Subjective Values Theory:

The Psychophysics of Psychological Value

Amanda R. Cromley & Dale J. Cohen

The University of North Carolina Wilmington

Word Count: 10,118

Author Note

Amanda R. Cromley & Dale J. Cohen, Department of Psychology, University of North Carolina Wilmington. We 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.

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

People perceive the Psychological Value of all stimuli. Cohen and colleagues (Under Review; 2016) have measured perceived Psychological Values of a variety of stimuli and demonstrated that those measurements predict participants' RTs and response choices in preferential choice tasks. Here, we examine the psychophysical properties of perceived Psychological Value. The current work examines how the perceived Psychological Value of a group of items changes as a function of (a) the number of items, and (b) the perceived Psychological Values of the individual items in the group. If more is better, as the Axiom of Monotonicity assumes, then the perceived Psychological Value of the group should be a function of the sum of the perceived Psychological Values of the individual items in the group. Ensemble stimuli, in contrast, are generally averaged (termed perceptual averaging). If Psychological Value is perceived similar to other perceptual dimensions, then the perceived Psychological Value of the group should be a function of the average of the perceived Psychological Values of the individual items in the group. Using a magnitude estimation procedure, we collected perceived Psychological Values of individual items and groups of items. Results indicate that perceived Psychological Value of groups is well predicted by a function of the average of the perceived Psychological Values of the individual items in the group (i.e., perceptual averaging rather than summing), with some influence of the maximum valued item (i.e., attentional capture of high valued stimuli).

Keywords: Value, Preferential Choice, Psychophysical scaling; Subjective Values Theory, Ensemble representations, Attentional capture

Subjective Values Theory: The Psychophysics of Psychological Value

Recently, Cohen and colleagues (2016; Under Review) hypothesized that there is an underlying perceptual continuum which drives decision making behavior. The construct, termed *Psychological Value*, refers to the importance or worth or any stimulus to the observer. Across a series of experiments and presentation types, Cohen and colleagues have demonstrated that Psychological Value predicts preferential choice behavior (2016; Under Review). Until now, the psychophysical properties of Psychological Value have gone unstudied. Here, we use classic psychophysical methodology to empirically study how Psychological Value changes as a function of (a) the number of items in a group, and (b) the perceived Psychological Values of the individual items in the group.

What is Value?

Decision-making is ubiquitous, and most choices that humans make are preferential choices. A preferential choice refers to any decision in which one must choose between multiple alternatives based on one's subjective impression of each alternative. For example, choosing what to eat for a meal or what city to live in are preferential choices. Most decision scientists agree that "value" drives preferential choices (see Stigler, 1950 for review; or more recently, Simonson & Tversky, 1992; Rangel, Camerer, & Montague, 2008; Padoa-Schioppa & Assad, 2006; Kable & Glimcher, 2009; Rangel & Clithero, 2014; Colas, 2017).

Traditionally, value has been used as a synonym for preference (e.g., Kable & Glimcher, 2007; Colas, 2017), utility (e.g., Katahira, Yuki, & Okanoya, 2017; Simonson & Tversky, 1992), or likability (e.g., Colas & Lu, 2017). For example, Utility Theory explicitly equates value and preferential choice by assuming that an option in a multi-alternative choice context is preferable to another if, and only if, it has a higher value. In other words, decision makers choose their preferred option, and therefore, that option is more highly valued than any other available option. Because value is directly inferred from preferential choice, value and preferential choice are equated tautologically. Because this tautology is a critical assumption of Utility Theory, the validity of assumption cannot be assessed within the context of Utility Theory. Nevertheless, Utility theorists' significant success modeling preferential choice behavior suggests that there is some objective truth embedded in Utility Theory's conceptualization of value (e.g., Azari, Parks, & Xia, 2012; Bundorf, Mata, Schoenbaum, & Bhattacharaya, 2013; Hey & Orme, 1994).

Occasionally, researchers claim to measure the construct of value and use those measurements to predict preferential choice. Here, value is most commonly measured along a single dimension such as monetary value, likability, or desirability (e.g., Chib, Rangel, Shimojo & O'Doherty, 2009; Colas, 2017; Colas & Lu, 2017; Grueschow, Polania, Hare & Ruff, 2015; Gwinn & Krajbich, 2016; Gwinn, Leber, & Krajbich, 2018; Krajbich, Armel & Rangel, 2010; Lebreton, Jorge, Michel, Thirion & Pessiglione, 2009; Lim, O'Doherty, & Rangel, 2011; Mormann, Malmaud, Huth, Koch & Rangel, 2010; Smith & Krajbich, 2018). For example, in one prototypical experiment, Krajbich and colleagues (2010) asked participants to report how much they would like to eat a specific snack food following the experiment on a scale from -10 to 10. Subsequently, the same participants completed a decision-making task while being eye-tracked. For each participant, trials were generated by randomly selecting two snack foods the participant rated positively and asking the participant to choose which of the snack foods he or she would rather have.

Although the researchers claim to measure a general construct of "value," these stimulus specific preference ratings are theoretically distant from the hypothesized theoretical construct of "value." That is, if a stimulus independent preferential value construct exists, and this construct drives preferential choice, then any valid scale that measures this construct should be both (1) stimulus independent and (2) predict preferential choice. The stimulus specific rating scales described above fail the first criterion of such a scale. For example, one cannot predict preferential choices between snack foods using pleasantness ratings of faces. If researchers successfully measured value—the construct believed to drive preferential choice—the measurements should allow them to make predictions about preferential choice regardless of stimuli or context. Although researchers have identified stimulus specific features that predict preferential choice within a specific stimulus set to some degree (e.g., snack food; Armel, Beaumel, & Rangel, 2008; Colas, 2017; Colas & Lu, 2017; Gwinn, Leber, & Krajbich, 2019; Smith & Krajbich, 2018; Krajbich et al., 2010; Lim et al., 2011; Mormann et al., 2010; Mormann, Bausch, Knieling, & Fried, 2017; Polania, Woodford, & Ruff, 2019) we suggest that they are not measuring a stimulus independent preferential value construct that allows them to flexibly predict preferential choice across stimulus sets.

Neuroscientists have compiled a large body of literature supporting the existence of a stimulus independent, preferential value construct. For example, researchers have examined the neural correlates of preferential value during simple decision-making tasks (e.g., Bartra, McGuire, & Kable, 2013; Grueschow et al., 2015; Kable & Glimcher, 2007; Lebreton et al., 2009; Padoa-Schioppa & Assad, 2006), neural representations of the preferential values of rewards (e.g., Peters & Büchel, 2010), the relationship(s) between visual gaze, preferential value, and decision-making (e.g., Colas & Lu, 2017; Krajbich et al., 2010; Krajbich & Rangel, 2011; Lim et al., 2011; Smith & Krajbich, 2018), and the physical location of preferential value computation in the brain (e.g., Chib et al., 2009; Kable & Glimcher, 2007; Rangel et al., 2008). This work generally focuses on how the preferential value of stimuli is encoded and/or represented in the brain by examining the neural activity of participants as they make preferential choices. Although all stimuli are multidimensional (e.g. Ashby & Townsend, 1986), these data seem to suggest that preferential value is best represented as a unidimensional feature of a stimulus and is encoded in the brain as such. Because, this research only addresses the neural regions(s) associated with preferential value, it does not provide information about the features or function of preferential value. Similar to behavioral measures of preferential value, the work is interesting but insufficient to establish how preferential value drives preferential choice.

In contrast to past research, Cohen and colleagues (2016; Under Review) have reconceptualized, measured, and validated a value-based construct which drives preferential choice: termed *Psychological Value*. Below, we describe the theoretical construct of Psychological Value and demonstrate how it drives preferential choice. For clarity, we will refer to past conceptualizations of value as “preferential value” and refer to the continuum identified and validated by Cohen and colleagues (2016; Under Review) as “Psychological Value.”

Psychological Value as a Perception

Psychological Value refers to the “perception of the importance, worth, or usefulness of an item to the observer” (Cohen, Cromley, Freda, & White, Under Review). This definition makes clear that Psychological Value is experienced as a percentual event. To understand how Psychological Value is experienced as a perceptual event, it is important to distinguish between sensation, perception (for a review, see Coren, 2012), and conception. Briefly, sensations refer to the impressions that incoming perceptual information leave on receptor cells found on sense organs. For example, electromagnetic energy (i.e., light) originates in the external environment, and sensation refers to the contact of incoming electromagnetic energy on photoreceptor cells. In contrast, perception refers to one’s conscious awareness of the environment as a direct result of the processing and transformation of sensations by the perceptual system, but does not require deliberate, higher order thought process. The perceptual system, for example, gives rise to percepts of objects and space when one encounters electromagnetic energy. Here, it is important to avoid the incorrect assumption that every perceptual event is directly related to an associated sensation. For example, consider the location of furniture in a room. When one enters a room, one is immediately aware of the physical location of objects. There are no receptor cells on sense organs for physical location of objects in three-dimensional space (sensation), and locating objects in space does not require deliberate reasoning or thought. Thus, the awareness of location of objects in three-dimensional space is a perception. Conception is the result of a deliberate, higher order thought process. For example, the intentional manipulation and integration of information within working memory is required when one considers how to best arrange the furniture in the room for practical and aesthetic purposes.

We conceive of Psychological Value as a perception. Psychological Value is neither a sensation—because there are no receptor cells for Psychological Value—nor a conception—because it is unnecessary to intentionally manipulate and integrate information within working memory to derive Psychological Value. Psychological Value is one feature of a stimulus that is quickly apparent to an observer upon presentation of a stimulus. Just as the perceptual system uses sensations, evolved systems, and prior knowledge to derive and add spatial location to our awareness of objects, we propose that the perceptual system also uses sensations, evolved systems, and prior knowledge to derive and add Psychological Value to our awareness of objects. Like other percepts, Psychological Value should be efficient, automatic, and experienced from the self-view.

By specifying Psychological Value as a perception and identifying features it likely shares with other perceptual continua, we constrain the theoretical construct of Psychological Value in ways it has not been constrained in the past. Below, we quantify Psychological Value.

Specifying Psychological Value. All stimuli are multidimensional (e.g., Ashby & Townsend, 1986), and we propose that one such dimension is Psychological Value. Just as color can be decomposed into multiple features (e.g., intensity, saturation, hue), Psychological Value can also be decomposed into multiple features. Cohen et al. quantified Psychological Value as follows:

$$\Psi_V = \sum_{I=1}^{N} C_I P_I \tag{1}$$

where Ψ_V is the value of an any given thing, P_I are perceptual factors, and C_I are cognitive weights for each perceptual factor (Cohen & Ahn, 2016, pp. 1362). Perceptual factors may include influences of emotional connection, monetary value, religious beliefs, and many others (Cohen & Ahn, 2016). People may cognitively weigh perceptual factors sub optimally (Cohen & Lecci, 2001) and/or based on the context they are being considered in (Cohen & Ahn, 2016). For example, one's Psychological Value of a wedding dress may be influenced by the monetary worth of the garment, the ceremonial symbolism, and one's emotional connection to the marriage that the dress signifies.

We propose that Psychological Value adheres to the Axiom of Perceptual Variability. The Axiom of Perceptual Variability states that there is variability in every percept even when the stimulus and the observer are held constant (Ashby & Lee, 1993). Because Psychological Value is perceptually variable, Psychological Values are best represented as distributions that fall along a continuum (Cohen & Ahn, 2016). We therefore describe perceived Psychological Value as a distribution, $f(\Psi_V)$, rather than a point, Ψ_V. The distribution, $f(\Psi_V)$, represents likelihood of perceiving any specific Ψ_V at any instant in time.

If we are correct in assuming that Psychological Value is a perception, it should be measurable using psychophysical methodology. Cohen and colleagues (2016; Under Review) have measured Psychological Value using a magnitude estimation task. Magnitude estimation is a task in which participants assign numerical values to their perception of their sensory experiences (e.g., Marks & Algom, 1998). The most common magnitude estimation methodology involves presenting a participant with a standard that is assigned a numerical value by the experimenter and subsequently presenting a comparison stimulus (Stevens, 1956). The participant's task is to assign a numerical value to the comparison stimulus in relation to the standard. For example, an experimenter may sound a 50 dB tone and assign it a value of 100. They would then sound a different tone—the comparison stimulus—and have the participant assign it a number in relation to the standard. If the observer perceived the tone as five times as loud as the standard, they would assign it a value of 500. If they perceived it as one fifth the loudness of the standard, the observer would assign it a value of 20. This process is typically repeated with a wide array of comparison stimuli spanning from very low to very high stimulus intensity (e.g. barely audible tones to intense, startling tones).

Consistent with the traditional magnitude estimation procedure, participants assigned numerical values to their Psychological Value of various probes compared to a standard (Cohen & Ahn, 2016; Cohen et al., Under Review). Probes were heterogeneous and included people (e.g. a terrorist, an astronaut, your sibling), animals (e.g. a rabid possum, a frog, a pet dog), objects (e.g. a book, a kayak, a car), among others. The standard of "a chimpanzee" with a consistent standard value of 1,000 was always used. Participants could use any real numbers including decimals and negative numbers to indicate their responses. The order of probes was randomized to avoid possible order effects.

Psychological Value is validated by the predictive power of Subjective Values Theory (SVT), a fully specified and quantified mathematical model of preferential choice that makes point predictions about response choice and RT in preferential choice tasks, a priori. Cohen and colleagues (2016; Under Review) conducted a series of preferential choice experiments in which participants made two-alternative forced-choice decisions in various scenarios. In the scenarios, Item A would be destroyed unless the observer acts to save item A by allowing item B to be destroyed instead. Briefly, SVT predicts that the observer will always try to save the item with the greatest perceived Psychological Value. SVT predicts a strong positive relationship between distributional overlap, $f(\Psi_A \cap \Psi_B)$, and RT and a strong negative relationship between $f(\Psi_A \cap \Psi_B)$ and probability of selecting the option with the highest Psychological Value.

In a series of preferential choice experiments, Cohen and colleagues demonstrate that SVT predicts probability of response choice and RT across multiple presentation types including moral dilemmas, simplified scenarios which omitted all contextual information, and an RT pressure variation with nearly perfect accuracy, accounting for over 90% of variance in response choice and RT, respectively (2016, Under Review). For the sake of brevity, we have excluded many details that are extraneous to the current work but imperative to understanding the mathematics, precision, and accuracy associated with SVT. For a full explanation, please refer to Cohen et al. (Under Review). Having observed the accurate predictive power of SVT driven by Psychological Value, we seek to further understand psychophysical properties of Psychological Value.

The Psychophysics of Psychological Value

Most perceptual continuum have a nonlinear relationship with stimulus intensity. Nonlinearity of stimulus intensity and perception have been observed for brightness (e.g. Raab, 1962; Stevens & Stevens, 1963), loudness (e.g. Luce & Mo, 1965; Stevens, 1955/1956), heaviness (e.g., Luce & Mo, 1965), pitch (e.g., Beck & Shaw, 1961), tactile roughness (e.g., Verrillo, Bolanowski, & McGlone, 1999), taste (e.g., Mcburney, 1966), smell (e.g., Baird, Berglund, & Olsson, 1996), length (e.g., Miller, Pedersen, & Sheldon, 1970; Miller & Sheldon, 1969), area (e.g., Algom, 1991), vibrotactation (e.g., Verrillo, Fraioli, & Smith, 1969), electric shock (e.g. Sternbach & Tursky, 1964; Stevens, Carton, & Shickman, 1958), temperature (e.g., Banks, 1969; Stevens, 1975), and hue (e.g., Indow & Stevens, 1966). We can infer that perceived Psychological Value likely also has a nonlinear relationship with stimulus intensity.

Most researchers accept the *Axiom of Monotonicity* with reference to preferential value (e.g., Dickert, Västfjäll, Kleber, & Slovic, 2012; Shenhav & Greene, 2010; Tversky & Kahnemann, 1992). The Axiom of Monotonicity states that an increase in quantity produces in increase in preference, and thus, an increase in preferential value (Fishburn, 1970). That is, more is always better. If the Axiom of Monotonicity holds for perceived Psychological Value, we should observe a monotonically increasing growth function whereby increases in quantity of stimuli always produce an increase in preferential value.

To study Psychological Value's growth function, we must manipulate stimulus intensity. In psychophysical experiments such as those cited above, adjusting stimulus intensity is straightforward because the continua of interest fall on objective scales. In contrast, perceived Psychological Value is nonphysical, and therefore, it is more difficult to (1) adjust stimulus intensity, and (2) have objective knowledge of the magnitude of that adjustment. The most straightforward way to manipulate stimulus intensity and retain knowledge of the magnitude of the manipulation is to manipulate the quantity of a probe. For example, we can meaningfully compare the probe “1 wedding dress” and the probe “50 wedding dresses.” In this instance, by increasing the quantity of the stimulus fifty-fold, we have increased the stimulus intensity by 4,900%.

Recently, we conducted two preferential choice experiments with quantity manipulations. Prior to conducting the preferential choice experiments, we measured the Psychological Values of the probes and subsequently used these measurements to predict RT and response choice in the preferential choice experiments. Participants made two-alternative forced-choice decisions in various scenarios whereby Item A would be destroyed unless the observer acts to save item A by allowing item B to be destroyed instead. In these experiments, the Items were randomly assigned quantities denoted by a number symbol (e.g., 1, 3-9, 47-53; Cohen et al., Under Review). SVT accounted for over 90% of variance in RT and response choice . Across both experiments, results suggest that a quantity manipulation is almost irrelevant to decision making. These results suggest that Psychological Value is insensitive to changes in stimulus quantity (and thus stimulus intensity) when (1) the quantity of a single item changes (e.g., 1 nun vs 5 nuns) and (2) those changes are denoted by a number symbol.

Here, we perform a rigorous evaluation of the relationship between stimulus intensity and Psychological Value. To do so, we present ensemble stimuli—stimuli which have multiple distinguishable components. For example, an array of four various sized circles presented together would comprise an ensemble stimulus because each circle within the array is distinguishable from the others. In the current experiment, stimuli were heterogenous groups of people or objects (e.g., “a group of a nun, a judge, a convict, and a soldier”). In contrast to using number symbols, this type of stimulus intensity manipulation requires one to consider each individual component of a stimulus. Because each component is discrete, quantity is an analog stimulus (i.e., there are physically four things), rather than a symbolic stimulus (i.e., the quantity is represented by a single symbol, 4). This distinction is important because researchers debate whether quantity represented by symbols is processed the same as analog quantity (e.g., Sasanguie, De Smedt, & Reynvoet, 2017). Furthermore, because larger analog quantities are also physically larger on the computer display, any violation of the Axiom of Monotonicity could not be ascribed to the symbolic nature of the quantity manipulation. Below, we review how observers perceive ensembles.

Ensemble Representations. Research suggests that the perceptual system tends to extract a statistical summary of some property of ensemble stimuli. This singular representation of a set of properties of stimuli is referred to as an ensemble representation (e.g., Alvarez, 2011). Although an ensemble representation may theoretically be any statistic, it is most commonly the average of some property of a set. For example, Ariely (2001) found that when presented with a test set of various sized circles followed by a comparison circle, participants were able to accurately identify whether the comparison circle was larger or smaller than the mean size of the circles in the test set. However, when asked whether a comparison circle was novel or was present in the test set, participants responded at chance levels. These data suggest that observers likely retain a statistical summary of a set of stimuli (i.e., the average circle size) without retaining much or any information about individual objects within the set (Ariely, 2001). This ability of the perceptual system to rapidly extract the mean of some property of a set of stimuli is termed *perceptual averaging*. A replication and expansion of Ariely's (2001) work (Allik, Toom, Raidvee, Averin, & Kreegipuu, 2014) indicated that perceptual averaging is likely "obligatory" and requires very little awareness of the features of any individual element of a stimulus.

A significant body of research suggests that perceptual averaging occurs accurately across a range of stimuli. For example, perceptual averaging occurs regardless of the homogeneity or heterogeneity of the size of test set item size (e.g., Chong & Treisman, 2003) or the number of elements within a test set (e.g., Marchant, Simons, & Fockert, 2013). Perceptual averaging has also been observed for facial identity (e.g., de Fockert & Wolfstein, 2009), hue (e.g., Maule & Franklin, 2015), emotional expression of faces (e.g., Haberman & Whitney, 2009), average size of continuously changing visual stimuli (e.g., Albrecht & Scholl, 2010), color diversity of unattended stimuli (e.g., Ward, Bear, & Scholl, 2016), motion speed (e.g., Watamaniuk & Duchon, 1992), orientation of stimuli (e.g., Parkes, Lund, Angelucci, Solomon, & Morgan, 2001), visual stimuli presented for as little as 50 ms (e.g., Chong & Tresiman, 2003), and visual and auditory information presented simultaneously (e.g., Albrecht, Scholl & Chun, 2012).

Were the perceived Psychological Value of ensemble stimuli to give rise to perceptual averaging, rather than perceptual summing, then one may infer that Psychological Value is similar to other perceptual dimensions. Nevertheless, some research has demonstrated that high-valued stimuli within groups of stimuli draw an observer's attention. For example, in one experiment, participants completed a training phase in which they viewed visual arrays which contained various geographic shapes displayed in either red or green and each of which contained a short line segment (Anderson, Laurent, & Yantis, 2011). On each trial, participants reported the orientation of the line segment within the target stimulus. For half of participants, correct responses regarding the orientation of line segments within green stimuli were associated with an 80% chance of a high versus low monetary reward and vice versa. Next, participants completed a task which required them to identify a specific geographic shape presented within a visual array. Distractor stimuli—those presented in the color previously associated with high or low monetary rewards—were present on half of the trials. The target stimuli were never red or green. Results indicated that participants were significantly slower at identifying shapes when a high value distractor stimulus was present suggesting that previously rewarded stimuli capture attention even when they are irrelevant to the task at hand. When replicated without monetary rewards in the training phase, the effect of distractor stimuli completely diminished (Anderson et al., 2011). The tendency for high value stimuli to attract attention my skew the extracted perceptual average of ensemble value stimuli toward the maximum value of the group.

The Current Experiment

To study how perceived Psychological Value changes as a function of stimulus intensity, we replicated our past magnitude estimation experiments using heterogenous groups of human and object item probes. We accomplished this by first selecting a subset of single item probes (i.e., not preceded by a quantity greater than 1) that we collected the perceived Psychological Values of in previous experiments (Cohen et al., Under Review; Cohen & Ahn, 2016). We then manipulated stimulus intensity by creating groups of different sizes. Specifically, Group size—the number of individual items within each group—ranged from two to five. For example, "a group of a child and a coworker" has a group size of two, and "a group of a tree, a rock, and a car" has a group size of three. Here, we define the objective stimulus intensity of the group as the sum of the perceived Psychological Values of the individual items that compose the group. Any systematic variation from that objective intensity represents the perceptual bias specific to Psychological Value.

Although this work is exploratory, we can make several hypotheses based on the reviewed literature. Because Psychological Value is perceived, we hypothesize that the Psychological Value of a heterogenous group will be represented by some statistical summary of the properties of the group (i.e., an ensemble representation). The Axiom of Monotonicity predicts that the statistical summary will be the sum of the values. However, the extensive literature on the automaticity of perceptual averaging and our earlier findings that quantity has little to no influence on perceived Psychological Value suggests that the statistical summary will be the average. The tendency for high value stimuli to capture attention may skew the statistical summary of ensemble representations of Psychological Value. Specifically, stimuli with high Psychological Values may capture an observer's attention in the same manner that rewarded stimuli draw an observer's attention, and this attentional capture may correspond to an upward skew in Psychological Value of heterogenous groups.

Method

Participants

One hundred eighty-six naïve undergraduate students enrolled in an introductory psychology course volunteered and participated for partial course credit.

Materials

We identified 80 single items for which we had perceived Psychological Values collected from previous magnitude estimation experiments (Cohen et al., Under Review; Cohen & Ahn, 2016). Two lists were created: one that contained 40 human items and one that 40 object items. We chose to use both humans and objects to determine whether perceived Psychological Value of groups differs for humans and objects. Each list was designed to include items that span a very wide range of Psychological Values from very low to very high. All human items and object items used are presented in Tables 1 and 2, respectively.

Next, we created heterogeneous groups which served as probes in the current experiment. Group size—the quantity of individual items contained within each probe—was manipulated to range from 2-5. Group size was a between-subjects manipulation. Group category—whether the group consisted of humans or objects—was a within-subjects manipulation. A probe consists of a group of unique items of the same category (i.e., human or object). Probes never contained both human and object items, and there were never multiple instances of the same item within a probe[1].

To create the probes for each category, we generated all possible combinations (i.e., no duplicates of items, and order does not matter) of items in that category for each group size (i.e. 2, 3, 4, and 5). For each category, there were 780 possible combinations for a group size of 2, 9,880 possible combinations for a group size of 3 and so on. In total, we generated 760,058 possible probes for each category for a total of 1,520,116 unique stimuli.

We took care to avoid methodology that would encourage participants to use higher order strategies (i.e., mathematical estimation) during the task by manipulating group size between-subjects and never repeating items within a probe. By manipulating group size between-subjects, participants were prevented from noticing and explicitly considering group size variation in their responses. Such a consideration might lead to mental math, rather than a judgment of perceived Psychological Value. For example, if a participant recalled the value they provided for a probe of "an adult and a nun," they may later see the probe "an adult, a nun, and your father" and use a mathematical strategy to calculate one response based on the other.

For each participant, we randomly selected 300 probes of the appropriate group size without replacement from the list of all possible probes with that group size. As such, no two participants saw the same set of 300 probes. In addition to the 300 grouped probes, all participants, regardless of experimental condition, were presented with nine nonhuman animal probes with a group size of one. The probes were “a pet kitten,” “a frog,” “a lion,” “your pet dog,” “a worm,” “a zebra,” “a pet dog,” “your pet kitten,” and “a panda.” These probes were used to link the current dataset to past datasets (see Results for linking procedure). These probes were *only* used to link datasets, and the data collected for them was discarded before data analysis addressing the current research questions. Nonhuman animal probes were selected because this category was not used in any other trials. The list of possible probes for each experimental session totaled 309 trials (i.e., nine single animal probes, 300 combinations). The order of the 309 trials was fully randomized for each participant prior to the start of the experiment.

All stimuli were presented 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 was 1920 × 1200 pixels. Participants input responses on an Apple keyboard.

Procedure

Participants completed the experiment in a small, dark individual testing room containing a desk, a chair, and a computer. White noise was playing at a low volume from speakers in the ceiling to mask any ambient sound.

Consistent with the standard magnitude estimation procedure, participants were presented with a standard and a series of comparison probes. The standard was always “a chimpanzee” assigned a value of 1,000. The standard was constant across trials and across participants. The participant's task was to provide a number to represent their perceived Psychological Value for each probe in relation to the standard. They were instructed to assign values proportionally. Psychological Value was defined using the exact same phrasing from our past magnitude estimation experiments:

> "…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…"

(Cohen & Ahn, 2016; Cohen et al., Under Review). We referred to Psychological Value as "personal value" in the definition to facilitate participants' understanding of the term.

On each trial, a red fixation cross was presented for 500 ms. Next, a dialogue box containing the standard ("a chimpanzee = 1,000"), the probe, and a text box to input responses was presented (see Figure 1). The dialogue box was light gray and centered on the screen. The dialogue box subtended 7.16 degrees vertical visual angle and 19.36 degrees horizontal visual angle. The standard was presented in centered, gray 16-point Helvetica text. Probes were centered in the dialogue box with a black border around them. One individual item was presented on each line in left aligned text. The words "A group of" was presented above the border (see Figure 2). The probe and phrase "A group of" were presented in black 16-point Helvetica font.

Participants input responses by typing on a number pad on a Mac computer keyboard. Responses could be whole numbers or decimals. Responses were formatted in real time using thousand separators to aid participants in reading the numbers as they appeared in the response text box. For example, an estimation of 5000 would appear as “5,000.” After inputting a value, participants pressed the “return” key to advance to the next trial.

Participants completed eight practice trials and up to 309 experimental trials. Probes for the practice trials were created identically to experimental trials, with the exception that no single item animal probes were presented during practice trials. Probes in the practice trials never duplicated probes in the experimental trials. After participants completed the eight practice trials, a dialogue box appeared that read, “Please Click OK to start the experiment.” Upon clicking OK, participants began the first experimental trial. Participants completed experimental trials for the remainder of the experimental session. All trials were fully randomized.

A timer for fifty minutes began at the beginning of the practice trials. The experimental session ended when the timer ran out or the participant completed all 309 trials—whichever occurred first. A self-timed break was provided every 17 minutes. At the conclusion of the experimental session, a dialogue box appeared that thanked the participant for their participation and cued them to exit the testing room.

Results

Prior to analysis, we excluded two participants because they did not complete the experiment. The remaining participants completed an average of 301.25 trials (SD = 30.38). Similar to the procedure used by Cohen and Ahn (2016), two criteria were used to remove participants whose responses or RT were outliers. We excluded three participants because their median log response was an outlier (< 1). We removed 26 participants because their median RT was an outlier (i.e., median RT < 2750 ms) indicating that they likely were not putting effort into or paying attention to the experiment. We excluded all data from practice trials. The remaining data were analyzed.

Linking Datasets

Recall that our predictions for Psychological Value of groups are based on magnitude estimates of Psychological Value of single items collected by Cohen and colleagues (2016; Under Review). Because we are comparing the magnitude estimates collected in different experiments, and therefore under slightly different experimental conditions, we must assess whether there is a systematic variation between the two datasets. If there is, a simple power transformation will be applied to correct for that variation.

Theoretically, the distribution of perceived Psychological Values collected for any given probe in one experiment should overlap perfectly with that for the same probe collected in a different experiment. This would indicate that both experiments successfully measured the same concept under the same experimental conditions. However, due to perceptual variability, there is some stochasticity in the perception of Psychological Value and our sampling techniques. As such, a high (but likely imperfect) overlap between the distributions of the same probes across different experiments is a good indication that they accessed the same concept. To assess whether a systematic variation existed between the two datasets, we compared the 25^{th}, 50^{th}, and 75^{th} quantiles of a subset of probes, termed *common items*, which appeared in the two single item datasets collected by Cohen and colleagues (2016; Under Review). As stated in *Methods*, the common items were not used to assess the influence of grouping in the present experiment.

Prior to linking datasets, we transformed the magnitude estimates. Because we sought to constrict the highest response values and expand the lowest response values, a logarithmic transformation would be ideal (e.g., Stevens, 1955). Unfortunately, a logarithmic transformation is sub-optimal for our current datasets because, (1) the logarithm of zero is undefined (participants occasionally used zero as a response), (2) logarithms of fractions are negative and, once transformed, the negative values become outliers, and (3) and the logarithms of negative numbers are undefined (participants occasionally used negative numbers as a responses). For these reasons, we transformed the data by applying the following formula:

$$\Psi_{TV} = sign\ (|\Psi_{V}|)^{1/10} \tag{2}$$

where Ψ_{TV} is the transformed response, Ψ_{V} is the original, untransformed response, and *sign* is the sign (+/-) of the original response. Next, we linked three datasets: Cohen and Ahn's (2016) data, Cohen et al.'s data (Under Review), and the current data. These datasets will be referred to as Dataset 1, 2, and 3, respectively.

Dataset 1 and 2 were collected prior to the current research and will be used as the source of the baseline values of the single items. For each common item, the 25th, 50th, and 75th quartiles were calculated. To link the two datasets, these 84 (28 probes x 3 quartiles) data points in Dataset 1 ($\Psi_{V1common}$) were regressed on the same 84 data points in Dataset 2 ($\Psi_{V2common}$), using the following formula:

$$\Psi_{V1common} = a * \Psi_{V2common}^{beta}, \tag{3}$$

where *a* and *beta* are free parameters that are fit to the data. Then, all the datapoints in Dataset 2 were transformed using equation 3 with the best fit *a* and *beta* determined for the 84 common data points. This effectively linked Dataset 1 and Dataset 2 using the 28 common items. To ensure that this method was effective, the mean overlap of same probes was calculated using a bootstrap procedure (see Cohen & Ahn, 2016). A high mean overlap (M = 0.9, SD = .10) suggested that—accounting for stochasticity—the probes were properly transformed. We will refer to the combined dataset as the *Single Item* dataset. The transformed median values and interquartile ranges of human and nonhuman single probes are presented in Table 1 and 2, respectively.

Next, Dataset 3 (the data from the current experiment) was linked to the Single Item dataset. The same process was used with one exception: only nine common item probes—all of which were animals—were used to link Dataset 3 to the Single Item dataset. First, all responses for probes from the Group Size 2 condition were transformed and linked to the Single Item dataset. These data were then combined. Subsequently, all probes from the Group Size 3 condition was linked to the larger dataset and then combined. This process was repeated for the probes from the Group Size 4 and 5 conditions. Mean overlaps for same items for group sizes 2 (M = .88, SD = 0.07), 3 (M = 0.92, SD = 0.04), 4 (M = 0.89, SD = 0.06), and 5 (M = 0.93, SD = 0.06) were calculated, respectively.

After successfully linking datasets, all common items were excluded from subsequent analyses. These probes were presented specifically for the purpose of transforming data, and data from these trials was not used to address the current research question.

Exploratory and Confirmatory Data Analysis

The purpose of the current analysis was to test models that could plausibly account for psychophysical properties of Psychological Value. Therefore, we used half of the data to explore models and variables of interest and the remaining half of the data for confirmatory analyses. We accomplished this by randomly selecting all data for half of the participants of each Group Size condition. The exploratory and confirmatory datasets included 77 and 78 participants, respectively.

Exploratory Data Analysis. Here, we propose and test four distinct models. We will refer to these as (1) the Default Model, (2) the Perceptual Summing Model, (3) the Perceptual Averaging Model, and (4) the Perceptual Averaging Plus Model.

To identify the optimal predictors for each hypothesized model, we ran a series of linear mixed models. First, we ran a linear mixed model with the most theoretically relevant predictor variable. We noted the R^2 and BIC value for that model. We then ran a second linear mixed model in which we added the next most we theoretically relevant predictor variable to the previous model. Again, we noted the R^2 and BIC value for that model. We continued this process for all theoretically relevant predictor variables. As a general rule, predictor variables only remained in subsequent analyses if they had a significant effect on responses based on p-values. The hypothesized model with the greatest R^2 and lowest BIC value was chosen as the final model to assess fit for the confirmatory analysis.

Default Model. The Default Model posits that group size influences responses. There are two ways that group size might influence responses: (1) the perceived Psychological Value of the group is some function of the sum of the perceived Psychological Value of single items, or (2) the perceived Psychological Value of the group is only influenced by the quantity of items in each group, but not influenced by the perceived Psychological Value of the single items themselves (see Figure 3).

If the perceived Psychological Value of the group is some function of the sum of the perceived Psychological Value of single items, then group size will be related to participants perceived Psychological Value of the group simply because group sums increase with group size. In this case, then both the Default and the Perceptual Summing Models will account for a significant amount of variance. However, the Perceptual Summing Model will outperform the Default Model because it will include the perceived Psychological Value of the single items as a predictor. In contrast, if the perceived Psychological Value of the group is only influenced by the quantity of items in each group, then the Default will outperform all other models.

According to the Default Model, responses are driven by group size. Participants would assign higher values for probes with higher quantities and lower values to probes with lower quantities regardless of values of single items. A Linear Mixed Model with participants as a random factor and group size as a fixed factor, revealed that group size failed to predict responses, $\beta = 0.04$, $t(75) = 0.81$, $p = n.s.$ (see Table 3). The model, with a significant intercept of 2.20, had an $R^2 = 0.20$ and a BIC = 59144.63. Because group size did not have a significant effect on responses, it was excluded from all subsequent analyses.

The finding that group size alone does not account for a significant amount of variance in the data suggests that perceived Psychological Values of groups are not based on the sum of the perceived Psychological Values of individual items. Because all single items in our dataset had a positive perceived Psychological Value, the addition of an item to a group will inherently increases the sum of that group's Psychological Value.

Calculating Value-Dependent Variables. Next, we tested the models that *do* rely on the values of single items. To do so, we first calculated predictor variables that are used in the analyses below.

All probes were separated into their single items. Next, we identified the 25th, 50th, and 75th response value quantiles for each single item in each probe for every experimental trial. The median value (50th quantile) of each item in the Single Item Dataset was used as our measure of central tendency because it is insensitive to outliers and unaffected by transformation. Recall that these values originate from the Single Item Dataset.

To assess the predictions of the Perceptual Summing Model, we calculated the *Summed* ψ_V of each probe:

$$\sum \psi_V = Md_1 + Md_2 + Md_{\ldots n} \tag{4}$$

where $\sum \psi_V$ is the Summed ψ_V of single items—the sum of the median values for any given probe—and $Md_{1\ldots n}$ are the median values for each item within each probe.

To assess the predictions of the Perceptual Averaging Model, we calculated the *Average* ψ_V of each probe:

$$\overline{\psi_V} = (Md_1 + Md_2 + Md_{\ldots n})/n \qquad (5)$$

where $\overline{\psi_V}$ is the Average ψ_V, and $Md_{1\ldots n}$ are the median values of each item within each probe, and n is group size.

We rounded all Summed ψ_V and Average ψ_V values to the nearest 0.1 to eliminate false precision and the likelihood of fitting to experimental noise (e.g., Cohen, 1990). The Summed ψ_V values ranged from 0 to 15.1 (M = 7.07, SD = 2.71). Average ψ_V values ranged from 0 to 3.2 (M = 2.03, SD = 0.43). All future uses of the variable names "Summed ψ_V" and "Average ψ_V" refer to these rounded values.

Perceptual Summing Model. The Perceptual Summing Model posits that participants determine the Psychological Value of a group by summing the individual values of single items. Here, the predictor variable is Summed ψ_V (see Figure 3).

We calculated a Linear Mixed Model with participants as a random factor. As can be seen in Table 4, Summed ψ_V accounted for a significant amount of variance. When category was added, it also accounted a significant amount of variance. The final model, with Summed ψ_V, β = 0.17, $t(4579) = 31.89$, $p < .0001$, and category, β = 0.19, $t(17337) = 27.29$, $p < .0001$, had a significant intercept of 1.16, an $R^2 = 0.55$, and a BIC = 52610.53 (see Table 3).

The finding that Summed ψ_V accounts for a significant amount of variance in the data suggests that perceived Psychological Values of single items are driving participants' responses. However, recall that group size alone was not significant in the Default Model. Together, these two results suggest that perceived Psychological Values of single items are driving participants' responses, but participants are not summing items' values. This suggests that participants may be averaging the perceived Psychological Values of single items. In this instance, the significant effect of Summed ψ_V may arise because sums are linear transformations of averages. We explored the possibility that averaging rather than summing is occurring by testing the Perceptual Averaging Model.

Perceptual Averaging Model. The Perceptual Averaging Model posits that perceived Psychological Value is equal to the average of the median values of each individual item within each group (see Figure 3).

We used Average ψ_V as the main predictor variable to test this model. We calculated a Linear Mixed Model with participants as a random factor. As can be seen in Table 3, Average ψ_V accounted for a significant amount of variance. A subsequent Linear Mixed Model with Average ψ_V and category as predictor variables revealed that category accounted for an additional small, but significant amount of variance. The final model, with Average ψ_V, β = 0.71, $t(1458) = 27.82$, $p < .0001$, and category, β = 0.19, $t(20458) = 27.59$, $p < .0001$, were both included in the final Perceptual Averaging Model. The model, with a significant intercept of 0.91, had an $R^2 = 0.54$ and a BIC = 50707.39 (see Table 3).

The Perceptual Averaging Model accounted for a similar amount of variance and had a lower BIC than the Perceptual Summing Model. However, given that high valued items attract attention (e.g., Anderson et al., 2011; Anderson & Yantis, 2013), we suspected that the salience of the highest valued item within each probe may account for additional variance. To explore this hypothesis, we propose and test the Perceptual Averaging Plus Model.

Perceptual Averaging Plus Model. The Perceptual Averaging Plus Model posits that perceived Psychological Value of a group is subject to perceptual averaging with some greater influence of the highest valued item within a group. To test this model, we first identified the highest valued item within each probe and rounded its median to the nearest tenth to reduce the likelihood of fitting to experimental noise. This value—the Max(ψ_V)—ranged from 0 to 3.2 (M = 2.50, SD = 0.48). We used Average ψ_V and Max(ψ_V) as the main predictor variables to test this model (see Figure 3).

Table 3 presents the Linear Mixed Models with participants as a random factor. The final model with Average ψ_V, $\beta = 0.42$, $t(1458) = 14.77$, $p < .0001$, Max(ψ_V), $\beta = 0.50$, $t(5613) = 20.39$, $p < .0001$, and category, $\beta = 0.08$, $t(14844) = 20.39$, $p < .0001$, had a nonsignificant intercept, an R^2 value of 0.67, and a BIC value of 49223.85.

The Perceptual Averaging Plus Model unambiguously outperforms all other models. It accounts for the greatest percentage of variance in data with the lowest BIC value indicating a better fit than the Default Model, the Perceptual Summing Model, or the Perceptual Averaging Model. Therefore, we adopt the Perceptual Averaging Plus Model as the final accepted model.

For the final model, we remove the intercept because it was nonsignificant, and the category predictor variable because it only accounted for 0.1% of variance. The final Mixed Model with participants as a random factor and Average ψ_V, $\beta = 0.40$, $t(1459) = 15.13$, $p < .0001$, and Max(ψ_V), $\beta = 0.62$, $t(5614) = 30.97$, $p < .0001$, had an R^2 value of 0.67 and a BIC value of 49288.24 (see Table 4).

Variance Analysis. Next, we explored the spread of the data. Specifically, we explored standard deviation and interquartile range data to determine whether they supported our acceptance of the Perceptual Averaging Plus Model.

If participants sum the perceived Psychological Values of single items to arrive at their perceived Psychological Value of a group, the standard deviation of responses should increase with group size as predicted by the following formula:

$$SD_G = \sqrt{\sum_{I=1}^{n} SD_I^2} \tag{6}$$

where SD_{group} is the standard deviation of a group, and $SD_{I_{1\ldots n}}^2$ are the squared standard deviations of each item within a probe.

If, however, participants took an average of the perceived Psychological Values of single items to arrive at their perceived Psychological Value of a group, the standard deviation of responses should remain constant with group size as predicted by the following formula:

$$SD_{group} = \frac{\sqrt{\sum_{I=1}^{n} SD_I^2}}{n} \tag{7}$$

where SD_{group} is the standard deviation of a group, $SD_{I_{1\ldots n}}^2$ are the squared standard deviations of each item within a probe, and n is the number of items within a probe.

The data reveal that the standard deviations remain relatively stable across group sizes 2 (*SD* = 1.01), 3 (*SD* = 0.93), 4 (*SD* = 0.82), and 5 (*SD* = 1.26). This would favor a process of averaging versus summing occurring. However, because SD is not a robust statistic, it is easily influenced by outliers. Therefore, we did a similar analysis of spread of data using interquartile ranges.

To assess whether the spread of participants' responses increased with group size, we standardized the responses by the interquartile range and median of the single item values in that composed each group. Recall that a *probe* refers to a group of items in a single trial. First, we calculated the 25th quartile of every item in each probe using the Single Items Dataset. Next, for every probe, we calculated the mean of these 25th quartiles:

$$\overline{Q^{25}} = \frac{\sum_{i=1}^{n} Q_i^{25}}{n} \tag{8}$$

Where $\overline{Q^{25}}$ is the mean 25th quartile for each probe, Q_i^{25} is the 25th quartile for each item, and n is the number of items within a probe. We then repeated this procedure for the 75th quartile, $\overline{Q^{75}}$. Next, we calculated the interquartile range for each probe, Ψ_{IQR}, by subtracting $\overline{Q^{25}}$ from $\overline{Q^{75}}$. We then standardized the participants' responses by this value (which, recall, is unique for every probe) using the following formula:

$$QR = \frac{\Psi_{\mathrm{V}} - (\overline{Q^{25}} + 0.5 * \Psi_{IQR})}{\Psi_{IQR}} \tag{9}$$

where QR is the standardized interquartile range.

By standardizing by the interquartile range, we can assess whether spread increased with group size—as predicted by the Perceptual Summing Model—or spread stayed constant across group sizes—as predicted by the Perceptual Averaging Model and Perceptual Averaging Plus Model. If either the Default or Perceptual Summing Model sufficiently accounted for the data, QR values would increase as group size increased.

If spread stayed constant across group sizes, the transformed data will center around zero because the average of the single items in the probe will be equal to the participants' average response to that probe. Furthermore, if the spread of the single item values is equal to the spread of the participants' responses, then 50% of transformed data will fall between the values -0.5 and 0.5. We did observe this pattern with values skewed slightly high, and this pattern stayed stable across group size. As depicted in Figure 5 and Table 5, we did not observe this pattern. Standardized interquartile ranges by group size were all roughly 1 with the 25th and 75th quartiles skewed slightly high. The consistency of this pattern suggests that values are averaged with some influence of the Max(ψ_V) pulling the values slightly higher than one would expect if only averaging was occurring.

We will use the interquartile range analysis, rather than the standard deviation analysis, in our confirmatory data analysis as our analysis of the spread of our data because it is a more robust statistic.

Confirmatory Data Analysis. The purpose of the confirmatory data analysis was to replicate the success of the Perceptual Averaging Plus Model using the confirmatory dataset. The confirmatory dataset, recall, is a random subset of participants' data that was not used for the exploratory analysis. Replication would indicate that the exploratory analyses were meaningful and not simply fitting to noise. To attempt replication, we calculated that the final Perceptual Averaging Plus Model and our interquartile range analysis on the confirmatory dataset.

If we were correct in adopting the Perceptual Averaging Plus Model, we should observe significant effects of Average ψ_V and Max(ψ_V) (see Figure 3). Together, the variables should account for a high percentage of variance in responses. Furthermore, when analyzing the spread of the data using our interquartile range analysis, the transformed data should center slightly higher than zero, and 50% of transformed data should fall between the values slightly higher than -0.5 and 0.5. Recall the slight shift is the influence of the maximum perceived Psychological Value of the items in the group.

We calculated the Perceptual Averaging Plus Model using a linear mixed model with participants as a random factor (see Table 5). The results revealed significant effects of Average ψ_V, β = 0.27, $t(1501) = 12.88$, $p < .0001$, and Max(ψ_V), β =, β = 0.66, $t(5691) = 36.33$, $p <$

.0001 (see Figure 4). The model had an R^2 value of 0.71 and a BIC value of 43560.46. These results replicate those of the exploratory analysis.

Similar to the exploratory interquartile range analysis, about 50% of QR values fell between -0.5 and 0.5 with values skewed slightly high. This pattern stayed stable across group size (see Figure 5, Table 5). Standardized interquartile ranges by group size were all roughly 1 with the 25^{th} and 75^{th} quartiles skewed slightly high. The consistency of this pattern, once again, suggests that values are averaged with some influence of the Max(ψ_V) pulling the values slightly higher than one would expect if only averaging was occurring. The results of the confirmatory analysis are clear; they replicate the exploratory analysis and support the Perceptual Averaging Plus Model.

Discussion

For the past century, researchers have consistently asserted that there are values associated with each option in a multi-alternative choice, and a decision maker can make an optimal choice by choosing the option with the highest value (e.g., Simonson & Tversky, 1992; Rangel et al., 2008; Padoa-Schioppa & Assad, 2006; Kable & Glimcher, 2009; Rangel & Clithero, 2014; Colas, 2017). Cohen and colleagues have asserted that the value-based construct that predicts preferential choice is Psychological Value—an underlying perceptual continuum that has been identified, measured and validated (Cohen & Ahn, 2016; Cohen et al., Under Review). The purpose of the current research was to advance understanding of Psychological Value by studying its' psychophysical properties. Specifically, we used magnitude estimation to describe Psychological Value's growth function.

We conceive of Psychological Value as a perception and, therefore, we should be able to gather relatively consistent measurements of it across individuals and experiments. We tested this consistency by predicting the Psychological Values of groups in the current experiment using Psychological Value estimates of single probes collected in two separate past experiments (Cohen & Ahn, 2016; Cohen et al., Under Review). Using the single item probes from the past experiments, we generated over 1.5 *million* unique stimuli for the current experiment. Using these unique stimuli, we created unique probe lists for each participant. Next, we calculated value-dependent predictor variables (e.g., Summed ψ_V, Average ψ_V) for each stimulus. Thus, there was a novel prediction for every stimulus. The quantity of unique probes and predictions negates the possibility that chance responding could produce Psychological Values of groups that are predicted by the value-dependent variables.

Our results indicate that the perceived Psychological Values of individual items predict the perceived Psychological Values of groups. The best predictors of the Psychological Values of groups were the average of the Psychological Values of the items within each group and the Psychological Value of the highest-valued item within each group. Because value-dependent variables successfully predicted the Psychological Value of groups, we can unambiguously conclude that there are consistent perceptions of Psychological Value across individuals and experimental sessions. Critically, group size and category—variables that are *not* value-dependent—were poor predictors of the Psychological Values of groups.

Our data establish the reliability of our measurements of Psychological Value. Nevertheless, in the paper that he formalized magnitude estimation as a technique for estimating subjective sensory magnitudes, S.S Stevens stated,

> "It is always tempting, of course, to take as valid the measure that is the most reliable, but a measure can sometimes be reliable simply because it is biased by constraints of one sort or another. In short, we can never fully escape the uncertain task of deciding, without

> external criteria, that a measure does or does not assess the thing we are interested in. Fortunately, this judgment does not belong to any one scientist. In the long run, it is the scientific community that will decide the issue"

(Stevens, 1956, pp. 25). Stevens (1956) sentiment reminds us that the consistency of our measurements make them reliable but not necessarily valid. As discussed, a valid preferential value construct should be both (1) stimulus independent and (2) predict preferential choice. Our measurement of Psychological Value satisfies both of these criteria.

First, our results suggest that our measurements of Psychological Value are stimulus independent. The same model—the Perceptual Averaging Plus Model—successfully predicted the Psychological Values of groups regardless of probe category. This would suggest that magnitude estimation suffices as a method for measuring the Psychological Value of different types of stimuli. Further, the success of the Perceptual Averaging Plus Model suggests that the Psychological Value of humans and objects *are perceived using the same system*. Although it is tempting to assume that we value humans in some special way because they are sentient, capable of higher order thought, communicate verbally, feel emotion, etc., our data suggest that this is not the case. The Psychological Value of all stimuli fall on a common perceptual continuum. Though we may value some humans more than some objects, they are not valued in a fundamentally different way.

Second, Cohen and collogues have demonstrated that Psychological Value predicts preferential choice. Across a series of experiments and presentation styles, Cohen et al. (2016; Under Review) used SVT to validate the construct of Psychological Value by demonstrating that measurements of Psychological Value predict participants' RT and response choice with near perfect accuracy.

Having established both the reliability and validity of Psychological Value, here we ask: how does Psychological Value change as a function of stimulus intensity? We addressed this question by presenting participants with heterogeneous group probes consisting of individual items. Because we had previously collected Psychological Values of each individual item within the group probes, we maintained knowledge of the magnitude of our stimulus intensity manipulation.

Most percepts increase monotonically—though but nonlinearly—with stimulus intensity (e.g., Stevens, 1955/1956/1957/1960/1971). Taken together, classic psychophysics research and the Axiom of Monotonicity (Fishburn, 1970) suggests that Psychological Value should increase monotonically as well. If this was the case, the Psychological Value of a group should be well predicted by the sum of the Psychological Values of each item within the group. However, we observed surprising, yet unambiguous, evidence that Psychological Value *does not* follow this pattern. In fact, we found that the single best predictor of the Psychological Value of a group is the average of the Psychological Values of the items within the group. This finding is striking because the results of the current experiment directly demonstrate that Psychological Value—validated as the value-based construct that drives preferential choice—violates the Axiom of Monotonicity. These results contradict a large body of influential research which assumes that preferential value abides by the Axiom of Monotonicity (see Fishburn, 1970).

Our data revealed that observers perceived the Psychological Value of a group as the perceptual average of the single items within the group rather than the perceptual sum of the single items within the group. This finding may be a function of the perceptual system that gives rise to ensemble representations—statistical summaries of properties of ensemble stimuli (e.g., Alvarez, 2011). Although an ensemble representation can theoretically be any extracted statistic (e.g., sum, median, etc.), for most perceptual continua the extracted statistic is the mean (i.e., perceptual averaging). Though we do not have any evidence to explain why Psychological Value is averaged, we suspect that it may be the result of the evolved mechanism in the perceptual system that encodes ensembles. This mechanism may have some evolutionary advantage. For example, when one sees a herd of animals approaching, there may be some survival value associated with perceiving the average features of the individuals in the group (e.g., the average size, shape, etc.). This permits classification of the animals in the herd (e.g., predator, prey, etc.). There would be, in contrast, little evolutionary advantage to perceptually summing the features of the individuals in the group (e.g., the summed size, shape, etc.). Thus, if the brain uses a perceptual mechanism which produces ensemble representations and this mechanism defaults to perceptual averaging, Psychological Value may be perceptually averaged simply because it would be inefficient for the brain to use a novel mechanism rather than the pre-evolved mechanism. Of course, there are separate evolved numerical cognition mechanisms to identify quantity. Although numerical cognition mechanisms may be involved in perception of and interact with Psychological Value, we do not believe that they are identical.

In sum, we presented and rigorously evaluated four models capable of predicting the Psychological Value of heterogenous groups using an exploratory followed by a confirmatory analysis. We concluded that the Psychological Value of groups is well predicted by the Psychological Value of items within a group, Psychological Value is subject to perceptual averaging, and the salience of high-valued probes skews the Psychological Values of groups upward. Importantly, Psychological Values of groups *are not* well predicted by group size. These data lie at the intersection of psychophysics and decision-making research, and we believe that they have important implications for each. We are confident that not only is the current research the first attempt to study the psychophysical scaling of a value-based construct which drives preferential choice but also that it is the first attempt to study the psychophysical scaling of *any* nonphysical perceptual continuum. It is our hope that the current work will spur future research on psychophysical scaling of nonphysical continua and the perceptual continuum which drives preferential choice: Psychological Value.

Open Practices Statement

The data and r code for all experiments are available at https://github.com/ccpluncw/ccpl_data_valuePsychophysics2019/. The experiment was not pre-registered.

References

Albrecht, A. R., & Scholl, B. J. (2010). Perceptually averaging in a continuous visual world: Extracting statistical summary representations over time. *Psychological Science*, *21*(4), 560-567.

Albrecht, A. R., Scholl, B. J., & Chun, M. M. (2012). Perceptual averaging by eye and ear: Computing summary statistics from multimodal stimuli. *Attention, Perception, & Psychophysics*, *74*(5), 810-815.

Algom, D. (1991). Memory psychophysics for area: effect of length of delay. *Perceptual and motor skills*, *72*(1), 296-296.

Allik, J., Toom, M., Raidvee, A., Averin, K., & Kreegipuu, K. (2014). Obligatory averaging in mean size perception. *Vision Research*, *101*, 34-40.

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in cognitive sciences*, *15*(3), 122-131.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Learned value magnifies salience-based attentional capture. *PloS one*, *6*(11), e27926.

Anderson, B. A., & Yantis, S. (2013). Persistence of value-driven attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(1), 6.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological science*, *12*(2), 157-162.

Armel, K. C., Beaumel, A., & Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgment and Decision making*, *3*(5), 396-403.

Ashby, F. G., & Lee, W. W. (1993). Perceptual variability as a fundamental axiom of perceptual science. In *Advances in psychology* (Vol. 99, pp. 369-399). North-Holland.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological review*, *93*(2), 154.

Azari, H., Parks, D., & Xia, L. (2012). Random utility theory for social choice. In *Advances in Neural Information Processing Systems* (pp. 126-134).

Baird, J. C., Berglund, B., & Olsson, M. J. (1996). Magnitude estimation of perceived odor intensity: Empirical and theoretical properties. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(1), 244.

Banks, W. P. (1969). Temperature sensitivity: One subjective continuum or two?. *Perception & Psychophysics*, *6*(3), 189-192.

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage*, *76*, 412-427.

Beck, J., & Shaw, W. A. (1961). The scaling of pitch by the method of magnitude-estimation. *The American journal of psychology*.

Bundorf, M. K., Mata, R., Schoenbaum, M., & Bhattacharya, J. (2013). Are prescription drug insurance choices consistent with expected utility theory? *Health Psychology*, *32*(9), 986–994. https://doi-org.liblink.uncw.edu/10.1037/a0033517.supp (Supplemental)

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.

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision research*, *43*(4), 393-404.

Cohen, D. J., & Ahn, M. (2016). A subjective utilitarian theory of moral judgment. *Journal of Experimental Psychology: General, 145*(10), 1359-1381. http://dx.doi.org/10.1037/xge0000210

Cohen, D.J., Cromley, A., Freda, K., & White, M. (2019). Subjective values theory: the psychology of value and preferential choice. Manuscript submitted for publication.

Cohen, D. J., & Lecci, L. (2001). Using magnitude estimation to investigate the perceptual components of signal detection theory. *Psychonomic bulletin & review*, *8*(2), 284-293.

Cohen, J. (1990). Things I have learned (so far). *American psychologist*, *45*(12), 1304.

Colas, J. T. (2017). Value-based decision making via sequential sampling with hierarchical competition and attentional modulation. *PloS one*, *12*(10), e0186822.

Colas, J. T., & Lu, J. (2017). Learning where to look for high value improves decision making asymmetrically. *Frontiers in psychology*, *8*, 2000.

Coren, S. (2012). Sensation and perception. *Handbook of Psychology, Second Edition*, *1*.

de Fockert, J., & Wolfenstein, C. (2009). Short article: rapid extraction of mean identity from sets of faces. *Quarterly Journal of Experimental Psychology*, *62*(9), 1716-1722.

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.

Fishburn, P. C. (1970). *Utility theory for decision making* (No. RAC-R-105). Research analysis corp McLean VA.

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.

Gwinn, R., Leber, A. B., & Krajbich, I. (2019). The spillover effects of attentional learning on value-based choice. *Cognition*, *182*, 294-306.

Haberman, J., & Whitney, D. (2009). Seeing the mean: ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 718.

Hey, J. D., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica: Journal of the Econometric Society*, 1291-1326.

Indow, T., & Stevens, S. S. (1966). Scaling of saturation and hue. *Perception & Psychophysics*, *1*(2), 253-271.

Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature neuroscience*, *10*(12), 1625.

Kable, J. W., & Glimcher, P. W. (2009). The neurobiology of decision: consensus and controversy. *Neuron*, *63*(6), 733-745.

Katahira, K., Yuki, S., & Okanoya, K. (2017). Model-based estimation of subjective values using choice tasks with probabilistic feedback. *Journal of Mathematical Psychology*, *79*, 29-43.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature neuroscience*, *13*(10), 1292.

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.

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.

Lim, S. L., O'Doherty, J. P., & Rangel, A. (2011). The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *Journal of Neuroscience*, *31*(37), 13214-13223.

Luce, R. D., & Mo, S. S. (1965). MAGNITUDE ESTIMATION OF HEAVINESS AND LOUDNESS BY INDIVIDUAL SUBJECTS: A TEST OF A PROBABILISTIC RESPONSE THEORY 1. *British Journal of Mathematical and Statistical Psychology*, *18*(2), 159-174.

Marchant, A. P., Simons, D. J., & de Fockert, J. W. (2013). Ensemble representations: Effects of set size and item heterogeneity on average size perception. *Acta psychologica*, *142*(2), 245-250.

Marks, L. E., & Algom, D. (1998). Psychophysical scaling. In *Measurement, judgment and decision making* (pp. 81-178). Academic Press.

Maule, J., & Franklin, A. (2015). Effects of ensemble complexity and perceptual similarity on rapid averaging of hue. *Journal of vision*, *15*(4), 6-6.

Mcburney, D. H. (1966). Magnitude estimation of the taste of sodium chloride after adaptation to sodium chloride. *Journal of Experimental Psychology*, *72*(6), 869.

Miller, A. L., Pedersen, V. M., & Sheldon, R. W. (1970). Magnitude estimation of average length: A follow-up. *The American Journal of Psychology*, 95-102.

Miller, A. L., & Sheldon, R. (1969). Magnitude estimation of average length and average inclination. *Journal of Experimental Psychology*, *81*(1), 16.

Mormann, F., Bausch, M., Knieling, S., & Fried, I. (2017). Neurons in the Human Left Amygdala Automatically Encode Subjective Value Irrespective of Task. *Cerebral Cortex*, *29*(1), 265-272.

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. *Judgment and Decision Making*, *5*(6), 437-449.

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*(7090), 223.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature neuroscience*, *4*(7), 739.

Peters, J., & Büchel, C. (2010). Neural representations of subjective reward value. *Behavioural brain research*, *213*(2), 135-141.

Polania, R., Woodford, M., & Ruff, C. C. (2019). Efficient coding of subjective value. *Nature neuroscience*, *22*(1), 134.

Raab, D. H. (1962). Magnitude estimation of the brightness of brief foveal stimuli. *Science*, *135*(3497), 42-44.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature reviews neuroscience*, *9*(7), 545.

Rangel, A., & Clithero, J. A. (2014). The computation of stimulus values in simple choice. In *Neuroeconomics* (pp. 125-148). Academic Press.

Sasanguie, D., De Smedt, B., & Reynvoet, B. (2017). Evidence for distinct magnitude systems for symbolic and non-symbolic number. *Psychological research*, *81*(1), 231-242.

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.

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of marketing research*, *29*(3), 281-295.

Smith, S. M., & Krajbich, I. (2018). Attention and choice across domains. *Journal of Experimental Psychology: General*.

Sternbach, R. A., & Tursky, B. (1964). On the. Psychophysical power function in electric shock. *Psychonomic Science*, *1*(1-12), 217-218.

Stevens, J. C., & Stevens, S. S. (1963). Brightness function: Effects of adaptation. *JOSA*, *53*(3), 375-385.

Stevens, S. S. (1955). The measurement of loudness. *The Journal of the Acoustical Society of America*, *27*(5), 815-829.

Stevens, S. S. (1956). The direct estimation of sensory magnitudes: Loudness. *The American journal of psychology*, *69*(1), 1-25.

Stevens, S. S. (1957). On the psychophysical law. *Psychological review*, *64*(3), 153.

Stevens, S. S. (1960). The psychophysics of sensory function. *American scientist*, *48*(2), 226-253.

Stevens, S. S. (1971). Issues in psychophysical measurement. *Psychological review*, *78*(5), 426.

Stevens, S. S. (1975). Psychophysics New York Wiley.

Stevens, S. S., Carton, A. S., & Shickman, G. M. (1958). A scale of apparent intensity of electric shock. *Journal of Experimental Psychology*, *56*(4), 328.

Stigler, G. J. (1950). The development of utility theory. I. *Journal of political economy*, *58*(4), 307-327.

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

Verrillo, R. T., Bolanowski, S. J., & McGlone, F. P. (1999). Subjective magnitude of tactile roughness. *Somatosensory & Motor Research*, *16*(4), 352–360. https://doi-org.liblink.uncw.edu/10.1080/08990229970401

Verrillo, R. T., Fraioli, A. J., & Smith, R. L. (1969). Sensation magnitude of vibrotactile stimuli. *Perception & Psychophysics*, *6*(6), 366-372.

Ward, E. J., Bear, A., & Scholl, B. J. (2016). Can you perceive ensembles without perceiving individuals?: The role of statistical perception in determining whether awareness overflows access. *Cognition*, *152*, 78-86.

Watamaniuk, S. N., & Duchon, A. (1992). The human visual system averages speed information. *Vision research*, *32*(5), 931-941.

Footnotes

[1] If an observer was asked to provide their Psychological Value of "a group of an adult and an adult," they are likely to multiply their perceived Psychological Value of one adult by two and provide a strategically calculated number. This data would reflect more about number cognition than it would the psychophysical properties of Psychological Value. Additionally, Psychological Value is a perception, and therefore, should not be transformed using math that takes place under cognitive control.

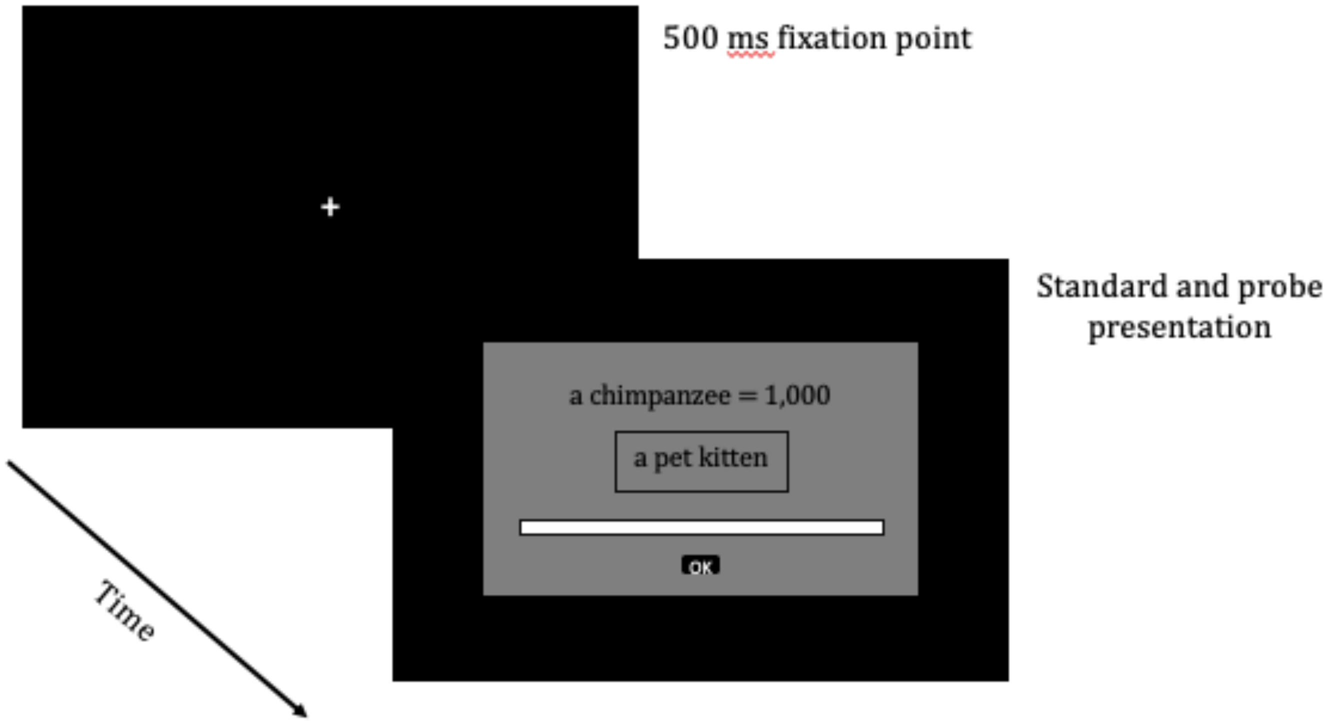

Figure 1: Each trial consisted of a fixation cross presented for 500 ms followed by a dialogue box containing the standard, the probe, and a text box to input responses. The probe depicted (a pet kitten) has a group size of one. Each participant was presented with nine probes with a group size of 1.

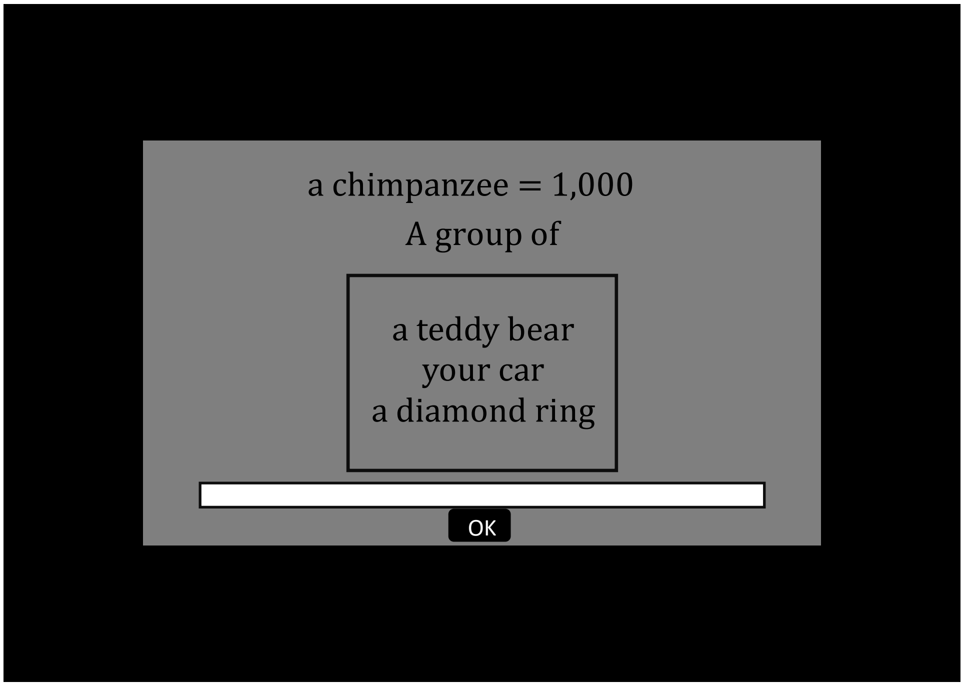

Figure 2: Probes with a group size of 2-5 were presented with one item on each line and the words “A group of” was included above the black border.

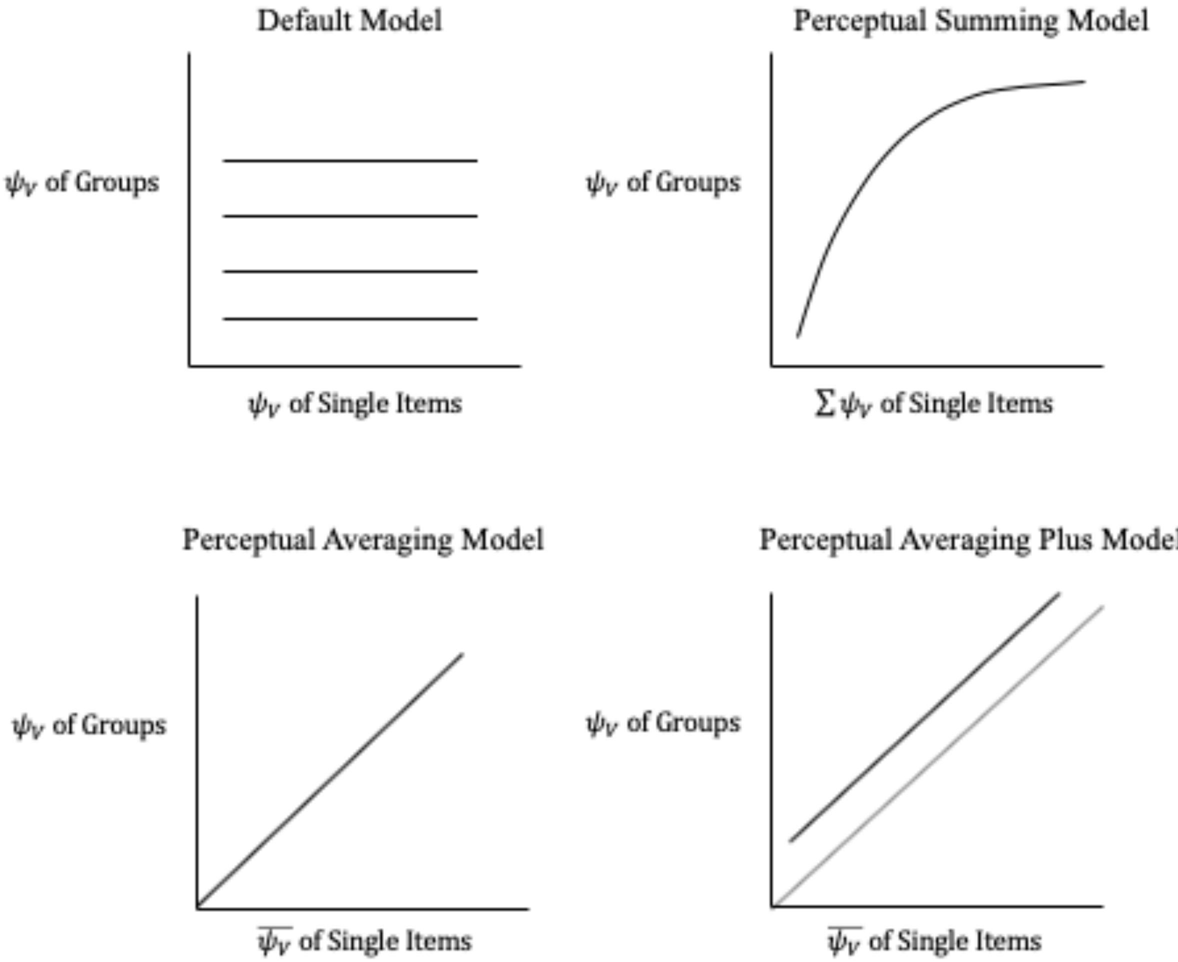

Figure 3: Predicted patterns of data for each model. If the Default Model (top left) is accurate, the ψ_V of Single Items should be a poor predictor of the ψ_V of Groups. If the Perceptual Summing Model (top right) is accurate, we should observe a negatively accelerating function between ψ_V of Groups and the $\sum \psi_V$ of Single Items within each group. If the Perceptual Averaging Model (bottom left) is accurate, there should be a linear relationship between ψ_V of Groups and $\overline{\psi_V}$ of Single Items. Lastly, if the Perceptual Averaging Plus Model (bottom right) is accurate, there should be a linear relationship between ψ_V of groups and $\overline{\psi_V}$ of Single Items with data skewed upwards due to the influence of Max(ψ_V) of Single Items. The gray diagonal line represents a perfect linear relationship, and the black line represents the predicted data assuming an influence of Max (ψ_V) of Single Items.

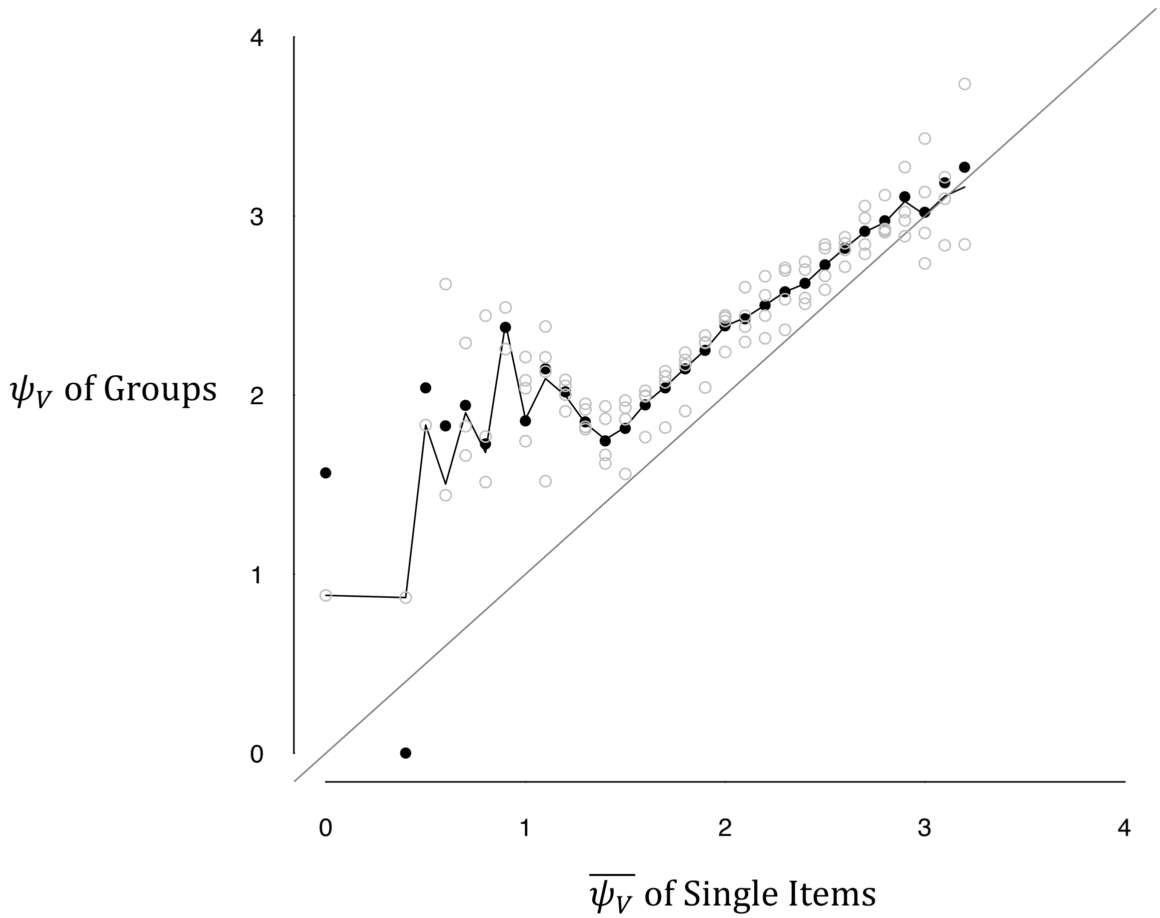

Figure 4: The relationship between the Psychological Value of single items and the Psychological Value of groups. The x-axis represents the Average ψ_V for probes in the confirmatory dataset. The y-axis represents the mean transformed responses for probes in the confirmatory dataset. Gray data points are plotted by group size, and black data points are collapsed across group size. The thin diagonal gray line with an intercept of 0 and a slope of 1 reflects the predicted model fit if was equal to the Average ψ_V. As reflected on the graph, the datapoints and model fit fall slightly above this line demonstrating the influence of both the Average ψ_V and the Max(ψ_V). If the Perceptual Summing Model was accurate, the lowest value of the y axis should start at the highest value of the x-axis and increase. We did not observe this pattern.

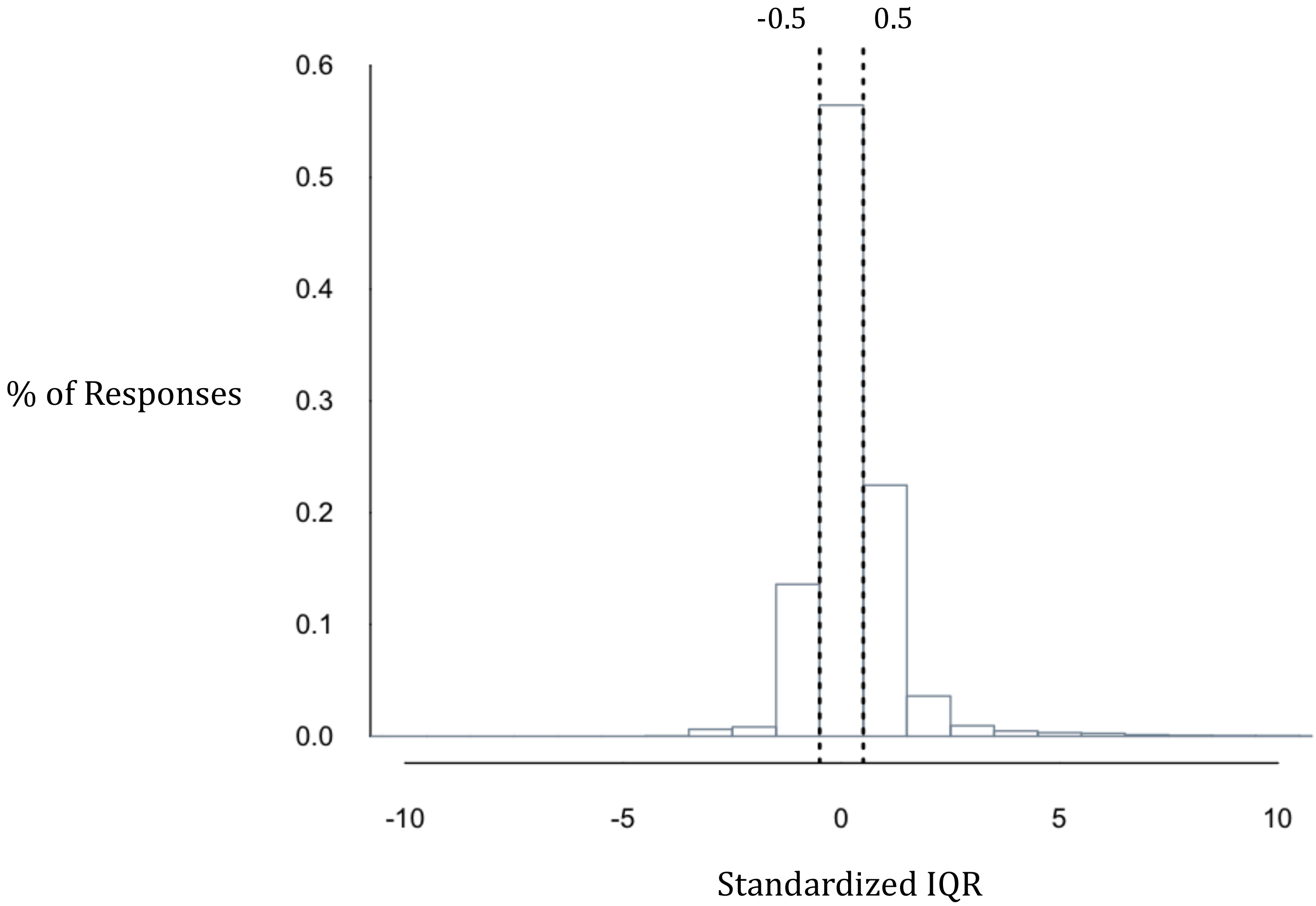

Figure 5: Histogram of standardized interquartile range values collapsed across group size. The x-axis represents standardized interquartile range values, and the y-axis represents percentage of responses. The black dotted lines are plotted at -0.5 and 0.5. If perceptual averaging occurs, then about 50% of the QR values should fall between -0.5 and 0.5. The data confirm this hypothesis: about 55% of QR values fell between -0.5 and 0.5. Importantly, the values skewed slightly high which is consistent with the predictions of the Perceptual Averaging Plus Model. Here, it is evident that standardized interquartile range values *do not* reliably increase with group size.

Table 1

Quartiles of transformed response data for human single item probes

Probe	25th Quartile	Median	75th Quartile
a pedophile	0	0	1.26
a terrorist	0	0	1.58
a rapist	0	0	1.58
an assassin	0	1.26	2.00
a thief	0	1.58	2.00
a dead body	1.11	1.69	2.12
a gang member	1.00	1.70	2.00
a convict	1.00	1.91	2.08
an addict	1.58	2.00	2.34
a Congressman	1.77	2.14	2.51
a celebrity	2.00	2.23	2.69
a waitress	2.00	2.23	2.56
a billionaire	2.00	2.29	2.95
an astronaut	2.14	2.34	2.72
a homeless adult	2.00	2.34	2.51
a judge	2.00	2.34	2.51
a coworker	2.14	2.42	2.69
a college student	2.14	2.46	2.95
a teenager	2.14	2.46	2.95
a person with cancer	2.25	2.48	2.85
a blind person	2.12	2.48	2.80
an adult	2.12	2.48	2.85
a toddler	2.34	2.51	3.16
an elderly person	2.15	2.51	3.13
a kindergartener	2.23	2.51	3.15
a police officer	2.14	2.51	3.07
a teacher	2.23	2.51	3.16
a soldier	2.39	2.56	3.78
a child	2.34	2.67	3.16
an infant	2.34	2.69	3.71
your aunt	2.42	2.75	3.98
your uncle	2.42	2.84	3.98
your cousin	2.42	2.85	3.70
a newborn baby	2.42	2.89	3.98
your friend	2.42	2.95	3.98
your sibling	2.51	3.09	3.98
your father	2.51	3.16	5.01
your grandfather	2.51	3.16	3.98
your grandmother	2.51	3.16	4.90
your mother	2.62	3.16	6.31

Table 2

Quartiles of transformed response data for object single item probes

Probe	**25th Quartile**	**Median**	**75th Quartile**
a rock	0.80	1.11	1.39
a pencil	1.00	1.21	1.48
a baseball	1.00	1.26	1.58
a candle	1.00	1.26	1.58
a stapler	1.12	1.31	1.58
a chocolate bar	1.12	1.33	1.65
a toothbrush	1.07	1.33	1.86
an umbrella	1.17	1.35	1.58
a water bottle	1.26	1.48	1.86
a mailbox	1.26	1.48	1.70
a pair of sunglasses	1.26	1.48	1.74
a key	1.26	1.51	1.86
a raincoat	1.35	1.58	1.86
a teddy bear	1.26	1.58	1.86
a bag of old clothes	1.15	1.58	1.86
a book	1.35	1.58	1.93
a bottle of wine	1.26	1.58	1.86
a pair of shoes	1.41	1.58	1.86
a graduation cap	1.26	1.69	2.12
a diary	1.26	1.70	2.00
a kayak	1.48	1.77	1.95
a baby blanket	1.41	1.86	2.14
a birthday present	1.45	1.86	2.00
a brand new bike	1.53	1.87	2.12
a motorcycle	1.58	1.90	2.14
a passport	1.58	1.97	2.34
your wallet	1.66	1.98	2.34
an antibiotic	1.86	2.00	2.51
your family photograph	1.82	2.00	2.51
a diamond	1.86	2.00	2.34
an artificially intelligent computer	1.69	2.12	2.34
your cellphone	1.87	2.12	2.48
a hydroelectric plant	1.69	2.12	2.59
your laptop	1.93	2.18	2.59
a wedding ring	2.00	2.34	2.59
a year’s supply of antibiotics	2.12	2.34	2.74
your car	2.06	2.34	2.74
a college diploma	2.10	2.51	3.16
$1,000,000 cash	2.18	2.67	3.25
a town’s water reservoir	2.23	2.70	2.95

Table 3

Exploratory Analysis Model Fits

	Predictor	***BIC***	***R²***	***p***
Default Model				
	Group Size	59144.63	0.20	0.42
Perceptual Summing Model				
	Summed ψ_V	53321.83	0.54	< .0001
	+ Category	52610.53	0.55	< .0001
Perceptual Averaging Model				
	Average ψ_V	51437.66	0.52	< .0001
	+ Category	50707.39	0.54	< .0001
Perceptual Averaging Plus Model				
	Average ψ_V	51437.66	0.52	< .0001
	+ Max(ψ_V)	49301.41	0.67	< .0001
	+ Category	49223.85	0.67	< .0001

Table 4

Final, Accepted Model Summary

	Predictor	β	SE	t	*p*	*BIC*	R^2
Exploratory						49288.24	0.67
	Average ψ_V	0.40	0.03	15.13	< .0001		
	Max(ψ_V)	0.62	0.02	30.97	< .0001		
Confirmatory						43560.46	0.71
	Average ψ_V	0.27	0.03	11.88	< .0001		
	Max(ψ_V)	0.66	0.02	36.33	< .0001		

Table 5

Standardized Interquartile Range Data by Group Size

	Group Size	25th Quartile	50th Quartile	75th Quartile	QR Range
Exploratory Analysis					
	2	-0.35	0.06	0.53	0.87
	3	-0.24	0.12	0.53	0.77
	4	-0.40	0.05	0.54	0.94
	5	-0.17	0.24	0.69	0.86
	Total	-0.29	0.12	0.57	0.86
Confirmatory Analysis					
	2	-0.40	-0.002	0.44	0.84
	3	-0.14	0.23	0.71	0.85
	4	-0.20	0.20	0.68	0.88
	5	-0.12	0.22	0.67	0.78
	Total	-0.22	0.16	0.63	0.84