Psychological Value Theory: Predicting Health-Seeking Behavior from Symptom Perception

G. Ryan Page1, Philip Quinlan2, Len Lecci1, and Dale J Cohen1

1 Department of Psychology, The University of North Carolina at Wilmington

2 Department of Psychology, The University of York, UK

Word Count: 10,506

**Declarations**: There are no real, potential, financial, or non-financial conflicts of interest that may impact the current research. The authors did not receive support from any organization for the submitted work. Experiments 1-3 were approved by the UNCW IRB, protocol# 16-0210. Informed consent was obtained from all individual participants included in the study. The data and R code for the analysis of Experiments 1- 3 can be downloaded at: https://github.com/ccpluncw/ccpl\_data\_health2024.git. The experimental presentation code is proprietary and is not open sourced. No experiment was pre-registered.

**Author Contributions**: **G. Ryan Page (0000-0001-8264-2637; ryan@findyourbase.com)**: Conceptualization, Data curation, Funding, Investigation, Methodology, Validation, Writing, Review; **Philip Quinlan (0000-0002-8847-6390; philip.quinlan@york.ac.uk)**: Conceptualization, Software, Writing, Review; **Len Lecci (0000-0003-2284-8320; leccil@uncw.edu)**: Conceptualization, Funding, Methodology, Administration, Resources, Supervision, Review; **Dale J. Cohen (0000-0003-3353-3340; cohend@uncw.edu)**: Conceptualization, Data curation, Analysis, Funding, Investigation, Methodology, Administration, Resources, Software, Supervision, Validation, Visualization, Writing, Review

**Author Note:** 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. Dale J. Cohen. We thank the members of the Cohen Cognition and Perception Laboratory for their assistance and support.

**Abstract**

We recruit Psychological Value Theory (PVT) to understand how symptom value influences health-seeking decisions. Estimates of the Psychological Value of relief from a particular symptom were previously collected and used to predict the speed of participants’ decision and the choice they make in three discrete choice experiments. Experiment 1 presented participants with a scenario and asked them to identify which of two symptoms they would seek healthcare services to treat. For each participant on every trial, two randomly chosen symptoms were inserted into the scenario. Experiment 2 addressed how the Psychological Value of a group of symptoms is predicted from the individual symptoms. Experiment 2 replicated Experiment 1 using groups of two symptoms, and predicted choice based on three grouping functions. Experiment 3 replicated Experiment 2 using a yes/no task, whereby participants were asked if they would pursue a health care visit for a single set of symptoms. The results showed that PVT accurately predicted speed and choice in all three experiments. The Psychological Value of relief from a symptom was the primary driver of choice along with a response bias in favor of avoiding symptoms labeled “severe.” Health-seeking decisions are well modeled by a general-purpose, value-based computational model (PVT), with the Psychological Value of relief from health symptoms as a primary driver of health-seeking behavior.

**Keywords**

health seeking decisions; sequential sampling; symptom appraisal; computational modelling; discrete choice

Psychological Value Theory: Predicting Decisions about Health-Seeking Behavior from Symptoms

Discrete choice experiments have gained traction in health services research over the last 30 years (see e.g., Ryan & Gerard, 2003). In discrete choice experiments, participants are asked to choose between a set of alternatives. For example, in a study by Schulpher et al. (2004), patients were quizzed about their preferences for treatments about prostate cancer. Each question was defined relative to an attribute, such as life expectancy. Patients were asked to choose between options differing on levels of the attribute – here the options were separated by two additional months of life. The results revealed that men exhibited a willingness to trade-off life expectancy against some of the perceived adverse side-effects of treatment (e.g., lack of physical energy). As noted by Ryan (2004), the benefits of using such a methodology are evident especially given the drive (at least in the UK) to encourage the use of patient centered evaluation of health technologies alongside clinical factors and cost effectiveness in determining treatments.

 Here we use a discrete choice paradigm to examine the cognitive processes that underlie healthcare decision making. More specifically, we will examine whether Psychological Value Theory (PVT), a computational cognitive model of value and choice, accurately models the cognitive processes that are assumed to underlie healthcare decision making.

**Symptoms and Value**

Each year, undiagnosed illnesses cause notable morbidity and decreases in quality of life through symptom experience (Beagle et al., 2014; Joseph et al.,1996; Song et al., 2017). In cases of transmissible illness, undiagnosed illness poses risk not only to the individual experiencing symptoms but also to others through potential community spread (Mohr et al., 2021). Whereas some cases of undiagnosed illness may be attributed to unrecognized or unperceived symptoms (Schultz, et al., 1969; Sheifer et al., 2001), other cases are likely attributable to individuals’ decisions to refrain from seeking healthcare for a recognized symptom (Taber et al., 2015).

Safer and colleagues (1979) have identified three distinct stages of healthcare delay: 1) appraisal delay: the time between an individual noticing symptoms and deciding they are ill, 2) illness delay: the time between when one decides they are ill and when they decide to seek professional help, and 3) utilization delay: the time between when the individual decides to seek care and when they are seen by a healthcare provider. At least the first two stages of delay are a direct result of an individual’s health decision-making approach with respect to symptom experience, while the third may be concerned with external factors related to healthcare (i.e., utilization delay may involve clinic scheduling, wait times, etc.). Retrospective reporting among individuals in four hospital clinic waiting rooms suggested symptom pain and bleeding were negatively correlated with appraisal delay while an individual reading about the symptom was positively correlated with appraisal delay. Furthermore, there is some evidence that symptoms that individuals perceive as more severe or having a stronger negative impact on their quality of life have been associated with higher likelihood of diagnosis (Gartland et al., 2019, Sayuk et al., 2017).

Such findings suggest symptom experience is likely a major factor in care-seeking behavior, particularly in situations where individuals have not been diagnosed and symptoms are ambiguous. To understand the role of symptom perception in health seeking behavior, we turn to Psychological Value Theory (PVT). PVT is a computationally explicit model of value and choice that defines value, measures value independent of choice, and uses those measurements to predict choice and the time to respond in discrete choice experiments (Cohen et al., 2022). It is our aim to apply PVT, to measure the value of a symptom and to assess whether that value predicts choices in health-seeking decisions.

Central to PVT is a new theoretical construct, termed *Psychological Value*. Similar to Lancaster (1966), Psychological Value (ψ*v*) is an emergent property of many attributes of a stimulus (see Cohen, et al., 2022 for more formal, quantitative definition). The significance of Psychological Value has emerged with the publication of a series of papers that demonstrate its explanatory power in a large range of discrete choice tasks. Initially the ideas were explored in terms of classical sacrificial moral dilemmas (Cohen & Ahn, 2016) and other forms of economic decisions (Cohen et al., 2022), but more recently the explanatory scope was expanded to encompass various facets of charitable giving (Cohen et al., 2023).

A key assumption in PVT is that the Psychological Value of all stimuli is perceived and like any perception, the perceived Psychological Value of a stimulus will vary from moment to moment – even when the stimulus and environment remains constant (Ashby & Lee, 1993). As such, the Psychological Value of a stimulus corresponds to a distribution of values, rather than a point estimate.

Cohen and colleagues measured Psychological Value using a magnitude estimation task. Magnitude estimation tasks have been shown to provide valid estimates of perceptual events, such as the perceived loudness of a tone (Stevens, 1956). In the magnitude estimation task, participants provide estimates of their perception of a stimulus relative to a standard stimulus. Here, Cohen and colleagues told participants that they were to estimate their personal value (not necessarily the same as monetary value) of various items (known as probes). To do so, the participants were asked to estimate their personal value of probes in relation a chimpanzee, whose personal value was assigned 1000. So, if the probe is twice as valuable as a chimpanzee, assign it a 2000; half as valuable, assign it a 500. Participants were permitted to assign any numbers they pleased, including decimals, fractions, etc.

For a particular probe, the estimate would be akin to a random draw from the participant’s perceived Psychological Value distribution for the corresponding item. If most peoples’ perceived Psychological Values of a given item are relatively consistent, then the responses from all participants would describe the distribution of Psychological Values that represent that item. Although theories in economics assume there is significant between person variability in the value of an item, Cohen et al, 2023 have shown that estimates of Psychological Value distributions from US participants predicts choice for participants in both UK and mainland China.

 Once the Psychological Values of a set of items are estimated, these can then be used to predict choice in discrete choice experiments (see e.g., Cohen & Ahn, 2016; Cohen et al., 2022). To predict choice, PVT models the cognitive decision process computationally as a sequential sampling procedure (henceforth SSPs) (Busemeyer et al., 2019; Krajbich, 2019; Krajbich et al., 2010) based around a random walk process (see e.g., Cohen & Quinlan, 2016; Link & Heath, 1975; Ratcliff, 1978; Ratcliff & Rouder, 1998). More particularly, Cohen et al., (2022) developed a so-called Robust Random Walk model (henceforth RRW) as a (mostly) non-parametric, value-based sequential sampling procedure (VSSP) that instantiates the assumptions of PVT (see Cohen et al., 2022). Similar to other SSPs, evidence in favor of each option accumulates in an incremental stepwise fashion. Each option has a corresponding evidence threshold and the time it takes for a threshold to be reached is taken to reflect the time it takes the observer to reach a decision (it is taken as a proxy for their reaction time, henceforth, RT). Which threshold is reached is also taken to reflect the chosen option.

PVT assumes that when forced to choose between two options, participants will attempt to choose the option associated with highest Psychological Value. Therefore, the options are referred to the higher valued option (HVO) and the lower valued option (LVO) as defined by the estimates of each option’s Psychological Value. The RRW simultaneously predicts the probability of choosing the higher valued option, *p*(HVO), the time of response when the higher valued option is chosen, RTHVO, and the time of response when the lower valued option is chosen, RTLVO. Like traditional random walk processes, the RRW accumulates evidence over time until one of two thresholds is reached. In contrast to traditional random walk processes, the RRW uses the Psychological Value distributional overlap of two options as a direct measure of drift rate (rather than estimating drift rate from the data). This provides strong, independent constraints on the model’s ability to fit data.

When the distributional overlap of two options is large, PVT predicts that participants’ responses will be slow and error prone: error prone in the sense that there will be an increased tendency to choose the LVO rather than the HVO. In contrast, when the distributional overlap of two options is small, PVT predicts that participants’ responses will be fast and accurate. More critically PVT predicts the functional forms of both RT and the probability of choosing the HVO (i.e., *p*(HVO)) (see e.g., Cohen et al, 2022).

In PVT (and other SSPs), choice is not a direct function of value. Rather response choice is influenced *both* by the overlap of the Psychological Value distributions and *response biases*. One particularly influential response bias is the position of the start point. A start point bias models a pre-choice tendency to default to one of the two options. For example, a participant may tend to choose the option presented on the left. In this instance, evidence accumulation begins closer to the left option. This manifests as faster left option responses (because less accumulation is required) and more errors when the lower valued option is on the left. As such, the start point bias influences the pattern of response choices but does not influence the latent value distributions.

PVT is detailed, highly constrained, and provides excellent concurrent fits of the time to respond to the chosen option and the probability of choosing that option. For instance, in a recent paper on sacrificial moral dilemmas, Cohen et al., (2023) showed that estimates of Psychological Value taken from USA undergraduates were able to predict with high accuracy moral judgments made, separately, by UK and by Chinese undergraduates. PVT was also successfully employed to predict charitable giving (Cohen, et al., 2023) and economic decisions (Cohen et al, 2022).

Here, we assess whether PVT can provide insights into how symptom perception influences healthcare seeking behavior. To predict health seeking behavior a priori, we measure the Psychological Value of symptoms. We propose that the Psychological Value associated with health symptoms lies not with the symptom itself, rather, the value is associated with *not having* the symptom. Suppose, for example, one is experiencing a headache. The headache itself is not valued; rather, any value lies in relief from the headache. As such, we propose that the value of not having a health symptom, in part, drives the decision to seek care. To test our proposal, we first measure the value of not having a symptom independent of choice, and then use that measurement to predict choices about seeking care.

The current study comprises discrete choice experiments that challenge healthy participant with hypothetical scenarios; they do not examine the actual decisions patients make regarding healthcare (cf. Pieterse & de Vries, 2013). Of course, the aspiration is that the results of the experiments can inform more real-world forms of medical decision making. Our basic claim is that fundamentally, decisions regarding healthcare choices are driven by a desire to achieve the option most highly valued by the person making the decision. The current work is offered as an initial attempt to understand the basic cognitive decisional processes underlying medical decisions. As a first step in this endeavor, discrete choice laboratory experiments stand as a useful means by which the preliminary relevant data can be collected.

First, we collected estimates of the Psychological Value for *not* having various symptoms (see Supplemental Information for details). In Experiment 1, we use these estimates to predict performance in a timed discrete choice task. Specifically, participants were presented, on each trial, with a scenario asking them for which of two symptoms they would seek medical care. Both the speed of response and the chosen option were modelled by PVT.

Experiment 2 assesses how multiple symptoms influence health seeking behavior by presenting participants with choices involving sets of differing numbers of symptoms. The Axiom of Monotonicity (Cohen et al., 2022) would hypothesize that the value of a group is the sum of the individual values (or some function of the sum). Therefore, the Psychological Value of not having two symptoms would be greater than the Psychological Value of not have one of those symptoms. However, Barberio and Cohen (Under Review) have shown in other contexts that the Psychological Value of a group is the average of the individual symptoms plus some extra influence of the highest valued symptom. Therefore, when a high valued symptom is paired with a lower valued symptom, the value of the group is less than that of the highest valued symptom.

In Experiments 1 and 2, we presented participants with forced choice experiments. However, people are more often faced with a symptom and must decide whether to seek care for the symptom. In Experiment 3, we assess whether PVT can predict health seeking behavior when participants are asked, on each trial, whether or not they would seek help given the presence of a particular symptom set.

**Experiment 1**

 In preliminary research, we collected estimates of the Psychological Value of 82 health symptom probes from a large sample of undergraduate participants (see Supplemental Information). The primary aim in Experiment 1 was to assess whether these estimates accurately predict choice. Novel participants were presented with a scenario asking them which of two symptoms would they seek medical care for. For each participant on every trial, two randomly chosen symptoms were inserted into the scenario. PVT predicts a positive linear relationship between distributional overlap and RT, and a negative exponential relationship between distributional overlap and probability of seeking care for the symptom probe with the higher Psychological Value (termed *p*(HVO)).

**Methods**

***Participants***

One hundred and fifteen naive undergraduate students participated in exchange for course credit. After Cohen and Ahn (2016), we set a minimum sample size of 100, estimated the time necessary to collect the number of participants, posted available experimental time slots to match the estimated time, and ran all participants who signed up (i.e., there were no exclusion criteria).

***Stimuli***

 In a preliminary study, 41 health symptoms were derived from the “Symptoms A-Z” page on webmd.com. These 41 symptoms were paired with two magnitude modifiers: “mild” or “severe,” producing a total of 82 probes. For example, this created two levels for “difficulty seeing”: “mild difficulty seeing” and “severe difficulty seeing.” Finally, because we were interested in the Psychological Value of not experiencing the symptom in their day-to-day life, we added the modifier, “not having (a/an)” to all probes (e.g., “not having mild difficulty seeing”).

 Two hundred and forty-six participants provided estimates of Psychological Value of the health probes (see Supplemental Information for details). Briefly, participants were instructed that they would be presented with a standard and a probe in each trial and to estimate the personal value (i.e., psychological value) of the probe. Consistent with the magnitude estimation procedure, participants were told to assign a value in relation to the standard (a chimpanzee, personal value = 1000).

 The health probes ultimately used in Experiment 1 had to meet two criteria: (i) the health probes had to be less than 22 characters in length to facilitate reading time, and (ii) the health probes had to be logically appropriate for the scenario. Fifty-six of the of the 82 health probes met both criteria. Of the 82 health probes, approximately 10% were excluded because they did not meet Criteria 1 and approximately 20% were excluded because they did not meet Criteria 2.

 To assess the influence of Psychological Value on health decisions, the current study asked participants to make a choice between seeking treatment for one of two symptoms. To control for factors which may influence health decisions, the cost, time required, and efficacy of the health behavior were held constant for all probes. The scenario presented is as follows:

*Assume the following symptoms are unrelated:*

Symptom A

*and*

Symptom B.

*Medical care for each symptom*

*has a 90% success rate, costs $25, and takes 1 hour.*

*Would you be more likely to seek medical care for*

Symptom A

*or*

Symptom B?

Similar to Cohen and Ahn (2016), we reduced the influence of reading time from participants’ RT by initially presenting the scenario in full, with Symptom A and Symptom B masked. Symptoms A and B were masked by replacing each character in Symptom A with a teal “+,”and each character in Symptom B with a teal “=.” The masks remained visible for 300 ms per word in the scenario in the practice trials and 150 ms per word in the scenario in the experimental trials. On the left of the scenario was a vertical progress bar. The progress bar served as a countdown timer that filled as 12.6 seconds time completed. The masked scenario was intended to allow participants to read and understand the scenario, without being able to make a choice because the symptoms were not visible.

Once the appropriate time elapsed, the progress bar turned from white to blue and the masked probes were replaced by the health symptom probes. The participant was to indicate, as quickly as possible, which probe they would more likely seek treatment with a button press.

***Procedure and Design***

 Participants were tested in a small, dark room containing a Dell LCD monitor controlled by a Dell Optiplex 9010 PC running Windows 10. . All instructions were presented in text form on the computer. Participants were provided instructions that they would be presented the masked version of the scenario for a fixed amount of time before it was unmasked. Participants were instructed to read the scenario prior to unmasking, after which they would respond as quickly as possible following the unmasking of the symptom probes.

 Each trial consisted of a 500ms red fixation cross point, followed by the masked scenario. When the masking time ran out, the items in the scenario was unmasked and the entire unmasked scenario remained visible until the participants indicated his or her response by pressing a keyboard key. For half of the participants, the “d” key indicated they would seek treatment for the top symptom in the scenario and “k” indicated they would seek treatment for the bottom symptom in the scenario, with the keys being reversed for the remaining participants. RT was calculated from the unmasking of the items to the button press. The next trial was shown immediately after participants indicated their response for a given trial.

 The experiment consisted of eight practice trials and as many experimental trials as could be completed in 45 minutes. Practice trials were identical to experimental trials except they allowed more reading time before unmasking, so participants could become familiar with the scenario. After the practice trials, participants were presented with a dialogue box reminding them which key indicated the top symptom and which key indicated the bottom symptom. When participants clicked “ok” on this box, a second dialogue box indicated experimental trials were about to begin. During experimental trials, participants were prompted with dialogue boxes every fifteen minutes indicating they may take a self-timed break during which they could resume experimental trials as soon as they wished.

 During experimental trials, Symptoms A and B were randomly selected for each trial from the list of 56 health symptom probes. A dialogue box thanked participants for their participation when the experiment was complete, and all participants were given experimental credit in exchange for their time.

**Results**

***Psychological Value Symptoms***

 For every possible pair of the 56 health symptoms used in Experiment 1, we calculated distributional overlap using the same non-parametric bootstrap procedure as Cohen and Ahn (2016). Overlap has a range of 0 to 1, with 0 indicating no overlap and 1 indicating complete overlap. Overlap for the symptom set, ranged from 0.44 to 1.0, with a median of 0.86, with (not having) “severe” symptoms typically being valued higher than “mild” symptoms, paired t(27) = 6.06, p < 0.001.

 We also calculated Directional Overlap, which identified the HVO in relation to the LVO. Directional Overlap ranges from -1 to 1. When comparing Symptom A with Symptom B, -1 indicates there is no overlap and Symptom A is the LVO. A 0 indicates Symptom A and Symptom B overlap completely. A 1 indicates there is no overlap and Symptom A is the HVO. Directional Overlap is a continuum, so intermediate values indicate intermediate overlaps. For each symptom, we calculated their average DO across all other symptoms (Figure 1).



**Figure 1.** Average Directional Overlaps for Mild vs. Severe Symptom Probes as Compared to All Other PVT Probes. The symptoms on the *x*-axis are arranged from lowest average Psychological Value on the left to highest average Psychological Value on the right. As can be seen, not having a “severe” symptom has greater Psychological Value than not having a “mild” symptom.

***Choice***

 Three participants were removed prior to analysis because they were not fluent English speakers/readers, or they indicated that they did not understand the instructions. For the remaining analyses, we followed the same analysis pipeline as Cohen et. al. (2022). For every participant, we calculated the RT Coefficient of Variation (CVRT = sRT/mRT). CVRT is a measure of variability in responding and high values indicate inattention. We automatically removed the 5% of the participants who had the greatest CVRT (six participants). We also removed participants who responded at or below chance, indicating a lack of attention or confusion concerning the key mapping (twelve participants). The remaining 94 participants were analyzed.

 We filtered individual outlier trials by automatically removing the fastest and slowest 2.5% of trials for each level of Overlap. This process does not assume a distributional shape. Finally, for each participant, we removed the effects of practice from RTs by taking the residuals of an exponential decay function fit to log(RT) by trial number. We used these residuals (RTres) as our measure of RT. On average, participants completed 233 (*SD* = 37) experimental trials.

 To provide enough data to stabilize RT and *p*(HV0), distributional overlap was rounded to the nearest 0.05 (O0.05). Cohen and colleagues (2022) typically rounded distributional overlap to the nearest 0.1. However, because the range of distributional overlaps of the health symptoms was truncated (between 0.425 and 1.0), we reduced the bin size to 0.05.

 Prior to analyzing the data, we coded a variable, termed *Severity-HVO* that identified whether the HVO was severe or mild (see Supplemental Information). We then summarized the data by calculating the *p*(HVO), mean RTHVO, and mean RTLVO for Overlap X Severity-HVO. Conditions with fewer than 40 trials were excluded to remove estimates with uncertain reliability.

 To assess whether PVT accurately predicts participants’ healthcare seeking behavior, we modeled participants’ responses using the RRW. We compared how well four RRW models predicted the choice data. Across the models we examined changes in the start point, changes in the values of the symptoms and a combination of both. All four models included the following four free parameters: the boundary separation (*b*), the SD of the added noise (*NSD*), the degree of primacy/recency influence (*dB*), and the non-decision time, (*TER*) (see Supplemental Information). The models differed with respect to free parameters (effect coding for each parameter identified below is described in Supplemental Information).

* **Simple Model**: The simple model contained only the four free parameters described above. The remaining parameters of the model were fixed at default values (e.g., 0 which excludes the parameter from the model).
* **Severity Bias Model**: This model is identical to the simple model, with the exception that it adds a free parameter that captures a start point bias as a function of symptom severity.
* **Severity Value Model**: This model is identical to the simple model, with the exception that it adds a free parameter that captures value changes as a function of symptom severity.
* **Severity Bias + Value Model**: This model is identical to the simple model, with the exception that it adds two free parameters: one that captures start point effect as a function of symptom severity and one that captures value changes as a function of symptom severity. This model contained the effect coded variables present in the Severity Bias Model and the Severity Value Model.

A Smart Grid Search algorithm was used to optimize the parameters (see Cohen et al., 2022). Because the RRW is stochastic, different model runs result in slight variability to fits. As such, the best fit model was run 40 times and an average predicted RT and *p*(HVO) for each Overlap together with the *r2* and BIC of that average predicted fit was calculated. To determine the best fit model, we compared the Bayesian Information Criterion (BIC) of each model (lower is better). When the BICs were essentially equivalent (within 1 or 2 points), we accepted the simpler model (i.e., with fewer free parameters) as the superior model based on parsimony.

The behavioral data are presented in Figure 2. The fit of each model is presented in Table 1 and the parameter values for the best fit model are presented in Table 2. As can be seen in Table 1, the Simple Model fits poorly because there was a large influence of symptom severity (which the Simple Model does not take into account). The best fit model accounts for the influence of symptom severity as a start point bias, the Severity Bias Model (r2 = 0.89, BIC = 70). The symptom severity start point parameter was quite large (ssev = 0.76), indicating that evidence begins to accumulate about 76% closer to the “severe boundary” than is the case with an unbiased start point. This indicates that participants are much more likely to go to a doctor when the symptom is labeled as severe even in those instances where not having the severe symptom is *less* valuable than not having the mild symptom (e.g., not having mild memory problems vs not having severe daytime drowsiness). Finally, because the best fit model did not include a symptom severity value change parameter, the influence of the severity label on value was likely well estimated by the magnitude estimation task.



Figure 2: Data points and model fits for p(HVO) (upper panel) and RTs (lower panel) as a function of Overlap for Experiment 1. Lower panel left shows data and fits for RTHVO and lower panel right shows data and fits for RTLVO.

**Table 1**

*The fit statistics for all PVT models fit to the data from Experiment 1*

|  |
| --- |
| **Experiment 1** |
| Model | FreeParameters | *r2* | BIC | Best Fit |
| Simple | 4 | 0.21 | 248 |  |
| *Severity Bias* | *5* | *0.89* | *70* | ***✓*** |
| Severity Value  | 5 | 0.85 | 87 |  |
| Severity Bias + Value  | 6 | 0.90 | 73 |  |

**Table 2**

*The free parameter values of the best fit PVT models for Experiments 1-3.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | **Experiment 1** |  |  |
|  |  | Free Parameters |  | Model Fit |
|  |  | Ter | b | dB | nSD | ssev  |  | r2 | BIC |
| Single Symptom |  | -0.52 | 127 | 0.18 | 1.32 | 0.76 |  | 0.89 | 70 |
|  |  | **Experiment 2** |  |  |  |
|  | With value correction (vc) parameter |  |  |  |
| Integration Algorithm |  | Ter | b | dB | nSD | ssev  | bSQ | vcSQ |  | r2 | BIC |
| Sum |  | -0.71 | 178 | 0.05 | 3.08 | 0.46 | 96 | 2.38 |  | 0.88 | 146 |
| Average |  | -0.69 | 133 | 0.03 | 3.00 | 0.39 | 58 | -0.38 |  | 0.89 | 195 |
| Biased Average |  | -0.82 | 156 | 0.04 | 2.39 | 0.39 | 71 | -0.28 |  | 0.88 | 211 |
|  | Without value correction (vc) parameter |  |  |  |
| Integration Algorithm |  | Ter | b | dB | nSD | ssev  | bSQ | vcSQ |  | r2 | BIC |
| Sum |  | -0.37 | 122 | 0.18 | 5.07 | 0.75 | 91 | - |  | 0.60 | 321 |
| Average |  | -0.67 | 160 | 0.03 | 2.98 | 0.45 | 79 | - |  | 0.73 | 382 |
| Biased Average |  | -0.78 | 151 | 0.06 | 2.20 | 0.46 | 76 | - |  | 0.77 | 352 |
|  |  | **Experiment 3** |  |  |  |
|  | Free Parameters |  |  |  |
|  |  | Ter | b | dB | nSD | ssev  | bSQ | vcSQ |  | r2 | BIC |
| Yes/No |  | -0.37 | 16 | 0.12 | 1.41 | 0.49 | 83 | - |  | 0.90 | 91 |

**Discussion**

The data from Experiment 1 demonstrates that PVT can accurately predict health decision-making. When faced with a choice to seek medical treatment for one of two symptoms, participants consistently indicated they would be more likely to seek treatment for the symptom from which relief had the highest Psychological Value. Consistent with the prediction of PVT, as the similarity of the Psychological Values of two symptoms increased, participants became less adept in identifying and choosing the more highly valued symptom. This was evidenced by increased RTs and decreased *p*(HVO).

In addition to the influence of Psychological Value, the data show that participants had a response bias towards the symptom with a “severe” label. This response bias was modeled in PVT by moving the starting point for the accumulation of evidence towards the symptom with the “severe” label (the start point was about 76% closer to the severe option than an unbiased start point). This bias is present prior to seeing the choice options (like a head start in a running race) and does not reflect a change in the value of the symptoms. This suggests that labeling a symptom as severe can prompt decision makers to seek medical help. We will discuss this in more detail in the General Discussion.

In Experiment 2, we assess whether the results of Experiment 1 generalize to cases where a person must decide about multiple symptoms simultaneously.

**Experiment 2**

Most health events are composed of a combination of health symptoms. It is the combination of health symptoms that serves as the foundational perceptual event indicating the increased likelihood of an illness that is in need of care (Rosenstock, 1974; Safer et al., 1979). Experiment 2 is intended to assess whether (i) the Psychological Value of a set of symptoms is a function of the Psychological Value of the individual symptoms, and (ii) whether PVT can predict choice in multi-symptom cases.

For PVT to predict choice in a multi-symptom set, one must have an accurate estimate of the value of the set of symptoms. With human lives and objects as stimuli, Barberio and Cohen (Under Review) showed that the Psychological Value of the set can be well predicted by the Psychological Value of the items in the set. Here, we examine whether Barberio and Cohen’s (Under Review) finding generalizes to health decisions which demand consideration of competing sets of symptoms. Experiment 2 is identical to Experiment 1, with the exception that we ask participants to indicate for which of two symptom sets they would be more likely to seek care. Each symptom set is composed of either one or two symptoms: there were three types of trials - one symptom vs. one symptom, two symptoms vs two symptoms, and one symptom vs. two symptoms.

We estimated the Psychological Value of each set of symptoms using three grouping functions: *Summing*, whereby the sum of the Psychological Values of the items in the set is taken; *Averaging*, whereby the average of the Psychological Value of the items in the set is taken, and *Biased Averaging*, whereby the average of the Psychological Value of the items in the set is taken but the symptom in the set with the maximum Psychological Value is given added weight. We then used these estimates of Psychological Value of the set as input to the RRW. To estimate any degree of over/underestimation of the functions, we ran RRW models with a value correction parameter. The magnitude of the value correction parameter quantifies the over/under estimation of each function. Critically, though, the function that most accurately estimates the Psychological Value of the set of symptoms will fit the choice data without a value correction parameter.

**Method**

 **Participants**

In-lab data collection, in a similar environment to that used previously, began in the spring of 2020. However, face-to-face testing had to be abandoned due to COVID restrictions. The experiment therefore was redesigned for online data collection, using *testable.com*, and a bespoke javascript experiment, both of which participants were able to access through the university’s SONA pool. Prior to abandoning face-to-face testing, 95 naïve undergraduate participants completed an in-person version of the experiment. Subsequently, 149 naïve undergraduate participants volunteered their participation in exchange for course credit through the university’s SONA pool. The sample was approximately 70% female and 80% Caucasian.

 **Apparatus and Stimuli**

Initial data collection relied on apparatus and stimuli virtually identical to that used previously. For the on-line version of the experiment the stimuli were presented via weblink, accessed from the participants’ personal computer, which took them to a webpage hosting the experiment’s instructions and experimental trials. Participants were only able to access the link via laptop or desktop computer (i.e., they could not access the link via smartphone or any other internet-connected mobile device).

 Similar to Experiment 1, each trial presented the participant with two options. In Experiment 2, each option could consist of a single symptom or a pair of symptoms. As such, there were three types of trials: both options containing a single symptom, both options containing a pair of symptoms, and one option containing a single symptom and one option containing a pair of symptoms. For every trial of each participant, probes were selected without replacement from the 56 individual symptom probes from Experiment 1.

Participants were asked to make a choice between seeking care for one of two sets of symptoms that represent distinct groups of one or two individual health symptoms. The scenario presented was identical to that in Experiment 1, with the exception that the first statement was changed to “Assume the following symptom groups are unrelated.”

 **Procedure**

The in-lab procedure was identical to Experiment 1. There were two internet versions: one hosted on Testable.org, and one using a bespoke javascript platform. The internet versions were created to be virtually identical to the in-lab version. They are described in detail in Supplemental Information.

Each trial consisted of a 500ms red fixation cross point, intended to draw the participant’s attention to the center of their computer screen, followed by the scenario. For half of the participants, the “d” key indicated intention to seek treatment for the symptoms in the top group in the scenario and the “k” indicated intention to seek treatment for the symptoms in the bottom group, with the keys being reversed for the remaining participants. Reaction time was measured until the button press. The next trial was shown 500ms after participants indicated their response for a given trial.

The in-lab experiment consisted of 300 experimental trials or as many trials as the participant could complete in 45 minutes. The online experiments consisted of 140 trials. After five trials, participants were presented with a screen reminding them which key indicates the symptoms in the top group and which key indicates the symptoms in the bottom group. When participants indicated they understood the reminder they resumed the experiment.

During experimental trials each group of symptoms was randomly selected from the 3,248 possible symptom pair probes (and the single symptom groups were randomly selected from the 56 individual symptoms). Halfway through the experimental trials, participants were prompted with a screen encouraging them to take a self-timed break and resume the experiment when they were ready to continue. At the end of the experiment, a dialogue box appeared thanking participants for their participation and letting them know they will receive credit for their participation.

**Results**

 The same analysis pipeline used in Experiment 1 was performed for Experiment 2. We removed the data from 24 participants (9.84% of the sample) due to response sets indicating they did not follow the experimental instructions (i.e., responding at or below chance). In addition, the 5% of participants with the greatest CVRT were removed (i.e., 13). The remaining 207 participants were analyzed. We removed the highest and lowest 2.5% of trials in each Overlap condition. Finally, similar to Experiment 1, we removed the influence of learning. On average, participants completed 161.71 (*SD* = 32.38) trials in the allotted time.

**In lab vs online data analysis**

 The current study had the unique opportunity to compare in-person to online data. As such, data from in-lab, Testable-hosted, and in house javascript conditions were compared. This analysis can be found in the Supplemental Information. In sum, although online responses were overall slower than in-lab responses, the patterns of data were not significantly different. As such, we concluded that both testable.com and our inhouse javascript platform did not introduce excess noise or bias to the data. Therefore, we combined all the data for the remaining analyses.

**PVT Analysis**

We first estimated the Psychological Values of all multi-symptom sets using each of the three multi-symptom estimation functions (henceforth Grouping functions): Summing, Averaging, and Biased Averaging. We then calculated the Overlap of sets presented to participants using these estimates. We then used the estimated Overlap from each function to predict choice using the RRW model. The function that most accurately estimates the Psychological Value of a set of symptoms should fit the choice data well without the addition of a value change parameter. As in Experiment 1, we rounded Overlap to the 0.05.

For each Grouping function, we ran the following RRW analyses.

First, we coded two variables: *Severity\_HVO* and *Quantity\_HVO* (see Supplemental Information). *Severity\_HVO* identified the whether the HVO was severe or mild. *Quantity\_HVO* categorized each trial by the number of symptoms in each set (1 or 2) and whether each set was the higher valued option (HVO) or lower valued option (LVO) in the decision. Then we conducted a series of RRW models (effect coding for each parameter identified below is described in Supplemental Information).

 **RRW Models**

Four model variants of PVT were compared.

* **Severity Bias Model**: Because Experiment 1 identified a strong start point bias for Symptom Severity, we adopted the Severity Bias Model from Experiment 1 as our base model.
* **Severity Bias + Symptom Quantity Value Model**: This model is identical to the Severity Bias Model, with the exception that it adds a free parameter that captures value changes as a function of the number of symptoms in the set. This model will identify how accurately the Grouping function estimated the Psychological Value of multiple symptoms. Here, a positive vc value indicates that the Psychological Values were overestimated, a negative vc value indicates that the Psychological Values were underestimated, and a small or 0 vc value indicates that the Psychological Values were accurate. Therefore, the closer the value change parameter is to zero, the more accurately the Grouping function estimated the value of multiple symptoms.
* **Severity Bias + Symptom Quantity Boundary Model**. Here, we add a free parameter to assess whether there is a shift in the boundary as the number of symptoms increases. In addition to adding time to read and integrate the presented symptoms, a shift in the boundary is assumed to reflect a change in weighing speed against accuracy with wider boundaries suggesting the participants are weighing accuracy over speed (Cohen et al., 2022).
* **Severity Bias + Symptom Quantity Value + Boundary Model**: This model is identical to the Severity Bias Model, with the exception that it adds two free parameters: the Symptom Quantity value change parameter and the Symptom Quantity boundary parameter**.**

Prior to running the RRW models, we collapsed the data across participants and calculated *p*(HVO) and RT for each Quantity\_HVO type X Severity\_HVO X Overlap bin. As in Experiment 1, we used the Smart Grid Search algorithm to optimize the parameters and removed conditions with fewer than 40 trials. Finally, the best fit model was run 40 times and an average predicted RT and *p*(HVO) for each Overlap together with the *r2* and BIC of that average predicted fit was calculated. Within each Grouping function, the model with the lowest BIC was identified as the best fit model. If multiple models were equivalent with respect to *r2* and BIC values, the simpler model was favored based on parsimony.Importantly, one cannot compare BIC values *across* Grouping functions because the BIC is sensitive to the number of data points and there is no guarantee that the number of data points will be constant across Grouping functions.

The value correction parameter for symptom quantity (vcSQ) is the primary parameter used to identify the accuracy of a Grouping function. An accurate Grouping function will require little or no value correction for symptom quantity. Therefore, the best fit model of the most accurate Grouping function will have a value change parameter for symptom quantity close to zero (vcSQ ≈ 0). Because an accurate Grouping function will need little or no value correction, the most accurate Grouping function should also explain the most variance (r2) in the best fit model that *does not* contain a value change parameter for symptom quantity.

For each Grouping function, Table 3 presents the fit statistics for each PVT model (see Supplemental Information for figures displaying the model fit and data). For each Grouping function, the best fit model is the Severity Bias + SQ Value + SQ Boundary model. This model was an excellent fit (r2 ≈ 0.88) for all three Grouping functions. Table 2 presents the best fit parameter values for each model. In addition to the Severity Bias, all Grouping functions revealed an influence of symptom quantity on the boundary parameter. The boundary parameter describes the amount of information required before making a decision. This is manifested in the data as longer RTs and slightly more accurate responses. The “*b*” parameter specifies the boundary spread when there are three symptoms to process (1 vs 2). The *bSQ* parameter identifies the shift in the boundary as a function of total number of symptoms to process. The value is added to the *b* parameter when there are 4 symptoms to process and subtracted when there are only 2 symptoms to process. For all three Grouping function models, the *bSQ* parameter was about 50% of the *b* parameter.

The value change parameter (vcSQ) quantifies the over/underestimation of the associated Grouping function. The value change parameter assumes Psychological Value distribution of each group of symptoms is Gaussian with equal variance. Its value describes the change in the overlap of the two distributions required to fit the data well. The value change parameter is scaled in SD units. A positive value change parameter indicates that the Grouping function overestimated the Psychological Value of the two-symptom group. The value change parameter corrects this overestimation by shifting the two distributions together. A negative value change parameter indicates that the Grouping function underestimated the Psychological Value of the two-symptom group. The value change parameter corrects this underestimation by shifting the two distributions apart.

**Table 3**

*The fit statistics for all PVT models fit to the data for each Grouping Function in Experiment 2*

|  |
| --- |
| **Experiment 2****Summation Grouping Function** |
| Model | FreeParameters | *r2* | BIC | Best Fit |
| Severity Bias | 5 | 0.20 | 462 |  |
| Severity Bias + SQ Boundary | 6 | 0.60 | 321 |  |
| Severity Bias + SQ Value  | 6 | 0.43 | 420 |  |
| *Severity Bias + SQ Value + SQ Boundary* | *7* | *0.88* | *146* | ***✓*** |
| **Average Grouping Function** |
| Model | Free Parameters | *r2* | BIC | Best Fit |
| Severity Bias | 5 | 0.45 | 559 |  |
| Severity Bias + SQ Boundary | 6 | 0.73 | 382 |  |
| Severity Bias + SQ Value  | 6 | 0.59 | 492 |  |
| *Severity Bias + SQ Value + SQ Boundary* | *7* | *0.89* | *195* | ***✓*** |
| **Biased Average Grouping Function** |
| Model | Free Parameters | *r2* | BIC | Best Fit |
| Severity Bias | 5 | 0.50 | 554 |  |
| Severity Bias + SQ Boundary | 6 | 0.77 | 352 |  |
| Severity Bias + SQ Value  | 6 | 0.58 | 508 |  |
| *Severity Bias + SQ Value + SQ Boundary* | *7* | *0.88* | *211* | ***✓*** |

The data reveal that the Summing Grouping function was the least accurate and the Biased Average Grouping function was the most accurate. The value change parameter, vcSQ, indicated that the Summing Grouping function overestimated the Psychological Value of the two-symptom group by about 2.4 SDs (vcSQ = 2.38), the Average Grouping function underestimated the Psychological Value of the two-symptom group by about 0.4 SDs (vcSQ = ‑0.38), and the Biased Average Grouping function underestimated the Psychological Value of the two-symptom group by about 0.3 SDs (vcSQ = ‑0.28).

These results were confirmed by examining the amount of variance accounted for by the best fit model that does not contain a value change parameter. For all three Grouping functions, the best fit model that does not contain a value change parameter was the Severity Bias + SQ Boundary model. Again, the Summing Grouping function performed the worst (r2 = 0.6) and the Biased Average Grouping function performed the best (r2 = 0.77).

**Discussion**

 The present data provide further evidence showing that when faced with the choice to seek medical treatment for one of two sets of symptoms, participants’ choices were well-predicted by the relative Psychological Value of symptom set. Furthermore, the data also were well-accounted for by the Biased Average Grouping Function (Barberio & Cohen, under review). The most accurate model (i.e., the Severity Bias + SQ Value + SQ Boundary model) replicates the evidence in Experiment 1 for a large severity start point bias (the Severity Bias parameter). The presence of the SQ Value parameter provides a correction to the model such that without it the model slightly underestimated the Psychological Value of the two-symptom group. Finally, the model also includes a boundary parameter that takes account of the amount of information to be considered reflected in the number of symptoms in the set. Time is added as the number of symptoms increases in a way that it is independent of the valuation process. Increasing the boundary separation as a function of the number of symptoms that need to be considered suggests that participants are less inclined to respond quickly and more inclined to respond accurately (cf. Cohen et al., 2022, p. 2041) as the number of symptoms increase. This may well explain some of the mixed findings regarding the degree to which symptom number influences health behavior (Altice & Madigan, 2012; Elnegaard, et al., 2015; Heaton et al., 1992; Ludwig & Gibson, 1992; Macfarlane et al., 2003).

**Experiment 3**

 In Experiments 1 and 2, participants were asked to choose between which of two sets of symptoms they would be more likely to seek healthcare. People, however, are not typically confronted by such a choice. Rather, people are more likely to consider whether to seek care given the presence of a particular set of symptoms. Experiment 3 addresses this issue by assessing whether PVT can predict “yes/no” healthcare decisions when confronted with different kinds of sets of symptoms. The same set of probes from Experiment 2 were used here.

**Method**

 **Participants**

Seventy-eight naïve undergraduate participants volunteered their participation in exchange for course credit through the university’s SONA pool. The sample was approximately 70% female and 80% Caucasian. The number of timeslots were posted; any participant who signed up was run (i.e., there were no exclusion criteria).

 **Apparatus and Stimuli**

The protocol was virtually identical to that used in Experiment 2’s online bespoke javacript procedure. For each trial x participant, the symptom set comprising an individual symptom or a pair of symptoms pairs that were randomly selected without replacement from the list of all possible probes.

 Participants were asked to make a choice as to whether or not they would seek care for the symptom set. The scenario presented was as follows:

*You have the following symptoms:*

Symptom set

*Medical care for the symptom(s) has a 90% success rate,*

*costs $25, and takes 1 hour.*

*Would you seek medical care for*

Symptom set*?*

 **Procedure**

The experiment was carried out using the internal javascript protocol described in Experiment 2. Participants accessed the study through a weblink provided through the university’s student research participation portal or an emailed weblink distributed to those in summer courses interested in participating in exchange for extra credit. The procedure was identical to that outlined in Experiment 2 with the exception that participants were presented a scenario with a single symptom set and asked whether they would seek medical care for that symptom set. Participants were randomly assigned to one of two conditions, one in which the “d” key indicated a “yes” response and the “k” key indicated a “no” response; in the second condition in which the keys were reversed. The experiment consisted of 200 experimental trials.

**Results**

The same analysis pipeline described in Experiment 2 was used here. The 5% of participants with the greatest CVRT were removed (i.e., 4) and two participants were removed because they responded at or below the chance (i.e., indicating insufficient effort or difficulty in understanding the experimental task). The remaining 72 participants were analyzed (i.e., 92.3% of the total sample). Additionally, 5% of individual trials (i.e., the highest and lowest 2.5% of trials for each overlap bin) were removed. Finally, as before, we removed the influence of learning. Given the results of Experiment 2, we estimated the Psychological Value of multi-symptom sets using the Biased Average Grouping function.

Within the context of PVT, people make yes/no decisions by comparing the Psychological Value of the probe to a threshold. When the Psychological Value of the probe is greater than the threshold, a “yes” response is produced. When the Psychological Value of the probe is less than the threshold, a “no” response is produced. Unlike SDT, the threshold in PVT is described as a distribution. As such, choice and RT are a function of the Overlap between the probe and the threshold distribution. To identify the latent threshold distribution we generated a threshold distribution and used the value correction parameter to shift it relative to the Psychological Values of the probes. To create the threshold distribution, we created two vectors: the median and the standard deviation of the Psychological Value distributions of every health symptom item used in the experiment. The threshold distribution was generated by randomly selecting values from a Gaussian distribution whose mean and SD was the median of those two vectors (respectively).

For every set of symptoms, we calculated distributional overlap between that set of symptoms and the threshold distribution using the same procedure as Experiment 1. Similar to Experiments 2 and 3, distributional overlap was rounded to the nearest 0.05 (O0.o5). We then coded three new variables:, Severity\_HVO, and Quantity\_HVO, and Reference\_HVO (see Supplemental Information). Severity\_HVO and Quantity\_HVO were coded similar to Experiment 2. Reference\_HVO identifying whether the threshold distribution was the HVO (*refHVO*) or LVO (*refLVO*) for each trial. This variable is used to modify a value correction parameter that estimates whether threshold distribution is over or underestimated.

With these variables in hand, we conducted a series of RRW models.

**RRW Models**

Four primary model variants of PVT were compared. These were the Severity Bias Model, the Severity Bias + SQ Value model, the Severity Bias + SQ Boundary model and the Severity Bias + SQ Value + SQ Boundary model as used in Experiment 2. To assess whether the value of the reference distribution was accurate, we ran the above four models with and without a value change parameter dummy coded for the reference distribution (Reference Value).

Prior to running the RRW models, we collapsed the data across participants and calculated p(HVO) and RT for each Reference\_HVO X Severity\_HVO X Quantity\_HVO X Overlap bin. As before, we used the Smart Grid Search algorithm to optimize the parameters and removed conditions with fewer than 40 trials. Finally, the best fit model was run 40 times and an average predicted RT and *p*(HVO) for each Overlap together with the *r2* and BIC of that average predicted fit was calculated. The model with the lowest BIC was identified as the best fit model. If multiple models were equivalent with respect to *r2* and BIC values, the simpler model was favored based on the law of parsimony.

Table 4 presents the fit statistics for each PVT model (see also Figure S3 in Supplemental Information). The best fit model was the Severity Bias + SQ Boundary model. This model was an excellent fit (*r2* ≈ 0.90). In addition to the Severity Bias, there was an influence of symptom quantity on the boundary parameter. Here, the “*b*” parameter specifies the boundary spread when there is only one symptom to process. The *bSQ* parameter identifies the shift in the boundary when there are two symptoms to process. As can be seen, the participant can make a judgment concerning a single symptom with a small boundary separation. However, when there are two symptoms, the participant requires much more evidence before the decision is made. Importantly, the best fit model did not include a value change parameter for symptom quantity. This indicates that the Biased Average Grouping function accurately estimated the value of the two-symptom group in this yes/no experiment.

**Table 4**

*The fit statistics for all PVT models fit to the data in Experiment 3*

|  |
| --- |
| **Experiment 3** |
| Model | FreeParameters | *r2* | BIC | Best Fit |
| Severity Bias | 5 | 0.50 | 266 |  |
| Severity Bias + SQ Value | 6 | 0.48 | 267 |  |
| Severity Bias + SQ Boundary | 6 | 0.90 | 91 | **✓** |
| Severity Bias + SQ Value + SQ Boundary | 7 | 0.90 | 92 |  |
| Ref Value + Severity Bias  | 6 | 0.51 | 271 |  |
| Ref Value + Severity Bias + SQ Value | 7 | 0.43 | 274 |  |
| Ref Value + Severity Bias + SQ Boundary  | 7 | 0.90 | 96 |  |
| Ref Value + Severity Bias + SQ Value + SQ Boundary | 8 | 0.87 | 126 |  |

**Discussion**

Real-world decisions on seeking healthcare are most often marked by the decision to seek care or abstain from seeking care when experiencing symptoms. In Experiment 3, PVT modeled this process. Similar to Experiments 1 and 2, the best fitting model (i.e., Severity Bias + SQ Boundary model) contained a Severity Bias parameter that accounted for the tendency to weight more severe symptom sets higher than less severe symptom sets. Similar to Experiment 2, the model also included the SQ Boundary parameter that accounted for the influence of having to consider symptom sets that contained more than one instance.

In Experiment 2 the best fitting account contained an additional parameter that modulated the effect of having to consider multi-instance symptoms sets, whereas Experiment 3 does not require such a parameter. This suggests that the Biased Average Grouping Function provides, in general, a reasonable estimate of the Psychological Value of a group of symptoms. As such, very good data fits arise when PVT is implemented in computational terms and differing variants of the general theory are pitted against one another.

**General Discussion**

Here we have examined how PVT, a general-purpose, value-based mechanism, can be used to explain healthcare seeking decisions. We posited that relief from a health-related symptoms has Psychological Value and such value predicts health-seeking behavior. In Experiment 1, we presented two symptoms and asked participants to choose the symptom that they would seek healthcare for. The data revealed that the overlap of the Psychological Value distributions of the options accurately predicted the time it takes to choose (RT) and choice itself (r2 ≈ 0.9). Participants revealed a strong response bias (a start point effect) favoring the symptom labeled as “severe.” In Experiment 2, we replicated Experiment 1 with the exception that we presented multiple symptoms. Again, PVT accurately predicted RT and choice (r2 ≈ 0.9). The data replicated the response bias favoring the “severe” label and revealed that the Psychological Value of the group was well estimated by the Biased Average Grouping Function. Finally, in Experiment 3, we replicated Experiment 2, with the exception that we presented one set of symptoms and asked the participants whether or not they would seek healthcare for treatment. Again, PVT accurately predicted RT and choice (r2 ≈ 0.9). The data again revealed the response bias favoring the “severe” label and that the Psychological Value of the group was well estimated by the Biased Average Grouping Function. These data support the conclusion that PVT can accurately model healthcare seeking decisions from the Psychological Value of relief from the symptoms.

Pre-diagnostic symptom recognition is seen as one of the primary drivers for care-seeking behavior and delay for a range of serious illnesses (Altive & Madigan, 2012; Mor et al., 1990). A significant proportion of the total number of estimated illnesses go undiagnosed (Beagley et al., 2014; Song et al., 2017). Thus, understanding the factors involved in individuals’ health decisions when faced with undiagnosed symptoms prior to their initial presentation to a health professional, could eventually be integral in creating effective public health interventions.

PVT offers insight into how symptoms themselves influence health seeking decisions. Traditional models of health behavior are largely applicable to situations in which illness perceptions are present, and this research has typically focused on care-seeking behavior related to specific illnesses (Carpenter, 2010; Janz & Becker, 1984). PVT provides evidence that individuals’ perceived relief from symptoms can influence health behavior intention (Safer et al., 1979). Experiments 1-3 revealed that symptom perceptions (i.e., the Psychological Value of symptom relief) predict response choices in a health-seeking decision scenario.

In all three experiments, there was a large response bias that favors seeking healthcare for “severe” symptoms. This bias manifested as a start point effect. One way to understand the start point effect is to consider a footrace. In this analogy, a runner’s speed is analogous to the Psychological Value of the symptom. The speed of the runner is typically the most important factor as to who will cross the finish line first. However, consider a case where one of the runners gets a head start. In this instance, both the speed of each runner and the size of the head start will influence who crosses the finish line first. A fast runner may beat a slow runner who has a small head start, but if the head start is large enough, the slow runner will win the race. Here, the head start is analogous to a start point response bias. Participants gave symptoms with the label “severe” a very large head start when evaluating the value of symptoms. So, even when a mild symptom has a high Psychological Value (e.g., not having mild memory problems), a lower valued severe symptom (not having severe eye irritation) may be chosen.

Because researchers have demonstrated that start point biases are relatively easy to manipulate (Ratcliff & Rouder, 1998), such response biases may prove useful when educating the public about health issues. For example, one can consider the recent COVID-19 pandemic, during which the U.S. government widely provided vaccines at no cost to consumers (Centers for Disease Control and Prevention, 2021). In the vast majority of COVID-19 cases many individuals who contracted the virus displayed relatively mild symptoms or were even entirely asymptomatic (Chowdhury & Oommen, 2020). From a PVT perspective, in such cases valuation of “not having” the symptoms (e.g., mild fever, mild cough, etc.) is likely to be relatively low, potentially even lower than a relatively insignificant cost (e.g., the time it takes to get a free vaccine). Critically, those individuals whose symptoms would likely be more severe (e.g., the elderly) received the vaccine at a much higher rate (NCRID, 2022).

Experiments 2 and 3 revealed that the Psychological Value of a group of symptoms can be predicted from the Psychological Value of the individual symptoms in the group using the Biased Average Grouping Function. The Biased Average Grouping Function posits that groups of symptoms equal the average Psychological Value of the individual symptoms in the group plus some extra influence of the symptom with the maximum Psychological Value. As such, more symptoms will not lead to a greater likelihood of healthcare seeking behavior. Indeed, people are more likely to seek healthcare for single symptom with a high Psychological Value than the same high value symptom paired with a symptom with a low Psychological Value. This seemingly counterintuitive finding predicts that the symptom with the highest Psychological Value will have disproportional influence over the total number of symptoms, per se. This is a result that should be explored in more detail in the future.

PVT has been successfully applied to a wide range of complex social decisions and yields large effect sizes (see e.g., Cohen et al., 2022; Cohen et al., 2023). The current research results confirm that PVT applies to health care-seeking decisions equally well as it applies to other complex social decisions. This suggests that the mechanism involved in health care-seeking decisions is likely the same as that involved in decision-making in general. As a consequence, PVT offers a marked departure from how health behavior has historically been studied. Models like the Health Belief Model (Janz & Becker, 1984) have proposed health behavior decisions are unique and, thus, require a standalone model to predict decisional outcomes. Such health behavior models suggest it would be difficult to compare between health and non-health-related variables in understanding how individuals make health decisions. PVT shows, in contrast, that Psychological Value constitutes a type of “common currency” by which people can and do compare health and non-health related factors in their health decisions. This approach allows for the fact that decisions in any context are rarely segregated into a single conceptual category (e.g., health), as other categories (e.g., finances, time, etc.) are necessarily implicated. PVT’s increased generalizability, paired with its superior ability to predict RT and choice, provide strong and compelling evidence for its applicability to health decision-making.

**Implications, Limitations, and Future Research**

 Understanding the relation between symptom perception and health-seeking behavior can aid in patient communication and education. By recognizing which symptoms patients perceive as most distressing or severe, healthcare providers can tailor their communication strategies to address these concerns directly. For example, medical professionals could provide education on the importance of symptoms which may be less salient or perceived as being more benign, effectively working with the individual to adjust the perceived Psychological Value of not having such symptoms. Similarly, medical providers could draw connections between symptoms that are less influential in evoking health behavior and those that are more likely to elicit health behavior (e.g., a doctor describing how a symptom with a lesser Psychological Value is often a precursor for a symptom with a higher Psychological Value). Educating patients about the importance of seeking timely care for these symptoms can lead to better health outcomes and increased patient satisfaction. Moreover, this understanding can inform the development of public health campaigns and interventions that resonate more deeply with patient needs, ultimately fostering a more responsive and patient-centered healthcare system.

Existing campaigns and interventions may be improved by ensuring they are more aligned with patient priorities and experiences. For instance, in a recent study by Chan et al. (2024) the aim was to understand better how different vaccine characteristics affect public preferences for vaccines in New Zealand. Over 600 participants across New Zealand took part in an online discrete choice experiment in which over each of a sequence of trials they were asked to choose between two different options regarding alternative vaccines given competing attributes. The resulting findings showed that the risk of severe adverse effects, vaccine effectiveness and duration of protection were rated of most importance by the New Zealand public. As Chan et al. (2024) conclude, information regarding such factors should feature prominently “when developing, introducing, and promoting future vaccines to the public.” (Abstract). An implication from the present study is that more generally gathering data about patients’ preferences should also feature heavily in the health care decisions that are made by practitioners.

This research has practical implications for both individual clinical work as well as broader, public health initiatives. From the clinical perspective, multiple individual clinical interventions such as motivational interviewing and Acceptance and Commitment Therapy, emphasize the client's goal and values, as well as discrepancies between those values and their actions (Van Horn, Wenzel, & Britton, 2021; Rahal & Gon, 2020). The direct comparison approach centered on psychological value provides a direct way to consider the relative worth of the client's goals for, and barriers against, clinical change. These discrete assessments have the potential to inform both the health care professional and the client to establish achievable therapeutic goals and practical plans to achieve them by increasing insight around the ways symptom perception influences both avoidance- and values-based behavior.

Despite its strengths and diverse potential implications, the current research presents some of the same limitations faced by other discrete choice research. These provide grounds for future research in more diverse samples, for a wider variety of decision-making factors, and in settings with greater ecological validity. The current research relies entirely on an undergraduate sample and, to be enrolled in the university, students are required to maintain health insurance. As such, the current sample likely represents a relatively homogenous and healthy population (Cooper et al., 2011; Greene et al., 2016; Wensing et al., 2001). Future PVT research could benefit from application to more diverse populations and, in contrast to the current research, particularly those with more complex patterns of health behavior and health vulnerabilities. Future research might also assess how factors outside of symptom experience (i.e., financial cost, one’s valuation of themselves) play a role in influencing health decisions.

The current studies examine health-related behavior in a lab setting. This is a necessary first step before conducting experiments in health care settings, and future research should assess the validity of PVT in more applied settings. Future research might address this issue through the implementation of a pilot program in clinical settings. Such a program could test the feasibility and effectiveness of integrating Psychological Value Theory (PVT) into routine healthcare practices. By collecting real-world data on how PVT influences patient decision-making and treatment outcomes, researchers can refine the model and identify best practices for its application. This pilot program could also provide valuable insights into patient and provider experiences, highlighting potential barriers and facilitators to implementation. Additionally, the findings from this pilot could inform larger-scale studies and support the development of guidelines for incorporating PVT into various clinical and public health interventions, ultimately enhancing patient-centered care and improving health outcomes.

**Conclusion**

 In conclusion, the model fitting we have explored accommodates the typical kind of data (RTs and accuracy) that are collected in tightly controlled experimental settings. This provides a detailed proof of principle that the underlying theory, namely PVT (Cohen & Ahn, 2016; Cohen et al. 2022; Cohen et al. 2023), provides an impressive account of performance in these kinds of choice tasks. Of course, it remains to be seen the degree to which the theory can accommodate the kinds of data that are typically collected in healthcare settings – that is, the actual choices people make when confronted by health and healthcare seeking events. What emerges from the current work is a very clear prediction that such choices will reflect appraisals that factor in symptom severity, the number of symptoms and, fundamentally, the psychological value that the person attributes to those symptoms.

References

Altice, N. F., & Madigan, E. A. (2012). Factors associated with delayed care-seeking in hospitalized patients with heart failure. *Heart and Lung, 41*(3), 244-254. doi: <http://dx.doi.org/10.1016/j.hrtlng.2011.09.007>

Ashby, F. G., & Lee, W. W. (1993). Perceptual variability as a fundamental axiom of perceptual science. *Advances in psychology, 99*, 369-399. doi: [http://dx.doi.org/10.1016/S0166-4115(08)62778-8](http://dx.doi.org/10.1016/S0166-4115%2808%2962778-8)

Beagley, J., Guariguata, L., Weil, C., & Motala, A. A. (2014). Global estimates of undiagnosed diabetes in adults. *Diabetes Research and Clinical Practice, 103*(2), 150-160. doi: <http://dx.doi.org/10.1016/j.diabres.2013.11.001>

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, 23*(3), 251-263. doi: <https://doi.org/10.1016/j.tics.2018.12.003>

Carpenter, C. J. (2010). A Meta-Analysis of the Effectiveness of Health Belief Model Variables in Predicting Behavior. *Health Communication, 25*(8), 661-669. doi: 10.1080/10410236.2010.521906

CDC. (2021). Key Things to Know About COVID-19 Vaccines. doi: <https://stacks.cdc.gov/view/cdc/108920>

Chan, A. H. Y., Tao, M. R., Marsh, S., & Petousis-Harris, H. (2024). Vaccine decision making in New Zealand: a discrete choice experiment. *BMC Public Health, 24*(1). doi: 10.1186/s12889-024-17865-8

Chowdhury, S. D., & Oommen, A. M. (2020). Epidemiology of COVID-19. *Journal of Digestive Endoscopy, 11*(1), 3-7. doi: 10.1055/s-0040-1712187

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

Cohen, D. J., Campbell, M. K., & Quinlan, P. T. (2023). Psychological value theory: A computational cognitive model of charitable giving. *Cognitive Psychology, 145*, 101593. doi: <https://doi.org/10.1016/j.cogpsych.2023.101593>

Cohen, D. J., Cromley, A. R., Freda, K. E., & White, M. (2022). Psychological value theory: The psychological value of human lives and economic goods. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48*(12), 2015–2050. doi: <http://dx.doi.org/10.1037/xlm0001047>

Cohen, D. J., & Quinlan, P. T. (2016). How numbers mean: Comparing random walk models of numerical cognition varying both encoding processes and underlying quantity representations. *Cognitive Psychology, 91*, 63-81. doi: <http://dx.doi.org/10.1016/j.cogpsych.2016.10.002>

Cohen, D. J., Quinlan, P. T., & Liu, X. Y. (2023). Moral judgments are value-based decisions driven by culturally stable valuations and culturally variable decision biases. *Social Psychological and Personality Science*, 1-11. doi: 10.1177/19485506231199527

Cooper, C. A., McCord, D. M., & Socha, A. (2011). Evaluating the college sophomore problem: The case of personality and politics. *Journal of Psychology, 145*(1), 23-37. doi: 10.1080/00223980.2010.528074

Elnegaard, S., Andersen, R. S., Pedersen, A. F., Larsen, P. V., Sondergaard, J., Rasmussen, R. P., . . . Jarbol, D. E. (2015). Self-reported symptoms and healthcare seeking in the general population -exploring "The Symptom Iceberg". *BMC Public Health, 15*, 685. doi: <http://dx.doi.org/10.1186/s12889-015-2034-5>

Gartland, N., Long, H., & Skevington, S. M. (2019). Undiagnosed cancer symptoms in the community: does poor quality of life influence the decision to seek help? *Quality of Life Research, 28*(5), 1327-1335. doi: 10.1007/s11136-018-2088-9

Green, M. A., Little, E., Cooper, R., Relton, C., & Strong, M. (2016). Investigation of social, demographic and health variations in the usage of prescribed and over-the-counter medicines within a large cohort (South Yorkshire, UK). *Bmj Open, 6*(9). doi: 10.1136/bmjopen-2016-012038

Heaton, K. W., O’Donnell, L. J., Braddon, F. E., Mountford, R. A., Hughes, A. O., & Cripps, P. J. (1992). Symptoms of irritable bowel syndrome in a British urban community: consulters and nonconsulters. *Gastroenterology, 102*(6), 1962-1967. doi: [http://dx.doi.org/10.1016/0016-5085(92)90320-X](http://dx.doi.org/10.1016/0016-5085%2892%2990320-X)

Janz, N. K., & Becker, M. H. (1984). The Health Belief Model: A decade later. *Health Education Quarterly, 11*, 1-47. doi: <https://doi.org/10.1177/109019818401100101>

Joseph, C. L. M., Foxman, B., Leickly, F. E., Peterson, E., & Ownby, D. (1996). Prevalence of possible undiagnosed asthma and associated morbidity among urban schoolchildren. *Journal of Pediatrics, 129*(5), 735-742. doi: 10.1016/s0022-3476(96)70158-0

Krajbich, I. (2019). Accounting for attention in sequential sampling models of decision making. *Current Opinion in Psychology, 29*, 6-11. doi: <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. doi: <https://doi.org/10.1038/nn.2635>

Lancaster, K. T. (1966). A new approach to consumer theory. *Journal of Political Economy, 74*, 132-157. doi: <http://www.jstor.com/stable/1828835>

Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika, Psychometrika*(1), 77-105. doi: <http://dx.doi.org/10.1007/BF02291481>

Macfarlane, T. V., Blinkhorn, A., Davies, R. M., Kincey, J., & Wothington, H. V. (2003). Factors associated with health care seeking behaviour for orofacial pain in the general population. *Community Dental Health, 20*(1), 20-26.

Mohr, N. M., Harland, K. K., Krishnadasan, A., Ten Eyck, P., Mower, W. R., Willey, J., . . . Project, C. E. D. N. (2021). Diagnosed and Undiagnosed COVID-19 in US Emergency Department Health Care Personnel: A Cross-sectional Analysis. *Annals of Emergency Medicine, 78*(1), 27-34. doi: 10.1016/j.annemergmed.2070.12.007

Mor, V., Masterson-Allen, S., Goldberg, R., Guadagnoli, E., & Wool, M. S. (1990). Pre-diagnostic symptom recognition and help seeking among cancer patients. *Journal of Community Health, 15*, 253-266. doi: <http://dx.doi.org/10.1007/BF01350291>

NCIRD. (2022). COVID-19 Vaccination and Case Trends by Age Group, United States. doi: <https://data.cdc.gov/Vaccinations/Archive-COVID-19-Vaccination-and-Case-Trends-by-Ag/gxj9-t96f/about_data>

Pieterse, A. H., & de Vries, M. (2013). On the suitability of fast and frugal heuristics for designing values clarification methods in patient decision aids: a critical analysis. *Health Expectations, 16*(3), E73-E79. doi: 10.1111/j.1369-7625.2011.00720.x

Rahal, G. M., & Gon, M. C. C. (2020). A systematic review of values interventions in acceptance and commitment therapy. *International journal of psychology and psychological therapy*, *20*(3), 355-372.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59-108. doi: <http://dx.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*(5), 347-356. doi: <https://doi.org/10.1111/1467-9280.00067>

Rosenstock, I. M. (1974). Historical origins of the health belief model. *Health Education and Behavior, 2*, 328-335. doi: <https://doi.org/10.1177/1090198174002004>

Ryan, M. (2004). Discrete choice experiments in health care. NICE should consider them for patient centred evaluations of technologies. *British Medical Journal, 328*, 360-361. doi: <http://dx.doi.org/10.1136/bmj.328.7436.360>

Ryan, M., & Gerard, K. (2003). Using discrete choice experiments to value health care programmes: current practice and future research reflections. *Applied Health Economics and Health Policy, 2*(1), 55-64.

Safer, M. A., Tarps, Q. J., Jackson, T. C., & Leventhal, H. (1979). Determinants of three stages of delay in seeking care at a medical clinic. *Medical Care, 19*(1), 11-29. doi: <http://dx.doi.org/10.1097/00005650-197901000-00002>

Sayuk, G. S., Wolf, R., & Chang, L. (2017). Comparison of Symptoms, Healthcare Utilization, and Treatment in Diagnosed and Undiagnosed Individuals With Diarrhea-Predominant Irritable Bowel Syndrome. *American Journal of Gastroenterology, 112*(6), 892-899. doi: 10.1038/ajg.2016.574

Schulpher, M., Bryan, S., Fry, P., de Winter, P., Payne, H., & Emberton, M. (2004). Patients' preferences for the management of non-metastatic prostate cancer: discrete choice experiment. *British Medical Journal, 328*(7436), 382-384A. doi: 10.1136/bmj.37872.497234.44

Schultz, T. J., Jacobs, B., & Patterson, R. L. (1969). Unrecognized dislocations of the shoulder. *Journal of Trauma and Acute Care Surgery, 9*(12), 1009-10023. doi: <http://dx.doi.org/10.1097/00005373-196912000-00005>

Sheifer, S. E., Manolio, T. A., & Gersh, B. J. (2001). Unrecognized myocardial infarction. *Annals of Internal Medicine, 135*(9), 801-811. doi: 10.7326/0003-4819-135-9-200111060-00010

Song, R. G., Hall, H. I., Green, T. A., Szwarcwald, C. L., & Pantazis, N. (2017). Using CD4 Data to Estimate HIV Incidence, Prevalence, and Percent of Undiagnosed Infections in the United States. *Jaids-Journal of Acquired Immune Deficiency Syndromes, 74*(1), 3-9. doi: 10.1097/qai.0000000000001151

Stevens, S. S. (1956). The direct estimation of sensory magnitudes-loudness. *The American Journal of Psychology, 69*(1), 1-25. doi: <http://www.jstor.org/stable/1418112>

Taber, J. M., Leyva, B., & Persoskie, A. (2015). Why do people avoid medical care? A qualitative study using national data. *Journal of General and Internal Medicine, 30*(3), 290-297. doi: <http://dx.doi.org/10.1007/s11606-014-3089-1>

Van Horn, D. H. A., Wenzel, A., & Britton, P. C. (2021). Motivational interviewing. In A. Wenzel (Ed.), Handbook of cognitive behavioral therapy: Overview and approaches (pp. 313–347). American Psychological Association.

Wensing, M., Vingerhoets, E., & Grol, R. (2001). Functional status, health problems, age and comorbidity in primary care patients. *Quality of Life Research, 10*(2), 141-148. doi: 10.1023/a:1016705615207