

Are some attitudes more self-defining than others? Assessing self-related attitude functions and their consequences

Peter V. Zunick, Jacob D. Teeny, and Russell H. Fazio

The Ohio State University

Personality and Social Psychology Bulletin, in press

Acknowledgements: We would like to thank Matthew Rocklage, Elise Bui, Aaron Hatchett, Javier Granados Samayoa, Lindsey Johnston, Andrew Luttrell, and Richard Petty for their helpful comments at various stages of this project.

Word count: 10,999

Abstract

Attitudes serve multiple functions, some related to the self-concept (Katz, 1960; Smith, Bruner, & White, 1956). We call attitudes that help people define who they are “self-defining.” Across four studies, we tested a brief self-report measure of the extent to which an attitude is self-defining. Studies 1 and 2 showed that self-defining attitudes tend to be extreme, positive, and unambivalent. Studies 3 and 4 produced two main findings. First, self-definition was related to, but not redundant with, a number of other characteristics of the attitude (e.g., attitude certainty). Second, self-definition predicted intentions to spontaneously advocate and, in Study 4, reactions to an opportunity to advocate behaviorally (i.e., writing about their attitude in an optional response box) following a self-threat. Overall, the results highlight the utility of this approach and, more broadly, demonstrate the value of considering the role of the self in attitudinal processes, and vice versa.

Keywords: attitudes, attitude functions, self-concept, identity, advocacy

Theorists have long held that attitudes (i.e., evaluations) serve multiple functions. Beyond the obvious object-appraisal function of directing people toward positive and away from negative outcomes, attitudes also express important aspects of one's self-concept and identity (Katz, 1960; Smith, Bruner, & White, 1956). For instance, someone's attitude toward a political cause, sports team, or clothing brand might communicate a lot, both to that person and to the world, about what kind of person he or she is. This simple idea, that attitudes can help us define who we are, has the potential to link together two vast literatures, those pertaining to attitudes and the self-concept. Unfortunately, methodological issues have made it difficult to empirically examine this self-concept-related attitude function. We introduce an efficient method of measuring the extent to which an attitude serves this self-related function for an individual, and we test several novel hypotheses to demonstrate the potential of this approach to studying the relationship between attitudes and the self.

We call attitudes that help people define who they are *self-defining*. This term does not necessarily reflect a departure from previous theoretical approaches. Rather, we hope to clarify and expand upon a theme that has been touched upon in prior work, both implicitly and explicitly, yet at the same time conflated with similar but, in our eyes, distinct constructs, such as Katz's (1960) value-expressive function and Smith and colleagues' (1956) social-adjustment function. We view self-defining attitudes as those that answer the question, "Who am I?" They help people define themselves, both for their own self-knowledge (Campbell et al., 1996) and to communicate their identities and self-concepts to others. And because a large part of one's identity comprises not just who one is and what one values (Schwartz, 1992) at the moment, but also what one aspires to be in the future (Markus & Nurius, 1986), we include these aspects in our conception of what makes an attitude self-defining.

The current work has two primary goals: first, to provide a new, easy-to-use measure of a construct that has proven difficult to operationalize in the past; and second, to use this measure to explore the characteristics and consequences of self-defining attitudes. A third, more abstract goal, is to highlight the role of the self in attitude functions, as well as the role of attitudes in creating, supporting, and defining the self-concept, thus demonstrating the theoretical advantages of linking these two literatures together.

Past theory and research

The idea that some attitudes help people define who they are has a rich foundation in social psychological theory. Early theorists interested in attitude functions proposed that some attitudes serve value-expressive (Katz, 1960) or social-adjustment (Smith et al., 1956) functions. Our conception of self-defining attitudes incorporates Smith and colleagues' (1956) focus on attitudes as tools for, and signals of, group affiliations, as well as Katz's (1960) emphasis on the inherent satisfaction of confirming one's identity. However, as we have described it above, the self-defining function does not perfectly overlap with either of these functions. For instance, with regard to the value-expressive function, consider two attitude objects, gay marriage and running. A person might support (or oppose) marriage equality based on values (e.g., equality, tradition) without that attitude necessarily defining their self-concept. Conversely, an avid runner (or hopeless couch potato) might consider his or her attitude toward running highly self-defining, but for reasons unrelated to values (e.g., self-perception of one's daily activities, beliefs about one's self-control or overall fitness). The two constructs, then, are conceptually related but nonetheless distinct. Self-defining attitudes communicate "who I am" but need not focus on a value per se, and attitudes that express a value need not be self-defining. Likewise, with regard to the social-adjustment function, an attitude may indicate a group affiliation (e.g., a rider of municipal buses)

without that affiliation being central to one's self-concept, and conversely, not every self-defining attitude necessarily expresses group affiliation.

Beyond the umbrella of functional theory, researchers have examined several additional attributes of attitudes that are somewhat related to the self-defining function upon which we focus. For example, ego-involvement is conceptually similar to self-definition, at least in one classic formulation (Sherif & Cantril, 1947), but is problematic due to the wide variety of ways in which it has been used in different areas of research. Greenwald (1982) identified three distinct meanings of ego-involvement in the literature, only one of which was even related to attitudes, and Johnson and Eagly (1989) further subdivided the attitudinal sense of ego-involvement into three different meanings (value-relevant, impression-relevant, and outcome-relevant involvement; cf. Petty, Cacioppo, & Haugtvedt, 1992). Thus, the very meaning of the term is clouded, and ego-involvement cannot be equated solely with self-definition. Additionally, personal relevance (Petty & Cacioppo, 1986) and importance (Boninger, Krosnick, & Berent, 1995; Judd & Krosnick, 1982) of an attitude appear similar to self-definition, but are conceptually distinct. Although self-defining attitudes should be personally important and relevant, the inverse is not true. For example, a person might support a tax cut that would benefit them, such that the attitude is personally relevant and important, but this attitude need not define the person's sense of self. Self-defining attitudes are conceptually distinct enough from these existing attitudinal constructs to warrant closer examination.

Another relevant line of work, situated within the marketing and consumer psychology literature, explores associations that form between brands and consumers' self-concepts, a phenomenon termed self-brand connection (e.g., Escalas & Bettman, 2003, 2005). Although the two constructs are related, there are several important distinctions between self-brand

connections and the self-defining attitude function. First and foremost, although the self-brand connection literature suggests brands can elicit self-defining attitudes, our conception of self-defining attitudes broadly encompasses many potential attitude objects – self-definition need not be limited to brands. In addition, self-definition need not necessarily be limited to positive attitudes. By focusing on links between the self and an attitude (regardless of its valence), our conceptual framework allows for the possibility that dislikes can be self-defining. In contrast, focusing on connections between the self and a brand (i.e., an attitude *object*) necessitates a positive attitude toward that brand.

In short, although the self-defining attitude function is distinct from existing theoretical constructs, it is not entirely unrelated to these constructs. After all, our ideas grew out of careful readings of classic attitude functions theory, especially regarding Katz's (1960) value-expressive and Smith and colleagues' (1956) social adjustment functions. However, the extent to which an attitude helps one define his or her self-concept does not perfectly overlap with these functions. More specifically, we do not view self-definition as either a superordinate or subordinate category to value expression or social adjustment, or to the other related constructs discussed above, though future work might explore such hierarchical relationships. We instead consider self-definition to be one of the most intriguing reasons why one might be motivated to engage in value-expression or social adjustment, or, for that matter, why an attitude might be personally important or relevant, or why a brand might be connected to the self. Our goal is simply to see what can be gained by focusing explicitly on the role of attitudes in defining one's self-concept.

An additional goal of the current research is to provide an efficient methodological tool to advance the study of self-defining attitudes. Although the early functional theorists inspired much subsequent research, operationalizing the various attitude functions has proven difficult.

One approach relied on individual difference measures. For example, in one set of studies, researchers explored the effects of matching persuasive appeals to recipients based on their level of self-monitoring, assuming that high self-monitors' attitudes tend to be more social-adjustive than those of low self-monitors (Petty & Wegener, 1998; Snyder & DeBono, 1985). The basic idea was to identify, via an individual difference measure, inter-person variation in people's average attitude functions across objects. Other work instead took an inter-object approach, exploring the idea that certain attitude objects lend themselves to certain functions across people (Shavitt, 1989, 1990). A comprehensive approach to measuring attitude functions must account for attitude functions that differ between people *and* objects (i.e., Person A may have a self-defining attitude toward Object A but not Object B, whereas Person B may show the opposite pattern).¹

Current research

In the current work, we propose a brief new measure of the extent to which any given attitude fulfills a self-defining function for an individual. This approach allows for the simultaneous assessment of inter-person and inter-object variability in attitude functions. In Studies 1 and 2, we created and validated the measure, and also tested a novel hypothesis regarding the predominant valence of self-defining attitudes. In Study 3, we tested whether self-definition predicts spontaneous advocacy intentions. In Study 4, we examined whether self-definition predicts compensatory advocacy following threat (Gal & Rucker, 2010). In Studies 3 and 4, we also included measures of the aforementioned personal importance and relevance, along with other indicators of attitude strength (Petty & Krosnick, 1995), for two reasons. First,

¹ Some past work has focused on inter-person variation in attitude functions for one specific attitude object (e.g., developing object-specific scales, coding free response essays; Herek, 1987). However, this approach is impractical for researchers interested in more than one attitude object, and theoretically only represents half of the current person-by-object approach.

we expected self-definition to be related to these variables. Second, we expected the effects of self-definition to not be simply due to shared variance with one or more of these variables. In other words, we included these other attitude-related variables to help establish both convergent and divergent validity regarding the self-definition measure and construct.

Study 1

Our first task was to develop and validate a self-report measure of the extent to which an attitude is self-defining. In developing the measure, we sought to cover the key theoretical aspects of what makes an attitude self-defining while being brief enough to allow for inclusion in surveys measuring attitudes toward multiple objects. In validating the measure, we pursued two strategies. First, we identified attitude objects for which we expected participants' attitudes would be more self-defining (e.g., the local university football team, feminists) vs. less self-defining (e.g., an irrelevant non-rival football team, a coffee brand). We formed these expectations based both on past work research on between-object differences in attitude functions (Shavitt, 1989, 1990) and on our intuitions regarding which objects might tend to elicit attitudes more or less related to participants' self-concepts. Second, we expected that extreme attitudes would be more self-defining than moderate attitudes, reasoning that people might prefer to define themselves by their passions rather than their slight preferences, ambivalent attitudes, or non-attitudes. Moreover, self-definition might also promote attitude polarization, either directly or indirectly via the behavior it promotes, raising the possibility that extremity and self-definition influence each other in a recursive manner.

In addition, for exploratory purposes, we planned to test whether positive or negative attitudes might, on average, be more self-defining. Because people tend to associate good things with the self (e.g., Pelham, Mirenberg, & Jones, 2002; Sedikides & Strube, 1998; Taylor &

Brown, 1992) as well as approach and, hence, experience liked objects more frequently than disliked ones (Fazio, Eiser, & Shook, 2004), we thought that positive attitudes might be more self-defining than negative attitudes. On the other hand, because negatives often carry more psychological weight than positives (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001), particularly in the attitudinal domain (Bizer, Larsen & Petty, 2011; Shook, Fazio, & Eiser, 2007), we were open to the alternative that negative attitudes would be more self-defining, or that neither valence bias would occur.

We also asked participants to provide their own idiosyncratic self-defining attitudes. This method could potentially be useful for researchers concerned not with specific attitude objects per se, but rather with whatever is highly self-defining for specific participants.

Method

Participants

Fifty-nine undergraduates (43 female, 16 male) participated as part of an introductory psychology course requirement.

Attitude measurement

Participants rated each of 16 attitude objects on a 7-pt scale from -3, “Dislike very much,” to +3, “Like very much.” We selected these objects to elicit wide variation in attitude and self-definition ratings and to represent different types of attitude objects, including brands, activities, groups, concepts, and places.

Idiosyncratic self-defining attitudes

Participants then listed four attitudes that were self-defining for them personally (two positive and two negative). Specifically, we asked them to list two things they were proud of liking (disliking) and would want other people to know that they liked (disliked).

Self-defining self-report measure

Participants then responded to two statements with regard to each of the 16 attitude objects and their idiosyncratically self-defining attitude objects: “My evaluation of <object> reflects the kind of person I am or aspire to be,” and “My evaluation of <object> says something, both to myself and others, about who I am as an individual.”² The appropriate attitude object appeared in place of <object>, and participants responded on a 7-pt Likert scale from 1, “Strongly disagree,” to 7, “Strongly agree.”³

Results

Participants’ attitudes to the 16 objects varied widely (across all participants and objects: $M = .87$, $SD = 1.69$). We averaged each participants’ responses to the two self-definition self-report items for each attitude (across all participants and objects, $r = .70$, $p < .001$) to produce 16 self-definition scores for each participant ($M = 4.25$, $SD = 1.62$).

Object ranking

We averaged self-defining ratings across participants for each of the 16 attitude objects separately and then rank-ordered them from most to least self-defining (Table 1). The results largely confirmed our expectations, with participants’ attitudes toward their hometowns, feminism, and the home team among the most self-defining, and their attitudes toward unremarkable brands of coffee and grocery stores, along with irrelevant non-rival sports teams,

² Our goal to create an extremely efficient measure that could be widely used even in studies with multiple attitude objects resulted in each item being somewhat “double-barreled” (i.e., tapping into two distinct constructs). To assess whether this feature was problematic, in a separate study ($N=101$) we split the two items into four (i.e., whether an attitude “reflects the kind of person *I am*,” “reflects the kind of person *I aspire to be*,” “says something *to others* about who I am as an individual,” and “says something *to myself* about who I am as an individual.”) Across four attitude objects (Apple products, running/jogging, environmentalism, and professional football), these four items showed high internal consistency (mean $\alpha = .95$) and loaded onto a single factor (mean variance accounted for = 87.82%; see the online supplement for more details). Thus, unless researchers have specific hypotheses requiring otherwise (e.g., different predictions for actual vs. possible selves or self-knowledge vs. social communication), the two-item “double-barreled” measure appears to effectively assess the general construct of self-definition.

³ We report the primary measures of interest for all four studies in the main text. Descriptions of exploratory measures not central to the hypotheses appear in the online supplement.

among the least. Very near the midpoint were two attitudes that we expected would be self-defining for some people but not others: eating steak and Nike apparel.

Table 1
Ranking of attitude objects by mean self-defining ratings in Study 1

Object	Self-definition mean
Hometown	5.64
Running / jogging	5.20
Feminism	5.09
Ohio State Buckeyes	5.04
Organic food	4.84
Democrats	4.80
Capital punishment	4.50
Apple products	4.47
McDonalds	4.24
Eating steak	4.05
Nike apparel	3.97
Tea party	3.86
Giant Eagle Groceries	3.64
Notre Dame Fighting Irish	3.31
Arkansas Razorbacks	2.97
Folgers coffee	2.44

Note: Self-definition scores could range from 1 to 7.

Extremity effect

Testing whether more extreme attitudes were more self-defining required a statistical model that would account for the dependencies within the data (i.e., ratings by the same participant might be more similar than those made by different participants, and ratings of the same object might be more similar than ratings of different objects). Indeed, two simple random effects models, one with an intercept randomly varying between participants and the other between objects, revealed exactly this type of non-independence (ICCs: .19 and .27). For the

extremity analysis, we used a random-effects model crossing participants with objects (Judd, Westfall, & Kenny, 2012), allowing both intercepts to randomly vary⁴, and including as predictors fixed linear and quadratic terms for attitude ratings. This model is analogous to a linear regression model with a quadratic term, but it accounts for non-independence among observations (Judd et al., 2012). The quadratic term was significant, $b = .12$ (95% CI: .10, .15), $t(906.71) = 10.05$, $p < .001$, $R^2_{\beta} = .10^5$, such that more extreme attitude ratings predicted higher self-defining ratings (Figure 1).

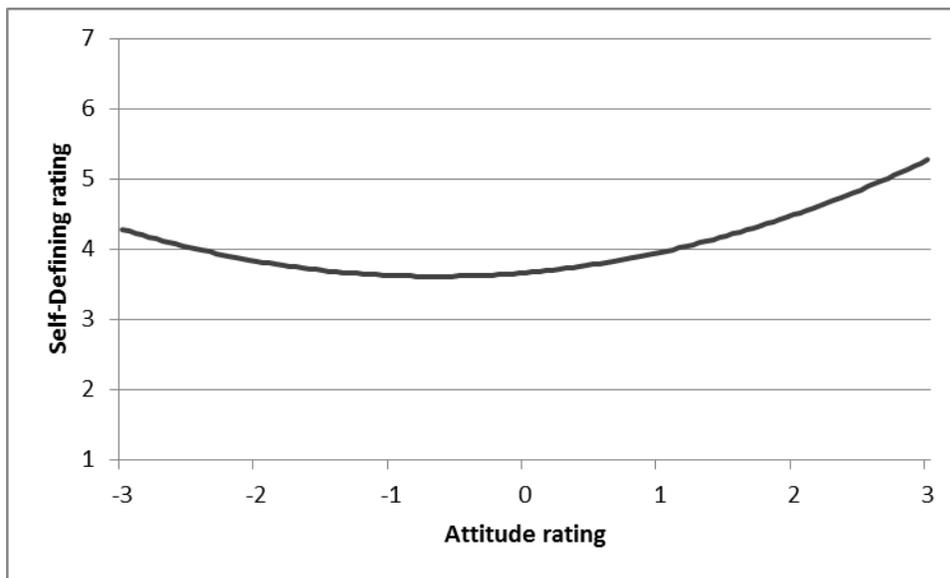


Figure 1. Self-defining ratings predicted by attitude ratings (represented in the model with both a linear and quadratic effect) in Study 1.

Valence effect

⁴ For the sake of simplicity, the slopes in this model (and others reported in this paper) do not randomly vary. However, including random slopes results in very similar estimates of the fixed effects in all models reported.

⁵ R^2_{β} is a pseudo semi-partial R^2 statistic designed for mixed models to mimic the interpretation of R^2 for a single regression coefficient in simple linear regression (i.e., proportion of variance explained uniquely by a predictor; Edwards, Muller, Wolfinger, Qaqish, & Schabenberger, 2008).

To test whether positive or negative attitudes tended to be more self-defining, we examined the linear term in the above model. This term was positive and significant, $b = .16$ (95% CI: .11, .22), $t(901.854) = 5.98$, $p < .001$, $R^2_{\beta} = .04$, indicating that, even controlling for extremity, positive attitudes were more self-defining than negative attitudes.

Idiosyncratic self-defining attitudes

The idiosyncratic self-defining attitudes participants provided varied widely, and included activities, values, groups, individuals, items, and concepts (e.g., likes: yoga, dogs, Aerosmith, Lake Huron, science, religion, family, chocolate; dislikes: politics, Wal-Mart, rape, writing, racism, carrots, LeBron James⁶, airplanes). Unsurprisingly, participants rated these attitudes as highly self-defining, although in line with the above valence effect, likes ($M = 5.89$, $SD = 1.14$) were more self-defining than dislikes ($M = 5.02$, $SD = 1.53$), mean difference = .86 (95% CI: .55, 1.21), $t(58) = 5.07$, $p < .001$.

Discussion

Study 1 produced a number of interesting findings. The two-item measure of self-definition, designed to be brief and face-valid, seemed effective at measuring the construct. The measure differentiated between attitude objects that we expected to be more and less self-defining for our sample (e.g., attitudes toward the home football team were more self-defining than attitudes toward an irrelevant football team). The measure also confirmed our hypothesis regarding an extremity effect, such that more extreme attitudes tended to be more self-defining, in line with both intuition and previous work on related constructs (i.e., attitudes regarding personally important issues also tend to be extreme; Judd & Krosnick, 1982).

⁶ Data collection occurred in Ohio in early 2014, while James was playing for the Miami Heat and before his return to his native Ohio later that year (and his historic NBA championship with the Cleveland Cavaliers in 2016).

The self-definition measure also revealed a positivity effect, such that positive attitudes tended to be more self-defining than negative attitudes. This finding fits well with the broad literature on self-enhancement showing that people like to associate the self with positive things (e.g., Sedikides & Strube, 1998; Taylor & Brown, 1992). In addition, because positive attitudes lead to approach behavior and negative attitudes avoidance (Fazio et al., 2004), positive but not negative attitudes are likely to result in frequent experiences with the favored object. These frequent experiences may create associations between the self and the object or attitude, and in the end these experiences may come to define the self much more than equivalent avoidance behavior stemming from negative attitudes.

Broadly speaking, our general framework for thinking about and measuring self-defining attitudes proved valuable. The new measure allows one to easily assess the extent to which multiple attitudes of interest are self-defining for an individual. Indeed, our random-effects models revealed plenty of variation in self-defining ratings between both objects and people, and participants' idiosyncratic self-defining attitude objects also varied considerably. These results strongly support the utility of treating attitude functions as the product of person-by-object interactions rather than as differing only between people (Snyder & DeBono, 1985) or objects (Shavitt, 1990).

Study 2

The primary aim of Study 2 was to replicate the findings from Study 1. We made several minor alterations to the method, the most notable of which was the inclusion of a language-based attitudinal measure drawing from an "evaluative lexicon" of adjectives people use to describe their attitudes (Rocklage & Fazio, 2015). This measure provides valence, extremity, and emotionality scores based on the adjectives participants choose from a list. We hoped to confirm

the extremity and positivity effects from Study 1 with both this measure and traditional attitude ratings. Given that the evaluative lexicon adjective task produces an emotionality score, we also planned to explore whether attitudes characterized by more emotionality might be more self-defining, or whether instead attitudes can be equally self-defining regardless of the extent to which they involve emotional reactions. Finally, the evaluative lexicon adjective task allows one to assess attitudinal ambivalence. We predicted that univalent attitudes would be more self-defining, on average, than ambivalent attitudes, both because people would presumably be reluctant to stake their identities on an issue toward which they have strongly mixed feelings (Fazio, 1995), and because self-definition might motivate people to resolve whatever ambivalence they might experience.

Method

Participants

Sixty-seven undergraduates (37 female, 29 male, 1 unspecified) participated as part of an introductory psychology course requirement.

Attitude measurement

Attitude measurement was similar to Study 1 and included the same attitude objects, with the following exceptions. We reworded several items for clarity (e.g., added “football team” to the end of “The Ohio State Buckeyes”). To avoid ideological asymmetry we added “The Republican party.” To ensure that the positivity effect in Study 1 was not an artifact due to a lack of objects eliciting positive evaluations but low self-defining ratings, we included “Mini-golf.” Finally, due to time constraints, we removed three attitude objects that seemed redundant.

Idiosyncratic self-defining attitudes

Participants provided one personally self-defining like and one dislike following a procedure similar to Study 1. Unlike in Study 1, participants then gave attitude ratings for these two objects so that we could include them in the extremity and valence analyses.

Evaluative lexicon adjective task

Next, participants familiarized themselves with a checklist of 42 adjectives (e.g., amazing, desirable, foolish, objectionable). Each adjective had previously been normatively rated on the extent to which it implied a positive vs. negative evaluation and an evaluation based on emotion (Rocklage & Fazio, 2015). For each of the 17 attitude objects (including the idiosyncratic ones), participants checked the two to five adjectives that best described their evaluation.

Self-defining self-report measure

Participants then saw the 17 attitude objects along with the two statements from Study 1 (e.g., “My evaluation of _____ reflects the kind of person I am or aspire to be”) and rated their agreement to each on the same 1 to 7 scale.

Results

Participants’ evaluations of the 17 attitude objects (15 provided, 2 idiosyncratic) varied widely (across all participants and objects: $M = .66$, $SD = 1.9$). We computed self-definition scores by averaging the two self-report items (across all participants and objects, $r = .71$, $p < .001$) to produce 17 self-definition scores for each participant (1 to 7; $M = 4.26$, $SD = 1.78$).

Object ranking

As in Study 1, to test our predictions about which attitudes would be most vs. least self-defining, we computed average self-defining ratings for each of the 17 attitude objects and ranked them from most to least self-defining. The results largely confirmed our expectations and

resembled the results from Study 1. For example, attitudes toward one's hometown and feminism were among the most self-defining, and attitudes toward Folgers coffee and an irrelevant sports team were among the least self-defining (see online supplement).

Extremity effect

To test whether more extreme attitudes were more self-defining, and to account for the non-independent data structure, we ran a random effects model predicting self-defining ratings from attitude ratings with both a linear and a quadratic term, and with intercepts randomly varying both between participants and objects. Unlike in Study 1, we collected attitude ratings for the idiosyncratic likes and dislikes and were able to include them in this analysis. As predicted, the quadratic term was significant, $b = .15$ (95% CI: .12, .17), $t(1094.24) = 12.01$, $p < .001$, $R^2_{\beta} = .12$, such that more extreme attitude ratings predicted higher self-defining ratings, in a pattern resembling the results from Study 1 (see online supplement for graph).

Valence effect

As predicted, the linear term in the above model was positive and significant, $b = .17$ (95% CI: .11, .22), $t(1071.15) = 6.09$, $p < .001$, $R^2_{\beta} = .03$, indicating that controlling for extremity, positive attitudes tended to be more self-defining than negative attitudes.

Idiosyncratic self-defining attitudes

As in Study 1, participants' self-provided idiosyncratic self-defining likes ($M = 5.97$, $SD = 1.42$) were more self-defining than the idiosyncratic self-defining dislikes ($M = 4.94$, $SD = 2.18$), mean difference = 1.03 (95% CI: .47, 1.59), $t(66) = 3.65$, $p = .001$. Since, unlike in Study 1, we also collected attitude ratings for these participant-provided objects, we recoded those ratings into extremity (distance from the scale midpoint) and found that the idiosyncratic likes

($M = 2.94$, $SD = .24$) were also slightly more extreme than the idiosyncratic dislikes ($M = 2.82$, $SD = .39$), mean difference = .12 (95% CI: .01, .23), $t(66) = 2.20$, $p = .03$.

Because valence and extremity could be confounded in these analyses, we ran a mixed-effects model predicting self-defining ratings from both valence (effect-coded) and extremity, with the intercept randomly varying between subjects. This model revealed both an extremity effect, $b = 1.28$ (95% CI: .34, 2.23), $t(125.87) = 2.69$, $p = .008$, $R^2_{\beta} = .05$, and a valence effect, $b = .44$ (95% CI: .16, .72), $t(67.76) = 3.16$, $p = .002$, $R^2_{\beta} = .13$. Critically, the idiosyncratic self-defining likes tended to be more self-defining than the idiosyncratic self-defining dislikes even controlling for extremity, providing strong additional evidence for the overall valence effect.

Evaluative lexicon analyses

For each participant, we averaged the implied evaluation and emotionality of the adjectives they chose for each of the 17 attitude objects (as in prior work, values were based on judges' normative ratings of the adjectives; see Rocklage & Fazio, 2015). We also coded each attitude as either univalent (only positive or only negative adjectives selected, 79% of responses) or ambivalent (both positive and negative adjectives, 21% of responses).

First, to replicate the extremity and valence effects found with the self-report attitudinal measure, we ran a similar mixed model predicting self-defining ratings from implied evaluation, with both a linear and quadratic term. As predicted, both the quadratic term, $b = .08$ (95% CI: .05, .10), $t(1098.88) = 6.70$, $p < .001$, $R^2_{\beta} = .04$, and the linear term, $b = .07$ (95% CI: .03, .12), $t(1079.08) = 3.48$, $p = .001$, $R^2_{\beta} = .01$, were significant, providing additional evidence for the extremity and positivity effects found in both studies using attitudinal ratings.

Next, we assessed whether implied emotionality predicted self-definition. Following prior work (Rocklage & Fazio, 2015), we controlled for implied extremity, running a mixed-

effects model similar to the previous one predicting self-defining ratings from implied emotionality and implied extremity. Although extremity was significant, $b = .66$ (95% CI: .38, .94), $t(1078.80) = 4.61$, $p < .001$, $R^2_{\beta} = .02$, emotionality was not, $b = .06$ (95% CI: -.08, .19), $t(1081.27) = .81$, $p = .42$, $R^2_{\beta} = .0006$, suggesting that attitudes characterized by greater emotionality are not necessarily more self-defining than less emotional attitudes.

Finally, we assessed whether ambivalent attitudes were less self-defining than univalent attitudes. In a mixed-effects model similar to the ones described above and controlling for extremity, $b = .70$ (95% CI: .50, .91), $t(1092.78) = 6.73$, $p < .001$, $R^2_{\beta} = .04$, ambivalence (dummy-coded) significantly predicted lower self-definition, $b = -.22$ (95% CI: -.43, -.01), $t(1090.10) = -2.02$, $p = .04$, $R^2_{\beta} = .004$.

Discussion

Study 2 replicated the findings of Study 1 using the same and different measures, and added several new findings. The two-item self-defining measure once again differentiated between attitude objects that we expected to be more and less self-defining. We again found an extremity effect, such that more extreme attitudes were more self-defining than less extreme attitudes, and a positivity effect, such that positive attitudes were more self-defining than negative attitudes. In addition, we confirmed both of these effects using a language-based attitudinal measure (Rocklage & Fazio, 2015). However, this measure showed no evidence of an emotionality effect, consistent with the idea that attitudes can be self-defining regardless of the extent to which they involve emotional reactions. For instance, attitudes toward participants' hometowns and feminism were similarly self-defining ($M_s = 5.25$ and 5.00 , respectively), but while attitudes toward hometowns were characterized by relatively high implied emotionality (.69 SDs above the mean), attitudes toward feminism tended to be less emotional (-.71 SDs

below the mean). On the other hand, as expected, ambivalent attitudes were less self-defining than univalent attitudes. In sum, Study 2 confirmed, using multiple methods, the findings from Study 1, while adding several new elements to the emerging picture of what self-defining attitudes look like.

Study 3

In Studies 1 and 2, we tested a new measure of the extent to which an attitude is self-defining and explored some of the characteristics of self-defining attitudes. In Study 3, we sought to explore the consequences of self-definition. Specifically, we reasoned that people should be more likely to advocate in favor of a self-defining attitude than an attitude less closely linked to one's self-concept. Actively trying to persuade others to one's point of view is an important outcome linked to various characteristics of attitudes (e.g., Cheatham & Tormala, 2015). In one particularly interesting set of studies, participants expressed greater intentions to advocate in favor of their attitudes following a self-threat, suggesting a link between the self, advocacy, and the attitudes used in those studies (and in fact, the researchers explicitly argued that their effects occurred because attitudes are related to the self-concept, although they did not attempt to measure that relationship; Gal & Rucker, 2010).

However, before including a self-threat in our experimental design (Study 4), we first wanted to test whether there was any simple relationship between self-defining attitudes and advocacy. Thus, in Study 3 we tested whether self-definition was associated with advocacy intentions, and more specifically, intentions to spontaneously advocate (rather than upon request) in favor of one's position. The distinction (Teeny & Petty, 2017) between spontaneous and requested advocacy (i.e., sharing your views unprompted vs. sharing your views when asked) has received surprisingly little research attention, despite its potential importance. Spontaneous

advocacy should be much more effective at spreading ideas and opinions than requested advocacy, but should also be more effortful and thus difficult to motivate –people may only spontaneously advocate regarding topics closely tied to their sense of self (Fazio, 1995).

We expected self-definition to correlate with a number of other attitude characteristics, such as certainty (Tormala & Rucker, 2007), so we included several such attitude-related measures. Besides further validating the self-definition construct by providing evidence of convergent validity, there was another reason to measure these variables: certainty predicts advocacy intentions (Akhtar, Paunesku, & Tormala, 2013; Cheatham & Tormala, 2015), and other attitudinal variables might do so as well. It was therefore important to determine whether any relationship between self-definition and advocacy intentions was simply due to shared variance between self-definition and another attitude-related variable.

Method

Participants

One hundred and twenty-one undergraduates (63 female, 56 male, 2 unspecified) participated online in exchange for partial fulfillment of an introductory psychology course requirement. Nine failed an attention check, leaving a final sample of 112 (61 female, 49 male, 2 unspecified).

Attitude measurement

Participants rated their attitude toward environmentalism on a scale from 1, “Strongly oppose environmentalism,” to 9, “Strongly support environmentalism.”

Spontaneous advocacy intentions

Participants then responded to five scenarios involving advocacy in favor of their stance toward environmentalism (e.g., providing arguments supporting their opinion to a fellow student,

speaking up in a group discussion). Participants indicated how likely they would be to spontaneously advocate in each situation (i.e., unprompted, of their own accord) on a scale from 1, “Not at all likely,” to 7, “Very likely.”

Attitudinal variables

Participants then responded to questions regarding their opinion toward environmentalism. Among these was the two-item self-definition measure (“My opinion of environmentalism reflects the kind of person I am or aspire to be,” and “My opinion of environmentalism says something, both to myself and others, about who I am as an individual”) with a 5-pt scale from 1, “Strongly disagree,” to 5, “Strongly Agree.” Other attitudinal variables included certainty (3 items, e.g., “How certain are you of your opinion on environmentalism?”), issue importance (“How important to you is your opinion on environmentalism?”), personal relevance (2 items, e.g., “How relevant is your opinion on environmentalism to your life?”), pride (“I am proud of what I consider to be my opinion on environmentalism”), morality (“I consider my stance on environmentalism to reflect a component of my morality”), perceived elaboration (“How deeply have you thought about environmentalism?”), perceived knowledge (“How knowledgeable do you feel about environmentalism in general?”), objective ambivalence (a score derived from separate ratings of positivity and negativity toward environmentalism), and perceived affective and cognitive bases of the attitude (e.g., “How much of your opinion on environmentalism do you believe is based upon your feelings and emotions?”).

Results

Participants, on average, had positive attitudes toward environmentalism ($M = 7.0$, $SD = 1.42$). We recoded these ratings into extremity (0 to 4). We then computed means for spontaneous advocacy intentions ($\alpha = .76$, $M = 3.44$, $SD = 1.32$) and self-definition ($r = .77$, $M =$

3.18, $SD = .87$). Self-definition correlated significantly with every other attitudinal variable (Table 2, second column).

The extent to which participants' attitudes toward environmentalism were self-defining significantly predicted spontaneous advocacy intentions, $r = .40, p < .001$. To assess whether this effect was due to a third attitudinal variable that was related to both self-definition and spontaneous advocacy intentions, we ran a series of partial correlations between self-definition and spontaneous advocacy intentions, controlling for each other attitudinal variable in turn (Table 2, fourth column). None of these variables fully explained the relationship between self-definition and spontaneous advocacy intentions.

Table 2

Partial correlations between self-definition and spontaneous advocacy intentions, controlling for other attitudinal variables

Control Variable	r with self-definition	r with spontaneous advocacy intentions	Partial r of self-definition with spontaneous advocacy intentions
Self-definition	N/A	.40***	N/A
Affective basis	.43***	.31**	.31**
Certainty	.24*	.20*	.37***
Cognitive basis	.24*	.22*	.37***
Elaboration	.47***	.31**	.31**
Extremity	.51***	.30**	.31**
Importance	.59***	.39***	.23*
Knowledge	.42***	.16	.37***
Morality	.67***	.28**	.31**
Objective ambivalence	-.22*	-.15	.38***
Pride	.64***	.35***	.25**
Relevance	.61***	.41***	.21*

Note. * $p < .05$, ** $p < .01$, *** $p < .001$

Discussion

We expected, and found, that the more self-defining an attitude, the more participants would intend to spontaneously advocate for that attitude, that is, to share their views without being asked. Partial correlations showed that this relationship was not due simply to the relationship of both variables to a third variable, such as attitude extremity, certainty, or personal relevance.

A secondary result of Study 3 was that self-definition was related, in expected and sensible ways, to other attitudinal variables. At the same time, the magnitude of these relationships, and the fact that self-definition uniquely predicted spontaneous advocacy intentions, suggests that self-definition is not redundant with these variables. In other words, the second column of Table 2 suggests both convergent and divergent validity of the self-definition scale.

Study 4

Having established in Study 3 that self-definition predicts intentions to spontaneously advocate in favor of an attitude, we next sought to explore advocacy under threat. We reasoned that threats to the self should make any link between an attitude and the self-concept especially impactful. Specifically, self-threats often produce compensatory attempts to re-assert the self or stances important to the self (e.g., Gao, Wheeler, & Shiv, 2009; Landau, Greenberg, Sullivan, Routledge, & Arndt, 2009; McGregor, Zanna, Holmes, & Spencer, 2001; Sherman & Gorkin, 1980), and specifically, to increased advocacy or intentions to advocate (Festinger, Riecken, & Schachter, 1956; Gal & Rucker, 2010).

In Study 4, we led pro-environment participants to believe they had not consistently acted in such a manner. Without considering the self at all, one might expect that this attitudinal threat

would lead participants to conclude they are not very pro-environment after all, and thus to advocate less (i.e., a self-perception effect, Bem, 1972). However, if and only if being pro-environment is self-defining, one would expect the opposite, compensatory advocacy to reaffirm the sense of self. Indeed, one might consider the scenario of having acted out of line with one's attitude and responding by reasserting one's commitment to that attitude in terms of cognitive dissonance, which may be most impactful when the attitude is self-defining (e.g., Sherman & Gorkin, 1980). In other words, self-definition should moderate the effect of attitudinal threat on advocacy.

We included measures of various other attitudinal variables, as in Study 3, to assess whether any self-definition effects might be explained by these variables. Also, in addition to measuring spontaneous advocacy intentions, we gave participants an opportunity to advocate their position behaviorally within the context of a survey response.

Method

Participants

One hundred and forty-eight users of Amazon's Mechanical Turk (64 female, 84 male) participated online in exchange for \$.75. Of these, two failed an attention check, one reported not taking the study seriously, and three expressed suspicion regarding the manipulation (six total: two from the bolster condition, four from the threat condition). We further narrowed the dataset to participants who were unambiguously positive toward helping the environment, removing 32 participants (16 from each condition).⁷ The final dataset included 110 participants (50 female, 60 male).

⁷ We deliberately chose an attitude object ("helping the environment") we expected would elicit positive attitudes so that we could design the manipulation to threaten positive attitudes. Indeed, 77.5% of participants who passed the initial inclusion criteria unambiguously favored helping the environment (i.e., averaged at least 6 out of 7 on the two attitudinal items; 6 was labeled "Positive" and "Like" and 5 "Somewhat positive" and "Somewhat like"). The

Attitude measurement

Participants rated their attitude toward helping the environment on two items, from 1 (“Very negative” and “Strongly dislike”) to 7 (“Very positive” and “Strongly like”).

Attitudinal variables

Participants then completed a 2-item self-definition scale similar to that used in Studies 1-3 but with regard to “helping the environment” and using a 5-pt response scale. They also completed measures of other attitudinal variables similar to those in Study 3 (see online supplement for details).

Attitudinal threat manipulation – biased questionnaire

Participants completed one of two versions of a questionnaire containing 24 true/false items regarding their pro-environment behaviors. The two versions paired the same 24 statements with different qualifiers that made them either easy (sometimes, occasionally) or difficult (frequently, almost always) to agree with (Salancik & Conway, 1975). In the threat condition, these pairings occurred such that in 19 of the 24 statements the pro-environment behaviors were difficult to endorse (e.g., “I frequently shop at a local farmer’s market”) and the anti-environment behaviors easy to endorse (e.g., “If it’s inconvenient, I sometimes won’t go out of my way to recycle”). The other 5 items were paired in the opposite way to avoid suspicion (e.g., “I sometimes vote for politicians who support the environment”, “I frequently litter”). In the other condition, rather than threaten participants’ pro-environment attitudes, we bolstered them: 19 of the 24 items were worded so that pro-environment behaviors were easy to endorse and anti-environment behaviors difficult, with 5 items showing the opposite pattern to avoid suspicion. The intention was to push participants in the threat condition toward reporting fewer

responses below 6 were very skewed, making the alternative approach, moderating by attitudes rather than excluding based on them, statistically untenable.

pro-environment and more anti-environment behaviors, and vice versa for participants in the bolster condition.

Advocacy behavior

Next, participants saw an optional response box and were told they were “free to provide us with anything else you would like us to know about your views on helping the environment,” but if they had nothing else to share, they could leave the box empty and continue with the survey.

Spontaneous advocacy intentions

Participants then saw the five advocacy scenarios from Study 3, but with regard to their opinion toward helping the environment, and for each scenario rated their likelihood of spontaneously advocating on a scale from 1, “Not at all likely,” to 7, “Very likely.”

Results

We averaged the two attitude items ($r = .82$) to create an attitude score and the two self-defining items ($r = .72$) to create a self-definition score. Unsurprisingly, given our exclusion criteria, attitudes were very positive on average ($M = 6.46$, $SD = .48$), and they were relatively self-defining as well ($M = 4.05$, $SD = .73$). We also calculated means for spontaneous advocacy intentions ($\alpha = .87$, $M = 4.73$, $SD = 1.37$).

To test our hypothesis that self-definition would moderate the effect of attitudinal threat on spontaneous advocacy intentions, we ran a regression predicting standardized spontaneous advocacy intentions from standardized self-definition, threat (effect-coded), and their interaction term (Figure 2). The interaction was significant, $b = .22$ (95% CI: .04, .40), $t(106) = 2.45$, $p = .02$, R^2 change = .05. Simple effects analyses revealed that self-definition significantly predicted

spontaneous advocacy intentions in the threat condition, $b = .50$ (95% CI: .27, .73), $t(106) = 4.29$, $p < .0001$, but not in the bolster condition, $b = .05$ (95% CI: -.23, .33), $t(106) = .38$, $p = .71$.

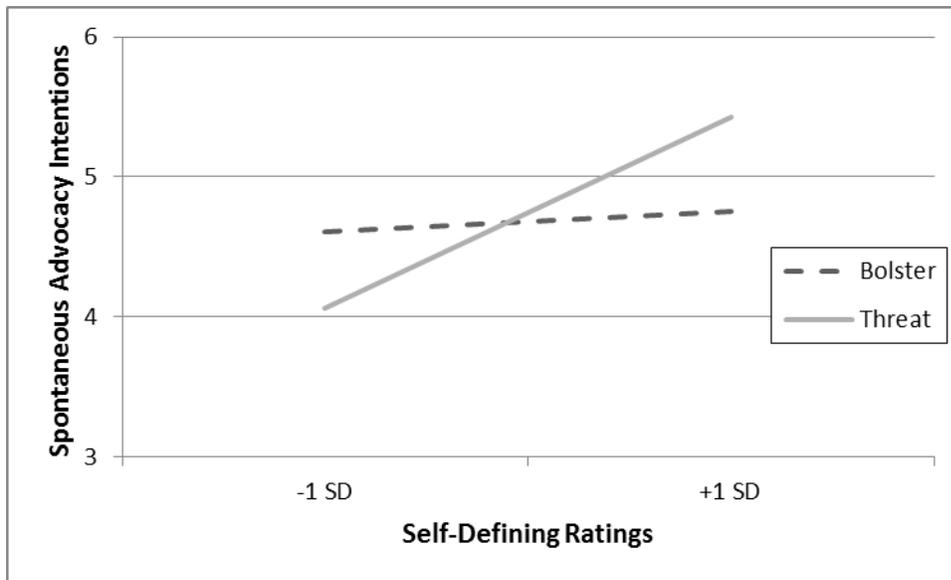


Figure 2. Spontaneous advocacy intentions predicted by condition, self-definition, and their interaction term in Study 4.

We adopted a similar strategy as in Study 3 to assess whether the condition by self-definition interaction effect predicting spontaneous advocacy intentions was due to a third attitudinal variable that was related to both self-definition and advocacy. Specifically, we ran a series of regressions similar to the one described above, but each also including an attitudinal variable and its interaction with condition. Overall, these covariates left the critical condition by self-definition interaction largely intact, and, in none of the models did the control variable by condition interaction explain more variance in spontaneous advocacy intentions than the condition by self-definition interaction (see online supplement for more details).

Behavioral advocacy – dichotomous outcome

Fifty-four participants wrote in the free response box (27 bolster, 27 threat). Of these, 8 were irrelevant (e.g., “no comments”) and 46 were relevant (21 bolster, 25 threat). We ran a logistic regression to test whether the condition by standardized self-definition interaction that predicted spontaneous advocacy intentions would also predict writing something relevant in the response box. This interaction was non-significant, $b = -.14$ (95% CI: $-.63, .28$), Wald $\chi^2 = .44$, $p = .51$.

However, in a separate logistic regression, standardized spontaneous advocacy intentions significantly predicted whether the participant offered a relevant response, $b = .65$ (95% CI: $.26, 1.20$), Wald $\chi^2 = 8.12$, $p = .004$, and since indirect effects can occur in the absence of a total effect (Hayes, 2009; Rucker, Preacher, Tormala, & Petty, 2011), we tested whether the condition by self-definition interaction affected relevant responses indirectly via spontaneous advocacy intentions. We ran a conditional process model (Hayes, 2013) in which effect-coded condition interacted with standardized self-definition to predict relevant responses (direct effect, c') and standardized spontaneous advocacy intentions (indirect effect, a), which in turn predicted relevant free responses controlling for the direct effect (indirect effect, b). The direct effect of the interaction was not significant, $b = -.31$ (95% CI: $-.76, .13$), Wald $\chi^2 = 1.90$, $p = .17$. However, consistent with the results reported thus far, both the a and b paths of the indirect effect were significant: the condition by self-definition interaction significantly predicted spontaneous advocacy intentions, $b = .22$ (95% CI: $.04, .40$), $t(106) = 2.45$, $p = .02$, and spontaneous advocacy intentions significantly predicted relevant responses controlling for the direct effect, $b = .73$, Wald $\chi^2 = 8.17$, $p = .004$. As a whole, the indirect effect (ab path) of the condition by self-definition interaction on relevant responses via spontaneous advocacy intentions was significant, $b = .32$ (95% CI: $.06, .76$). Conditional indirect effects analysis revealed that self-definition

increased the likelihood of a relevant response indirectly via spontaneous advocacy intentions in the threat condition, $b = .36$ (95% CI: .12, .72) but not in the bolster condition, $b = .04$ (95% CI: -.20, .39).

Behavioral advocacy – response ratings

Beyond the dichotomous outcome of whether or not participants wrote something relevant in the response box, there was considerable variation in the content of what participants wrote. Since our theoretical approach suggests that people might advocate for a self-defining attitude following threat in order to reassert their identity, we coded the responses for identity-related content. Two coders who were blind to condition read each response and answered the question, “To what extent does this person seem to be communicating that he or she is an environmentalist; that is, someone for whom helping the environment is a part of his or her identity?” on a 1 (“Not at all”) to 5 (“Very much”) scale. They then discussed any responses for which their ratings diverged by more than one scale point (4 cases) and could change one or both of their ratings (final $r = .79$, $p < .001$), which we averaged to produce a single mean score. The responses varied widely on this dimension (e.g., “I don’t drive, so none of the driving-related questions really applied. I have never driven and don’t know how to drive,” earning a score of 1, vs. “I’m so extreme I grow my own food in my petrol-free garden. No trucked-in starts or fertilizers. My own compost and seeds,” earning a score of 5).

We then standardized these identity ratings and predicted them from effect-coded condition, standardized self-definition, and their interaction term. The interaction was significant, $b = .40$ (95% CI: .12, .68), $t(42) = 2.84$, $p = .007$, R^2 change = .15 (Figure 3). Self-definition

significantly predicted identity ratings in the threat condition, $b = .48$ (95% CI: .10, .85), $t(42) = 2.58$, $p = .01$, but not in the bolster condition, $b = -.31$ (95% CI: -.73, .11), $t(42) = -1.51$, $p = .14$.⁸

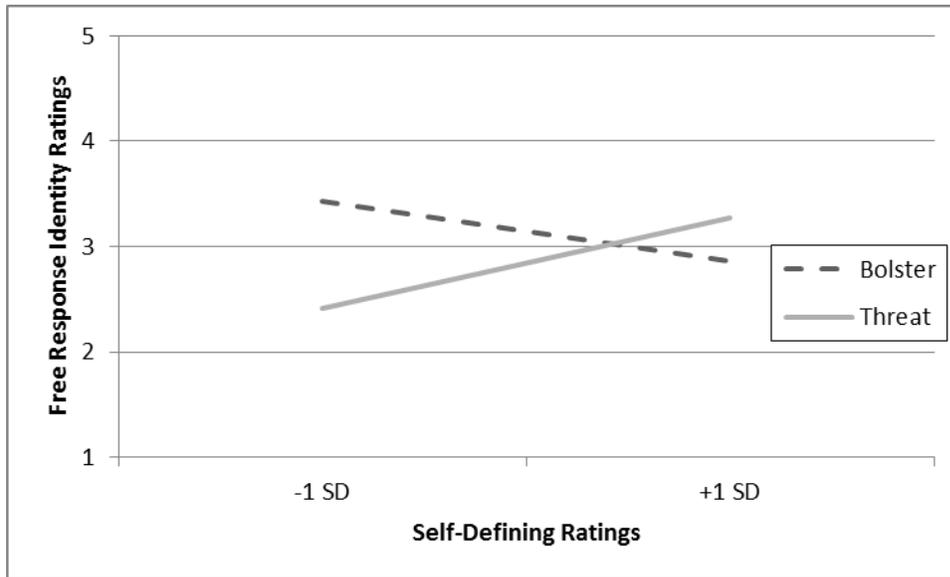


Figure 3. Identity ratings for free responses predicted by condition, self-definition, and their interaction term in Study 4.

Discussion

Study 4 showed that self-defining attitudes play a role in reactions to threat, at least with regard to advocacy. Specifically, the extent to which an attitude was self-defining predicted advocacy following attitudinal threat, but not following attitudinal bolstering. Study 4 also moved beyond spontaneous advocacy intentions to actual behavior. Although the threat by self-definition interaction did not predict the dichotomous outcome of whether or not participants wrote about their attitudes in an optional response box, spontaneous advocacy intentions did,

⁸ As with the other analyses, we were interested in whether these results were primarily driven by self-definition or a third variable related to both self-definition and the identity ratings, so we ran a series of additional models each controlling for another attitudinal variable and its interaction with condition. In no instance did the competitor interaction term explain more variance than the condition by self-definition interaction (see online supplement for more details).

resulting in a significant indirect effect of the interaction on the dichotomous response box outcome via intentions. Specifically, under threat, the more one's pro-environment attitude was self-defining, the greater one's intentions to spontaneously advocate for that attitude, which in turn led to a higher likelihood of writing something about one's attitude in the response box. This indirect path was not significant in the bolster condition. The fact that the indirect effect occurred in the absence of a total effect is not particularly troubling, since tests of indirect effects often have greater power than tests of total effects (Kenny & Judd, 2014; Rucker et al., 2011). The current results may also have been due to a suppression effect, in which the positive indirect effect we observed was canceled out by one or more unknown, unmeasured indirect effects in the opposite direction, resulting in no total effect (Rucker et al., 2011). This possibility is not surprising when one considers the myriad factors that might influence participants' choice to fill out an optional response box in an online survey.

Of course, treating the free responses as dichotomous ignores a great deal of information regarding what people actually wrote. By coding responses, we determined that the same condition by self-definition interaction that predicted spontaneous advocacy intentions (and, indirectly, whether or not participants wrote in the response box) also predicted the extent to which participants' responses asserted an environmentalist identity. Specifically, self-definition predicted identity assertions under threat, but not in the bolster condition.

Interestingly, self-definition did not predict advocacy intentions in the bolster condition, as it did in Study 3, which had no manipulation. This reinforces the fact that the bolster condition was not simply a control condition. Rather, it bolstered participants' perceptions of having acted in line with their attitude, which may have rendered further attempts to assert the self-concept via advocacy superfluous. Although our predictions in Study 4 hinged primarily on the relative

difference in slopes between the bolster and threat conditions, future work could include a true control condition to further clarify these results.

General discussion

Psychology has a long and rich tradition of discussing how attitudes help people understand who they are (e.g., Katz, 1960; Smith et al., 1956). Researchers have measured the extent to which an attitude fulfills functions related to the self-concept by creating new scales specific to an attitude object of interest (Herek, 1987), using individual difference scales (Petty & Wegener, 1998; Snyder & DeBono, 1985) or attitude extremity (Sherman & Gorkin, 1980) as proxies for the construct, measuring participants' values (Hullett & Boster, 2001; Maio & Olson, 2000), or identifying, *a priori*, objects thought to fulfill this function (Shavitt, 1989). However, none of these approaches has produced an easy-to-use self-report measure that can simultaneously assess both inter-person (same attitude object across people) and inter-object (different attitude objects within people) variability. In the current work, we have attempted to create and validate such a measure, and we have begun to use it to explore hypotheses concerning attitudes and the self.

We use the term “self-defining” to describe attitudes that help people define their self-concepts, both to themselves and others, and that are relevant to both who they are and who they aspire to be. In other words, self-defining attitudes help people answer the question, “Who am I?” Across four studies, we both validated this new measure of self-definition and used it to test hypotheses concerning attitudes and the self. In Studies 1 and 2, we confirmed several key expectations regarding particular attitudes we expected would be more vs. less self-defining on average. We also confirmed that self-defining attitudes tend to be more extreme than attitudes less closely related to the self. Interestingly, we also found a valence effect in both studies, such

that positive attitudes tended to be more self-defining than negative attitudes. This finding fits well with the literature on positivity biases regarding the self (e.g., Taylor & Brown, 1992). Moreover, positive attitudes lead to approach behavior (Fazio et al., 2004) and thus to frequent experiences with the object in question, experiences that involve attitude activation, expression, and relevant behavior, all of which may create or strengthen the association between the attitude and the self.

Adding to the characterization of self-defining attitudes as relatively extreme and more often positive than negative, in Studies 3 and 4 we measured a wide array of attitudinal variables we expected would relate to self-definition, including variables traditionally considered to represent attitude strength (Petty & Krosnick, 1995). In our research, self-defining attitudes were more likely to be important, relevant, certain, well thought-out (perceived elaboration and knowledge), un-ambivalent, and morally based than non-self-defining attitudes. Given these relationships to indicators of attitude strength, one might wonder whether self-definition is itself a determinant of attitude strength, in the sense that self-defining attitudes may be more stable, predictive of behavior, and resistant to persuasion (Petty & Krosnick, 1995). In the current work we were less concerned with characterizing self-definition as a determinant of attitude strength than with showing that self-defining attitudes have unique properties that cannot be fully explained by existing strength variables. However, future research should test whether self-definition similarly impacts stability, attitude-behavior consistency, and resistance to counterinfluence.

One question that may arise when considering self-defining attitudes is whether they are associated with more bipolar distributions of attitude scores, i.e., distributions centered on the scale's neutral point, or more unipolar distributions primarily involving a single valence (e.g.,

Pratkanis & Greenwald, 1989). Since controversial, value-charged issues (e.g., abortion) tend toward bipolar attitude distributions, one might similarly expect self-defining attitudes to have bipolar distributions. However, the current data suggest otherwise. For instance, in Study 1, although attitudes toward feminists were highly self-defining (on a 1 to 7 scale, $M = 5.09$) and bipolar (on a -3 to +3 attitude scale, $M = .58$, $SD = 1.49$), attitudes toward one's hometown were also highly self-defining (on a 1 to 7 scale, $M = 5.64$), yet largely unipolar ($M = 2.17$, $SD = 1.25$). Although future work might explore these questions in more detail, it does not appear that self-defining attitudes are solely associated with one type of distribution or the other.

Future research should examine the causal relationships between self-definition and other variables. For example, Studies 1 and 2 revealed a robust relationship between attitude extremity and self-definition but were necessarily mute on that relationship's causal direction. Does extremity make an attitude more self-defining, or does self-definition polarize attitudes? Or, as suggested earlier, might both causal paths operate in a recursive fashion? Future work should examine these questions using longitudinal designs (in which one variable predicts subsequent changes in another that are consistent with causality) or, ideally, experimental manipulations (e.g., of self-definition).

The main goal of Studies 3 and 4 was to test whether self-defining attitudes motivate attitudinal advocacy (Cheatham & Tormala, 2015; Gal & Rucker, 2010), and specifically, *spontaneous* advocacy; that is, the expression of one's views that occurs unprompted, without solicitation. This sort of advocacy should be more important (i.e., as an instrument of persuasion and social change) than requested advocacy as well as more difficult to motivate – most people will share an opinion if asked, but fewer will spontaneously express an opinion in an unprompted attempt to change a conversation partner's views. Study 3 demonstrated that self-definition is

related to intentions to spontaneously advocate. Study 4 showed that this effect of self-definition is strongest when the attitude in question has just been attacked, using a biased questionnaire to lead people to feel as if they had failed to act in line with their attitudes. Under this attitudinal threat, the extent to which participants were motivated to spontaneously advocate, thus reasserting the identity associated with the attitude, depended on whether the attitude was self-defining.

This “reaction to threat” aspect of Study 4 was especially interesting, as many have suggested that the extent to which the self is involved is important in determining reactions to threats of one form or another (e.g., Aronson, 1969; Gal & Rucker, 2010; McGregor et al., 2001; Sherman & Gorkin, 1980; Steele, 1988). Much of this work that has dealt with compensatory attitudinal processes has assumed, sometimes explicitly, that the attitudes in question are connected to participants’ self-concepts (Gal & Rucker, 2010; McGregor et al., 2001), and the work we are aware of that has measured this connection to the self did so in a way that is confounded with attitude extremity (Sherman & Gorkin, 1980). Future research should test these assumptions by measuring self-definition. For instance, the effects of making salient one’s past hypocritical behaviors on subsequent pro-attitudinal actions or judgments (e.g., Sherman & Gorkin, 1980; Stone, Aronson, Crain, Winslow, & Fried, 1995) should depend on the extent to which the attitude in question is self-defining.

In 1935, Gordon Allport famously lauded attitudes as “the most distinctive and indispensable concept in contemporary social psychology” (1935, p. 798). However, not long afterward, he less famously heralded a “recentering” of psychological theory towards the self (1943, p. 453) and predicted that the study of the self would “flourish increasingly” in the twentieth century (1943, p. 476). We believe both attitudes and the self are central concepts to

social psychology, and that our field cannot fully understand the one without taking the other into account. The role that attitudes play in defining the self and the consequences of any such linkages for information processing and behavior certainly merit theoretical and empirical attention.

References

- Akhtar, O., Paunesku, D., & Tormala, Z. L. (2013). Weak > strong: The ironic effect of argument strength on supportive advocacy. *Personality and Social Psychology Bulletin, 39*(9), 1214–1226.
- Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.), *A handbook of social psychology* (1st ed., pp. 798–844). Worcester, MA: Clark University Press.
- Allport, G. W. (1943). The ego in contemporary psychology. *Psychological Review, 50*(5), 451.
- Aronson, E. (1969). The theory of cognitive dissonance: A current perspective. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 4, pp. 1–34). New York: Academic Press.
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*(4), 323–370.
- Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 7, pp. 1–62). New York: Academic Press.
- Bizer, G. Y., Larsen, J. T., & Petty, R. E. (2011). Exploring the valence-framing effect: Negative framing enhances attitude strength. *Political Psychology, 32*(1), 59–80.
- Boninger, D. S., Krosnick, J. A., & Berent, M. K. (1995). Origins of attitude importance: Self-interest, social identification, and value relevance. *Journal of Personality and Social Psychology, 68*(1), 61–80.
- Campbell, J. D., Trapnell, P. D., Heine, S. J., Katz, I. M., Lavalley, L. F., & Lehman, D. R. (1996). Self-concept clarity: Measurement, personality correlates, and cultural boundaries. *Journal of Personality and Social Psychology, 70*(1), 141–156.

- Cheatham, L., & Tormala, Z. L. (2015). Attitude certainty and attitudinal advocacy: The unique roles of clarity and correctness. *Personality and Social Psychology Bulletin, 41*(11), 1537–1550.
- Edwards, L. J., Muller, K. E., Wolfinger, R. D., Qaqish, B. F., & Schabenberger, O. (2008). An R^2 statistic for fixed effects in the linear mixed model. *Statistics in Medicine, 27*(29), 6137–6157.
- Escalas, J. E., & Bettman, J. R. (2003). You are what they eat: The influence of reference groups on consumers' connections to brands. *Journal of Consumer Psychology, 13*(3), 339–348.
- Escalas, J. E., & Bettman, J. R. (2005). Self-construal, reference groups, and brand meaning. *Journal of Consumer Research, 32*(3), 378–389.
- Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247–282). Mahwah, NJ: Erlbaum Associates.
- Fazio, R. H., Eiser, J. R., & Shook, N. J. (2004). Attitude formation through exploration: Valence asymmetries. *Journal of Personality and Social Psychology, 87*(3), 293–311.
- Festinger, L., Riecken, H. W., & Schachter, S. (1956). *When prophecy fails*. Minneapolis, MN: University of Minnesota Press.
- Gal, D., & Rucker, D. D. (2010). When in doubt, shout! Paradoxical influences of doubt on proselytizing. *Psychological Science, 21*(11), 1701–1707.
- Gao, L., Wheeler, S. C., & Shiv, B. (2009). The “shaken self”: Product choices as a means of restoring self-view confidence. *Journal of Consumer Research, 36*(1), 29–38.

- Greenwald, A. G. (1982). Ego task analysis: A synthesis of research on ego-involvement and self-awareness. In A. H. Hastorf & A. M. Isen (Eds.), *Cognitive social psychology* (pp. 109–147). New York: Elsevier/North-Holland.
- Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, 76(4), 408–420.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford Press.
- Herek, G. M. (1987). Can functions be measured? A new perspective on the functional approach to attitudes. *Social Psychology Quarterly*, 50(4), 285–303.
- Katz, D. (1960). The functional approach to the study of attitudes. *Public Opinion Quarterly*, 24(2), 163–204.
- Hullett, C., & Boster, F. (2001). Matching messages to the values underlying value-expressive and social-adjustive attitudes: Reconciling an old theory with a contemporary measurement approach. *Communication Monographs*, 68(2), 133–153.
- Johnson, B. T., & Eagly, A. H. (1989). Effects of involvement on persuasion: A meta-analysis. *Psychological Bulletin*, 2, 290–314.
- Judd, C. M., & Krosnick, J. A. (1982). Attitude centrality, organization, and measurement. *Journal of Personality and Social Psychology*, 42, 436–447.
- Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, 103(1), 54–69.

- Landau, M. J., Greenberg, J., Sullivan, D., Routledge, C., & Arndt, J. (2009). The protective identity: Evidence that mortality salience heightens the clarity and coherence of the self-concept. *Journal of Experimental Social Psychology, 45*(4), 796–807.
- Maio, G. R., & Olson, J. M. (2000). What is a “value-expressive” attitude? In G. R. Maio & J. M. Olson (Eds.), *Why we evaluate: Functions of attitudes*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist, 41*(9), 954.
- McGregor, I., Zanna, M. P., Holmes, J. G., & Spencer, S. J. (2001). Compensatory conviction in the face of personal uncertainty: Going to extremes and being oneself. *Journal of Personality and Social Psychology, 80*(3), 472–488.
- Pelham, B. W., Mirenberg, M. C., & Jones, J. T. (2002). Why Susie sells seashells by the seashore: Implicit egotism and major life decisions. *Journal of Personality and Social Psychology, 82*(4), 469–487.
- Petty, R., & Cacioppo, J. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 123–205). New York: Academic Press.
- Petty, R. E., Cacioppo, J. T., & Haugtvedt, C. P. (1992). Ego-involvement and persuasion: An appreciative look at the Sherifs’ contribution to the study of self-relevance and attitude change. In *Social judgment and intergroup relations: Essays in honor of Muzafer Sherif*. New York: Springer/Verlag.
- Petty, R. E., & Krosnick, J. A. (Eds.). (1995). *Attitude strength: Antecedents and consequences*. Mahwah, NJ: Erlbaum Associates.

- Petty, R. E., & Wegener, D. T. (1998). Matching versus mismatching attitude functions: Implications for scrutiny of persuasive messages. *Personality and Social Psychology Bulletin*, 24(3), 227–240.
- Pratkanis, A. R., & Greenwald, A. G. (1989). A sociocognitive model of attitude structure and function. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 22, pp. 245–285).
- Rocklage, M. D., & Fazio, R. H. (2015). The Evaluative Lexicon: Adjective use as a means of assessing and distinguishing attitude valence, extremity, and emotionality. *Journal of Experimental Social Psychology*, 56, 214–227.
- Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass*, 5(6), 359–371.
- Salancik, G., & Conway, M. (1975). Attitude inferences from salient and relevant cognitive content about behavior. *Journal of Personality and Social Psychology*, 32(5), 829–840.
- Schwartz, S. (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). New York: Academic Press.
- Sedikides, C., & Strube, M. (1997). Self-evaluation: To thine own self be good, to thine own self be sure, to thine own self be true, and to thine own self be better. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 209–269). San Diego: Elsevier Academic Press Inc.

- Shavitt, S. (1989). Operationalizing functional theories of attitude. In A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *Attitude structure and function* (pp. 311–337). Mahwah, NJ: Lawrence Erlbaum Associates.
- Shavitt, S. (1990). The role of attitude objects in attitude functions. *Journal of Experimental Social Psychology, 26*(2), 124–148.
- Sherif, M., & Cantril, H. (1947). *The psychology of ego-involvements*. Hoboken, NJ: John Wiley & Sons.
- Sherman, S., & Gorkin, L. (1980). Attitude bolstering when behavior is inconsistent with central attitudes. *Journal of Experimental Social Psychology, 16*, 388–403.
- Shook, N. J., Fazio, R. H., & Richard Eiser, J. (2007). Attitude generalization: Similarity, valence, and extremity. *Journal of Experimental Social Psychology, 43*(4), 641–647.
- Smith, M., Bruner, J., & White, R. (1956). *Opinions and personality*. New York: John Wiley.
- Snyder, M., & DeBono, K. G. (1985). Appeals to image and claims about quality. *Journal of Personality and Social Psychology, 49*(3), 586–597.
- Steele, C. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261–302). New York: Academic Press.
- Stone, J., Aronson, E., Crain, A. L., Winslow, M. P., & Fried, C. B. (1994). Inducing hypocrisy as a means of encouraging young adults to use condoms. *Personality and Social Psychology Bulletin, 20*, 116–128.
- Taylor, S., & Brown, J. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*(2), 193–210.

Teeny, J., & Petty, R. E. (2017). *The role of perceived attitudinal bases on spontaneous and requested advocacy*. Manuscript in preparation.

Tormala, Z. L., & Rucker, D. D. (2007). Attitude certainty: A review of past findings and emerging perspectives. *Social and Personality Psychology Compass*, *1*(1), 469–492.