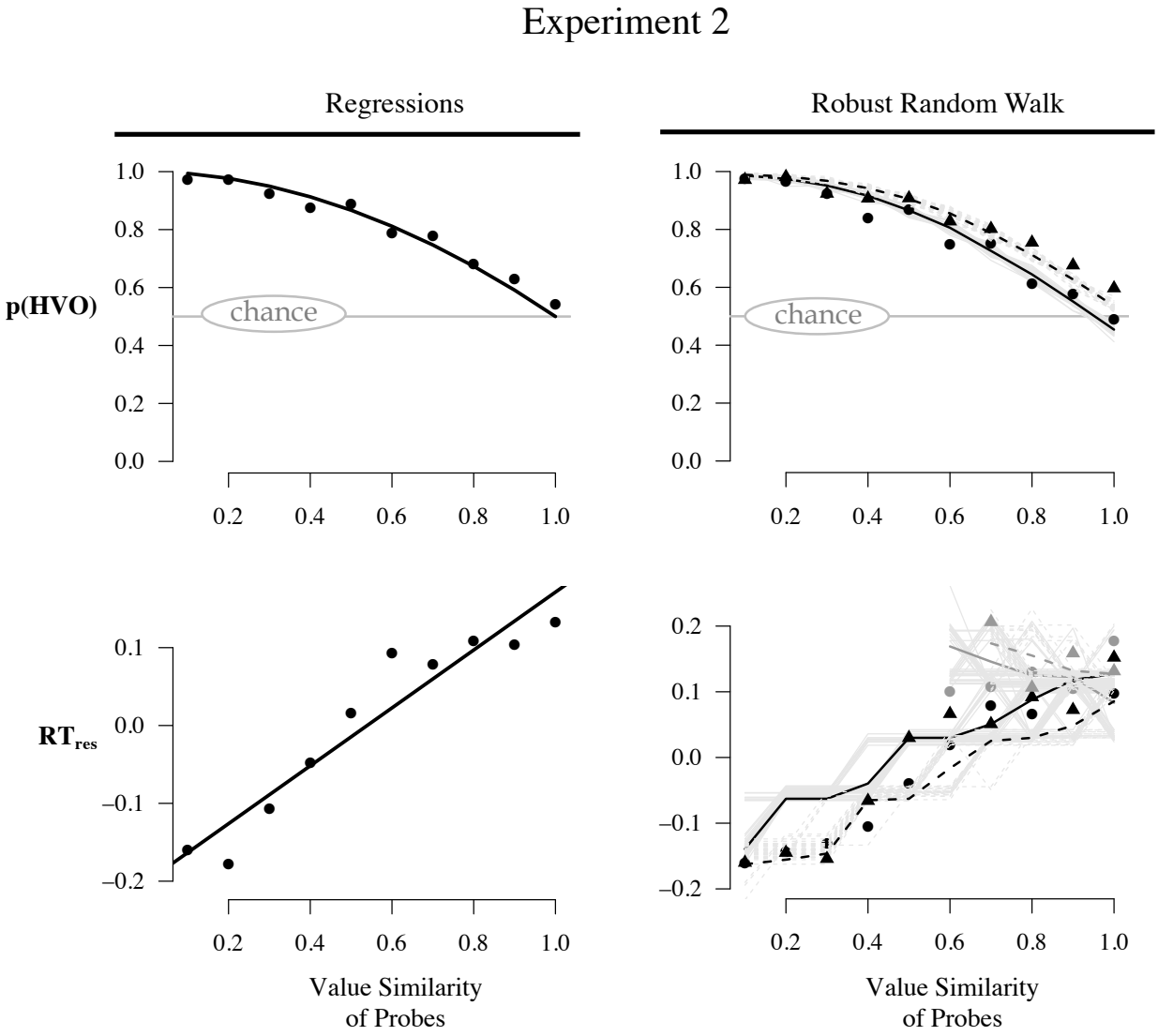

Figure 8. The data and model fits for Experiment 2. In the left column, the probability of choosing the higher valued option, termed p(HVO), and RT are fit with separate regression functions. The dots represent the data, the solid lines represent the best fit function. In the right column, p(HVO), the RT to save the HVO (black), and the RT to save the LVO (lower valued option, medium gray) are all fit simultaneously with the RRW. The dots represent the data. The thin, light gray lines represent 40 individual runs of the RRW with the best fit parameters. They provide a visualization of the variations that result from the stochasticity built into sequential sampling models. The thicker black and medium gray lines represent the average predicted fit. Value similarity of the probe, $f(\Psi v_1 \cap \Psi v_2)$, was directly measured in Experiment 1. The circles and solid lines are the data and predicted fit for trials when the HVO was killed by default, respectively. The triangles and the dashed lines are the data and predicted fit for trials when the LVO was killed by default, respectively. The data reveal a choice response bias that is modeled by a shift toward the "no action" boundary.

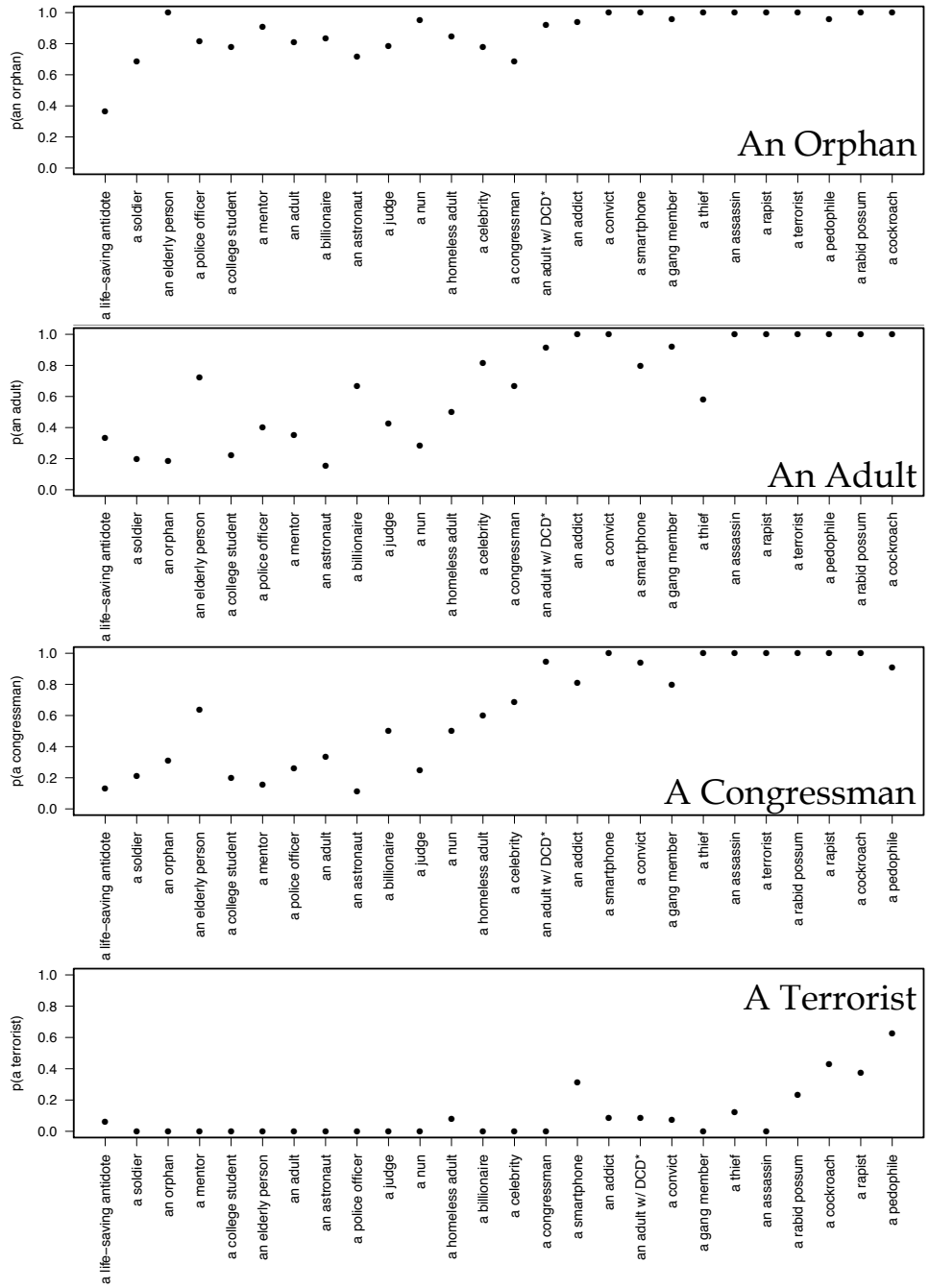

Figure 9. The probability of choosing an individual probe (shown in the left side of each graph) relative to the other comparison probes for Experiment 2. For each plot, the order of probes on the x-axis is determined separately depending on O_D. The label "an adult w/ DCD*" represents the probe "an adult with a deadly contagious disease." See text for a full explanation.

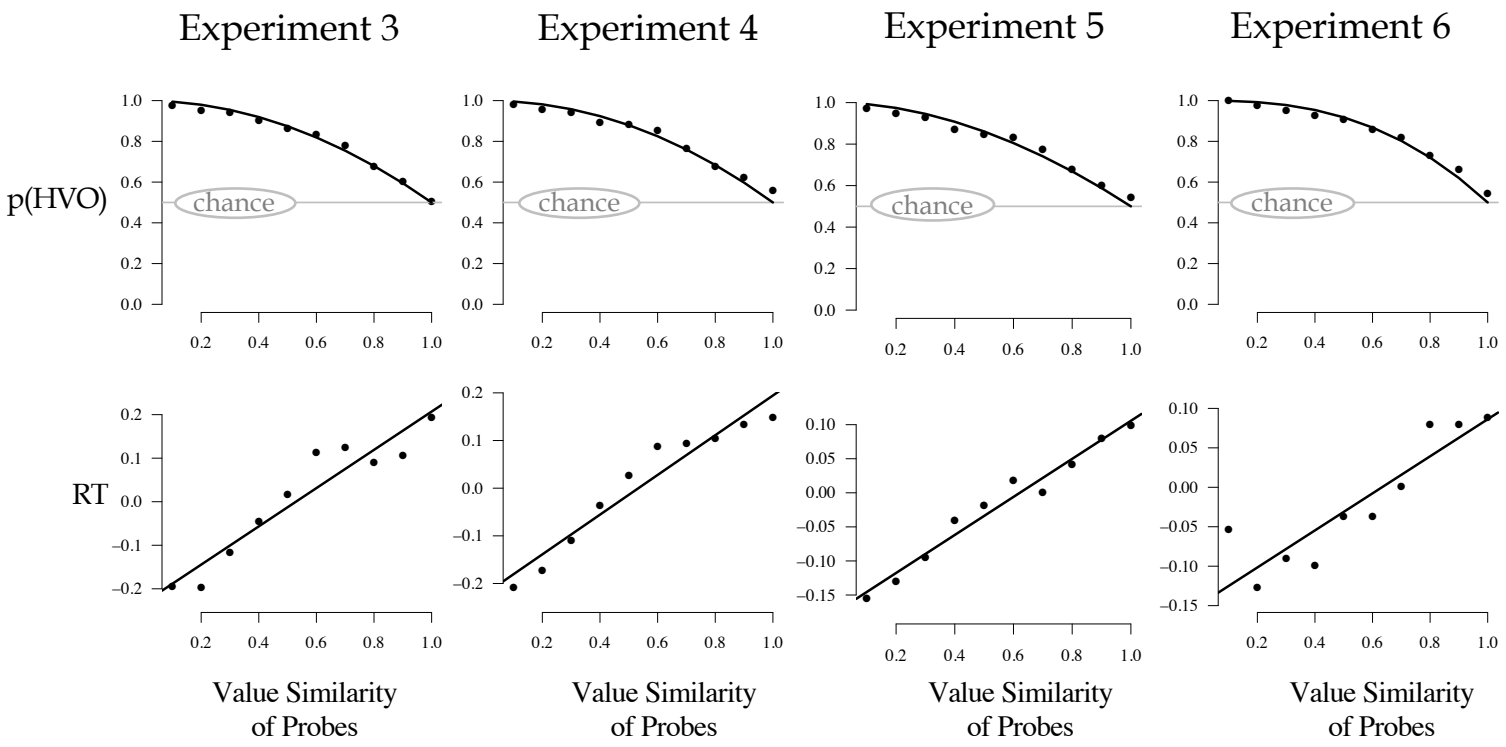

Figure 10. The data and regression fits for Experiment 3-6. In the top row shows p(HVO) and the bottom row shows RT. Each are fit with separate regression functions. The dots represent the data, the solid lines represent the best fit function. For each of the four experiments, the probability of choosing the higher valued option is almost perfect when the values of the options are most dissimilar and decreases monotonically until it reaches chance when the values of the options are equal to one another. Similarly, the time to make the decision increases as the similarity in the values of the two options increases.

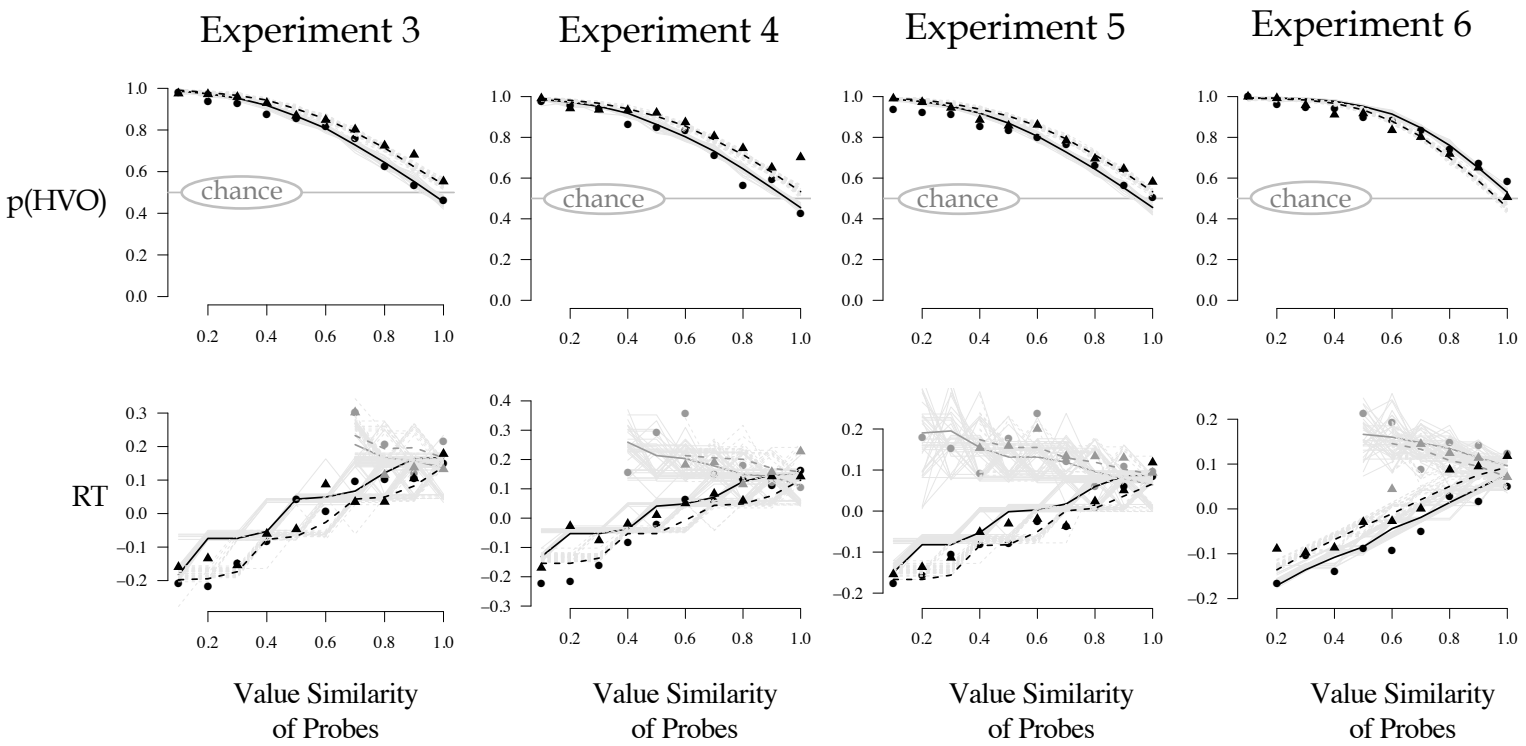

Figure 11. The data and the RRW model fits for Experiment 3-6. p(HVO), the RT to save the HVO (black), and the RT to save the LVO (medium gray), are all fit simultaneously with the RRW. The RRW models a simple value-based decision process whereby people attempt to save the higher valued probe. The RRW used the Psychological Values collected in Experiment 1 as input for a sequential sampling procedure. See the caption of Figure 8 for details. The parameter values of the RRW in Experiment 3-5 (people) were fixed to the best fit values of Experiment 2 and revealed a bias toward the "no action" boundary. In contrast, the participants in Experiment 6 (economic goods) revealed a bias toward the "action" boundary. The data fit extremely well regardless of whether the scenario was a sacrificial moral (Experiments 3-5) or economic dilemma (Experiment 6) and regardless of the complexity of the scenario (Experiment 2 vs 3-6) and the number of people being saved vs sacrificed (Experiments 2-4 vs 5).

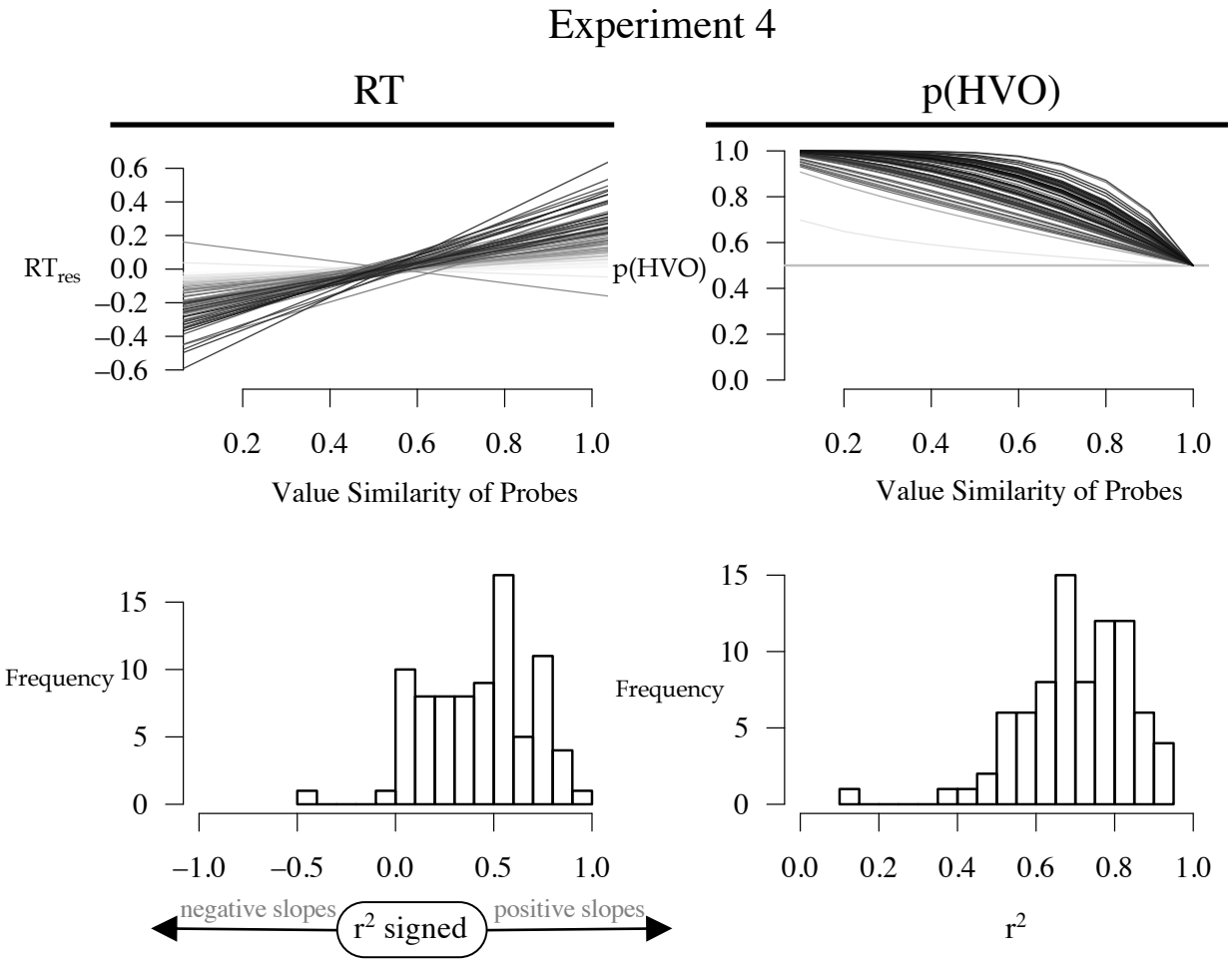

Figure 12: The data of individual participants in Experiments 4 as predicted by Psychological Value Theory. The graphs on the top present each participant's best fit function for RT (left) and p(HVO) (right). The shade of grey of the line indicates r^2 value whereas darker grey corresponds to higher r^2. The graphs on the bottom present the r^2 of each participants' best fit function for RT (left) and p(HVO) (right). The r^2 for RT is signed whereby the r^2 is assigned the same sign as the slope.

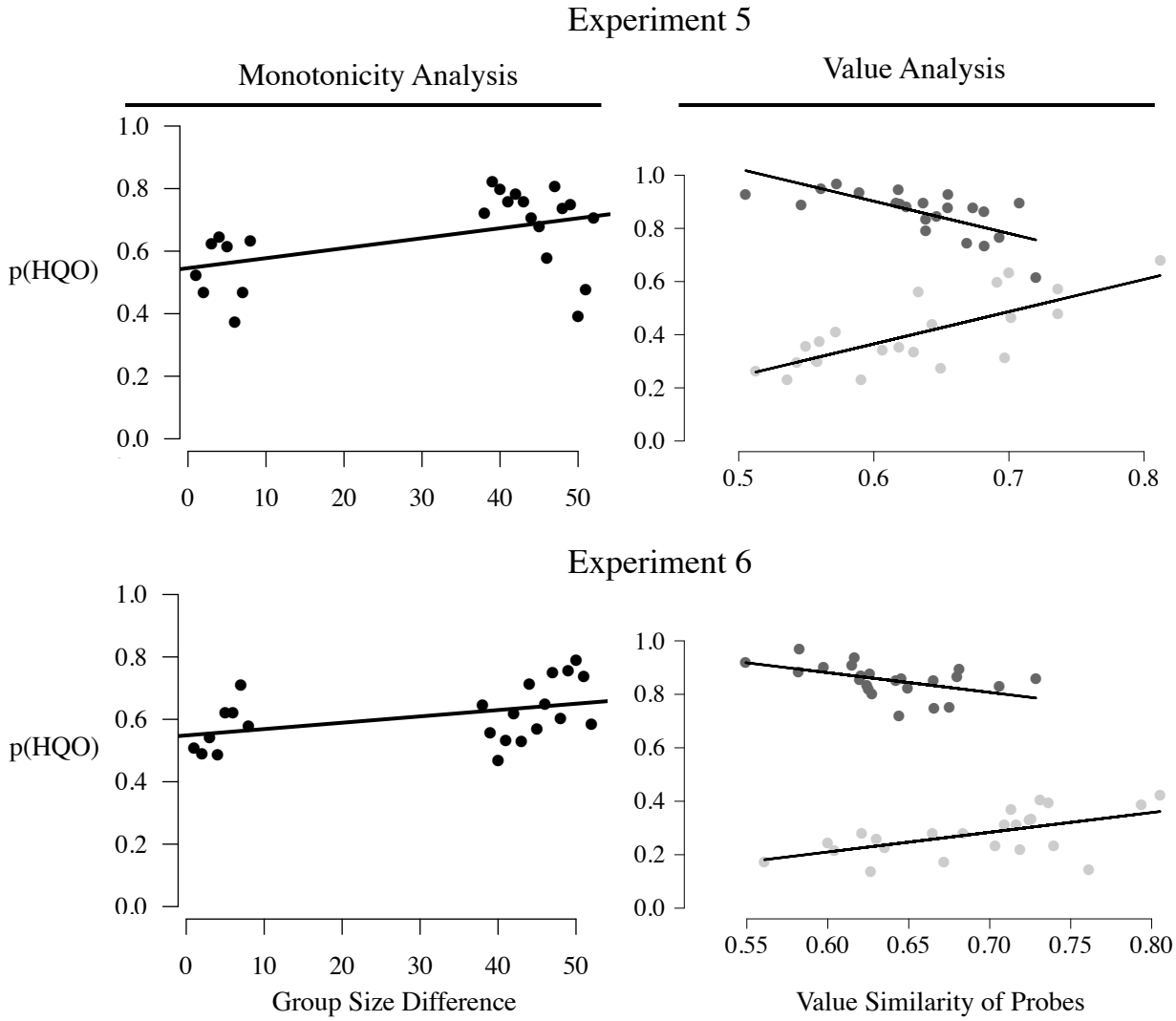

Figure 13: The quantity and values-based analyses for Experiments 5 (top) and 6 (bottom). The quantity analysis is in the left column. It shows the probability of choosing the to save the group with the largest number of people, p(HQO). The values-based analysis is in the right column. It shows p(HQO) as a function of the value similarity of the probes, $f(\Psi v_1 \cap \Psi v_2)$. Inconsistent trials are in light grey and Consistent trials are in dark grey (see text for details). The black lines fit with a single equation (Equation 4) are the predictions of Psychological Value Theory.

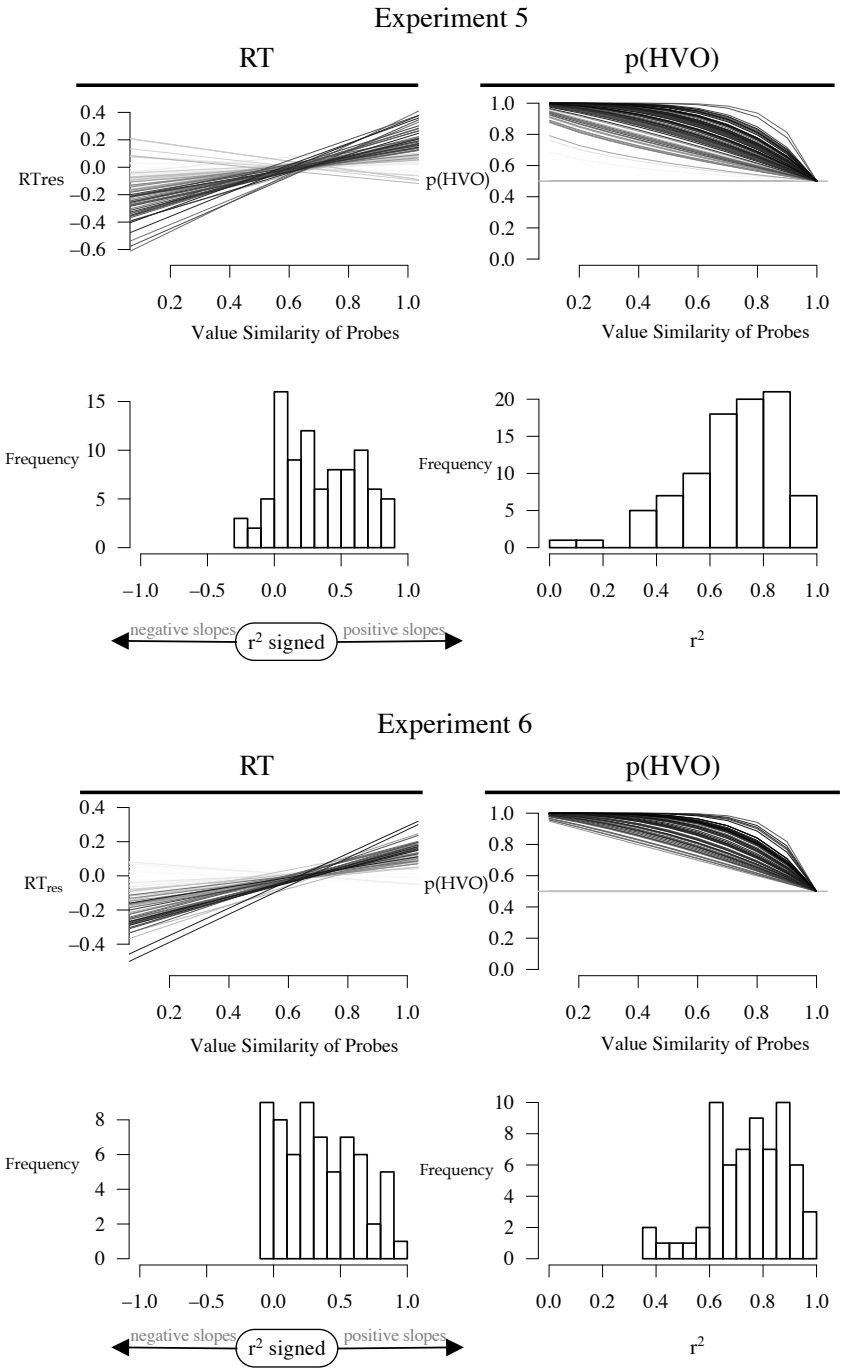

Figure 14: The data of individual participants in Experiments 5 and 6 as predicted by Psychological Value Theory. The graphs on the top present each participant's best fit function for RT (left) and p(HVO) (right). The shade of grey of the line indicates r^2 value whereby darker grey corresponds to higher r^2. The graphs on the bottom present the r^2 of each participants' best fit function for RT (left) and p(HVO) (right). The r^2 for RT is signed whereby the r^2 is assigned the same sign as the slope.

Appendix A

The Robust Random Walk

Here, we introduce a Robust Random Walk (RRW) to model the parameters of Psychological Value Theory. The RRW is different from traditional VSSPs in two ways. First, the RRW does not estimate drift rate. Within the context of Psychological Value Theory, Psychological Value is measured. We do not assume that the measurements provide data that inherently contains a true zero and equal intervals. Rather, we use the measurements to derive a non-parametric measure of $f(\Psi_{V1} \cap \Psi_{V2})$. Specifically, we use a bootstrap procedure, whereby we calculate the probability that a sampled value from Ψ_{V1} is greater than a sampled value from Ψ_{V2}, $p(\Psi_{V1} > \Psi_{V2})$. We convert this probability into a measure of *overlap*, such that

$$Overlap = 1 - \left(\frac{abs(p(\psi_{V1} - \psi_{V2}) - 0.5)}{0.5}\right) \quad \text{(A5)}$$

Overlap ranges from 0 to 1. Overlap = 0 indicates that $p(\Psi_{V1} > \Psi_{V2}) = 0$ or 1, and Overlap = 1 indicates that $p(\Psi_{V1} > \Psi_{V2}) = 0.5$. Furthermore, when Overlap < 1, we note which item has the higher $p(\Psi_{V1} > \Psi_{V2})$.

The second way that RRW differs from traditional VSSPs is that it does not assume the shape of the distributions driving the random walk. The most common VSSPs assume that the latent psychological distributions driving decision choice have a specific distributional shape, such as Gaussian (e.g., Ratcliff, 1978). The RRW, in contrast, makes no such assumptions. Rather than estimate drift rate (i.e., Overlap) as a free parameter, the RRW takes Overlap as a predictor variable. Because Overlap is equivalent to the $p(\Psi_{V1} > \Psi_{V2})$, the RRW uses $p(\Psi_{V1} > \Psi_{V2})$ to drive the accumulation of evidence. Specifically, for each step, V_i, of the RRW equals:

$$V_i = \left(\left(\begin{cases} 1, & if\ p(\psi_{V1} > \psi_{V2}) \\ -1, & if\ p(\psi_{V1} < \psi_{V2}) \\ 0, & otherwise \end{cases} \right) + e_i \right) * IAB \quad \text{(A6)}$$

where

$$e_i \sim N(0, \sigma) \quad \text{(A7)}$$

and

$$IAB = (1 - d_A) * \exp(d_B * i) + d_A \quad \text{(A8)}$$

Briefly, with each sample of the random walk, V_i is assigned a 1 with the probability of $p(\Psi_{V1} > \Psi_{V2})$ and a -1 with the probability $p(\Psi_{V1} < \Psi_{V2})$. These probabilities are taken directly from the estimate of perceived Psychological Value (or any other construct that one measures). To that, we add noise sampled from a normal distribution with a mean of 0 and a standard deviation of σ. The noise distribution adds a parametric component because its' shape is assumed to be Gaussian. Gaussian noise is generally an uncontroversial assumption. However, if one wishes to remove the noise, simply set $\sigma = 0$. We generally fix $\sigma = 1$, but σ can be a free parameter in the model. *IAB* stands for *Information Accrual Bias*. The IAB is an exponential function that weights the influence of the sample, V_i, by the time since the start of the process (see for other weighting formulas, see Ashby, Jekel, Dickert, & Glöckner, 2016). If d_B is negative, then the early samples have more influence on the random walk than the later samples. This corresponds to a primacy of information bias. Conversely, if d_B is positive, then the later samples have more influence on the random walk than the early samples. This corresponds to a recency of information bias. The d_A and d_b parameters work in unison to describe the shape and strength of this bias. If either $d_B = 0$ or $d_A = 1$, then IAB = 1, so there is not information accrual bias. Therefore, fixing either $d_B = 0$ or $d_A = 1$ removes IAB from the model and neither d_B nor d_A will be free parameters in the model. If one wishes to include IAB as part of the model, then

d_B and/or d_A may be free parameters in the model. In general, as d_B moves away from 0 (positive or negative), the influence of IAB increases. When $d_B < 0$, d_A acts as a boundary such that IAB will not decrease below its' value. When $d_B > 0$, d_A acts as a moderator such that as d_A approaches 1, it will reduce the influence of IAB. Both d_A and d_B can be free parameters. We often fix $d_A = 0.2$ and make d_B a free parameter.

The IAB is a theoretically informative mechanism with interesting predictive properties. For example, assume that one holds d_A constant such that $d_A = 0.2$. Here, d_B is the single parameter that drives the IAB. When the d_B is positive, it indicates that perceptually recent information carries more decisional weight than perceptually distant information. Theoretically, such a finding would be consistent with Jones, Love, and Maddox (2006) who identified the presence of both perceptual and decisional recency effects in a categorization task. Although their findings applied to successive trials, we suspect that the perceptual recency effects apply to evidence accumulation within a trial as well. When d_B is negative, it reduces the amount of evidence that can accrue over time. Therefore, a negative d_B will reduce the influence of boundary separation on accuracy for a given overlap. This is theoretically interesting because standard VSSPs will increase accuracy with increasing boundary separation ad infinitum (see Link & Heath, 1975 for a detailed explanation).

From a predictive standpoint, the IAB will decouple correct RTs from incorrect RTs. Correct and incorrect RTs are typically symmetric in VSSPs. To model asymmetric correct and incorrect RTs, researchers have introduced two free parameters: variance in the start point and variance in the drift rate (Ratcliff & Rouder, 1998). The IAB naturally produces slow and fast errors under a variety of conditions. For example, slow errors can result from positive d_B and a

larger boundary separation whereas fast errors can result from a negative d_B and a smaller boundary separation.

In addition to the free parameters described above, the RRW includes the following free parameters:

- *Boundary* distance from a 0 start point, b. Within the RRW, the distance between boundaries is symmetric around a start point of 0. Within the RRW, the positive boundary corresponds to the correct response to the higher valued option whereas the negative boundary corresponds to the incorrect response to the higher valued option. Furthermore, when estimating the probability distribution for crossing the boundaries, for some sets of parameters, the accumulated evidence will stubbornly not cross a boundary even with a very large number of samples. In this instance, if the probability of crossing both of the two boundaries remains stable for a consecutive 10% of sample values (when estimating the probability for crossing the boundaries for n samples), then the boundary distance is reduced by 0.1. Collapsing boundaries are not uncommon in VSSPs (e.g., Bhui, 2019; Hawkins, Forstmann, Wagenmakers, Ratcliff, & Brown, 2015) and are implemented here because the RRW is significantly more constrained than traditional VSSPs. In the RRW, the latent distributional overlap is directly input rather than estimated. Therefore, it is possible for an experimental condition to exist whereby the overlap = 1.0 (complete overlap of the distributions). When this is the case, the VSSP can drift for extended time. When presented with such a problem, humans likely give up and make a decision based on the time constraints (see Bhui, 2019). This corresponds to a collapsing boundary. Therefore, we implemented the collapsing boundaries for exactly this case.

- *Start point* as a signed proportion of the boundary distance from 0, s. If $s > 0$, it represents a bias towards the positive boundary with the starting position equaling s*b. If $s < 0$, it represents a similar bias towards the negative boundary. s must be between -1 and 1. Of course, if $s = 1$ or -1, then the boundary is crossed before the first sample is drawn. Therefore, in practice, $-1 < s < 1$.
- Finally, we include the non-decision time, T_{er}, to represent what is generally measured as the intercept of the function. It is hypothesized to include all ancillary processes that are not associated with the variable of interest.

The RRW also has a scaling parameter that scales the number of steps in the RRW into RT. We do not count this scaling parameter as a free parameter because it is a simple linear transformation of one relatively arbitrary scale into another.

The RRW is designed to take advantage of Psychological Value Theory. As such, the RRW does not fit the shape of a distribution for a single condition. Rather, the RRW fits the mean (or median) RTs and proportion of trials that cross each boundary simultaneously for a range of values of distributional overlaps. Once one understands how perceived Psychological Value influences preferential choice across the entire range of distributional overlaps, one can assess how variables of interest change that pattern. These changes can be modeled within the RRW simply by effect coding the conditions and specifying whether the effect codes influence any subset of the parameters such as the start point (i.e., a choice response bias).

Appendix B

Why currency is a poor measure of Psychological Value

Although there is an influence of quantity on perceived Psychological Value of economic goods, that influence is quite small (e.g., Figure 7). There is one exception: items related to currency (dollars, pennies, etc.). Experiments 1 and 6 revealed quantity had a relatively large influence of the perceived Psychological Value of pennies and dollars. As such, there appears to be a "currency" exception to our finding that the Axiom of Monotonicity is a weak effect. This suggests that currency is a unique item type. Currency is unique because it can be (a) easily converted into other required items, and (b) transported, protected, and stored at a very low cost (e.g., banks do it for a negative cost).

Economists often use currency as a surrogate for value because currency (a) is assumed to adhere to the Axiom of Monotonicity, (b) appears to equate a variety of items on a value scale, and (c) provides a mechanism for people to demonstrate preferential choice. Despite currency adhering to the Axiom of Monotonicity, we believe that currency is a sub-optimal method for measuring Psychological Value.

Because currency, ¤, does not have any inherent value, it derives its' Psychological Value indirectly from the relation between (a) the prices of items in the environment and (b) the perceived Psychological Value of those same items. Because the prices of items are set by society (e.g., the cost of goods, manufacturing, demand, etc.) rather than by the observer, there likely exist cases whereby the perceived Psychological Value and monetary value are non-transitive. More formally, the monetary value of items A and B, I_A and I_B, are set by society (e.g., the cost of goods, manufacturing, demand, etc.) rather than being directly set by the

observer. Therefore, the association between items and their monetary values are learned by the observer:

$$I_A = ¤_A \quad \text{(B1)}$$

$$I_B = ¤_B \quad \text{(B2)}$$

We assume there is a function, f_I, that converts each item, I_A and I_B, into the Psychological Value of the Item, Ψ_{VI}:

$$\Psi_{VIA} = f_{IA}(I_A) \quad \text{(B3)}$$

$$\Psi_{VIB} = f_{IB}(I_B) \quad \text{(B4)}$$

We also assume that, initially, the monetary value of each item inherits the Psychological Value of that item. Therefore, we apply the same conversion function to Monetary Value:

$$\Psi_{VIA} = f_{IA}(¤_A) \quad \text{(B5)}$$

$$\Psi_{VIB} = f_{IB}(¤_B) \quad \text{(B6)}$$

This process occurs for some subset of the items in society for sale (i.e., goods and services) that the observer encounters. We assume the conversion function for each item is different:

$$f_{IA} \neq f_{IB} \quad \text{(B7)}$$

As such, there likely exist cases where perceived Psychological Value and monetary value are non-transitive. That is, Item A may have a greater Psychological Value than Item B, but also have a smaller Monetary Value than Item B:

$$\exists: \Psi_{VIA} > \Psi_{VIB} \;\&\; ¤_A < ¤_B \quad \text{(B8)}$$

These non-transitivities reveal that monetary value is not a monotone transformation of Psychological Value. Because monetary value is not a monotone transformation of Psychological Value, monetary value is a sub-optimal measure of Psychological Value.

Appendix C

Auxiliary Properties of Psychological Value

The auxiliary properties described below are derived from the defining property that Psychological Value is a *perception* rather than a *sensation* or a *conception*. In a very loose sense, one can view sensation, perception, and conception on a continuum of the immediacy of one's knowledge of the physical environment. On one end of the continuum is sensation which is the activation of receptor cells by the energy in the environment. On the other end of the continuum is *conception* which is abstract knowledge of the world that is derived from the integration of a collection of facts or ideas (often in working memory). Perception, in contrast, falls within the continuum. Perception is one's conscious awareness of the world (for a review, see Coren, 2012). It includes the qualities of sensation as well as the associated attributes of the stimulus including some influence of one's knowledge (e.g., spatial placement, quantity, a sense of "mine," etc.; see e.g., Helmholtz, 1924; Rock, 1997). For example, the perception of a photograph of one's mother consists of the immediate sensory stimulation, emotional arousal, attraction/aversion, etc. We propose that Psychological Value is a fundamental attribute of this perceptual process. As such, it should not require higher thought processes and should be available to non-human animals. Like the perception of color, Psychological Value cannot exist without a biological being to perceive it.

Similar to other perceptual events, we propose that Psychological Value is perceived *automatically, efficiently, and pre-attentively*. In more precise terms, we are proposing that Psychological Values are processed in an unlimited (or practically so) capacity system. As such, the primary limits to perceiving Psychological Value will be the perceptual encoding of the stimuli rather than recovery of the associated value information. Because we propose that

Psychological Value is perceived, we do not believe that the active manipulation of semantic information (e.g., integrating information in working memory) is necessary to derive value. Indeed, we suspect that the active manipulation of semantic information in working memory may produce biased estimates of Psychological Value. *As such, we predict that a cognitive load task that occupies the phonological loop of working memory should have minimal influence on the perception of Psychological Value*.

Perceptual events are sensed from the perspective of the perceiver, termed *self-view*. Therefore, Psychological Value Theory predicts that people will perceive Psychological Values with respect to their wants and needs. If this is the case, two observers who have different wants and needs with respect to a stimulus will have different perceived Psychological Values of that stimulus. *One consequence of self-view is that the observer has immediate information about the stimulus relative to themselves but has to infer the stimulus information as perceived by others* (presumably a slow, effortful process; Rock, Wheeler, & Tudor, 1989). The perceptual nature of Psychological Value—and thus its' inherent self-view—is a model-based explanation of the proposal that "utils" are inherent to the decision-maker (i.e, subjective expected utility theory).

The perceptual system generally does not discriminate between objects and people. Although there is evidence that special facial recognition systems exist, the perceptual system does not discriminate between people and objects when identifying positions in space, size of the objects, trajectory of motion, etc. *Similarly, we propose that the perceived Psychological Values of people and objects are processed in the same system, in the same way*. Our data provide some evidence for this proposal. Participants estimated the perceived Psychological Values of objects and people in Experiment 1. These estimates predicted preferential choices in Experiments 2-6—even when objects and people were pitted against one another. Thus, these data suggest that

objects and people are valued on the same scale and because these values predict choice, this scale is not arbitrary. Of course, it is possible that the similarities between people and objects are simply coincidence. This explanation, however, is not parsimonious and low probability. We hope to further explore this line of research in the future.

Perceptual events tend to be relatively consistent between individuals. For example, people tend to perceive changes in brightness (e.g., Jameson, & Hurvich, 1961) and loudness similarly (e.g., Stevens, 1955). Although color perception (e.g., Özgen, 2004) and facial attractiveness (e.g., Fink & Penton-Voak, 2002) appear to be influenced by culture, these influences are small relative to the amount of variance explained by the population (e.g., Cunningham, Roberts, Barbee, Druen, & Wu, 1995; Özgen, 2004). *Similarly, we propose the perceived Psychological Value will be relatively consistent between individuals*. Importantly, this does not mean that you will value my child as much as I value my child. It means that people tend to value their children more than other people's children. Interindividual consistency is likely restricted to basic level categories rather than specific exemplars. There are important examples of large individual differences in perception. These examples, however, are generally present in multi-stable stimuli such as the Necker Cube and the Vase-Face stimuli (e.g., Attneave, 1971; Leopold, & Logothetis, 1999; Girgus, Rock, & Egatz, 1977). *If Psychological Values are perceived, it is likely that researchers will identify items with multi-stable perceptions of Psychological Value*. That is, a subset of the population would perceive the item as having high Psychological Value whereas a subject of the population would perceive the item as having a low Psychological Value. Such examples are likely apparent in politics.

Footnotes

[1] Because Utility Theory infers value from choice, one cannot gain understanding of one (i.e., choice) from the other (i.e., value; Popper, 1968).

[2] Because we propose that Psychological Value is a perception, we frame it within the context of decision/perception science (e.g., Signal Detection Theory, General Recognition Theory, Sequential Sampling) rather than economic theory (e.g., Utility Theory and Multi-Attribute Utility Theory).

[3] Our purpose here is not to debate the merits of the RRW relative to the numerous proposed VSSPs (for a review, see Busemeyer, et. al, 2019). We developed the RRW to instantiate the assumptions of Psychological Value Theory, not to compete with other VSSPs.

[4] Values calculated based on formulaic algorithms (e.g., 5 * the value of a single person) likely arise from a semantic system, rather than a system that activates quantity information (Lemer, Dehaene, Spelke, & Cohen, 2003).

[5]The pattern of these results did not change when quantity was log transformed or when social status/item was treated as a random variable in a mixed effects model. Our finding that quantity has minimal influence on estimates of perceived Psychological Value for human lives may strike some as counterintuitive. Therefore, we calculated a Bayes factor analysis that tested the relative likelihoods of two models: The Small Effect model ($-0.2 < D_O < 0$) vs the Larger Effect model ($-1 < DO < -0.2$). We tested both against an alternative hypothesis at the boundary condition, H_a $D_O = -0.2$ (rscale = "ultrawide"). The results revealed decisive evidence in favor of the Small Effect model for both human lives, $\log_{10}(BF) = 67.6$, and economic goods, $\log_{10}(BF) = 65.2$. Thus, our data provide extremely strong evidence against a substantial effect of quantity on Psychological Value.

[6] Again, the patterns of these results did not change when quantity was log transformed or when social status/item was treated as a random variable in a mixed effects model. Again, the Bayes factor analysis that tested the relative likelihoods of the Small Effect model ($-0.2 < D_O < 0$) vs the Larger Effect model ($-1 < DO < -0.2$) revealed decisive evidence in favor of the Small Effect model for both human lives, $\log_{10}(BF) = 3.09$, and economic goods, $\log_{10}(BF) = 3.55$. Thus, our data provide extremely strong evidence against a substantial effect of quantity on Psychological Value.

[7] Several of these participants spontaneously reported confusion regarding the response keys after completing the experiment. This hypothesis is supported by the data: the average p(HVO) of these 12 participants was 0.29, suggesting a reversal of the response keys.

[8] We rounded to the nearest 0.1 (rather than another value) to (a) be consistent with Cohen and Ahn (2016), and (b) to provide enough data points per bin in the "error" choices (i.e., save lower value item) for analysis while simultaneously providing enough bins to produce a clear function. We found that 0.1 worked well.

[9] Because RT data has more data points than p(HVO), without averaging, the fit would be weighted toward the RT fit by the number of datapoints.

[10] The change in the dialogue box appears to have mitigated this problem. There are about half as many discarded participants for this criterion as in Experiment 2.

[11] The fit did not meaningfully improve when log(group size difference) was used as the predictor variable.

[12] A linear regression was also significant, $F(1, 4) = 46.28$, $p = .002$, $r^2 = .92$.

[13] The fit did not meaningfully improve when log(group size difference) was used as the predictor variable.

[14] A linear regression was also significant, $F(1, 5) = 31.77$, $p = .002$, $r^2 = .86$.