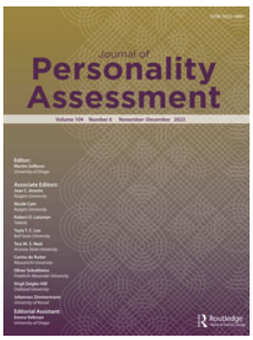

Routledge
Taylor & Francis Group

Journal of Personality Assessment

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hjpa20

Identifying Core Values with a Hierarchical, Ipsative, Preference Assessment

Harold Stanislaw & Jamie McCreary

To cite this article: Harold Stanislaw & Jamie McCreary (2022): Identifying Core Values with a Hierarchical, Ipsative, Preference Assessment, Journal of Personality Assessment, DOI: 10.1080/00223891.2022.2090369

To link to this article: https://doi.org/10.1080/00223891.2022.2090369

Published online: 12 Jul 2022.

JOURNAL OF PERSONALITY ASSESSMENT
https://doi.org/10.1080/00223891.2022.2090369

Identifying Core Values with a Hierarchical, Ipsative, Preference Assessment

Harold Stanislaw and Jamie McCreary

Department of Psychology and Child Development, California State University, Stanislaus, Turlock, CA, USA

ABSTRACT

Values provide a foundation for decision making, behavior, and emotional reaction; they are often used by practitioners to design effective interventions for self-awareness and personal growth. This report describes a novel, user-friendly method that identifies core values with a hierarchical ipsatization procedure (HIP) that is transparent and efficient. Response bias, validity, and user satisfaction were examined in a study in which 602 respondents completed a survey asking them to rate 80 values and use HIP to identify their 4 most inspiring and motivating values. HIP enabled selecting these 4 core values from 80 candidates in 5–7 min, with minimal evidence of response bias. The selections made during HIP were consistent with the rating data, providing evidence for HIP's concurrent validity. 88% of the respondents felt the 4 values identified by HIP inspired and motivated them more than any other values they could think of. These findings suggest HIP is a useful tool for identifying core values, especially in applied settings.

ARTICLE HISTORY
Received 20 August 2021
Accepted 31 May 2022

> Not to arrive at a clear understanding of one's own values is a tragic waste. You have missed the whole point of what life is for.
>
> – Eleanor Roosevelt, *You Learn by Living*

A basic tenet of psychology is that people possess a distinct set of core values – personal judgments about what is important in life – that define who they are. Schwartz (2011) described values as "trans-situational goals, varying in importance, that serve as guiding principles in the life of a person" (p. 464). Schwartz noted further that values may have both individual and cultural determinants. Cultural norms often favor particular values, but individuals may endorse their culture's values to varying degrees and assign different priorities and levels of importance to different values. This study focuses on the values that individuals endorse most strongly and consider most important: their core values.

The importance of core values is clear. Core values guide critical life decisions (Schwartz, 2010) and impact well-being (Sagiv et al., 2015; Schwartz & Sortheix, 2018; Sheldon & Elliot, 1999). Feather (1995) summarized the centrality of values in observing that they influence behavioral choices, emotional reactions to activities and their outcomes, and the way events and individuals are perceived.

There are also links between values and personality (Bilsky & Schwartz, 1994), as behavior is predicted by both values (Bardi & Schwartz, 2003) and personality traits (Horstmann et al., 2021). However, the relationship between values and behavior is under cognitive control: Individuals are typically aware of their values and may consciously choose behaviors consistent with those values (McClelland et al., 1989; Roccas et al., 2002). By contrast, individuals may be unaware how their personality affects their behavior, but nevertheless act in ways consistent with that personality (which is often how personality traits are inferred). For example, someone may not realize they are extroverted, but provide evidence of this personality trait by actively seeking social interactions. That same person may consciously prioritize compassion – a value – and purposely behave in a compassionate manner.

Even so, correlations between values and behaviors tend to be only weak to moderate (Lee et al., 2021). This is not unexpected, given that many other factors influence behavior, including situational constraints, one's role in a group (Jex et al., 2003), and normative pressure (Bardi & Schwartz, 2003). Furthermore, individuals regard some values as more central than others; values that people consider particularly important tend to influence their behavior more strongly than other values they also endorse but regard as less central (Lee et al., 2021).

McClelland (1985) noted that "an understanding of how implicit and self-attributed motives function … has important practical implications for psychological adjustment" (1985, p. 700). Accordingly, practitioners have leveraged personal values to design highly effective interventions. For example, self-affirmation (reflecting on one's values) is associated with a variety of positive outcomes. It promotes self-compassion (Thomaes et al., 2012) and these feelings can foster more positive behavior toward others (Lindsay & Creswell, 2014). Self-affirmation also limits rumination after

CONTACT Harold Stanislaw HStanislaw@csustan.edu Department of Psychology and Child Development, California State University, Stanislaus, One University Circle, Turlock, CA 95382.

Supplemental materials for this article are available online at https://osf.io/9azvx/?view_only=88076e36103a48e59e3b00c3879339e1

failure (Koole et al., 1999) and can improve performance in a variety of potentially threatening situations (Sherman, 2013) by reducing defensiveness and negative affect (Crocker et al., 2008; Emanuel et al., 2018). Asking university students to briefly reflect on their personal values can generate enduring improvements in academic performance (Miyake et al., 2010). Self-affirmation of core values can also foster prosocial behavior (Schwartz, 2010), creativity at work (Jiang, 2018), and pro-environmental attitudes and behavior (Graham-Rowe et al., 2019). Values-based interventions can improve health care follow-through (Epton et al., 2015) and promote healthy lifestyles more effectively than alternative approaches (e.g., Anshel et al., 2010; Hardcastle et al., 2015).

The impact of self-affirmation goes beyond self-protection: Focusing on core values also enhances psychological well-being more directly (Howell, 2017; Nelson et al., 2014; Schüz & Schüz, 2017). In one study, people who engaged in self-affirmation for 2 weeks showed increased need satisfaction and meaning in their lives (Nelson et al., 2014). Spontaneous self-affirmation – the habitual tendency to reflect on one's own values and strengths – is associated with higher levels of optimism and happiness and lower levels of sadness and anger (Emanuel et al., 2018). These findings confirm McClelland's (1985) suggestion that reflecting on one's values can improve psychological adjustment.

Practitioners also use values to assist clients with goal setting. Goals that are consistent with underlying values can be described as self-concordant (Sheldon & Elliot, 1999). People work harder to achieve self-concordant goals, are more likely to attain them, and experience greater psychological well-being when they are attained (Bono & Judge, 2003). This effect is evident across a variety of cultures, suggesting it may be universally important to feel that one's goals are consistent with one's personal values (Sheldon et al., 2004).

Despite the importance of values and their potential utility in applied settings, assessment of values remains challenging. Most formal methods are designed to *measure* values by locating them in a mathematically defined, multidimensional theoretical framework, where the assessment process may be intermediate to the primary goal of testing the theoretical framework or using that framework to better understand how values relate to other psychological constructs (e.g., Gouveia et al., 2014; Schwartz et al., 2012). This approach focuses on assigning numerical scores that describe how strongly individuals endorse different values; accordingly, we refer to this as the *quantitative* approach. By contrast, practitioners are more often interested in *identifying* values, which requires determining the verbal labels that best characterize the values an individual endorses most strongly. We label this the *qualitative* approach. It is often used in applied settings, where the process of assessing values may itself be an important intervention feature that provides an opportunity to engage individuals for positive change (e.g., The Values Project, 2020).

In this report we describe a novel, user-friendly method for identifying core values and present evidence for its validity. The method is particularly suitable for applied use, but is more structured and efficient than the exploratory approaches often used by practitioners. The method can also be used for theoretical work, as it shares features with several approaches developed for that purpose. We review these extant methods next.

Established methods for assessing values

Vernon and Allport (1931) proposed an early method that assumed an individual's values could be described by six numerical scores – one each for dimensions they labeled theoretical, economic, esthetic, social, political, and religious. Assessing these six scores required individuals to spend approximately 30 min reading 45 brief vignettes. The vignettes were followed by two or four response options, each of which was linked to a value dimension. The individuals sorted these response options in order of preference. The preference data were then weighted and combined across vignettes to yield the six value scores.

This method was revised and updated as recently as 1970 (Allport et al., 1970), but an alternative with a markedly different theoretical orientation was proposed a few years later by Rokeach (1973). Rather than using a quantitative approach to describe individuals with a set of numerical scores, Rokeach advocated a qualitative approach that describes individuals with the values they find most relevant, motivating, or inspiring. The Rokeach method is ipsative; it assumes that some values apply strongly to a given individual, while others apply weakly or not at all and can thus be omitted when describing that individual's values. Vernon and Allport assumed individuals can score high on all six dimensions; Rokeach, by contrast, assumed individuals who strongly endorse some values must reject others.

To identify value preferences using Rokeach's (1973) method, individuals see two lists of 18 values. One list of "terminal" values describes life goals toward which respondents might be working; the other list of "instrumental" values describes behaviors respondents might use to pursue their terminal values. A brief description of each value is presented below the corresponding value's label. Individuals complete the instrument by sorting all 18 values within each list from most to least personally important or relevant. Since all the items within a list are sorted, a researcher or practitioner can select as many of the highest-ranked items as desired (such as the top 3–5 items in each list) and assume these constitute an individual's core values.

Card sorts (e.g., Slaney & MacKinnon-Slaney, 2000; Tyler, 1961) share similarities with Rokeach's approach. Values are presented on cards that respondents sort into groups based upon their self-relevance. Optionally, the cards may be further subsorted within each group. The values themselves may be predefined as in Rokeach's approach, or developed by the practitioner or theoretician who is implementing the card sort.

Sorting tasks are popular, but as Braithwaite and Law (1985) noted, they can also be challenging. The values they present are positive and most are likely to be regarded as at least somewhat desirable; sorting them requires a significant

cognitive investment that individuals might be unwilling or unable to make. Situational information is not provided to contextualize or simplify the sorting task. Furthermore, all values must be ranked – even ones that individuals do not endorse. To address these concerns and enable expanding the number of values from 36 to 55, Braithwaite and Law proposed *rating* values rather than *ranking* them, thereby changing the assessment from ipsative to normative. Their recommended 7-point rating scale is asymmetrical, with more positive than negative response options. This mitigates the potential for ceiling effects that can result when responses cluster in a narrow region at the positive end of a symmetrical scale (Gorsuch, 1970).

It is not clear how respondents perceive and utilize asymmetric response scales; even minor alterations to more familiar, symmetric scales can trigger different responses (e.g., Cabooter et al., 2016). Response set may also be a concern, with individuals exhibiting acquiescence bias (agreeing with items regardless of content; Winkler et al., 1982) or failing to read instructions and items carefully (Brosnan et al., 2019). As well, the length of the Braithwaite and Law (1985) instrument may tax the attention of users, leading to careless responding – especially for the values presented later in the instrument, when response fatigue may set in (Porter et al., 2004).

Similar challenges with cognitive load exist for an instrument proposed by Schwartz (1992; Schwartz & Boehnke, 2004). Individuals see 56 or 57 items – many taken directly from Rokeach (1973) – and rate each as "a guiding principle in my life" using an asymmetrical, 9-point response scale. The ratings are then combined to yield scores for 10 "basic values" that Schwartz suggests are universal (common to greater or lesser degrees in all cultures studied) and organized in a circular space resembling a circumplex model. A more recent version proposes 19 basic values that are conceptually related to the original 10 basic values; these retain a circular relational structure but are nested within four higher-order values (Schwartz et al., 2012).

Respondents may find it tedious to sort, rank, or rate large numbers of items that consist solely of value labels and their accompanying descriptions. An alternative that attempts to address this problem uses "portraits," which briefly describe scenarios (Feather, 1995) or individuals (Schwartz et al., 2001) that embody particular values without actually naming them. Respondents rate how strongly they connect with the hypothetical situation or person in each portrait – and thus the value it incorporates.

The portrait approach has the potential to improve respondent engagement with the instrument. Cognitive load may be reduced as well, because the task requires concrete rather than abstract thinking (Schwartz et al., 2001). However, if the individual does not identify with the situation or person described in a particular portrait, the response it elicits will not provide insight into that individual's personal values. This issue is particularly concerning because the narratives for portraits tend to be more verbose and require more processing time than the value definitions used in more conventional prompts. Thus, in order to reduce instrument completion time, researchers may need to limit the number of portraits they present. For example, Schwartz et al. (2001) found participants could rate only 29 portraits (containing an average of 20 words each) in about the same time needed to rate 57 abstract labels accompanied by descriptions (containing an average of 12 words each). Accordingly, portrait-based instruments have been developed to present only 10 or 20 values (Sandy et al., 2017).

Engagement during value identification can also be improved by discussing values with others. For example, "contemplation and conflict" asks participants to consider what they believe to be true about themselves, and then share and defend their choices in a group setting (Brown & Crace, 1996). Similarly, individuals may be asked to explain the origin of their values, choose a symbol (such as a tattoo) to represent their values, or anticipate obstacles they might encounter in pursuing value-based goals (Bronk et al., 2019). Self-reflective journaling can also improve engagement (Mosconi & Emmett, 2003).

A final method, proposed by Lee et al. (2008, 2019), assesses values using an ipsative, "best-worst scaling" approach (BWS; Louviere et al., 2015). This involves presenting groups of values and asking individuals to identify within each group the value they consider the best or most relevant to themselves, and the value they consider the worst or least relevant. All values appear equally often and are paired equally often with each other. BWS shares features with paired comparison approaches but can accommodate more values. For example, Lee et al. (2019) assessed 20 values with BWS by presenting groups of five values 21 times; a paired-comparison approach would have required presenting pairs of values 190 times.

BWS has demonstrated test-retest reliability. Lee and colleagues also claim it eliminates response biases common to rating scales, which is a benefit of most ipsative methods (Chan, 2003). However, respondents may disengage with BWS because it requires them to consider the same value multiple times, even if that value was identified as the worst in a previous list. Fatigue may also set in, compromising the data. The method only works with specific numbers of items, which may require padding the values of interest with artificial items. Furthermore, BWS becomes intractable as the number of items increases. For example, the application we describe below assesses 80 values; using BWS would require adding one artificial item and asking individuals to make best-worst decisions for 90 groups containing 10 values each (Louviere et al., 2015, p. 19).

General features of assessment methods

The methods reviewed above are differentiated largely by several features. One is the type of prompt, which is either an explicit but abstract value label (typically accompanied by a description of the value), or a concrete vignette or portrait that embodies the value implicitly. Another difference is whether the method involves ranking (or sorting), rating, best-worst scaling, or open responding (e.g., journaling). These elements may be mixed. For example, Döring et al.

Table 1. Two categories used in the assessment, and the values they contain.

Category	Values
Learning You love questioning, learning, applying wisdom and gaining mastery	**Curiosity** You are most inspired when asking questions and seeking new experiences from which to learn **Learning** You are most inspired when learning, teaching, or supporting education **Mastery** You are most inspired when moving toward a masterful level of skill and understanding **Wisdom** You are most inspired when using your ability to integrate knowledge and insight into your life experiences
Teamwork You love connecting and working with others	**Collaboration** You are most inspired when collaborating with others to accomplish something **Connection** You are most inspired when connecting with others **Synergy** You are most inspired when working in a group that accomplishes more together than alone **Teamwork** You are most inspired when working in a group for a common purpose

(2010) assessed values in middle-school children by combining abstract labels with pictures (which are simple, concrete vignettes), while Lee et al. used BWS with both abstract value descriptions (2008) and concrete portraits (2019). Similarly, Rokeach's values could be rated instead of sorted, contemplation and conflict could follow a card sort, and a card sort could be based on tattoos or other symbols rather than verbally articulated values.

A third general difference is whether the assessment produces *scores* or *labels*. Scoring requires a quantitative approach that is typically normative, while labeling is more qualitative and typically ipsative. However, this distinction is somewhat arbitrary. Normative and ipsative methods both yield data that are at least ordinal-level and appropriate for correlation-based analyses (although ipsative data require special precautions; see Chan, 2003). Similarly, qualitative descriptors may be extracted from quantitative data by identifying the value(s) receiving the highest rank or rating, or by establishing a cutoff that must be exceeded before a value is used to label an individual. However, ties are guaranteed with normative data whenever the number of options exceeds the width of the response scale. Ipsative approaches avoid ties because each value has unique ordinality, but this may not mean ipsative techniques are superior; just because respondents are forced to assign different rankings to two values does not imply that one describes them better than the other.

What is perhaps more important is that normative approaches may be particularly susceptible to socially desired responding, because the approaches involve comparing values to absolute standards that often reflect cultural norms (Meglino & Ravlin, 1998). Another potential issue with normative approaches is that they typically present a small number of choices, which may limit their utility in applied settings that focus on helping individuals explore their values. In these settings, it is critical that the assessment method incorporate a broad and diverse range of descriptors, so that all individuals can identify and select values that truly describe who they are.

Furthermore, in applications where values assessment serves as the beginning of a self-exploration journey, it seems essential that the process for selecting those values be fully transparent to the user. These situations seem better suited to ipsative rather than normative methods, because during ipsatization respondents are fully aware that they are identifying some values as more important or relevant than others. By contrast, respondents may not understand how their responses in a BWS or rating procedure led to the identification of their values. Portraits add further opacity to the selection procedure, since they embody values only implicitly. Transparency requires presenting value labels and descriptions explicitly, and reminding respondents throughout the assessment that they are selecting some values and rejecting others.

The method we describe next is highly transparent and provides an alternative to values assessment methods that are less well suited to applied settings. Furthermore, the new method allows users to select their core values from a very large number of options, thereby accommodating a broader range of individual differences than many other methods allow. Despite its breadth, the assessment can be completed rapidly and without the attentional demands that can prove challenging for more conventional approaches. The method also has minimal response bias.

A novel ipsative preference assessment

The method we describe here was created by Seity Health LLC (2021) for use in mobile and desktop apps. It rapidly identifies the four values that individuals find most inspiring and motivating from a set of 80 candidates. A key feature of the method is that the 80 values are grouped into 20 categories. Each category contains four values that share a commonality reflected in that category's label. Table 1 provides examples of two categories and the eight values they contain; a complete description of all the categories and values is available in the supplemental materials. The hierarchical structure is not unique to the approach described here; for example, hierarchical relationships between values have been proposed by Rokeach (1973) and Schwartz et al. (2012). However, unlike other approaches, the hierarchical structuring in the present assessment methodology guides ipsatization. Thus, we label this method the "hierarchical ipsatization procedure" (HIP).

Figure 1 illustrates how HIP enables users to select four core values from 80 candidates; an animated version is available at https://vimeo.com/716289585. The 80 candidate

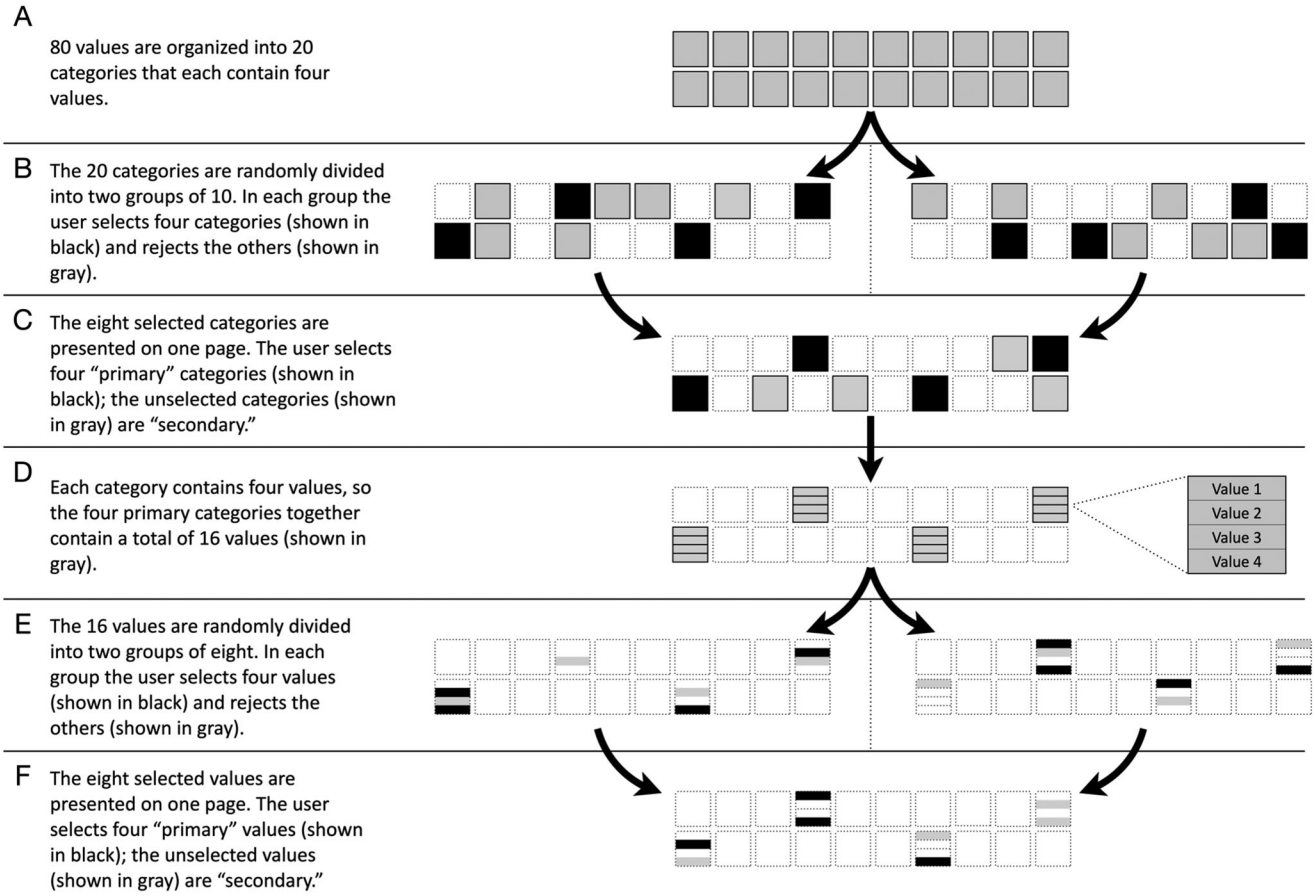

Figure 1. The steps used in HIP to identify four primary values from 80 candidate values. See https://vimeo.com/716289585 for an animated version.

values are organized into 20 categories, with four values in each category (Figure 1A). Ten categories are selected at random and presented to the user, who selects the four categories that motivate and inspire them most (Figure 1B). This is repeated with the remaining 10 categories. The eight categories thus chosen are then displayed together on one page (Figure 1C), where users select the four categories they find most motivating and inspiring – their "primary categories" (shown in black in Figure 1C) – and the four that are less motivating and inspiring – their "secondary categories" (shown in gray). The remaining 12 categories are "unselected categories."

The four primary categories incorporate 16 separate values (Figure 1D). Eight of these values are randomly selected, and users choose the four that motivate and inspire them most (Figure 1E). This is repeated for the eight remaining values. Finally, users are shown the eight values that were selected during this ipsatization process and asked to reduce the list to four values (Figure 1F). As with the categories, this yields four "primary values" (shown in black in Figure 1F), four "secondary values" (shown in gray), and eight "unselected values" for each user. The four primary values are assumed to be the user's core values.

HIP requires users to evaluate only 20 categories (shown in black and gray in Figure 1B) and 16 values (shown in black and gray in Figure 1E), even though they are effectively choosing between 80 values. This efficiency results from the hierarchical structure of the assessment. With values organized into categories, HIP involves only six decision-making steps. Two of these require selecting four items from a list of 10 (panel B in Figure 1), while the other four steps each require selecting four items from a list of eight (panels C, E, and F). The initial selection of four categories is the key to HIP's efficiency: By rejecting 16 categories, users are rejecting the 64 values they contain without even seeing any of those values. Thus, a critical requirement for establishing the validity of HIP is demonstrating that the unpresented values would likely have been rejected, had they actually been presented.

We examined this functionality in a study that asked participants to rate all 20 categories and all 80 values. The same participants also used HIP to select the four values they found most inspiring and motivating. We were particularly interested in determining how the value ratings compared with the HIP selections. We also examined completion times, tested for biased responding, and assessed whether respondents felt the values they selected with HIP were actually their four most inspiring and motivating values.

Materials and methods

Participants

Participants were recruited through Amazon's Mechanical Turk (MTurk), which is an online crowdsourcing site restricted to individuals who are at least 18 years old. The

Table 2. Participant characteristics.

Characteristic	Participant group		
	Abandoned ($n = 166$)	Excluded ($n = 70$)	Retained ($n = 602$)
Age, years			
M ± *SD*	32.45 ± 9.00	33.59 ± 9.05	34.45 ± 10.56
Mdn (Range)	30 (18–67)	31 (18–58)	32 (19–84)
Not provided	$n = 2$	$n = 0$	$n = 1$
Gender			
Male	57.9%	67.1%	49.3%
Female	41.5%	32.9%	50.4%
Other	0.6%	0.0%	0.3%
Not provided	$n = 2$	$n = 0$	$n = 1$
Marital status			
Single	54.3%	52.2%	53.3%
Married or domestic partner	37.0%	39.1%	38.7%
Separated, divorced, or widowed	8.6%	8.7%	7.5%
Other	0.0%	0.0%	0.4%
Not provided	$n = 4$	$n = 1$	$n = 3$
Highest level of education			
High school or less	22.7%	23.2%	33.3%
2-year degree	26.4%	15.9%	18.2%
4-year degree	44.2%	44.9%	34.9%
Graduate degree	5.5%	15.9%	11.7%
Other	1.2%	0.0%	1.8%
Not provided	$n = 3$	$n = 1$	$n = 4$

Note. Respondents who chose to not provide demographic information were excluded from the percentage calculation for that demographic.

survey posting stated participants "must live in the United States." The study was advertised as providing compensation of US $1.00 for completion, with another US $1.50 available in bonuses for "reading and answering the survey carefully." Thus, participants could receive as much as $2.50 for completing the survey.

The survey was started by 954 individuals, with 947 proceeding past the consent form. Seventy-three were subsequently prevented from completing the survey because they did not identify as female when only female responses were sought. Another 36 participants were excluded because their IP addresses resolved to locations outside the United States. Of the remaining 838 respondents, 166 abandoned the survey prior to completion. This 19.8% attrition rate is typical of MTurk studies (Zhou & Fishbach, 2016). Another 70 participants missed more than one attention check for the rating task or more than two attention checks for HIP (attention checks are described below in Materials), or provided the same rating for all 20 categories or values on at least one rating task page ("straightlining"; Schonlau & Toepoel, 2015). These participants were excluded – as is common practice for MTurk studies (Aguinis et al., 2021; Curran, 2016) – because their behaviors suggest they misunderstood the survey instructions or were otherwise not fully engaged. As a result, 602 participants were retained for final analysis.

Table 2 summarizes respondent characteristics. Those who identified as male were less likely to complete the survey and be retained than those who identified as female, which is consistent with findings that males tend to be less engaged with online surveys than females (Berry et al., 2019). However, gender was balanced among the retained respondents, in part because one MTurk posting was available only to females. No information regarding race or ethnicity was requested, but almost all of the retained participants indicated they spoke English often at home; 3 indicated they did not and 2 others did not answer this question.

Materials

The survey was administered online using Qualtrics. The landing page welcomed participants to the "California State University, Stanislaus Personal Values Survey" and presented an informed consent form. This was followed by demographic questions that could be skipped; all other survey items required a response unless they were open-ended. The demographic questions were followed by one section that assessed values using a rating method and another that implemented HIP (described in the survey as the "drag and drop method"). These two sections were presented in random order.

The rating section of the survey began with a page that presented all 20 categories in random order, and asked participants "How well does each of these statements describe you?" A 5-point response scale was used, ranging from *Does not describe me* (1) to *Describes me extremely well* (5). This was followed by four pages that each presented 20 randomly selected values, using the same instructions and response scale as the categories page. The one-word label for each category or value was presented in bold, with text describing that category or value following in plain text, as in Table 1.

The HIP section used identical text formatting for the categories and values, but presented these items differently. The first page randomly selected 10 of the 20 categories and displayed them on the left side of the screen. Participants were instructed to drag the four categories that most motivated and inspired them into a box on the right side of the screen. A second page repeated these instructions for the 10 remaining categories. The eight selected categories were then

Table 3. Mean ± SD completion times (in seconds) for the value rating section of the survey, and for the HIP section (excluding pages repeated for failed attention checks).

Section completed first	Rating section	HIP section
Rating ($n = 322$)	402.02 ± 349.38 s	350.54 ± 295.76 s
HIP ($n = 270$)	360.50 ± 237.34 s	399.56 ± 251.21 s

shown together on a third page, which now had two boxes on the right side – one above the other. Participants were instructed to drag the four categories that motivated and inspired them the most into the upper box and the remaining categories into the lower box.

The resulting four primary categories were linked to 16 values (four for each category). Eight of the 16 linked values were randomly selected and shown on the left side of a new page. Participants were instructed to drag the values that most inspired and motivated them into the upper box, and drag the remaining items (which *don't motivate and inspire me as much*) into the lower box. A second page repeated this procedure for the remaining eight values. Finally, a page showed the eight selected values on the left side and instructed participants to use the drag-and-drop method to indicate which four values motivated and inspired them the most, and which were less motivating and inspiring. Participants were then shown their primary values and asked "Do these four values motivate and inspire you more than any other values you can think of?" Responses were made on a 5-point scale ranging from *Definitely not* (1) to *Definitely yes* (5).

Each of the five survey pages that asked participants to rate categories or values included one attention check, while five of the six HIP pages included an attention check. Each attention check began with a label for an artificial value that was formatted to resemble the other items in that section of the survey. However, the artificial value label was not followed by a value description; instead, the text following the value label instructed the participant how to respond. For example, the rating section of the survey included an attention check that read "**Compliance** – Select 'Describes me very well' for this item." In the HIP section of the survey, three of the attention checks asked participants to drag an item to a specific location (e.g., "**Alertness** – Move this item into the bottom box"), while the other two asked participants to leave the item in its original location (e.g., "**Obedience** – Leave this item on the left side; do not move it into the box").

After values were assessed using both the rating method and HIP, participants rated their agreement with three statements: "This survey taught me a lot about my personal values," "I enjoyed exploring my personal values with this survey," and "I know more about myself now that I've completed the survey." The 5-point response scale for these items ranged from *Definitely not* (1) to *Definitely yes* (5). Next, participants indicated their preference for the two values assessment methods, using a 5-point response scale that ranged from *I liked the drag and drop method much less than the rating method* (1) to *I liked the drag and drop method much more than the rating method* (5). Finally, an open-ended question allowed participants to provide additional comments about the survey. This was followed by a debriefing page.

Procedure

The study protocol was approved by the Psychology Institutional Review Board at California State University, Stanislaus (P-16-72). The survey was posted on MTurk daily from March 5–9, 2017, and again on July 9, 2017. By March 7 the sample was skewing toward male respondents; to balance gender, quotas were used to oversample females on March 8, and the March 9 posting was for females only. All responses were obtained within 24 hr of the corresponding posting, which described the survey as an opportunity to "Answer questions about the values that inspire and motivate you to earn $1.00, with a bonus of up to $1.50 for reading and answering the survey carefully."

Participants who responded inappropriately to an attention check saw an error message that read "You performed an attention check incorrectly – please read all the items carefully." If the participant was completing the HIP section of the survey, the page was presented again without the attention check to ensure that participants selected only actual categories or values. Correct responses to attention checks provided a bonus compensation of $0.10 each in the rating section of the survey and $0.20 each in the HIP section.

Data were analyzed using IBM SPSS version 28 and Microsoft Excel version 16. Response biases for HIP were tested for statistical significance with chi-square analyses. These could not use standard significance tables because the chi-square independence assumption was violated through the selection of four values by each participant, rather than just one value. Accordingly, Monte Carlo techniques were used to estimate *p*-values by simulating the selections made on each HIP page of the survey 100,000 times. The simulations were conducted using custom software written and compiled using Absoft Fortran version 21. The Fortran code, together with analyses and aggregated data summaries, are provided in the supplementary materials.

Results

Completion time

Table 3 shows the estimated time needed to complete the two survey sections in a normal setting, which would exclude the category ratings and HIP attention checks. The median completion time was 5.07 min for HIP ($M = 6.21$, $SD = 4.63$) and 5.20 min for the rating section ($M = 6.39$, $SD = 5.08$). An analysis of variance (ANOVA) found no significant difference in the mean time needed to complete the two survey sections ($p = .643$) or in the mean time required by respondents who completed HIP first and respondents who completed the rating section first ($p = .849$). However, the interaction was statistically significant ($p < .001$, partial $\eta^2 = .019$); respondents completed the second section of the survey about 40–50 s faster than the first section.

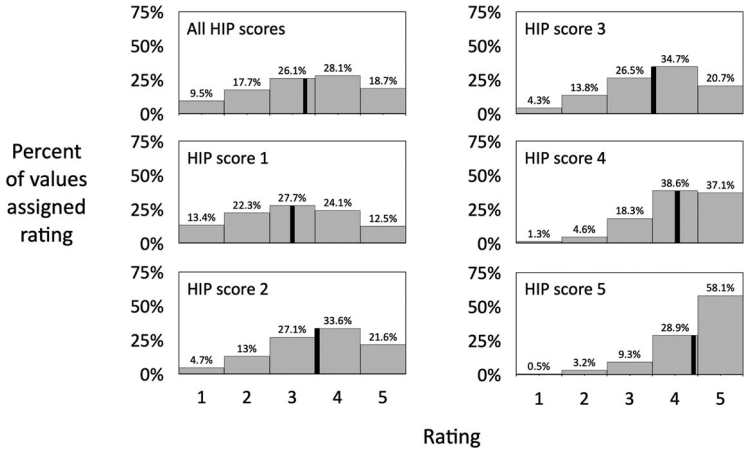

Figure 2. Distribution of ratings assigned to values with each of the five possible HIP scores. The black bar within each histogram illustrates the 95% confidence interval for the mean rating assigned to values with the corresponding HIP score.

Response bias

The mean rating for all 80 values was 3.29 (mean within-subject $SD = 1.04$), but there was evidence for acquiescence bias as the ratings were negatively skewed (see histogram in upper left of Figure 2): There were nearly twice as many ratings on the "agree" side of the response scale (47%) as on the "disagree" side (27%). This pattern was not surprising – most values tend to be regarded as positive by most people – and resulted in many values receiving the same high rating. Only 4.5% of respondents gave their highest rating to exactly four values (the number identified by HIP); 15.0% gave fewer than four values their highest rating, while 80.6% gave their highest rating to more than four values. On average, respondents assigned their highest rating to a median of 13 values ($M = 17.17$, $SD = 14.38$).

Response bias was also evident for HIP, but to a lesser degree. Categories and values presented at the top of each page were sometimes chosen more often than expected by chance, with a corresponding underselection of items near the bottom of each page. This pattern was observed more often when participants completed the HIP section *after* the rating section rather than before (see Table 4). Figure 3 illustrates how selection rates varied with the position of items on the sixth HIP page, where response bias was strongest. This page presented eight values, but the figure shows nine positions because the page also presented an attention check in a random position. (The attention check was ignored when calculating statistical significance.) The black horizontal bars in Figure 3 represent the expected probabilities in the absence of response bias. Overselection was strongest for the top item on the page (Position 1), which was selected 3.2% more often than expected by respondents who began the survey with the rating section, and 2.2% more often by respondents who began with HIP. The most extreme underselection was −2.1%, for Position 7 in respondents who began with the rating task and Position 8 in respondents who began with HIP.

Table 4. Response bias on each page of the HIP section of the survey.

	Bias (%)					
Position on page	Page 1	Page 2	Page 3	Page 4	Page 5	Page 6
HIP section of survey completed before rating section						
1	1.2	1.4	0.7	1.3	1.9	2.2
2	0.8	1.2	1.7	0.5	0.4	1.7
3	1.0	−0.3	0.5	1.5	0.9	1.0
4	−0.2	ac	0.2	0.1	0.6	0.2
5	1.2	−0.7	−1.5	−0.2	0.0	−0.4
6	−1.8	−0.9	−0.8	−1.5	−1.1	−1.2
7	−1.0	0.0	0.2	ac	−0.6	0.5
8	ac	0.3	−0.9	−1.3	−1.4	−2.1
9	−0.8	−1.1	nu	−0.3	−0.7	−1.9
10	0.5	−0.2	nu	nu	nu	nu
11	−1.0	0.3	nu	nu	nu	nu
Chi-square	11.22	6.69	6.36	6.98	9.10	18.25
p-value	.052	.350	.133	.094	.027	< .001
HIP section of survey completed after rating section						
1	1.2	3.0	2.9	1.2	1.3	3.2
2	−0.8	2.0	1.4	1.6	0.6	2.5
3	0.3	−0.1	0.8	1.8	0.6	0.6
4	−0.7	ac	0.4	−0.3	1.3	0.0
5	0.0	−0.8	−0.8	−1.0	−1.6	−0.2
6	−0.3	−0.4	−2.3	−0.4	−0.2	−1.0
7	0.1	−1.3	−1.4	ac	0.1	−2.1
8	ac	−0.3	−1.0	−2.4	−1.0	−1.4
9	−0.5	−1.4	nu	−0.4	−1.1	−1.7
10	0.3	−0.2	nu	nu	nu	nu
11	0.5	−0.4	nu	nu	nu	nu
Chi-square	4.53	23.20	20.75	15.47	6.25	32.93
p-value	.673	< .001	< .001	< .001	.147	< .001

Note. Positive entries for bias indicate selection of the corresponding position more often than expected by chance; negative values indicate selection of the position less often than expected by chance. *p*-values were estimated using Monte Carlo simulations. ac = Attention check, nu = Not used.

Validity

The concurrent validity of HIP was evaluated by comparing the ratings assigned to each value with the selection outcome for that same value in HIP. One analysis compared the mean rating assigned to each value by respondents who identified that value as primary, to the mean rating assigned to that same value by the remaining participants. Figure 4 illustrates these 80 pairs of means, sorted vertically by value popularity (the percent of respondents who selected the value as primary). For all 80 values, the mean rating was higher when that value was selected as primary than when it was not; 71 of these 80 differences were statistically significant (one-tailed t-test $p < .05$; see supplemental materials for complete details).

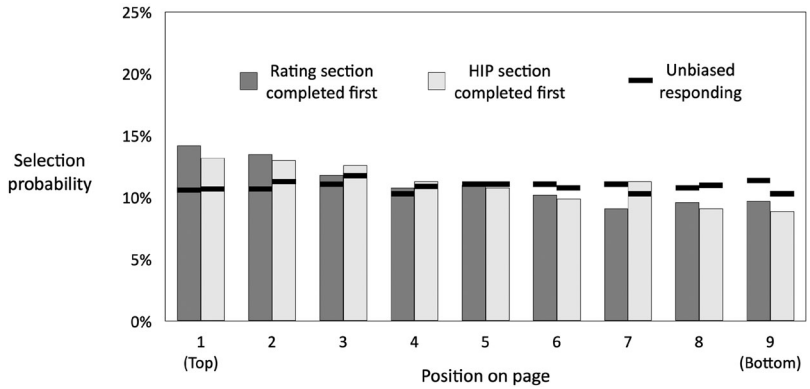

Figure 3. Probability of item selection as a function of page position. Unbiased responding would have resulted in a selection probability of about 11.1% for each position; the actual unbiased response probabilities (the black horizontal bars) varied between 10.8% and 11.6% because the attention check appeared more often in some positions than in others.

The t-tests described above were based on between-subject analyses that were not corrected for Type I error rate inflation, but a second analysis addressed both of these issues. This analysis began by using the method shown in Table 5 to assign each value an ordinal-level "HIP score," which increased as values survived further in the HIP selection. Orthogonal, planned contrasts were then used in a within-subjects, multivariate ANOVA to determine if values that survived further in the HIP selection process received higher ratings, on average, than items that were eliminated earlier in HIP. Figure 2 illustrates the distribution of ratings as a function of HIP scores. Black, vertical bars in the figure depict the 95% confidence interval for the mean rating assigned to values with each HIP score (see Table 5 for means and standard deviations). All of the contrasts were significant ($p < .001$) with large effect sizes. Values in primary and secondary categories (HIP score > 1) received higher ratings than values in unselected categories (HIP score $= 1$, partial $\eta^2 = .729$), and values in primary categories (HIP score > 2) received higher ratings than values in secondary categories (HIP score $= 2$, partial $\eta^2 = .489$). Primary and secondary values (HIP score > 3) received higher ratings than unselected values in primary categories (HIP score $= 3$, partial $\eta^2 = .567$), and primary values (HIP score > 4) received higher ratings than secondary values (HIP score $= 4$, partial $\eta^2 = .327$). Stated more simply, values tended to receive higher ratings as they survived further in the HIP selection.

Subjective impressions

When participants were asked whether the four values they had selected with HIP motivated and inspired them "more than any other values," 88.2% responded *Probably yes* (4) or *Definitely yes* (5). The mean response did not differ between participants who had already completed the rating section of the survey ($M = 4.19$, $SD = 0.80$, $n = 332$) and those who had not ($M = 4.15$, $SD = 0.73$, $n = 270$; t-test $p = .538$). This is notable because the rating section presented all 80 values, while the HIP section presented only 16 of those values – the other 64 values were "hidden" in the 16 categories that were presented but not selected. If the participants who completed the HIP section second had seen values in the preceding rating section that were especially motivating and inspiring, they should have disagreed with the "more than any other values" question if those same values did not appear in the HIP section. The low rate of disagreement by these participants (only 5.1% selected *Definitely not* or *Probably not*, compared to 3.0% for participants who completed HIP first; Fisher's exact $p = .221$) suggests no problems arose from HIP's presentation of only 16 of the 80 values.

More than half of the participants (56.0%) indicated they liked the HIP section of the survey *a little more* (4) or *much more* (5) than the rating section; only one quarter (25.4%) preferred the rating section (numerical score of 1 or 2), while 18.6% had no preference (numerical score of 3). Preference for one or the other method did not vary with the order in which the two sections were completed ($M = 3.50$, $SD = 1.42$ for completing ratings first; $M = 3.54$, $SD = 1.34$ for completing HIP first; t-test $p = .724$).

MTurk respondents rarely offer deep insights when invited to provide feedback at the end of a survey; most leave a "Thank you" or other brief message if they choose to comment at all (end-of-survey feedback was optional in this study). However, some participants volunteered more detailed feedback. These comments were insufficient in number to merit a formal content analysis, but several remarks were consistent with the notion that HIP offers cognitive and affective advantages over more conventional assessments: Respondents noted that, compared to the rating task, HIP "was far more engaging," "felt less monotonous, and more involved," and was "a lot more clear and a lot more fun." HIP may also prompt deeper reflection, as two participants acknowledged, despite favoring different methods. One preferred the rating method "because I didn't have to choose between important values, as so many of them are meaningful," while the other preferred HIP because it "forced me to really choose between what inspires me and what REALLY inspires me."

These anecdotal findings are bolstered by data that describe how participants felt about exploring their values (see Table 6). One-way ANOVAs revealed that participants who preferred HIP over the rating method agreed more

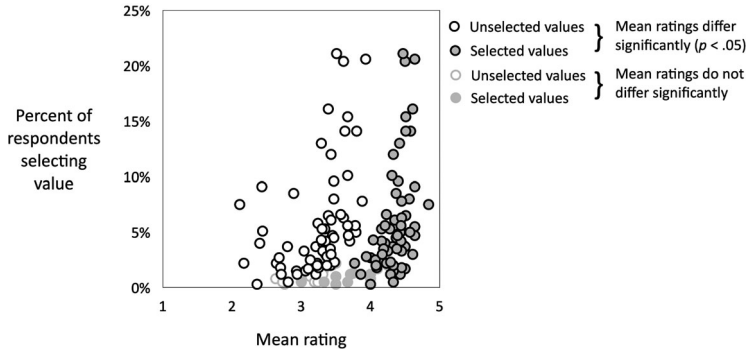

Figure 4. Mean rating assigned to each value. Separate means are shown for respondents who did and did not select the value as one of their four core values; black circles are used to indicate when these means were significantly different (t-test $p < .05$).

Table 5. Method for assigning scores to values based on their HIP selection, and descriptive statistics for the ratings assigned to values with each HIP score.

Category	Value	HIP score	Number of values with HIP score	Mean rating [and 95% CI] for values with HIP score
Unselected	Unselected	1	48	3.00 [2.95–3.05]
Secondary	Unselected	2	16	3.54 [3.49–3.60]
Primary	Unselected	3	8	3.54 [3.48–3.59]
	Secondary	4	4	4.06 [4.00–4.11]
	Primary	5	4	4.41 [4.36–4.46]
All	All	–	80	3.29 [3.24–3.33]

strongly that the survey taught them a lot about their personal values, and that they knew more about themselves as a result. Participants who preferred HIP also enjoyed exploring their personal values more than participants who preferred the rating method. The effect sizes for all of these differences were relatively small, but are still noteworthy given that participants probably chose to complete the survey not because they were especially keen to explore their values, but because they hoped to earn $2.50 for a task that could be completed in only a few minutes.

Discussion

Practitioners who wish to help clients explore their values in a variety of applications may seek an accurate, user-friendly tool for identifying core values – one that provides an alternative to existing approaches that emphasize quantitative rather than qualitative assessment. HIP appears to meet this need. The method is transparent to the user and highly efficient. Our data suggest that HIP enables the selection of four core values from 80 candidates in the same 5–7 min needed to rate those 80 values, even though selecting values with HIP prompts deeper reflection than rating the values. Eliminating the attention checks would reduce the HIP completion time further, as these required very close reading; some respondents described them as "sneaky." However, even with the attention checks, the time needed to select four values from 80 options in HIP compares favorably to other methods, such as the 7–10 min required for Schwartz's 29-portrait approach (Schwartz et al., 2001) and the 3 min needed for an 11-value assessment using BWS (Lee et al., 2008).

The response bias data also support the use of HIP. The ratings exhibited acquiescence bias that resulted in a majority of values receiving high scores. Many respondents assigned their highest rating to a large number of values, pointing to the need for additional – presumably time-consuming – steps should users wish to identify a smaller subset of core values. The 5-point, symmetrical scale we used undoubtedly contributed to this problem, forcing a large number of identical ratings (since the ratings for 80 values were distributed among only five response options). These issues could perhaps be reduced by adopting a 7- or 9-point, asymmetrical rating scale. Response option labels can also be designed to mitigate acquiescence; for example, Schwartz (1992) used a 9-point scale where responses could vary between "opposed to my values" and "of supreme importance." However, the more fundamental problem is that many values are highly regarded. Distinguishing the most relevant values from others that are also appealing but not quite as compelling is a challenging task that is easily bypassed when ratings are used.

Response bias in HIP appears as a tendency to overselect items presented near the top of the page and underselect items near the bottom. This argues for assigning values to page positions at random. However, the response bias we observed in HIP was weak; it attained statistical significance on several survey pages but its magnitude was small and its practical implications debatable. Furthermore, in applied settings HIP would be administered without the accompanying rating task. Thus, the most relevant data in our study were provided by participants who completed HIP before they began the survey's rating section. These participants exhibited statistically significant response bias only on the two final pages, when they had narrowed their choices considerably and were likely selecting between values that all seemed attractive. The difficult nature of these final decisions may have tempted some respondents to simply move the top items into the top box and the remaining items into the bottom box.

We evaluated HIP's concurrent validity by comparing the selections it produced to the ratings that were assigned to

Table 6. Mean ± SD responses to items asking about reactions to completing the survey.

Survey item (1 = *Strongly disagree*; 5 = *Strongly agree*)	Preferred method: Rating ($n = 153$)	Preferred method: No preference ($n = 112$*)	Preferred method: HIP ($n = 337$)	Linear contrast significance
This survey taught me a lot about my personal values	3.45 ± 1.11	3.66 ± 1.08	3.78 ± 0.90	$p = .001$ ($d = 0.335$)
I know more about myself now that I've completed the survey	3.37 ± 1.13	3.56 ± 1.08	3.69 ± 1.01	$p = .003$ ($d = 0.305$)
I enjoyed exploring my personal values with this survey	3.95 ± 1.10	4.15 ± 0.91	4.25 ± 0.82	$p = .003$ ($d = 0.327$)

*Two respondents in this group did not indicate if they enjoyed exploring their personal values.

values. Assessment of HIP's construct validity was precluded because the values examined in this study differ from the values used in other rating methods (e.g., Schwartz, 1992; Schwartz & Boehnke, 2004). Furthermore, ratings are not ideal criteria for assessing validity, given the tendency of participants to broadly endorse many values and avoid discriminating between them. Still, the HIP selections aligned with the ratings: Values received higher ratings when they were identified as primary than when they were not, and values that survived further in the HIP selection tended to receive higher ratings than values that dropped out of HIP earlier. There is only scant evidence that the unpresented values in HIP would have been selected had participants seen those values; 88% of the participants were satisfied with the value selections that emerged from HIP.

HIP was not universally preferred over the rating task, but respondents were twice as likely to favor HIP in comparison to the rating task. Those who preferred HIP also reported learning more about themselves than respondents who favored the rating task. We also found anecdotal evidence suggesting that HIP prompted deeper reflection than the rating task, which may be a general characteristic of ipsative tasks. Indeed, Saville and Willson (1991) suggested that one reason many practitioners prefer ipsative over normative methods is because "life is about choices" (p. 222).

Further evidence is needed to more firmly establish the psychometric foundations of HIP, particularly regarding test-retest reliability. It would also be desirable to establish criterion validity for HIP, perhaps by demonstrating that values selected by HIP are more effective in interventions or other applications than unselected values or values identified using other methods. However, the evidence presented here suggests that further consideration of HIP is warranted.

A limitation of HIP as we tested it is that situational information was not provided to facilitate the selection of values. This is true of most other methods, although approaches that use portraits, scenarios, or group discussions are exceptions. In future studies, it may be desirable to determine whether value preferences change in response to environmental or situational cues. For example, the "teamwork" category was rarely chosen by participants in our study, but could emerge as a selection more often with respondents from collectivist cultures rather than the individualist culture more typical of the United States (Kiffin-Petersen & Cordery, 2003). Similarly, it would be interesting to see if respondents who are employed in businesses that emphasize solving problems in teams select the teamwork category more often when they complete HIP at work than at home.

Of the existing methods used to assess values, card sorts are perhaps most similar to HIP. They have been used extensively for career counseling, where large numbers of options are routinely provided (e.g., Slaney & MacKinnon-Slaney, 2000; Tyler, 1961). At least one implementation for values assessment presents 100 options (Northrup-Snyder, 2021), but we are aware of no studies that have examined how well card sorts work when respondents are presented with so many options that their cognitive limits may be easily exceeded. As Braithwaite and Law (1985) observed, even sorting Rokeach's 36 values presents significant challenges – and those 36 values seem to omit several important areas. Thus, there is need for a method, like HIP, that can accommodate a large number of options without simultaneously overwhelming the respondent.

Our implementation of HIP assessed 80 values, but the method can be easily expanded to accommodate more values should gaps be identified, or contracted if redundancies are found. The approach can also assess preferences for constructs other than values, as long as these can be structured hierarchically, with broad categories that incorporate narrower instantiations of those categories. It is likely as well that the requirement to ensure each category contains the same number of items can be relaxed, although this should be confirmed by further research. Future studies could also explore whether four is the correct number of primary values to identify. HIP can be modified to produce a different number, but determining the optimal number might be challenging. Furthermore, in applied settings, it may not be particularly important whether three, four, or more core values are identified, or even that users feel the identified values are the *most* motivating and inspiring values possible. What may matter more is whether respondents recognize the primary values that emerge from HIP as highly relevant to themselves, and are willing to use those values as inspiration and motivation in achieving positive behavioral change. HIP seems well suited to this purpose.

Open scholarship and supplemental materials

Participants in this study agreed to share their responses publicly in aggregate form, but not individually. (The longer quotes near the end of the Results section were included by permission of the respondents.) Supplemental materials, together with statistical analyses and summaries of aggregated data that support this report, are openly available at https://osf.io/9azvx/?view_only=88076e36103a48e59e3b00c3879339e1.

Acknowledgments

The authors acknowledge the contributions of Chris Hawley, M.D. and Sam Romeo, M.D. of Seity Health LLC in developing the core value assessment procedure described in this paper. Portions of this research were presented at the 29th International Congress of Applied Psychology (Stanislaw & McCreary, 2018).

Declaration of interest statement

The Stanislaus County Office of Education provided funding for this paper. MTurk participants were compensated by Seity Health LLC of Turlock, CA. The authors are external consultants to, but have no financial interest in, Seity Health.

ORCID

Harold Stanislaw http://orcid.org/0000-0002-3214-598X

References

Aguinis, H., Villamor, I., & Ramani, R. S. (2021). MTurk research: Review and recommendations. *Journal of Management*, *47*(4), 823–837. https://doi.org/10.1177/0149206320969787

Allport, G. W., Vernon, P. E., & Lindzey, G. (1970). *Study of values* (3rd ed.). Houghton Mifflin.

Anshel, M. H., Brinthaupt, T. M., & Kang, M. (2010). The disconnected values model improves mental well-being and fitness in an employee wellness program. *Behavioral Medicine (Washington, D.C.)*, *36*(4), 113–122. https://doi.org/10.1080/08964289.2010.489080

Bardi, A., & Schwartz, S. H. (2003). Values and behavior: Strength and structure of relations. *Personality & Social Psychology Bulletin*, *29*(10), 1207–1220. https://doi.org/10.1177/0146167203254602

Berry, K., Rana, R., Lockwood, A., Fletcher, L., & Pratt, D. (2019). Factors associated with inattentive responding in online survey research. *Personality and Individual Differences*, *149*(15), 157–159. https://doi.org/10.1016/j.paid.2019.05.043

Bilsky, W., & Schwartz, S. H. (1994). Values and personality. *European Journal of Personality*, *8*(3), 163–181. https://doi.org/10.1002/per.2410080303

Bono, J. E., & Judge, T. A. (2003). Self-concordance at work: Toward understanding the motivational effects of transformational leaders. *Academy of Management Journal*, *46*(5), 554–571. https://doi.org/10.5465/30040649

Braithwaite, V. A., & Law, H. G. (1985). Structure of human values: Testing the adequacy of the Rokeach Value Survey. *Journal of Personality and Social Psychology*, *49*(1), 250–263. https://doi.org/10.1037/0022-3514.49.1.250

Bronk, K. C., Baumsteiger, R., Mangan, S., Riches, B., Dubon, V., Benavides, C., & Bono, G. (2019). Fostering purpose among young adults: Effective online interventions. *Journal of Character Education*, *15*(2), 21–38. https://www.proquest.com/scholarly-journals/fostering-purpose-among-young-adults-effective/docview/2309267095/se-2?accountid=147094

Brosnan, K., Babakhani, N., & Dolnicar, S. (2019). I know what you're going to ask me." Why respondents don't read survey questions. *International Journal of Market Research*, *61*(4), 366–379. https://doi.org/10.1177/1470785318821025

Brown, D., & Crace, R. K. (1996). Values in life role choices and outcomes: A conceptual model. *The Career Development Quarterly*, *44*(3), 211–223. https://doi.org/10.1002/j.2161-0045.1996.tb00252.x

Cabooter, E., Weijters, B., Geuens, M., & Vermeir, I. (2016). Scale format effects on response option interpretation and use. *Journal of Business Research*, *69*(7), 2574–2584. https://doi.org/10.1016/j.jbusres.2015.10.138

Chan, W. (2003). Analyzing ipsative data in psychological research. *Behaviormetrika*, *30*(1), 99–121. https://doi.org/10.2333/bhmk.30.99

Crocker, J., Niiya, Y., & Mischkowski, D. (2008). Why does writing about important values reduce defensiveness? Self-affirmation and the role of positive other-directed feelings. *Psychological Science*, *19*(7), 740–747. https://doi.org/10.1111/j.1467-9280.2008.02150.x

Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology*, *66*, 4–19. https://doi.org/10.1016/j.jesp.2015.07.006

Döring, A. K., Blauensteiner, A., Aryus, K., Drögekamp, L., & Bilsky, W. (2010). Assessing values at an early age: The picture-based value survey for children (PBVS-C). *Journal of Personality Assessment*, *92*(5), 439–448. https://doi.org/10.1080/00223891.2010.497423

Emanuel, A. S., Howell, J. L., Taber, J. M., Ferrer, R. A., Klein, W. M., & Harris, P. R. (2018). Spontaneous self-affirmation is associated with psychological well-being: Evidence from a US national adult survey sample. *Journal of Health Psychology*, *23*(1), 95–102. https://doi.org/10.1177/1359105316643595

Epton, T., Harris, P. R., Kane, R., van Koningsbruggen, G. M., & Sheeran, P. (2015). The impact of self-affirmation on health-behavior change: A meta-analysis. *Health Psychology : official Journal of the Division of Health Psychology, American Psychological Association*, *34*(3), 187–196. https://doi.org/10.1037/hea0000116

Feather, N. T. (1995). Values, valences, and choice: The influences of values on the perceived attractiveness and choice of alternatives. *Journal of Personality and Social Psychology*, *68*(6), 1135–1151. https://doi.org/10.1037/0022-3514.68.6.1135

Gorsuch, R. L. (1970). Rokeach's approach to value systems and social compassion. *Review of Religious Research*, *11*(2), 139–143. https://doi.org/10.2307/3510277

Gouveia, V. V., Milfont, T. L., & Guerra, V. M. (2014). Functional theory of human values: Testing its content and structure hypotheses. *Personality and Individual Differences*, *60*(2014), 41–47. https://doi.org/10.1016/j.paid.2013.12.012

Graham-Rowe, E., & Jessop, D. C., & Sparks, P. (2019). Self-affirmation theory and pro-environmental behaviour: Promoting a reduction in household food waste. *Journal of Environmental Psychology*, *62*(April), 124–132. https://doi.org/10.1016/j.jenvp.2019.02.003

Hardcastle, S. J., Hancox, J., Hattar, A., Maxwell-Smith, C., Thøgersen-Ntoumani, C., & Hagger, M. S. (2015). Motivating the unmotivated: How can health behavior be changed in those unwilling to change? *Frontiers in Psychology*, *6*, 835. https://doi.org/10.3389/fpsyg.2015.00835

Horstmann, K. T., Rauthmann, J. F., Sherman, R. A., & Ziegler, M. (2021). Unveiling an exclusive link: Predicting behavior with personality, situation perception, and affect in a preregistered experience sampling study. *Journal of Personality and Social Psychology*, *120*(5), 1317–1343. https://doi.org/10.1037/pspp0000357

Howell, A. J. (2017). Self-affirmation theory and the science of well-being. *Journal of Happiness Studies*, *18*(1), 293–311. https://doi.org/10.1007/s10902-016-9713-5

Jex, S. M., Adams, G. A., Bachrach, D. G., & Sorenson, S. (2003). The impact of situational constraints, role stressors, and commitment on employee altruism. *Journal of Occupational Health Psychology*, *8*(3), 171–180. https://doi.org/10.1037/1076-8998.8.3.171

Jiang, L. (2018). Job insecurity and creativity: The buffering effect of self-affirmation and work-affirmation. *Journal of Applied Social Psychology*, *48*(7), 388–397. https://doi.org/10.1111/jasp.12519

Kiffin-Petersen, S., & Cordery, J. (2003). Trust, individualism and job characteristics as predictors of employee preference for teamwork. *The International Journal of Human Resource Management*, *14*(1), 93–116. https://doi.org/10.1080/09585190210158538

Koole, S. L., Smeets, K., Van Knippenberg, A., & Dijksterhuis, A. (1999). The cessation of rumination through self-affirmation. *Journal of Personality and Social Psychology*, *77*(1), 111–125. https://doi.org/10.1037/0022-3514.77.1.111

Lee, J. A., Bardi, A., Gerrans, P., Sneddon, J., Van Herk, H., Evers, U., & Schwartz, S. (2021). Are value-behavior relations stronger than previously thought? It depends on value importance. *European Journal of Personality*, *36*(2), 133–148. https://doi.org/10.1177/08902070211002965

Lee, J. A., Sneddon, J. N., Daly, T. M., Schwartz, S. H., Soutar, G. N., & Louviere, J. J. (2019). Testing and extending Schwartz refined value theory using a best-worst scaling approach. *Assessment*, *26*(2), 166–180. https://doi.org/10.1177/1073191116683799

Lee, J. A., Soutar, G., & Louviere, J. (2008). The best-worst scaling approach: An alternative to Schwartz's Values Survey. *Journal of Personality Assessment*, *90*(4), 335–347. https://doi.org/10.1080/00223890802107925

Lindsay, E. K., & Creswell, J. D. (2014). Helping the self help others: Self-affirmation increases self-compassion and pro-social behaviors. *Frontiers in Psychology*, *5*(421), 421–429. https://doi.org/10.3389/fpsyg.2014.00421

Louviere, J. J., Flynn, T. N., & Marley, A. A. J. (2015). *Best-worst scaling: Theory, methods and applications*. Cambridge University Press.

McClelland, D. C. (1985). How motives, skills, and values determine what people do. *American Psychologist*, *40*(7), 812–825. https://doi.org/10.1037/0003-066X.40.7.812

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review*, *96*(4), 690–702. https://doi.org/10.1037/0033-295X.96.4.690

Meglino, B. M., & Ravlin, E. C. (1998). Individual values in organizations: Concepts, controversies, and research. *Journal of Management*, *24*(3), 351–389. https://doi.org/10.1016/S0149-2063(99)80065-8

Miyake, A., Kost-Smith, L. E., Finkelstein, N. D., Pollock, S. J., Cohen, G. L., & Ito, T. A. (2010). Reducing the gender achievement gap in college science: A classroom study of values affirmation. *Science (New York, N.Y.)*, *330*(6008), 1234–1237. https://doi.org/10.1126/science.1195996

Mosconi, J., & Emmett, J. (2003). Effects of a values clarification curriculum on high school students' definitions of success. *Professional School Counseling*, *7*(2), 68–78. https://www.jstor.org/stable/42732546

Nelson, S. K., Fuller, J. A., Choi, I., & Lyubomirsky, S. (2014). Beyond self-protection: Self-affirmation benefits hedonic and eudaimonic well-being. *Personality & Social Psychology Bulletin*, *40*(8), 998–1011. https://doi.org/10.1177/0146167214533389

Northrup-Snyder, K. (2021). *Value card sort online game*. https://motivationalinterviewing.org/value-card-sort-online-game

Porter, S. R., Whitcomb, M. E., & Weitzer, W. H. (2004). Multiple surveys of students and survey fatigue. *New Directions for Institutional Research*, *2004*(121), 63–73. https://doi.org/10.1002/ir.101

Roccas, S., Sagiv, L., Schwartz, S. H., & Knafo, A. (2002). The Big Five personality factors and personal values. *Personality and Social Psychology Bulletin*, *28*(6), 789–801. https://doi.org/10.1177/0146167202289008

Rokeach, M. (1973). *Nature of human values*. Free Press.

Sagiv, L., Roccas, S., & Oppenheim-Weller, S. (2015). Values and well-being. In S. Joseph (Ed.), *Positive psychology in practice: Promoting human flourishing in work, health, education, and everyday life* (pp. 103–120). Wiley. https://doi.org/10.1002/9781118996874.ch7

Sandy, C. J., Gosling, S. D., Schwartz, S. H., & Koelkebeck, T. (2017). The development and validation of brief and ultrabrief measures of values. *Journal of Personality Assessment*, *99*(5), 545–555. https://doi.org/10.1080/00223891.2016.1231115

Saville, P., & Willson, E. (1991). The reliability and validity of normative and ipsative approaches in the measurement of personality. *Journal of Occupational Psychology*, *64*(3), 219–238. https://doi.org/10.1111/j.2044-8325.1991.tb00556.x

Schonlau, M., & Toepoel, V. (2015). Straightlining in Web survey panels over time. *Survey Research Methods*, *9*(2), 125–137. https://doi.org/10.18148/srm/2015.v9i2.6128

Schüz, N., & Schüz, B. (2017). Self-affirmation: Protecting the self and protecting subjective well-being. In M. Robinson & M. Eid (Eds.), *The happy mind: Cognitive contributions to well-being*. Springer. https://doi.org/10.1007/978-3-319-58763-9_16

Schwartz, S. H., Melech, G., Lehmann, A., Burgess, S., Harris, M., & Owens, V. (2001). Extending the cross-cultural validity of the theory of basic human values with a different method of measurement. *Journal of Cross-Cultural Psychology*, *32*(5), 519–542. https://doi.org/10.1177/0022022101032005001

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 1–65). Elsevier. https://doi.org/10.1016/S0065-2601(08)60281-6

Schwartz, S. H. (2010). Basic values: How they motivate and inhibit prosocial behavior. In M. Mikulincer & P. R. Shaver (Eds.), *Prosocial motives, emotions, and behavior: The better angels of our nature* (pp. 221–241). American Psychological Association. https://doi.org/10.1037/12061-012

Schwartz, S. H. (2011). Values: Cultural and individual. In F. J. R. van de Vijver, A. Chasiotis, & S. M. Breugelmans (Eds.), *Fundamental questions in cross-cultural psychology* (pp. 463–493). Cambridge University Press. https://doi.org/10.1017/CBO9780511974090.019

Schwartz, S. H., & Boehnke, K. (2004). Evaluating the structure of human values with confirmatory factor analysis. *Journal of Research in Personality*, *38*(3), 230–255. https://doi.org/10.1016/S0092-6566(03)00069-2

Schwartz, S. H., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., Ramos, A., Verkasalo, M., Lönnqvist, J.-E., Demirutku, K., Dirilen-Gumus, O., & Konty, M. (2012). Refining the theory of basic individual values. *Journal of Personality and Social Psychology*, *103*(4), 663–688. https://doi.org/10.1037/a0029393

Schwartz, S. H., & Sortheix, F. (2018). Values and subjective well-being. In E. Diener, S. Oishi, & L. Tay (Eds.), *Handbook of well-being* (pp. 872–886). Noba Scholar. http://www.nobascholar.com/chapters/51

Seity Health LLC. (2021). *Unleash the potential in your people*. https://seityhealthinfo.com/

Sheldon, K. M., & Elliot, A. J. (1999). Goal striving, need satisfaction, and longitudinal well-being: The self-concordance model. *Journal of Personality and Social Psychology*, *76*(3), 482–497. https://doi.org/10.1037/0022-3514.76.3.482

Sheldon, K. M., Elliot, A. J., Ryan, R. M., Chirkov, V., Kim, Y., Wu, C., Demir, M., & Sun, Z. (2004). Self-concordance and subjective well-being in four cultures. *Journal of Cross-Cultural Psychology*, *35*(2), 209–223. https://doi.org/10.1177/0022022103262245

Sherman, D. K. (2013). Self-affirmation: Understanding the effects. *Social and Personality Psychology Compass*, *7*(11), 834–845. https://doi.org/10.1111/spc3.12072

Slaney, R. B., & MacKinnon-Slaney, F. (2000). Using vocational card sorts in career counseling. In C. E. Watkins, Jr., & V. L. Campbell (Eds.), *Testing and assessment in counseling practice* (2nd ed., pp. 371–429). Lawrence Erlbaum.

Stanislaw, H., & McCreary, J. (2018). Higher order thinking: An efficient method for ipsative preference assessments [Poster presentation]. *29th International Congress of Applied Psychology*.

The Values Project. (2020). *Let's reveal what matters to you*. https://www.thevaluesproject.com/

Thomaes, S., Bushman, B. J., de Castro, B. O., & Reijntjes, A. (2012). Arousing "gentle passions" in young adolescents: Sustained experimental effects of value affirmations on prosocial feelings and behaviors. *Developmental Psychology*, *48*(1), 103–110. https://doi.org/10.1037/a0025677

Tyler, L. E. (1961). Research explorations in the realm of choice. *Journal of Counseling Psychology*, *8*(3), 195–201. https://doi.org/10.1037/h0041019

Vernon, P. E., & Allport, G. W. (1931). A test for personal values. *The Journal of Abnormal and Social Psychology*, *26*(3), 231–248. https://doi.org/10.1037/h0073233

Winkler, J. D., Kanouse, D. E., & Ware, J. E. (1982). Controlling for acquiescence response set in scale development. *Journal of Applied Psychology*, *67*(5), 555–561. https://doi.org/10.1037/0021-9010.67.5.555

Zhou, H., & Fishbach, A. (2016). The pitfall of experimenting on the web: How unattended selective attrition leads to surprising (yet false) research conclusions. *Journal of Personality and Social Psychology*, *111*(4), 493–504. https://doi.org/10.1037/pspa0000056