The Preference for Attitude Neutrality

* Thomas I. Vaughan-Johnston[1]
Devin I. Fowlie[2]
Laura E. Wallace[3]
Mark W. Susmann[4]
Leandre R. Fabrigar[2]

* Corresponding author
[1] Cardiff University, Department of Psychology
[2] Queen's University, Department of Psychology
[3] University of Chicago, Booth School of Business
[4] Vanderbilt University, Department of Psychology

Note. Open data and syntax for all studies are available at https://osf.io/d5n6v/?view_only=3495c22f84ce43c899132cdbe035a53b. Materials are open in the online supplement (SOM).

Conceptualization: All authors; **Data Curation:** TIV; **Formal Analysis:** TIV, DIF; **Investigation:** All authors; **Methodology:** All authors; **Project Administration:** TIV; **Supervision:** LRF; **Validation:** DIF; **Visualization:** TIV, DIF; **Writing-Original:** TIV, DIF; **Writing-Review/editing:** All authors.

Abstract

Much research has noted people's tendency toward extremity. This work has made it clear that some people prefer to hold extreme views and might leave the impression that when biases and preferences occur, they primarily favor extremity. In contract, in the current work, we examine the possibility that some people prefer attitudinal neutrality across two pretesting samples, three main studies, and two supplementary studies (N_{total} = 1,873). The preference for neutrality is distinguished from low preference for extremity, as well as from an interest in collecting balanced information. We also show that the preference for neutrality is related to a sometimes uncritical and biased pursuit of attitudinal neutrality, paralleling effects found in the attitude extremity literature. The preference for neutrality is related to dispositional attitudinal neutrality and ambivalence, political centrism, a preference for other people with neutral versus extreme views, and biased responding to messages arbitrarily framed as "moderate" versus extreme. Implications for politically polarized attitudes, persuasion, and intellectual humility are discussed. The preference for neutrality may pose a substantial challenge for creating a shared understanding of the world and addressing pressing social issues.

Keywords: attitudes; extremity; individual differences; neutrality; politics

Public significance statement: Psychologists often caution the public that extreme attitudes have negative consequences, and the media often promotes a "balanced" presentation of information when presenting news stories. The public might reasonably interpret these phenomena as suggesting that extreme attitudes are "bad" and moderate attitudes (neutrality) are "good." In the present work we identify an individual difference, the *preference for neutrality*, that reveals how one can be "excessively neutral" in a fashion somewhat parallel to excessive

53 extremity, with important consequences for both political and non-political thoughts and
54 opinions.

55 **Manuscript Length:** 12,950

The Preference for Attitude Neutrality

"Everything in moderation, including moderation" – Oscar Wilde

Many social problems of our age—such as climate change, cancel culture and its opponents, resistance to COVID-19 vaccination—are often understood to be issues of attitudinal extremity. Much social science work has identified how people's psychological processes can lead them to problematic positions of extremity, and how extreme attitudes bias information processing (Fiske & Taylor, 1991; Lord et al., 1979; Ross & Nisbett, 1991; Tannenbaum, 1956), enhance attitude-behavior consistency (Wilson et al., 1989), create resistance to persuasion (Bassili, 1996), and drive polarized attitudes (Abelson, 1995; van Boven et al., 2012; Westfall et al., 2015). In short, extreme attitudes have been widely studied (also see Abelson, 1995; Brauer et al., 1995; Buder et al., 2021; Judd & Johnson, 1981; Tannenbaum, 1956). Relatedly, individual differences in the propensity to hold extreme attitudes has often been studied as dogmatism (Hanson, 1973; Leone, 1989), social vigilantism (O'Dea et al., 2018), the need to evaluate (Jarvis & Petty, 1996), low objectivism (Leone, 1996), and dispositional attitude extremity (DeMarree et al., 2020, supplemental materials). This prior research is incredibly valuable, and we agree that extremization and political polarization pose major threats to society (e.g., Abelson, 1995; Pretus et al., 2023; Toner et al., 2013; van Boven et al., 2012; Zmigrod et al., 2020). However, as noted in the epigraph above, paradoxically, excessive moderation might be possible. For example, it is not clear that the optimal position for an individual or society to take towards social problems like climate change, the civil rights of minoritized identities, or taking COVID-19 vaccinations (or historically, towards slavery in the leadup to the American Civil War), is a neutral attitude. Nonetheless, people may sometimes quite strongly prefer to maintain neutrality across a diverse array of political and other attitude objects. We raise some concerning findings about the likely problematic aspects of such an individual difference.

The preference for neutrality[1] is an epistemic position that commits some individuals to holding neutral positions due to their perceived greater "truthfulness." Seeing neutral attitudes as truer may also result in appraising neutrality as more functional, virtuous, and socially acceptable. At high levels, the preference for neutrality may become an intellectually arrogant and biased disposition towards attitudinally relevant information. Therefore, preference for neutrality does not merely reflect low levels of the various extremity-motivation variables we listed above because preference for neutrality is about the appeal of neutrality rather than the lack of appeal of extremity. For example, people could be low on both motivations, and therefore not think that holding neutral or holding extreme attitudes has special advantages. Alternatively, people could theoretically be high on both constructs, if they recognize the value of holding both extreme and neutral attitudes, holding a complex view of the value of different attitude positions. Preference for neutrality also differs from an open-minded interest in collecting information from "all sides" when forming attitudes, a hypothesis we test in the present work. Only preference for neutrality entails seeing neutral positions as *superior* to non-neutral positions, whereas open-mindedness entails learning all possible information regardless of valence to arrive at the best justified attitudinal position. Preference for neutrality also differs from low levels of extremity bias in the response set literature (also see Simonson et al.'s, 2017, characterization of extremity aversion as a "habit"), in that preference for neutrality involves a unified core of psychological beliefs rather than a habit or avoidance of effortful responding. In short, we test whether some

[1] We use the terms *preference for neutrality* and *preference for extremity* because we think "preference" captures the motivated nature of our constructs, and "neutrality" best captures the attitude goal that this disposition promotes. "Neutrality" is the antonym of "extremity" in the traditional attitudes literature (Abelson, 1995; Decker & Lord, 2022; deVries & Walker, 1987; Downing et al., 1992; Edwards & Ostrom, 1971; Thurstone, 1928), indicating an attitude position that avoids predominance of positive or negative reactions over the other. It does not presuppose controversy or the existence of defined sides, but simply characterizes a relative balance of positive and negative reactions, whether these reactions are few (indifference) or many (ambivalence).

people have sometimes powerful motivations to form and maintain neutral attitudes because they judge neutrality to be epistemically and functionally better than other attitude positions.

1.1. Attitude Neutrality

Attitudes are evaluations of objects (i.e., people, places, ideas) as good and/or bad (Eagly & Chaiken, 1993; Fabrigar et al., 2010). Although positive and negative evaluations can be conceptualized and measured independently (Cacioppo & Berntson, 1994; Pittinsky et al., 1994), attitudes are often conceptualized as falling on an evaluative dimension ranging from negativity to positivity, and usually are represented as a single numerical value which is presumed to reflect the valence and extremity of the evaluation. For example, an extremely positive attitude may be represented with a positive score, an extremely negative attitude with a negative score, and a neutral attitude with a moderate score near the conceptual midpoint (Abelson, 1995; Brauer et al., 1995; Krosnick & Petty, 1995). Other conceptual definitions of neutral attitudes have been offered. For example, neutral attitudes have sometimes been more narrowly defined as the absence of positive/negative reactions (in contrast to ambivalence, the presence of both positive/negative reactions; Cacioppo & Berntson, 1994; Cacioppo et al., 1997). Neutrality has also been proposed to be a property independent of attitudinal positivity and negativity, with neutrality therefore characterized as the degree to which one feels indifferent about an object (Hu & Gasper, 2022). Because in the present work we focus on people's beliefs and preferences about neutrality, we favor the "common sense" notion of neutrality as non-extreme (near the midpoint) evaluations (see Abelson, 1995), consistent with a lay understanding of the meaning of neutrality, and the majority of the attitudes literature.[2]

[2] "Neutrality" is sometimes used to characterize related but distinct constructs, such as a refusal to take sides, or a middle-ground position between two defined sides. Refusing to take a side is a behavior rather than an attitudinal position, and social perceivers distinguish between sides-refusing people who hold genuinely middle-ground opinions versus are being strategic in their refusal (Silver & Shaw, 2022). In

In contrast to the conceptual simplicity of extreme attitudes (i.e., either positive or negative reactions predominate over the other), neutral attitudes can entail several distinctive attitudinal structures. A neutral attitude may arise: (a) from having a balanced large number of positive and negative reactions, (b) from having a balanced small number of positive and negative reactions, or (c) from a set of only neutral reactions towards an object (Edwards & Ostrom, 1971). Attitudes scholars have focused on the first of these possibilities, an *ambivalent* attitude, defined as the holding of both strong positive and negative attitudes towards the same object (Kaplan, 1972; Scott, 1968; Thompson et al., 1995). Thus, psychologists have focused on extremity even here, in that the well-studied concept of ambivalence is the form of neutrality characterized by simultaneously holding extreme positive and negative reactions—whereas other structures supporting neutrality may include a balanced absence of reactions. Ambivalent attitudes are associated with "weakness": susceptibility to persuasion (Hodson et al., 2001; Maio et al., 1996), and attenuated attitude-behavior consistency (Conner et al., 2002; Costarelli & Colloca, 2004). People are usually motivated to avoid or reduce ambivalence (van Harreveld et al., 2009; Nordgren et al., 2006). However, because ambivalence is an attitudinal structure that produces neutrality (defined as near-midpoint attitudes), people high in preference for neutrality may more often cultivate attitudinal ambivalence (thus seeking and becoming more neutral about

turn, "middle ground" opinions are strictly distinct from neutrality in that they refer to positions in a social distribution rather than necessarily to an equality of positive and negative reactions. However, the distinction often dissolves empirically because a middle-ground (socially average) opinion will generally be closer to evaluative neutrality than a non middle-ground (socially extreme) opinion. This is because the social "sides" in social conflicts like "abortion rights," "COVID-19 vaccine laws," or "Black Lives Matters" are defined by a preponderance of positive over negative versus negative over positive reactions to these topics, and thus non-neutral attitudes. The "middle-ground" position for each of these topics would imply a more equal ratio of positive and negative reactions, whether few or many, thus entailing relative evaluative neutrality.

attitude topics, political positions, etc.). Despite the attention paid to ambivalence as a type of neutral attitude, much less attention has been paid to neutrality per se (Hu & Gasper, 2022).

Although some research has examined extremity avoidance as a response style (e.g., Simonson et al., 2017), this work usually concerns expressing neutrality as a satisficing habit. In contrast, we view preference for neutrality as a genuine veneration of neutrality as a superior position; these individuals do not merely express neutrality to avoid giving a thoughtful response, but believe that neutrality is generally more valid. The present work probes this research gap by testing people's preferences for attitude neutrality.

1.2. The Preference for Attitudinal Neutrality

Several motivational factors may lead people to pursue neutrality as a dispositional tendency.[3] First, people may desire to hold neutral attitudes insofar as neutrality may feel like a more epistemically defensible position. Often, a goal of efforts to address partisan animosity is not just to get people to understand other perspectives, but also to moderate their attitudes (Balietti et al., 2021; Tuller et al., 2015). Furthermore, mass media often presents social issues as comprising two relatively balanced sides (even in cases where scientific consensus strongly supports only one position; Brüggemann & Engesser, 2017; Dixon & Clarke, 2013; Koehler, 2016) which may train people to view neutral positions as more legitimate or scientifically justified. These findings may suggest that people feel that neutral attitudes are easier to defend than are extreme attitudes. This belief may lead neutrality preferring individuals to even like and use information of a fixed valence more when that information is simply labeled as neutral, much

[3] In this section we draw both from research in which the potential desirability of attitudinal ambivalence *or* the desirability of neutrality has been studied. This is because research seldom disentangles whether people are actually pursuing ambivalence per se, or are pursuing neutrality with ambivalence merely people's means of obtaining a desired neutral position.

like liberals/conservatives prefer political policies arbitrarily labeled as liberal/conservative (Cohen, 2003).

Second, neutral attitudes may be beneficial for appearing socially desirable to others. Anticipatory attitude change towards neutrality tends to occur when people anticipate social interactions about the attitude topic (Hass & Mann, 1976; Tetlock et al., 1986), an effect at least partially driven by self-presentational concerns (McFarland et al., 1984). For example, it can be socially or epistemologically beneficial to express ambivalent feelings about controversial topics (Pillaud et al., 2013; 2018; Reich & Wheeler, 2016). Indeed, people view others who express ambivalence (vs unambivalence) about controversial topics as more competent (Pillaud et al., 2018) and can appreciate their willingness to acknowledge a different position than their dominant one (Xu & Petty, 2022). Political parties that espouse moderate views are seen as more willing to compromise and in turn are more effective, an important tactic when dealing with divisive issues (Johns & Kölln, 2020). Importantly, such perceived competence is associated with a perception that the attitude holder was likely to have weighed both sides of the issue in coming to their neutral attitude (i.e., rather than simply assuming that a neutral attitude is appropriate). High preference for neutrality may increase the tendency to evaluate ambivalent or neutral others favorably (i.e., Pillaud et al., 2013, 2018) because seeing neutrality as epistemically and morally superior should lead to a veneration of those who actually hold such attitudes.

Believing that neutrality is epistemically and socially superior may lead high preference for neutrality individuals to believe that attitudinal neutrality is unquestionably right. This differs from a search for evaluatively balanced, issue relevant information. For instance, open-minded people may seek information on both sides of an attitudinal issue (Stanovich & West, 1997).

This should help people form well considered, knowledge-based attitudes, be intellectually humble, and understand other people from a variety of perspectives. However, such open-minded thinking is quite distinct from striving to reach a preordained neutral position as with the preference for neutrality in which a person has the goal to form and maintain a neutral position. Indeed, an interest in both sides should not necessarily produce attitudinal neutrality or ambivalence because it involves only *learning about* all sides of attitudinal conflicts, rather than internalizing each side.

1.3. The Present Research

As the previous section suggested, various strands of psychological literature indicate reasons people may want to hold neutral opinions. These possible benefits suggest that preference for neutrality is conceptually distinguishable from simply lacking motivation to be extreme and from a balanced search for all information on a topic across positive/negative valence lines. Measurement and theory are intertwined in attitudes research (Ostrom, 1989). Indeed, our interest in developing a self-report questionnaire concerning preferences for neutrality and extremity also is a theoretical claim. Specifically, we are arguing that people not only form attitudes about particular objects (e.g., “I really like social justice”; “I really dislike nuclear power”) but also evaluate those attitudes in terms of distinct metacognitive evaluative standards (e.g., “I dislike having strong opinions [about social justice, nuclear power, and anything else]”). Furthermore, we are arguing that people are sufficiently aware of those metacognitive standards that they can self-report at least what attitudes they hold about certain attitude positions. Indeed, although we were interested in developing a reliable, valid scale of the attitude preferences, our theoretical interest in exploring how attitude preferences might guide relevant person perception (Study 1), attitude structures (Study 2), and responses to persuasion (Study 3) were paramount.

In the present work, we assessed the viability of separate individual difference constructs for the goals of cultivating neutral opinions (*preference for neutrality*), knowing both sides of attitude topics (*interest in both sides*), and the goal of holding extreme opinions (*preference for extremity*). We then tested how these attitude preferences showed unique associations with other individual difference variables, attitude structures, person perception effects, and responses to persuasive messages. Our hypotheses are summarized in Table 1.

1.3.1. Hypotheses

Hypothesis 1 concerns the individual differences' links with a close-minded, intellectually arrogant thinking style. Interest in both sides should be negatively associated with this thinking style because these individuals are open-minded in their pursuit of information. Preference for extremity and preference for neutrality, however, each should be positively related to this thinking style because both posit a particular "state of reality" (i.e., extreme / neutral opinions are more correct) rooted in one's core values and epistemic views. A novelty of the present work is that just as holding extreme positions may entail an arrogant mindset (Lammers et al., 2017; Toner et al., 2013), so too might a strong preference for neutral attitudes.

Hypothesis 2 concerns the preference for moderate action, moderate consumption, and middle-of-the-road judgments (Drolet et al., 2021). Although preference for neutrality is specific to attitudinal preferences, because it involves recognition of the value of non-extremity it should also be related to a general preference for moderation. Preference for extremity would likely be negatively related to moderation in that extreme attitudes may be seen as a lack of moderation.

Hypotheses 3-5 are connected, each concerning people's anticipated degree of intensely evaluative thinking and attitude extremity. Hypothesis 3 focuses on the need to form strong, extreme opinions as a dispositional personality style (i.e., the need to evaluate); Hypothesis 4

targets political extremity (i.e., having very left-wing or very right-wing attitudes about political topics, landing one far from the political center); Hypothesis 5 targets dispositional attitude extremity (i.e., tending to have near-the-endpoint attitudes about most objects). Whereas preference for neutrality should be negatively related to all three of these phenomena because these individuals should venerate and pursue attitudinal neutrality, preference for extremity should be positively related to all three. Merely wanting information “on both sides” does not have a clear connection to attitude extremity.

Table 1

Hypothesized Relations Between Interest in Both Sides, Preference for Neutrality, Preference for Extremity, and Other Variables.

Hypothesis number	Other variable	Interest in both sides	Preference for neutrality	Preference for extremity
1	Close-minded, intellectually arrogant personality style	Negative	Positive	Positive
2	Moderation	No prediction	Positive	Negative
3	Need to evaluate	No prediction	Negative	Positive
4	Political extremity	No prediction	Negative	Positive
5	Dispositional attitude extremity	No prediction	Negative	Positive
6	Preference for attitude-neutral versus attitude-extreme others	No bias	Bias for neutral > extreme others	Bias for extreme > neutral others
7	Attitude ambivalence (structural and subjective).	No prediction	Positive	Negative
8	Persuasion from messages labeled as neutral over extreme.	No bias	Bias for neutral > extreme frame	Bias for extreme > neutral frame

We further anticipate that our individual differences will influence how people view other people based on those targets' tendencies towards neutrality or extremity (Hypothesis 6). People with an interest in both sides should like both neutral and extreme attitude holders alike, because both types of people can offer positive and negative information about topics (i.e., presumably "neutrals" will offer both positive and negative, whereas "extremes" will at least offer one side). However, for our preference variables, we anticipated bias effects: preference for neutrality (preference for extremity) should relate to seeing other people as more moral, competent, and likeable when those individuals have generally neutral (extreme) opinions.

Hypothesis 7 concerns attitudinal ambivalence. People high in preference for neutrality should be more willing to be ambivalent because ambivalence is one attitude structure that facilitates neutrality; thus, these individuals may be more structurally and subjectively ambivalent. Preference for extremity should be negatively related to ambivalence for the converse reason. Finally, people who eagerly seek information on both sides need not necessarily internalize either type of information, and so may not be more or less ambivalent.

Finally, Hypothesis 8 concerns how people may react to persuasive messages that are framed as neutral versus extreme. At least during attitude formation, people often may lack a clear objective frame of reference for what attitude positions might be considered "extreme" versus "neutral." Although people who simply want information on both sides should not be lured by such framings because they value all information in an open-minded way, both of our preference variables should react to such framings by showing preference effects (e.g., trusting or internalizing such information more) insofar as it matches their attitudinal preferences. Ironically, then, the exact same messages might be accepted versus rejected by people high in

preference for neutrality (preference for extremity) just because it is labeled as a neutral (extreme) viewpoint.

1.3.2. Transparency and Openness

For all studies, we made data and syntax openly available at https://osf.io/d5n6v/?view_only=3495c22f84ce43c899132cdbe035a53b. Materials are open in the online supplement (SOM-2). We report all manipulations, and all measures throughout. We explain how we determined our sample size for each study. The only data exclusions were in Study 1, as reported. Studies were not preregistered. We obtained ethics approval for all studies from the relevant institutions. For all studies we did not have specific expectations about effect size, so we used time-based stopping rules, checking sample size after each academic semester until a minimum sample was reached. However, we also provide sensitivity analyses to probe what effect size range we were powered to detect (details in SOM-3). See SOM-5 for additional JARS details.

2. Study 1

In Study 1 we developed items to constitute three main attitude preference factors (i.e., interest in both sides, preference for neutrality, and preference for extremity). Our first goal was therefore to determine whether three such factors emerged. Our second goal was to probe the unique nomological networks of our constructs (Hypotheses 1–4) and general patterns of attitude extremity versus neutrality (Hypothesis 5). Third, we examined person perception effects. That is, given profiles of people characterized as having a tendency towards extreme or neutral opinions, we tested whether attitude preferences predicted evaluations of these targets as moral, competent, and/or overall good (versus immoral, incompetent, and bad; Hypothesis 6).

2.1. Method

2.1.1. Participants

For Study 1, we collected data for two semesters towards a minimum sample of 400, selected because to reduce fatigue, some measures we only had about half of participants complete (see below). We recruited $N = 643$ Canadian undergraduate students to complete this study online for partial course credit (gender: 80.7% women, 16.0% men, 1.6% non-binary, 1.7% missing; $M_{age} = 19.2$, $SD_{age} = 4.2$; ethnicity: 77.0% White, 18.7% Asian, 3.3% Indigenous, 2.6% Black, 2.5% Latinx, 3.7% other; 59.7% liberal, 23.0% moderate, 17.3% conservative), after removing 51 participants who complete none of our attitude preference scale items and thus could not affect our results. Otherwise, participants were not removed in any studies. Study 1 has 80% power to find effects of $r > |.07|$ to $|.11|$ depending on the analysis, comparing favorably to average effect sizes in social psychology ($r_{mean} = .21$, Richard et al., 2003).

In Study 1, we refer briefly to results from two Supplementary Samples that are described more comprehensively in the Supplement (SOM4–5). We included Supplementary Sample 1 ($N = 350$ American Prolific participants) to obtain a more politically balanced (less left-leaning) sample for our analyses concerning political orientation in section 2.2.3, and Supplementary Sample 2 ($N = 185$ UK University students) to compare attitudes preferences against trait ambivalence (Schneider et al., 2021) in section 2.2.2.

2.1.2. Procedure

Participants first completed a series of six profile evaluations. For each participant, three profiles were presented of targets with generally neutral attitudes (e.g., “Charlie has moderate opinions on virtually everything,” “Charlie rarely [has] a strong opinion”), and another three profiles with generally extreme attitudes (e.g., “Charlie sees everything in the extreme,” “Charlie is someone that is pretty extreme”). We took steps to avoid confounding extremity of opinions

with being outspoken; for instance, one neutral profile states that one character "always speaks his mind about his centrism." Additionally, all profiles were supplemented with generically positive information (e.g., "engages energetically with his school work") to mask our hypothesis. We also avoided implying that the opinions in question were exclusively political or exclusively about 'high stakes' topics by either talking about several attitude objects that the target evaluated ("From his opinions that pepperoni is the ultimate pizza topping, to his…views on fossil fuels and climate change"), and/or by stating that the character has this disposition in a global way (e.g., "he sees nothing in the extreme"). To avoid confounding names, gender, or profile information with neutrality versus extremity of opinions, we created two between-participant set conditions, randomly assigning participants to one of these. For instance, half of participants read about a highly neutral-opinionated Robyn, and half read about a highly extreme-opinionated Robyn.

Next, participants completed the attitude preference questionnaire, described below, with items presented in randomized order. Finally, participants completed a subset of six measures from a large pool of possible questionnaires. We used this questionnaire subsetting method to protect against satisficing or fatigue. The questionnaire pool consisted of the measures listed in Table 2, as well as the dispositional attitudes measure (DAM; Hepler & Albarracin, 2013), and a modified DAM consisting of 16 politically important, controversial, and left/right polarized attitude objects ("abortion," "Black Lives Matter," "Cancel culture," "Laws mandating COVID vaccines," "Quotas for hiring women in STEM fields," etc.). The order of these measures was randomized, and item orders were randomized within each scale.

2.1.3. Materials

For each profile, participants rated their overarching opinion of the target person with eight unipolar items, each rated from 1 ("not at all") to 7 ("definitely"; Crites et al., 1994). These items consisted of both positive (e.g., "good," "like") and negative adjectives (e.g., "bad," "dislike"). Participant responses to each target had *M*s = 4.34–4.81 and *SD*s = .97–1.13 and demonstrated good reliability (α's = .91–.93). Additionally, participants rated a series of items (Wojciszke et al., 1998) assessing how *morally* good the target seemed (e.g., "fair," "truthful," "righteous") and how *competent* (e.g., "clever," "efficient," "gifted"), with labeled endpoints 0 (*not at all*) to 6 (*definitely*) but rescored 1–7 for interpretative ease. Participant responses to the moral items had *M*s = 4.36–4.93 and *SD*s = .97–1.10, and responses to the competency items with *M*s = 4.42–5.19, *SD*s = .93–1.05.

Our attitude preferences scale consisted of 52 items that were developed across several rounds of piloting, and intended to capture preferences for particular attitude positions (i.e., extreme, neutral) versus seeking information on both sides.[4] Items covered perceived advantages of holding a particular position, reactions to others holding a particular position, preferences for having a particular type of position, etc. (see SOM-2). Each item was rated 1 (*Strongly disagree*), 2 (*Somewhat disagree*), 3 (*Neither agree nor disagree*), 4 (*Somewhat agree*), or 5 (*Strongly agree*). We randomized item order.

For the questionnaire block, see Table 1 for the list of measures. These questionnaires were designed to map onto hypotheses 1–3. We also included several scales about attitude-relevant social interactions for exploratory reasons; for example, probing whether people higher in preference for neutrality were more likely to self-censor in avoidance of disputes, or self-

[4] We ran two pilot studies (*N*s = 131, 118) to begin the process of refining items before conducting the primary studies. Details are in SOM-1.

monitor (shift their attitudes to match people around them). Additional scale descriptives appear in SOM-3 for all studies (Table S3-6a, S3-6b, and S3-6c).

2.2. Results

2.2.1. Factor Analyses

For all three studies, we ran exploratory factor analyses using an identical set of statistical/interpretative procedures. We followed exploratory factor analysis procedures as outlined by Fabrigar and Wegener (2011), and considered multiple desiderata when evaluating possible factor solutions, including a scree plot of the reduced eigenvalue matrix, fit (RMSEA; Browne & Cudeck, 1992), parallel analysis performance, interpretability of resulting factors, and the number of items loading substantially and uniquely on common factors. Extractions were performed using maximum likelihood, and solutions with 2+ factors were rotated using direct oblimin, meaning that factors were permitted but not forced to be correlated with one another (Fabrigar & Wegener, 2011). For brevity's sake, the full statistics are available in SOM-3. All samples found three-factor solutions the most viable, additionally finding good absolute fit for a three-factor solution.

In Sample 1, three common factors emerged which were labeled interest in both sides (14 items, $M = 4.14$, $SD = .56$), preference for neutrality (21 items, $M = 2.92$, $SD = .70$), and preference for extremity (four items, $M = 2.70$, $SD = .93$; see Table S3-2 in the online supplement). This shows some initial evidence that our proposed psychological distinctions might represent distinct psychological constructs. Sample 2 added new items and discarded poorly-loading items as explained later; Sample 3 retained Sample 2's items. Given that all three studies support a three-factor structure, these analyses support the intended psychometric performance of our attitude preference scale.

In Study 1, three common factors emerged which were labeled interest in both sides (14 items, $M = 4.14$, $SD = .56$), preference for neutrality (21 items, $M = 2.92$, $SD = .70$), and preference for extremity (4 items, $M = 2.70$, $SD = .93$; see SOM-2 in the online supplement). These items were used for all Study 1 analyses concerning attitude preferences. This shows some initial evidence that our proposed psychological distinctions might represent distinct psychological constructs. Study 2 added new items and discarded poorly-loading items as explained later; Study 3 then retained Study 2's items. Given that all three studies (and Pretesting 2) support a three-factor structure, these analyses support the intended psychometric performance of our attitude preference scale.

Because Study 2 introduced changed items compared to Study 3, we only ran a confirmatory factor analysis for Study 3; details are in SOM-3, but the three-factor model showed good fit on most indicators (RMSR = .07, RMSEA = .05 [.05, .06], but TLI = .81), with consistently strong item-factor loadings.

2.2.2. Associations with Other Constructs

Table 2 presents associations between the attitude preferences and other individual differences. See supplementary Table S3-6 for the complete zero-order correlation matrix. A few macro-comments broadly characterize the results: (a) interest in both sides has a dramatically different nomological network from preference for neutrality, and these unique networks are consistent with our hypotheses from Table 1, and (b) preference for neutrality's correlation patterns are not simply the opposite of preference for extremity's.

First, preference for neutrality appears distinct from the interest in both sides in hypothesized ways. As per Hypothesis 1, people high in interest in both sides were more intellectually humble; open-minded in general, about politics and religion; higher in curiosity,

non-dogmatic, with more enjoyment of thinking. Preference for neutrality was related with less intellectual humility (although not dogmatism), and less need for cognition. These opposite loadings suggest that whereas interest in both sides reflects an epistemically open strategy, preference for neutrality is epistemically arrogant and unintellectual. As hypothesized, extremity preferring people were more dogmatic and close-minded, and less intellectually humble. This provides initial evidence that both "preferences" show some parallelism, although the effects for extremity preference are somewhat stronger and more robust.

Turning to Hypothesis 2, as predicted, moderation was correlated positively with preference for neutrality[5]; but to our surprise, moderation positively linked with preference for extremity, albeit very weakly. Interest in both sides was only weakly linked with moderation.

As per Hypothesis 3, preference for neutrality was lower in need to evaluate. This makes sense because the need to evaluate scale includes items referring both to a desire to evaluate at all (e.g., "I form opinions about everything"), and an aversion to neutrality (e.g., "It bothers me to remain neutral"). We would nonetheless argue that preference for neutrality may imply a kind of need to evaluate, in the more restrictive sense of a need to form *opinions that are neutral* (i.e., neither very positive nor very negative), but based on the traditional conceptualization of the need to evaluate this correlation is unsurprising. Preference for extremity was positively related, again conforming to our predictions.

[5] These constructs are nonetheless distinct, as the SOM-2 zero-sum matrix makes clear. For example, we suggested earlier that preference for neutrality has attitudinal relevance, and indeed preference for neutrality but not moderation was negatively linked to the need to evaluate. Preference for neutrality was negatively related to the need for cognition whereas moderation was positively related.

Table 2

Correlations Between Attitude Preferences and Other Individual Differences (Study 1)

Measure	*M* (*SD*)	Citation	Interest in Both Sides	Preference for Neutrality	Preference for Extremity
Interest in Both Sides	4.14 (.56)				
Preference for Neutrality	2.92 (.70)		.01		
Preference for Extremity	2.70 (.93)		-.21***	.18***	
Need to Evaluate	3.07 (.53)	Jarvis & Petty (1996)	.03	-.33***	.33***
Intellectual Humility	3.56 (.48)	Krumrei-Mancuso & Rouse (2016)	.40***	-.21***	-.50***
Dogmatism	4.32 (.70)	White et al. (2019)	-.41***	.07	.42***
Need for Cognition	3.13 (.55)	Cacioppo & Petty (1982)	.23***	-.12*	-.10
Curiosity	4.63 (.55)	Kashdan et al. (2018)	.39***	-.01	.10
Willingness to Self-Censor	3.21 (.68)	Hayes et al. (2005)	-.06	.20***	-.10
Self-monitoring	4.33 (.63)	Lennox & Wolfe (1984)	.41***	-.06	-.07
Moderation	4.18 (.91)	Drolet et al. (2021)	.15*	.53***	.14*
OMC General	4.79 (.92)	Price et al. (2015)	.43***	-.06	-.41***
OMC Politics	4.89 (1.02)	*As above*	.43***	-.08	-.19**
OMC Religion	5.09 (1.24)	*As above*	.35***	.11	-.20**
Need for Closure	3.60 (.29)	Webster & Kruglanski (1994)	.03	.11	.20**
Political extremity	1.99 (2.50)	Squared term of centered political position	.01	-.08*	.11**

Note. OMC indicates open-minded cognition. * $p < .05$, ** $p < .01$, *** $p < .001$

Exploratory associations with the scales of attitude-relevant social dispositions were also intriguing. People higher in interest in both sides were uniquely higher on self-monitoring, suggesting a greater tendency to express attitudes consistent with people around with them. However, individuals higher in preference for neutrality were uniquely higher on the willingness to self-censor, indicating a greater tendency to silence themselves when exposed to counter-attitudinal others. One interpretation is that knowing "both sides" of social issues makes it easier for people high in interest in both sides to express shared beliefs with other people, whereas preference for neutrality permits no such flexibility—these individuals identify with their neutral attitudes, and self-censor when exposed to people with strongly positive or negative views.

Finally, in a secondary analysis, we wanted to test for associations between the attitudes preferences and trait ambivalence, which represents people's dispositional experiences of subjective ambivalence (Schneider et al., 2021; Hohnsbehn et al., 2022). Our prediction would follow from Hypothesis 7: we would expect the preference for neutrality (extremity) to be positively (negatively) related to trait ambivalence. Supplementary Sample 2 examines this question, as we fully detail in SOM-5. Congruent with our hypothesis, we found a modest positive association between preference for neutrality and trait ambivalence, $r(183) = .30$, $p < .001$. We did not find an association between preference for extremity and trait ambivalence, $r(183) = -.01$, $p = .876$. Finally, we found an unexpected but weak association of interest in both sides with trait ambivalence, $r(183) = .17$, $p = .022$.

2.2.3. Political Beliefs

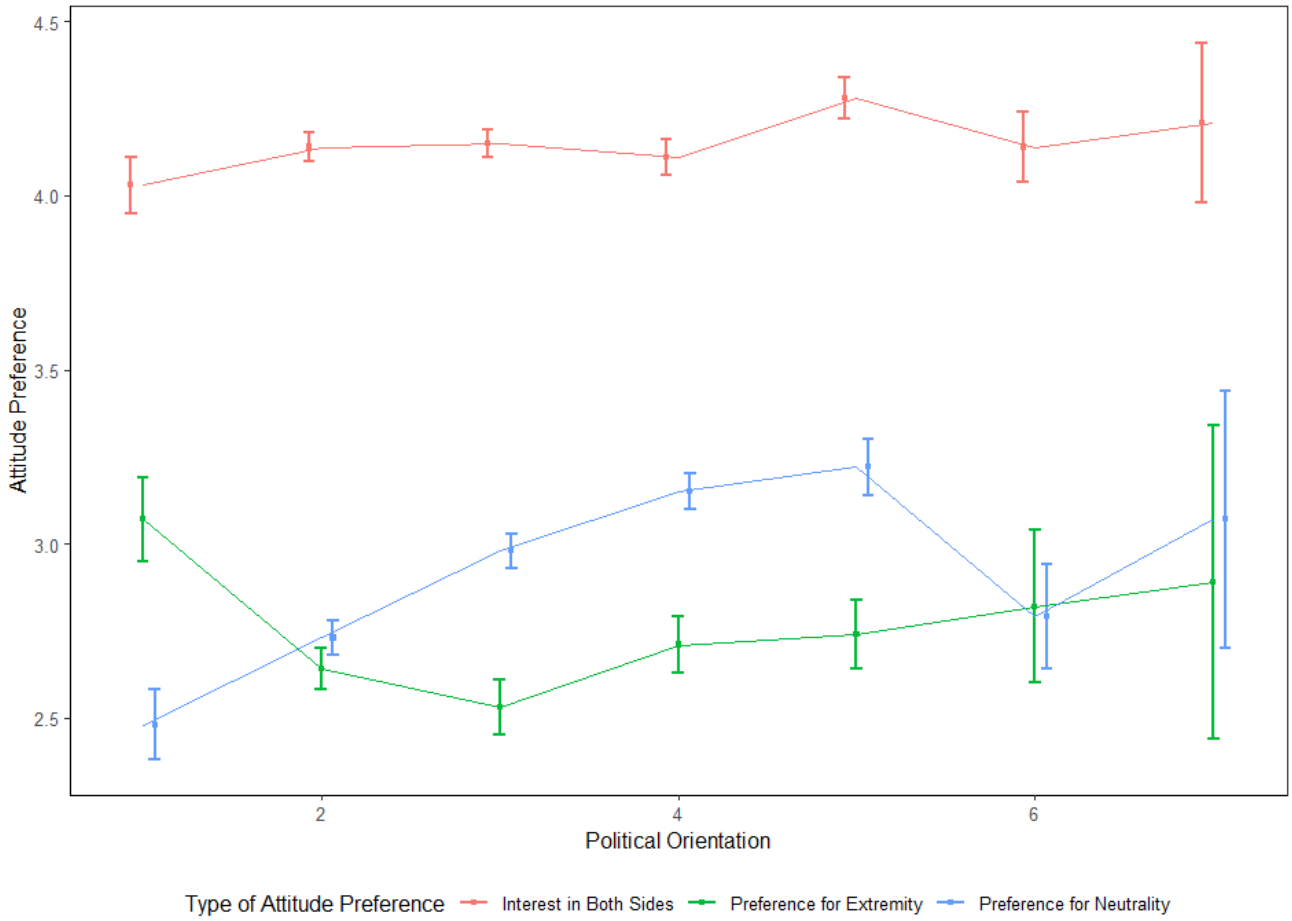
Hypothesis 4 suggested that preference for neutrality should be more common towards the political center, whereas preference for extremity should be linked with more political extremity. Both patterns emerged, albeit somewhat weakly (see Table 2). The pattern can be

viewed in greater detail in Figure 1a: interest in both sides is unrelated to politics, but preference for neutrality has a "reverse-U" shape (most politically centrist), and preference for extremity has a crude "U" shape (more politically extremist).

Some of these associations may have been weakened statistically by our left-leaning student sample, as suggested by the inflated standard error bars on the right side of Figure 1a (i.e., due to having so few conservatives among our undergraduates). We collected a supplementary sample of 350 US Prolific workers which balanced the political distribution (see SOM-4 for more detail, and Figure 1b). As seen in Figure 1b, preference for neutrality was again related to more political centrism and preference for extremity to more political extremism.

Figure 1a

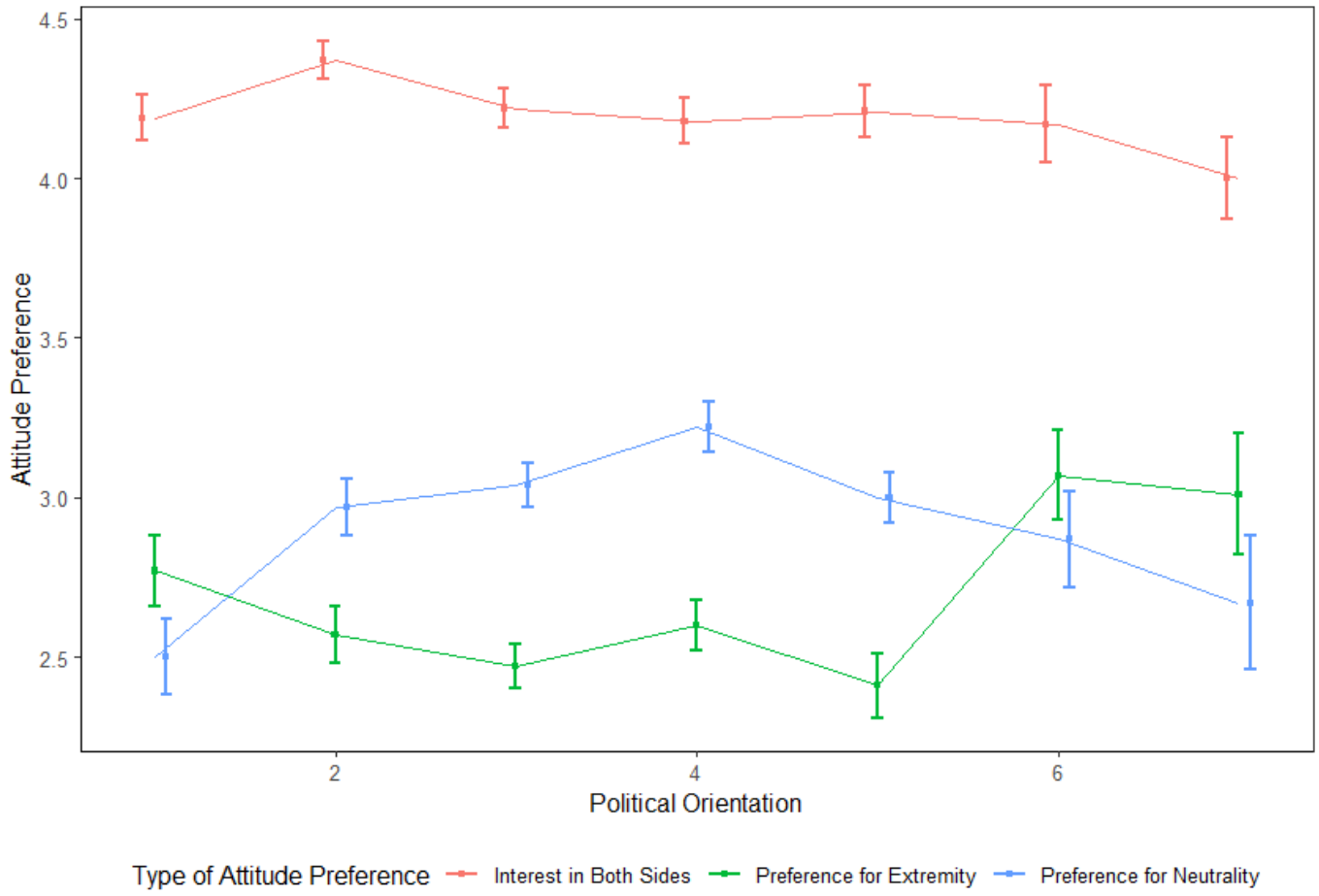
Political Associations with Attitude Preferences (Study 1, Canadian Students).



Note. Error bars indicate standard errors. Higher X-axis scores (political orientation scored 1–7) reflect more conservative political beliefs. Higher Y-axis scores (attitude preference subscales each scored 1–5) reflect more of each attitude preference (see legend).

Figure 1b

Political Associations with Attitude Preferences (Supplementary Data 1, American Prolific).



Note. Error bars indicate standard errors. Higher X-axis scores (political orientation scored 1–7) reflect more conservative political beliefs. Higher Y-axis scores (attitude preference subscales each scored 1–5) reflect more of each attitude preference (see legend).

Neutrality preference was positively related to political conservatism in Study 1, but as we discuss in SOM-3 and SOM-4, this was eliminated in the more politically balanced supplementary sample.

2.2.4. Dispositional Attitude Structures

Study 1 participants also completed the dispositional attitudes scale, a measure capturing dispositional tendencies in preferring to hold positive vs. negative attitudes (Hepler & Albarracin, 2013). Unlike the original scoring, we appraised how far from the conceptual midpoint each participants' rating was of each object, as a measure of general attitude extremity. Table 3 shows two regression analyses, in each of which one of the dispositional attitude indices was regressed on all three attitude preferences. The first analysis examined the original sixteen DAM items which concern non-controversial items like "bicycle." The second analysis examined sixteen more controversial items like "abortion." We anticipated that people preferring neutrality (extremity) should have non-extreme (extreme) attitudes.

Table 3

Attitude Preferences Relate to Dispositional Attitude Extremity (Study 1)

	Interest in Both Sides					Preference for Neutrality					Preference for Extremity					Model Statistics		
	b	*r*	*SE*	*t*	*p*	*b*	*r*	*SE*	*t*	*p*	*b*	*r*	*SE*	*t*	*p*	*F*	*p*	R^2
Dependent variable: Dispositional Attitude Extremity (Non-controversial)	.28	.31	.05	5.61	<.001	.00	.00	.04	-.08	.938	.05	.07	.03	1.36	.177	10.57	<.001	.10
Dependent variable: Dispositional Attitude Extremity (Controversial)	.28	.30	.05	5.56	<.001	-.07	-.09	.04	-1.66	.097	.07	.13	.03	2.33	.021	11.11	<.001	.10

For the non-controversial and again for the controversial objects, we found that interest in both sides was related to having more extreme opinions on both object sets. This is at least congruent with our theorizing that people interested in learning both sides should not automatically default into neutrality, but it was interesting that these individuals had substantially *more* extreme attitudes. One possibility is that these individuals are thinking more extensively about attitudinal issues (i.e., congruent with their heightened curiosity and need for cognition). Interestingly, then, despite these people also being more open-minded, their increased attention to both sides of attitudinal issues may lead to more extreme views either because they nonetheless showed biases like confirmation bias and motivated skepticism (Dawson et al., 2002; Ditto & Lopez, 1992; Edwards & Smith, 1996; Kunda, 1987). Indeed, when people think more (vs less) about attitude-relevant topics, they tend to show larger confirmation biases (Knobloch-Westerwick et al., 2020) and more attitude polarization (Clarkson et al., 2011; Tesser & Conlee, 1975). Stanovich and Toplak (2023) also have noted that open-minded thinking is not consistently related to reductions in the sorts of myside biases that fuel attitude polarization.

In contrast, however, those high in preference for neutrality showed a marginal *decrease* in attitudinal extremity as we had predicted, at least for the controversial objects. Finally, people high in preference for extremity reported more extreme opinions for controversial objects, confirming our expectations.[6]

[6] Our controversial objects differed substantially from the non-controversial original DAM items, as we tested through a series of paired-samples *t*-tests. We would expect controversial objects to be associated with more polarization and variance. First, opinions were on average further from the conceptual midpoint for the controversial ($M = 1.61$, $SD = .52$) versus non-controversial objects ($M = 1.41$, $SD = .49$), $t(126) = 3.51$, $p < .001$, $d = .31$ [.13, .49]. Second, people's opinions drifted further from the item-level medians for the controversial ($M = 1.43$, $SD = .41$) than for the non-controversial objects ($M = 1.29$, $SD = .39$), $t(126) = 4.56$, $p < .001$, $d = .40$ [.22, .58]. We did not have predictions for how our novel constructs would map onto overall liking of the attitude objects, but these analyses are provided in the supplemental Table S2-6 for interested readers.

2.2.5. Personality Profile Evaluations

Finally, we examined participants' evaluations of targets who were globally inclined towards attitudinal neutrality or extremity. People with neutrality (extremity) preferences should like the neutral (extreme) targets, and perhaps dislike the extreme (neutral targets). These patterns are key to our claim that both preferences represent core identity concerns, prompting in-group/out-group thinking towards people who fulfill/reject these attitudinal preferences. In contrast, interest in both sides should not predict much preference for extremists or neutrals. We performed regression analyses as displayed in Table 4. Each used overall attitudes towards targets as the dependent variable, but SOM-3 contains similar regressions using target morality and target competence as dependent variables, reaching very similar conclusions.

We first regressed ratings of the non-extreme targets onto our three attitude preferences, then regressed ratings of the extreme targets onto the same predictors (first two data rows of Table 4). Finally, we calculated the difference between evaluations of the two target sets (e.g., attitudes towards the three extreme targets minus attitudes towards the three non-extreme targets), and again regressed this onto our predictors (third data row). We found that interest in both sides and preference for neutrality were both linked with more liking of the neutral profiles (first data row), whereas preference for extremity was linked with less liking of these targets. However, when examining opinions of extreme-attitude targets (second data row), interest in both sides was marginally positive, suggesting that these individuals would like all profiles more, whereas preference for neutrality was not linked to positive evaluations of these targets. Preference for extremity was positively linked to judgments of the extreme targets, as hypothesized. Finally, we examined the difference scores to reveal how much pro or anti-extremity bias each attitude preference predicted (third data row). We found that whereas interest

in both sides was related only to a marginal bias against extreme-attitude targets, preference for neutrality was related to a strong anti-extremist bias, and preference for extremity was related to a strong pro-extremist bias.

Because we were interested in testing whether interest in both sides differs from neutrality preference, we wanted to test whether interest in both sides' marginally anti-extremist bias and preference for neutrality's significantly anti-extremist bias were significantly different from one another. We used the car package (Fox & Weisberg, 2019) in R (R Core Team, 2022) to test the (in)equality of these regression coefficients using an *F*-test. Indeed, preference for neutrality was *more* negatively related to extremist attitudes than was interest in both sides, $F(1, 610) = 4.10$, $p = .043$. This again demonstrates that preference for neutrality accounts for variance that is not captured by an open-minded interest in both sides, this time concerning the evaluations of other people. Furthermore, these results demonstrate individual differences moderating established effects in which either neutral/moderate or ambivalent people are evaluated differently than their extreme attitude-holding counterparts (e.g., Goldenburg et al., 2023; Han et al., 2023; Johns & Kölln, 2020; Pillaud et al., 2013, 2018; Siev et al., 2024; Toribio-Flórez et al., 2020).

538 **Table 4**

539 *Attitude Preferences Relate to Opinions of Target People Based on Targets' Attitudinal Dispositions (Study 1)*

Dependent Variable	Interest in both sides					Preference for neutrality					Preference for extremity					Model Statistics		
	b	*r*	*SE*	*t*	*p*	*b*	*r*	*SE*	*t*	*p*	*b*	*r*	*SE*	*t*	*p*	*F*	*p*	R^2
Attitudes Towards Neutrals	.24	.17	.06	4.43	<.001	.24	.21	.04	5.50	<.001	-.17	-.20	.03	-5.23	<.001	27.29	<.001	.12
Attitude Towards Extremists	.11	.08	.06	1.96	.051	-.07	-.06	.05	-1.48	.139	.19	.22	.04	5.50	<.001	10.36	<.001	.05
Relative Pro-Extremist Bias (Attitude)	-.13	-.07	.07	-1.96	.051	-.31	-.21	.05	-5.77	<.001	.37	.33	.04	8.98	<.001	38.30	<.001	.16

540
541

2.3. Discussion

The patterns shown by people who want to learn about both sides of issues differed detectably from the patterns shown by people who have neutrality preferences. These differences manifested in numerous ways: (a) unique factors emerged in factor analysis, consistent with the idea that these are conceptually distinct; (b) resultant factors showed near-zero associations with one another; (c) interest in both sides and preference for neutrality showed unique nomological networks in connecting with attitude-relevant constructs (need to evaluate, need for cognition), interpersonal constructs (self-censorship, self-monitoring), and more; (d) only preference for neutrality was related to political centrism; (e) interest in both sides and preference for neutrality showed different patterns of attitude structure, such that only interest in both sides predicted more extreme opinions; and (f) interest in both sides and preference for neutrality showed different patterns of evaluating people based on those targets' attitude patterns.

In short, Study 1 established the conceptual novelty of preference for neutrality, and supported most claims raised as Hypotheses 1–6. At high levels, preference for neutrality connects with other individual difference factors, shapes attitude structures, connects with political views, and guides social judgment.

Study 1 also validated the attitude preference scales. That people higher in the preference for neutrality (extremity) were each more intellectually arrogant, had overall less (more) extreme attitudes, were less (more) politically extreme, viewed neutral others more (less) favorably, all supports the validity of these critical novel measures. Similarly, that the interest in both sides was positively related to open-minded cognition, intellectual humility, and the need for cognition, and negatively related to dogmatism, supports the validity of that subscale.

3. Study 2

Study 1 supported our theorizing about how a preference for neutrality exists independent from an even-handed interest in "both sides" of attitudinal conflicts and from (low) preference for extremity. In Study 2 we explored other aspects of attitude structure more extensively. For example, the attitude measures used in Study 1 consisted of bipolar scales (e.g., rating "genetically modified organisms" from negative to positive) which made it impossible to examine participants' tendencies to have ambivalent (i.e., co-present positive *and* negative) opinions. Importantly, because ambivalence is one of several attitude structure types that can generate neutrality, people who prefer neutrality might be more willing to cultivate structural attitudinal ambivalence and feel ambivalent; people who prefer extremity should presumably avoid ambivalence and find it particularly aversive (Hypothesis 7). Interest in both sides also should link with less ambivalence: we have already observed that these individuals have more extreme attitudes in Study 1, and familiarity with both sides of issues may facilitate rather than inhibit polarization into strong attitude positions.

Another limitation of Study 1 was that preference for extremity was represented by only four items. To improve our measurement of the attitude preferences in general, we generated a substantial number of new items, resulting in a revised scale with 17 interest in both sides, 18 preference for neutrality, and 15 preference for extremity items (listed in SOM-2).

3.1. Method

3.1.1. Participants

We recruited 109 Prolific workers from the United Kingdom, compensating them £1.50 for 15 minutes of time. We selected the sample size based on budgetary constraints. Participants were 70.6% White, 7.3% Asian, 6.4% mixed 5.5% Black, 2.8% Other; M_{age} = 37.7, SD_{age} = 14.4; 58.4% were women, 37.6% men, 3.0% non-binary/third gender, 1% did not answer. No

participants were removed. Because of the repeated-measures design, we have 80% power to find effect of $r > |.15|$ to $|.16|$ (R^2s of .02-.03) per analysis (see SOM-3), again comparing favorably to Richard et al.'s (2003) average effect size in social psychology ($r = .21$ or $R^2 = .04$).

3.1.2. Procedure and Materials

Participants rated 10 attitude objects, ranging from commonplace brands and objects (the company Apple, vaping, electric cars, McDonalds, hairless cats) to politically important and controversial but not generally left/right polarized issues (veganism, genetically modified organisms, nuclear power, universal basic income, and capitalism). We randomized order of object presentation. First, participants rated the 10 objects in terms of their positive and negative characteristics, with positive/negative ratings in counterbalanced order. We asked participants to rate the object's positive aspects, while ignoring all negative aspects, on a scale from 1 (*not at all positive*) to 11 (*very positive*). We separately asked participants to rate the object's negative aspects, while ignoring all positive aspects, on a scale from 1 (*not at all negative*) to 11 (*very negative*). Structural ambivalence of each object could be calculated using the Griffin formula, (Positives – Negatives)/2 – ABS(Positives – Negatives), as described by Thompson et al. (1995). Participants structural ambivalence scores had *M*s = -.37–2.42, *SD*s = 3.52–4.02.

Second, participants viewed the 10 objects one at a time, and rated several subjective ambivalence items (from Priester & Petty, 1996): to what extent object made them feel conflict (1 = *feel no conflict at all*; 11 = *feel maximum conflict*), confusion (1 = *feel no confusion at all*; 11 = *feel maximum confusion*), indecision (1 = *feel no indecision at all*; 11 = *feel maximum indecision*), and mixed reactions (1 = *completely one-sided*; 11 = *completely mixed reactions*). These variables were averaged together such that higher scores indicated more felt ambivalence.

Participants felt ambivalence scores had *M*s = 2.98–5.01 and *SD*s = 2.07–2.73 and good reliabilities αs = .89–.94.

Finally, participants completed our revised (50-item) attitude preferences scale. As demonstrated in SOM-3, this item pool again produced clear evidence of a three-factor scale, with scales reflecting preference for neutrality (18 items; $M = 2.76$, $SD = .73$, $\alpha = .93$), interest in both sides (17 items; $M = 4.15$, $SD = .60$, $\alpha = .93$), and preference for extremity (15 items; $M = 2.81$, $SD = .75$, $\alpha = .92$). These same subscales were used in Study 3, both supplementary studies, and are the final version of the scale.

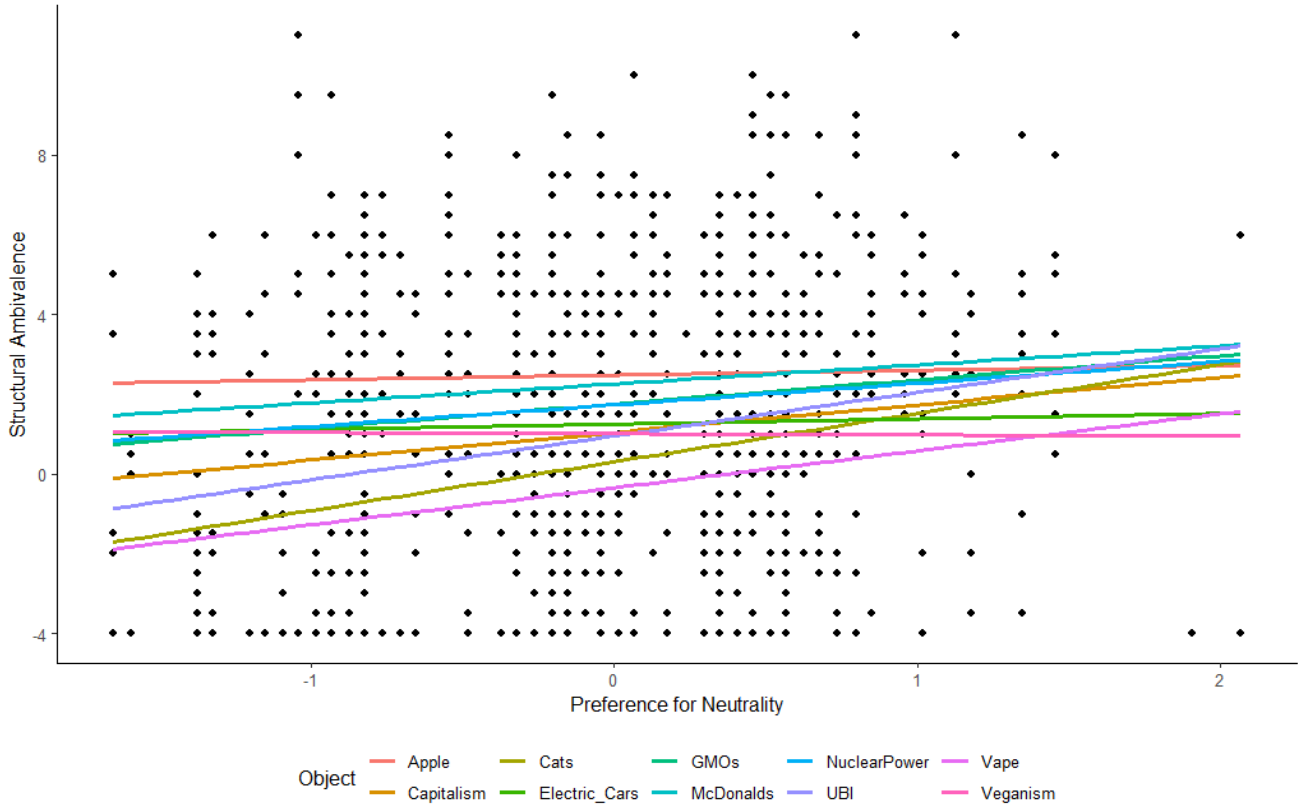
3.2. Results

3.2.1. Attitude Preferences as Predictors of Structural Ambivalence

We used multilevel models to assess how the attitude preferences were related to structural ambivalence across the 10 objects, because object-level ratings ($n = 1{,}008$) were nested within respondents ($N = 101$ for this analysis). A random slopes model failed to converge, so we used a random intercepts model. Regressing structural ambivalence on all three attitude preferences, we found that interest in both sides was unrelated to structural ambivalence, $B = -.45$ [-1.11, .20], $t(97) = -1.37$, $p = .175$, $R^2_c = .00$. However, preference for neutrality was related to marginally more structural ambivalence, $B = .54$ [-.01, 1.08], $t(97) = 1.95$, $p = .055$, $R^2_c = .01$, consistent with our reasoning. This can be seen in Figure 2a, where each colored line indicates one attitude object; more preference for neutrality, at least marginally, relates to more ambivalence across topics. Finally, preference for extremity was related to significantly less structural ambivalence, $B = -.57$ [-1.10, -.04], $t(97) = -2.13$, $p = .035$, $R^2_c = .02$, also consistent with our reasoning. This can be seen by the downward-sloping lines in Figure 2b.

Figure 2a

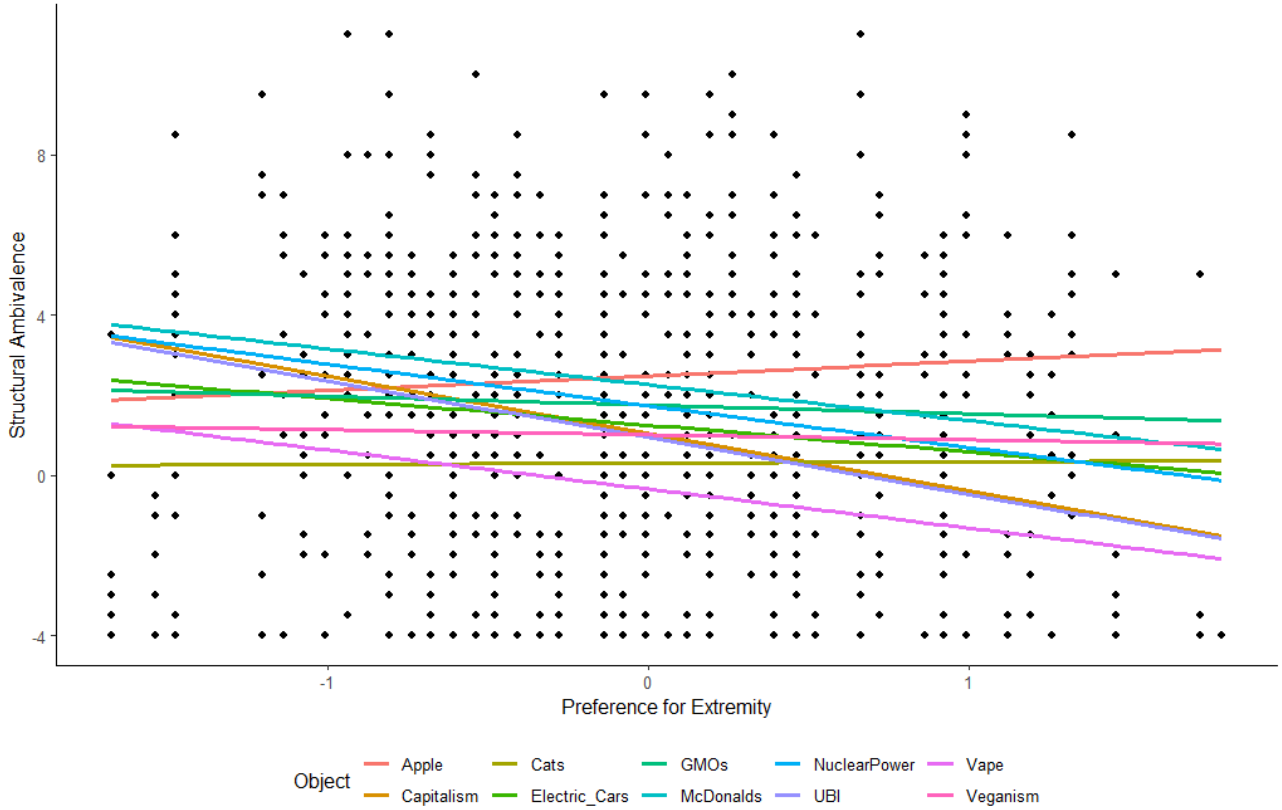
Association of Preference for Neutrality and Structural Ambivalence Across 10 Attitude Objects



Note. Individual lines represent object-specific associations of the variables noted above. "GMOs": Genetically Modified Organisms. "UBI": Universal Basic Income. X-axis scores are grand-mean centered. Higher Y-axis scores represent more structural ambivalence and are calculated through the Griffin formula (i.e., a difference score in which relatively dissimilar degrees of positive/negative reactions decrease measured ambivalence, and more overall reactions increase ambivalence).

Figure 2b

Association of Preference for Extremity and Structural Ambivalence Across 10 Attitude Objects



Note. Individual lines represent object-specific associations of the variables noted above. "GMOs": Genetically Modified Organisms. "UBI": Universal Basic Income. X-axis scores are grand-mean centered. Higher Y-axis scores represent more structural ambivalence and are calculated through the Griffin formula (i.e., a difference score in which relatively dissimilar degrees of positive/negative reactions decrease measured ambivalence, and more overall reactions increase ambivalence).

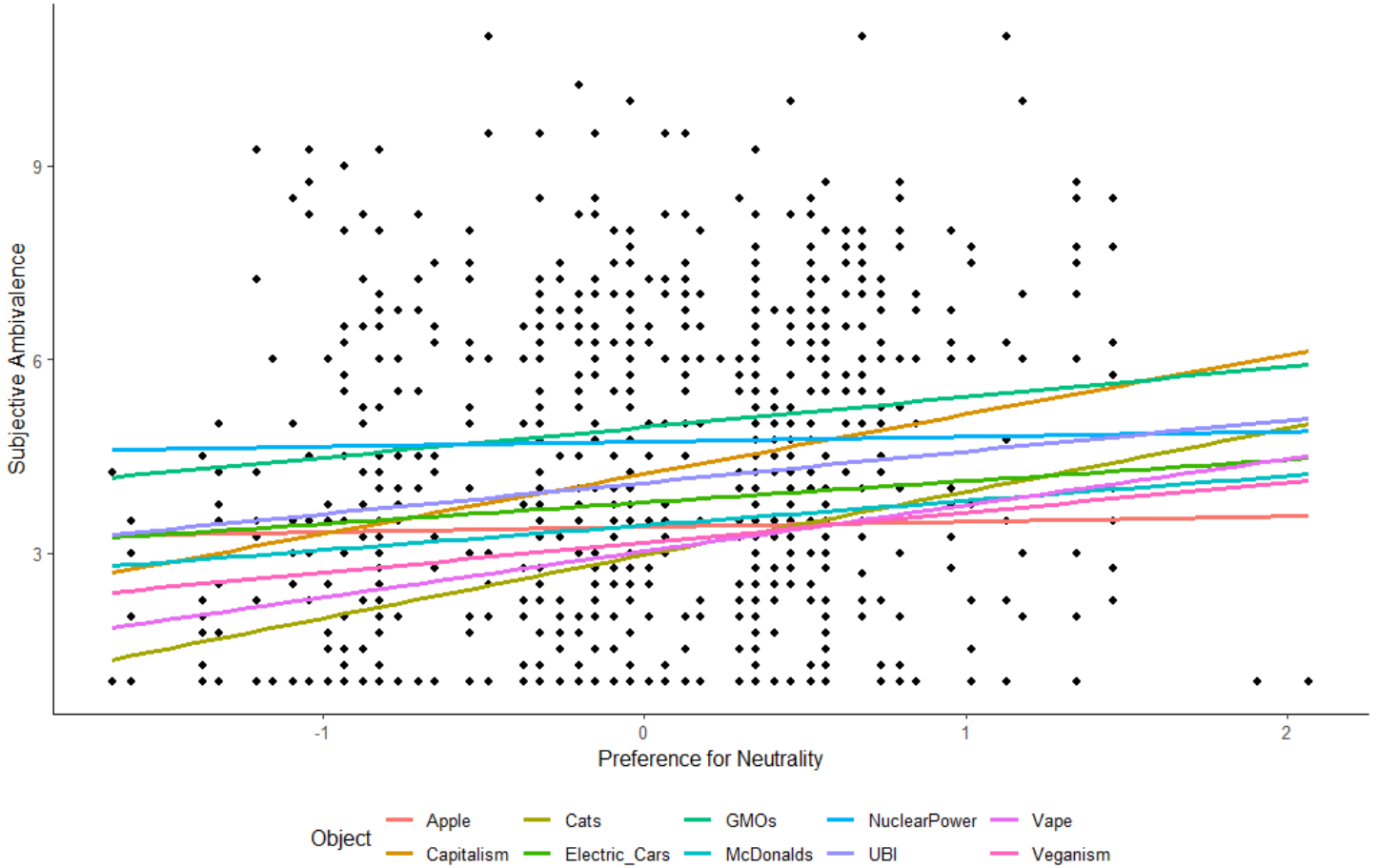
3.2.2. Attitude Preferences as Predictors of Subjective Ambivalence

We re-ran the analysis from the previous subsection, but analysing subjective ambivalence instead of structural ambivalence. Preference for neutrality was hypothesized and found to be connected with more subjective ambivalence, $B = .56$ [.16, .96], $t(97) = 2.80$, $p = .006$, $R^2_c = .03$. Once again this is seen with reasonable consistency across objects, as depicted in Figure 3a's consistently positive slopes, suggesting solid connections between higher preference for neutrality and higher felt ambivalence per object. This fulfills our expectation that people who prefer neutrality regularly feel mixed about attitude objects. Unlike for structural ambivalence, interest in both sides was hypothesized and found to be related to less subjective ambivalence, $B = -.54$ [-1.02, -.06], $t(97) = -2.25$, $p = .027$, $R^2_c = .01$. This is shown in Figure 3b, where the connection between interest in both sides and reduced subjective ambivalence is also quite consistent across objects. Finally, preference for extremity was unrelated to subjective ambivalence, $B = .03$ [-.35, .42], $t(97) = .18$, $p = .861$, $R^2_c = .00$.

These effects were very consistent when permitting slopes to be random across people (i.e., so that people could vary in their association of attitude preference to subjective ambivalence), the model for which converged this time: $B_{\text{both-sides}} = -.64$ [-1.11, -.18], $t(97) = -2.75$, $p = .007$, $R^2_c = .01$; $B_{\text{neutrality-preference}} = .56$ [.17, .95], $t(97) = 2.86$, $p = .005$, $R^2_c = .03$; $B_{\text{extremity-preference}} = .16$ [-.22, .53], $t(97) = .82$, $p = .413$, $R^2_c = .00$.

Figure 3a

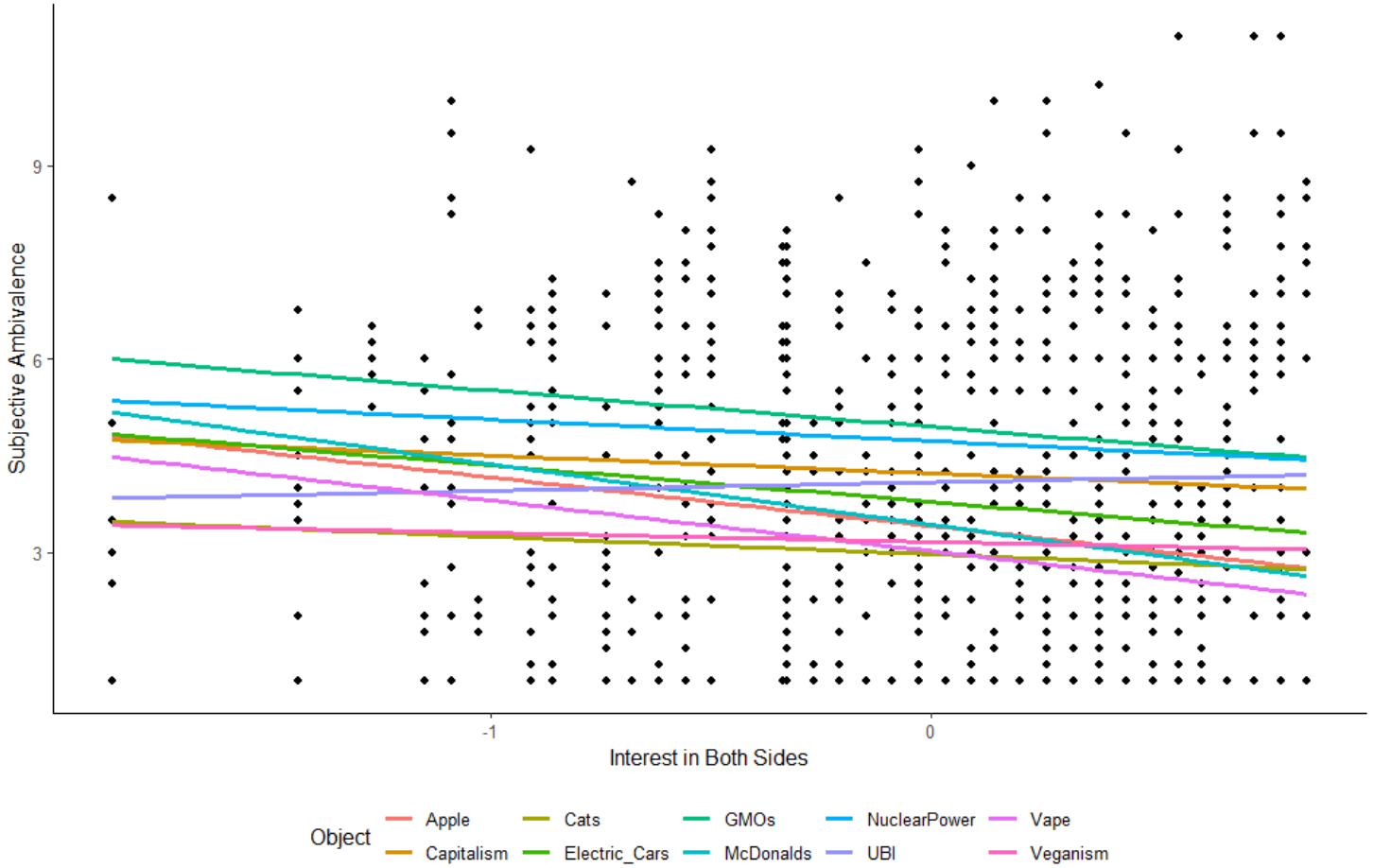
Association of Preference for Neutrality and Subjective Ambivalence Across 10 Attitude Objects



Note. Individual lines represent object-specific associations of the variables noted above. "GMOs": Genetically Modified Organisms. "UBI": Universal Basic Income. X-axis scores are grand-mean centered and trace participants' "preference for neutrality" scores. Higher Y-axis scores represent greater subjective ambivalence (e.g., feeling torn, divided about the object).

Figure 3b

Association of Interest in Both Sides and Subjective Ambivalence Across 10 Attitude Objects



Note. Individual lines represent object-specific associations of the variables noted above. "GMOs": Genetically Modified Organisms. "UBI": Universal Basic Income. X-axis scores are grand-mean centered and trace participants' "interest in both sides" scores. Higher Higher Y-axis scores represent greater subjective ambivalence (e.g., feeling torn, divided about the object).

3.3. Discussion

Study 2's more precise examination of attitude structure yielded some interesting discoveries about attitude preferences. First, it provided a double dissociation between interest in both sides (which was related to less subjective ambivalence) versus preference for neutrality (which was related to more subjective ambivalence), confirming Hypothesis 7. This makes sense because people with an interest in both sides of topics have no reason to feel more conflicted and torn about issues, because they do not necessarily internalize the positives and negatives of attitude objects—they examine all possible information about topics, and presumably show the same biases that most people show when weighing and integrating this information into attitudinal structures (Dawson et al., 2002; Ditto & Lopez, 1992; Edwards & Smith, 1996; Kunda, 1987). Indeed, we know from Study 1 that people high in interest in both sides tend towards more extreme opinions, and so it is unsurprising that they feel lower amounts of confusion and indecision. In strong distinction, people who prefer neutrality have more co-present positive and negative reactions across many attitude objects, so it follows that they are also feel more ambivalent.

4. Study 3

In Study 3, we wanted to examine whether attitude preferences contribute to the psychological processes underlying persuasion. This study builds on Studies 1–2 by testing whether preference for neutrality will guide responses to persuasive content, such that people high in preference for neutrality may be more likely to internalize information when that information is framed as "neutral." Furthermore, it develops the idea that preference for neutrality is an intellectually arrogant drive for neutrality (Hypothesis 1) by probing whether these individuals will be more persuaded by any information when it is flagged as neutral or

moderate *regardless of the actual valence of the information* (Hypothesis 8). This test can provide additional evidence that preference for neutrality is a fundamental commitment by showing a preference for information labeled as fulfilling the attitudinal goal, just as conservatives/liberals prefer information labeled as conservative/liberal even when the policy itself remains unchanged (Cohen, 2003). Ironically, this preference may not always lead to neutral attitudes, as we will demonstrate.

To do this, we introduced participants to an unfamiliar attitude object, lemphurs (an ostensibly real animal), and asked for their initial opinions about it. Then we exposed participants to different versions of an argument about why lemphurs are good creatures, which held all information constant, but merely labeled the argument's source as being "moderate," "extreme," or neither. In this way, we could examine whether the mere labeling of the source of information as conforming to people's attitude preferences could shape whether they accepted that information. That is, we hoped to create circumstances in which matching of information to recipients' underlying preferences might instigate a bias favoring persuasion (Luttrell et al., 2021; Petty & Wegener, 1998; Teeny et al., 2021). This study is also relevant to Hypothesis 6 because participants' judgments of a speaker whose position is arbitrarily labeled as extreme or neutral may depend on the participants' preference for neutrality or extremity.

4.1. Method

4.1.1. Participants

We followed a time-based rule: recruitment for two academic semesters, with a goal of reaching 100 participants for each of the three between-subject conditions. We stopped after the second semester, having reached approximately this number: 274. No participants were removed. We have 80% power to find (expected) crossover interactions with simple slopes of $r = +.25$ and

-.25 (see SOM-3 for details), slightly falling short of Richard et al.'s (2003) average effect size in social psychology. We recruited students from a Northern UK university to participate online for course credit; 67.4% were White, 24.4% Asian, .8% Black, and 7.4% mixed; M_{age} = 19.6, SD_{age} = 2.3; 82.4% women, 15.6% men, 1.2% non-binary/third gender, .8% preferred not to answer.

4.1.2. Procedure

In Study 3, we wanted to maximize the opportunity for our constructs to demonstrate their persuasive influence, so we introduced participants to an unfamiliar object, an ostensibly real creature called a lemphur (a common paradigm in persuasion psychology; Fabrigar & Petty, 1999; Guyer et al., 2018; Haddock et al., 2008; Rocklage & Fazio, 2015; See et al., 2013). We asked all participants to evaluate lemphurs, "even if you have not heard of them before," on eight items from Crites et al., 1994 (items rated 1 = *not at all* to 9 = *definitely*; M = 1.36, SD = 2.48, α = .94; order randomized). Four items were positive, four negative; we subtracted their negative from their positive reactions to form an attitudes index.

Next, we asked participants to read some positive information about lemphurs, describing lemphurs' high intelligence, low mortality rate, nutritious meat, et cetera. All participants received identical information about lemphurs, always provided by "Mr. Brown," a "hobbyist who maintains a blog about lemphurs." We framed it this way to make Mr. Brown's credibility somewhat ambiguous, so that the matching of the message's framing to the recipients' personality could hopefully exert a persuasive influence. Despite the information being held constant, we randomly assigned participants to one of three between-participant framing conditions. In the control condition, participants simply learned that Mr. Brown "is interested in [lemphurs]. In the neutrality condition, we told participants that Mr. Brown "has **very moderate opinions** about lemphurs" (emphasis in original). In the extremity condition, we told participants

that Mr. Brown "has **very extreme opinions** about lemphurs" (emphasis in original). Thus, the neutrality and extremity conditions' texts differed on only a single word, with the remaining 415 words remaining consistent (99.8% content overlap).

Once participants read the passage, we asked them to rate lemphurs again on the attitude items that they previously completed (M = 4.18, SD = 2.37, α = .95). We also asked them to rate the writer on a series of moral and competence evaluations as per Study 1 (Wojciszke et al., 1998; $M_{morality}$ = 4.42, SD = .95, α = .86; $M_{competence}$= 4.65, SD = .91, α = .85). Finally, we had them complete the attitude preferences scale ($M_{prefneutral}$ = 2.91, SD = .63; $M_{bothsides}$ = 4.25, SD = .51; $M_{extpref}$ = 2.66, SD = .59; α = .92–.94). Because preference for neutrality promotes a biased preference for neutral over extremist others (Hypothesis 6, Study 1), and a reduced tendency to hold extreme attitudes (Hypothesis 4–5, Study 1), we reasoned that preference for neutrality should increase positive evaluations and persuasion by arguments when they are framed as neutral versus extremist. The reverse should occur for preference for extremity.

4.2. Results

4.2.1. Persuasion Effects

We hypothesized that people would be more persuaded when a persuasive message was merely labeled as fulfilling their attitude preferences (Hypothesis 8). To this effect, we regressed several dependent variables onto a contrast code representing the degree to which the framing suggested extremity versus neutrality (neutrality label = -0.5, control group = 0, extremity label = +0.5), the three attitude preferences, and all possible two-way interactions between the contrast

code and the attitude preferences. Table 5 shows the effects from this regression analysis and the analyses corresponding to the next two subsections.[7]

First, we found that interest in both sides was related to more persuasion; this effect was not moderated by framing, and so is consistent with our conception of interest in both sides motivating an openness to all types of novel information. Also, unsurprisingly, pre-persuasion attitudes were associated with post-persuasion attitudes. More crucial was an interaction between preference for neutrality and framing. The influence of the extreme versus neutral label yielded no persuasive effect when people were low in preference for neutrality (-1 *SD*), $B = .13$, $SE = .47$, $t(229) = .27$, $p = .785$, but it significantly reduced persuasion at high (+1 *SD*) levels of preference for neutrality, $B = -1.25$, $SE = .49$, $t(229) = -2.55$, $p = .012$. Ironically, as the green "extremity frame" slope line in Figure 4a demonstrates, high preference for neutrality people (on the right) were much less neutral (specifically, they were more negative) than were moderate preference for neutrality people (in the center) given an extremity-framed message.

[7] Adding a second contrast code contrasting control (+.50) versus the two frame conditions (-.25, -.25) as a main and interactive effect (with the attitude preference measures) had no meaningful effect on the hypothesized results (see open syntax), so to reduce Table 5's complexity we used this simpler analysis.

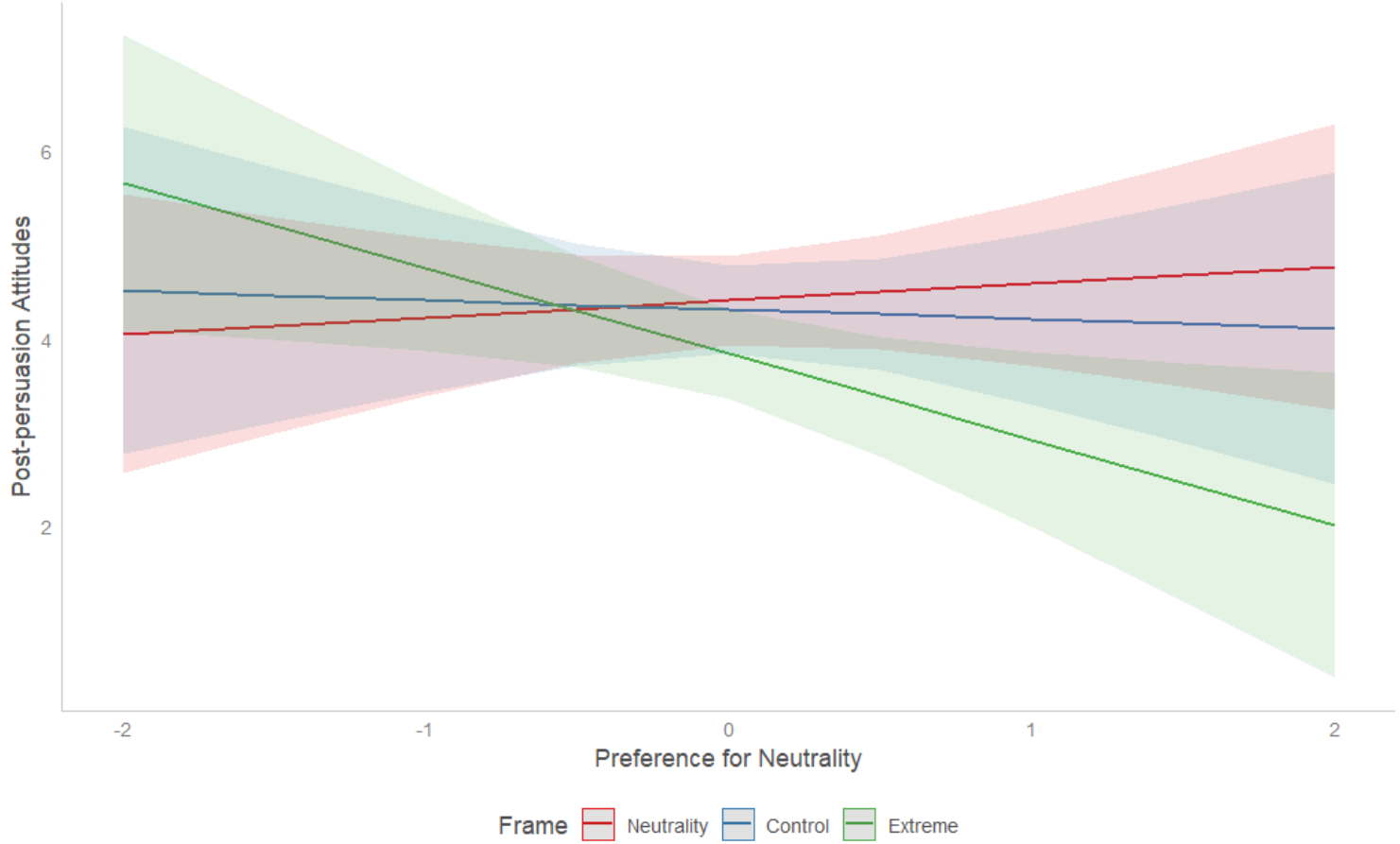
Table 5

Regression Analyses of Attitude Preferences and Framing Predicting Post-Persuasion Outcomes (Study 3)

	Dependent Variable: Persuasion					Dependent Variable: Perceived morality					Dependent Variable: Perceived competence				
	B	*r*	*SE*	*t*	*p*	*B*	*r*	*SE*	*t*	*p*	*B*	*r*	*SE*	*t*	*p*
A. Framing contrast (high = extreme)	-.56	-.10	.34	-1.65	.100	-.23	-.10	.15	-1.56	.121	-.15	-.07	.14	-1.07	.284
B. Interest in both sides	1.61	.34	.27	5.88	<.001	.29	.15	.12	2.42	.016	.32	.18	.11	2.80	.005
C. Preference for neutrality	-.29	-.08	.22	-1.32	.188	.01	.01	.10	.15	.885	.05	.03	.09	.49	.626
D. Preference for extremity	-.24	-.06	.25	-1.00	.321	.24	.15	.11	2.31	.022	.26	.16	.10	2.52	.012
A X B	-.08	-.01	.66	-.12	.902	-.43	-.10	.29	-1.49	.138	-.27	-.06	.28	-.96	.339
A X C	-1.08	-.12	.53	-2.04	.042	-.48	-.13	.23	-2.08	.039	-.46	-.13	.22	-2.10	.037
A X D	-.49	-.05	.58	-.83	.406	-.16	-.04	.25	-.63	.527	-.03	-.01	.24	-.12	.908
E. Time-1 measure of attitude	.30	.31	.06	5.41	<.001				N/A					N/A	
Model fit statistics	$F(8, 229) = 9.10$, $p < .001$, $R^2 = .24$					$F(7, 235) = 2.61$, $p = .013$, $R^2 = .07$					$F(7, 235) = 2.65$, $p = .012$, $R^2 = .07$				

Figure 4a

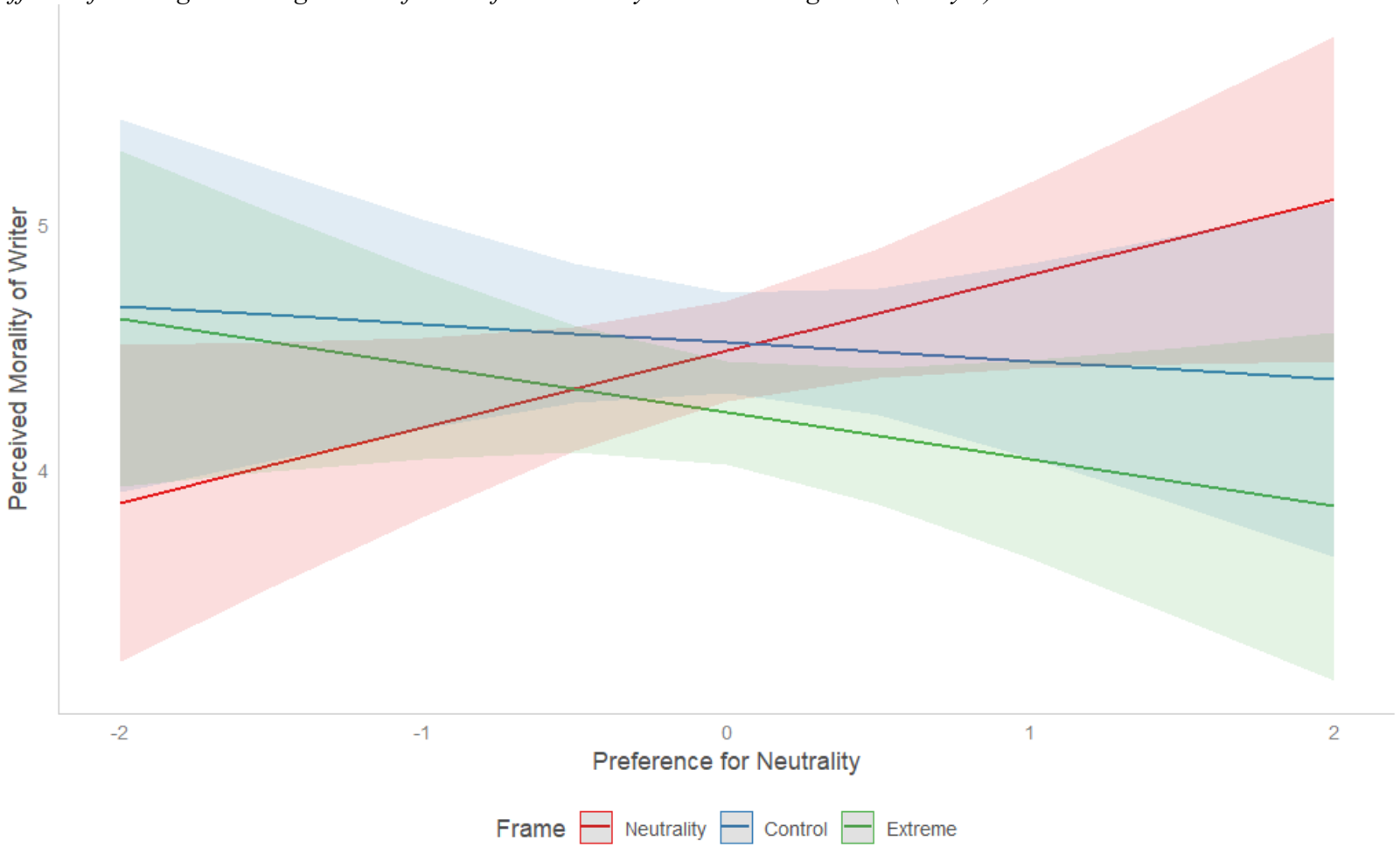
Effects of Message Framing and Preference for Neutrality on Persuasion (Study 3)



Note. Shaded region represents 95% confidence intervals. X-axis scores trace participants' preference for neutrality and are centered around zero. Y-axis scores represent attitudes after reading the persuasive message, with more positive scores indicating more persuasion.

Figure 4b

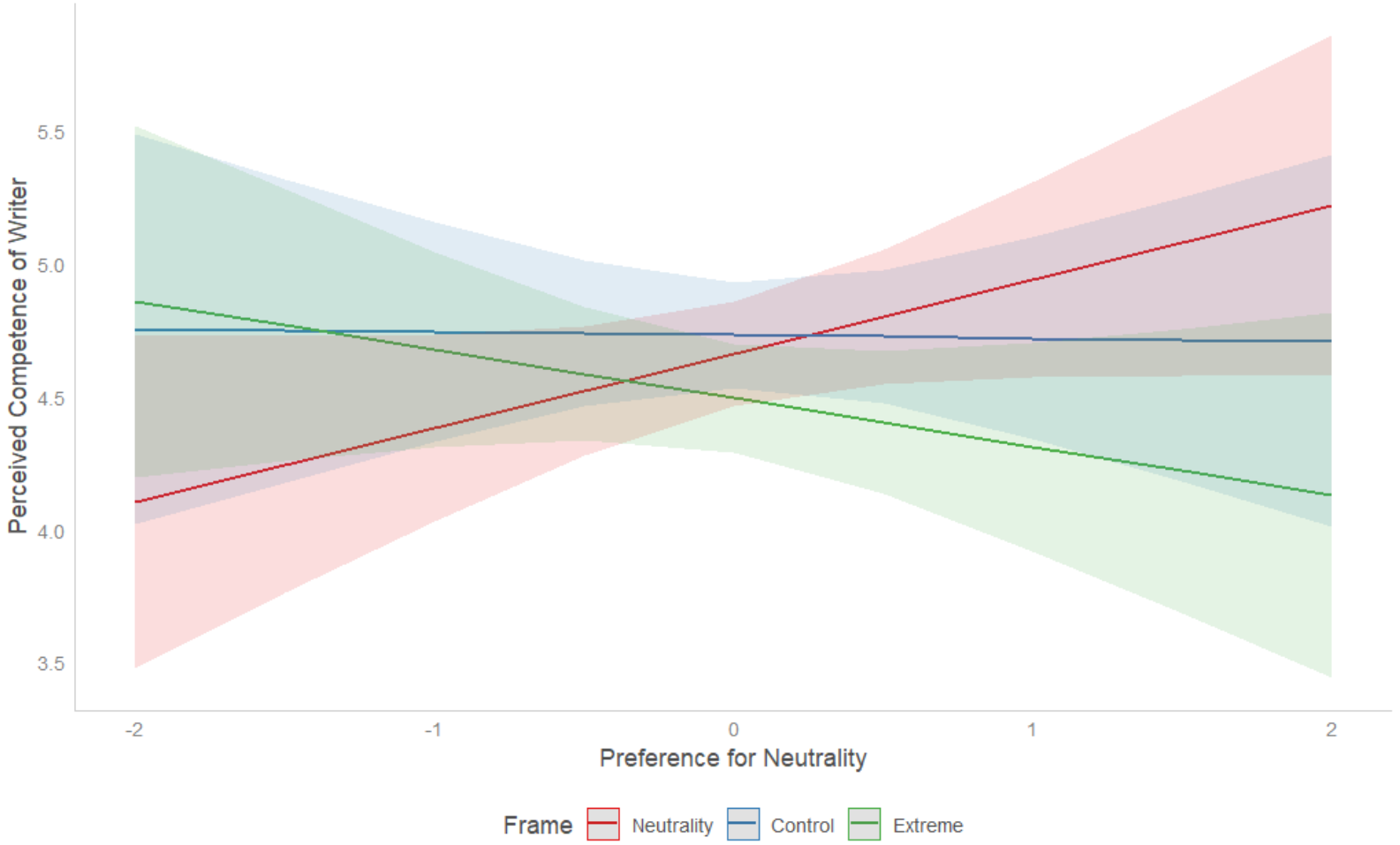
Effects of Message Framing and Preference for Neutrality on Moral Judgments (Study 3)



Note. Shaded region represents 95% confidence intervals. X-axis scores trace participants' preference for neutrality and are centered around zero. Y-axis scores represent judgments of the message writer's morality; higher scores suggest more positive views.

Figure 4c

Effects of Message Framing and Preference for Neutrality on Competence Judgments (Study 3)



Note. Shaded region represents 95% confidence intervals. X-axis scores trace participants' preference for neutrality and are centered around zero. Y-axis scores represent judgments of the message writer's competence; higher scores suggest more positive views.

4.2.2. Evaluations of the Writer

Finally, we examine evaluations made of the writer. Some of these effects parallel the persuasion effects we detected in section 4.2.1, in that just like interest in both sides was related to more persuasion, interest in both sides also was related to an increased perception that the source was more moral and competent, regardless of framing. Additionally, just as preference for neutrality was related to less persuasion given an extreme-labeled vs moderation-labeled source, preference for neutrality was related to disparagement of the source on both moral and competence dimensions in the same pattern (see Table 5, and Figures 4b–4c). This effect can also be seen as a replication of Study 1 (Hypothesis 6): preference for neutrality relates to more (less) liking of neutral (extremist) others. Specifically, at low levels of preference for neutrality (-1 *SD*), people did not see the source as more/less moral, $B = .07$, $SE = .20$, $t(235) = .37$, $p = .715$, or competent, $B = .14$, $SE = .19$, $t(235) = .73$, $p = .466$, based on framing. However, at high levels of preference for neutrality (+1 *SD*), people saw an extremism-labeled source as less moral, $B = -.53$, $SE = .21$, $t(235) = -2.51$, $p = .013$, and less competent, $B = -.44$, $SE = .20$, $t(235) = -2.19$, $p = .030$, based on the single word “extreme” versus “moderate” used to frame that target person.

4.3. Discussion

Hypothesis 8 was supported with higher preference for neutrality linked to more positive receptions of information framed as coming from a moderate versus extremist, not only in terms of moral and competence evaluations of the source labeled as such, but also in terms of persuasion to messages framed this way. This result is also interesting because the message was precisely identical in the various framing conditions. Such an effect shows the biased nature of preference for neutrality in that the mere label of neutrality is sufficient to drive substantial

persuasion differences. Although of secondary importance compared to preference for neutrality, we were surprised that a parallel framing effect did not emerge for preference for extremity as hypothesized. This could simply represent a Type II error, but another possible interpretation is that preference for extremity is not as responsive to mere labels of 'neutrality' and 'extremity' and instead requires information to be genuinely extreme. Alternatively, consistent with the stronger effects on attitude extremity for controversial versus non-controversial topics in Study 1, preference for extremity may more strongly predict reactions to controversial topics, a category that lemphurs would not fall into.

People high in preference for neutrality reacted so negatively to the extremity framed information that they ended up with less neutral (specifically, negative) attitudes than low preference for neutrality people in this condition. Because the persuasive message always offered positively-valenced information, these negative attitudes may have reflected a rejection of the "extreme" source's arguments—ironically, such an extreme rejection that they ended up less attitudinally neutral than people who prefer neutrality less.

5. General Discussion

5.1. Summary

Across three main and two supplementary studies, testing eight hypotheses, we examined distinct psychological orientations people have towards particular attitude positions, and how these orientations to attitude positions may have important implications for their reactions to other people, objects, and information. We found evidence for three independent constructs: a desire for learning about both sides of a topic, a preference for neutral opinions, and a preference for extreme opinions. These multiple motivations regarding attitude positions were distinct from one another and uniquely related to biased perceptions of others, the processing of evaluative

information, patterns of attitude position and structure, and more. We examined these constructs across multiple types of test (correlations with established measures in Study 1, associations with sets of attitude objects in Study 2, a persuasion paradigm in Study 3). Furthermore, the attitude objects represented across studies were diverse; some were politically trivial, uncontroversial, and non-polarized (Study 1's original DAM measure; Study 3's lemphur), others were politically important and controversial but not clearly left-wing/right-wing polarized (Study 2's objects like GMOs, nuclear power), and others were politically important, controversial, and polarized (Study 1's objects like abortion, BLM, quotas for women in STEM, abortion).

The first major theoretical advance is demonstrating that there are people who strongly prefer neutrality. We found evidence that these people are low in intellectual humility and are relatively disinterested in thinking. Their relatively neutral attitudes and high degree of structural and subjective ambivalence demonstrate successful cultivation of their desired attitude positions. They also showed a biased preference when evaluating others, holding positive views of others with more neutral positions, paralleling patterns typically associated with attitude-extreme individuals (van der Pligt et al., 1983). They engaged in biased information processing about a novel attitude object in which persuasive content arbitrarily framed as neutral (vs extreme) was more compelling to them. People preferring neutrality also were more likely to be political centrist in two samples. In sum, the neutrality preferences are linked with successful, relatively uncritical pursuit of neutrality.

Notably, preference for neutrality differed starkly from the concept of open-minded interest in learning about multiple sides in attitude controversies. The latter group was open-minded, curious, and intellectually humble, seeking to understand each topic by learning information regardless of valence or opposition to their views. Interestingly, this open-minded

gathering of information did not lead people to consistently take neutral positions. Those who scored high versus low on interest in both sides actually had more extreme attitude positions and felt less ambivalent. This makes sense given that open-minded thinkers were also likely to enjoy thinking more (i.e., higher on need for cognition): people tend to polarize when they think extensively about attitude objects, even when trying to be unbiased (Lord et al., 1984).

Finally, the notion that some people prefer extremity is relatively less novel, but it nevertheless was important to demonstrate that preference for neutrality was not simply the inverse of extremity preference, which we demonstrate across several outcomes. Extremity preferring people exhibit patterns resembling past work concerning dogmatism (Hanson, 1973), the need to evaluate (Jarvis & Petty, 1996), and studies of attitude extremity per se. They expressed a desire to form strong, extreme evaluations, and we found that they had more extreme and less ambivalent attitudes. We found evidence that they were less intellectually humble and more close-minded, suggesting an inflexible stance on these extreme positions. Those high on preference for extremity were dogged in their desire to form strong evaluations and this was reflected in a biased preference for attitude-extremists over attitude-neutrals. These findings echo the bulk of the literature, which identifies those with extreme attitudes as having strong attitudes (Abelson, 1995), distrustful of contrary information (Gunther, 1988), judgmental of counter-attitudinal others (van der Pligt et al., 1983), personally identifying with their extreme attitudes (Westfall & Van Boven, 2015), and prone to further polarization (Miller et al., 1993).

5.2. Implications and Future Directions

5.2.1. Preference for Neutrality

One major contribution of this work is preference for neutrality itself, an individual difference variable which intriguingly reveals how an excessive preference for attitude neutrality

can, perhaps ironically, become a close-minded form of intellectual arrogance itself. Although we agree that polarization concerning a range of important social issues is a serious and potentially growing problem in the Western world (Abelson, 1995; van Boven et al., 2012), we also think it is important not to overlook an unquestioned attraction to middle positions. There is nothing wrong with holding neutral attitudes in themselves, but many of our findings concerning preference for neutrality (i.e., a deep, uncritical belief that neutral positions are globally superior) may raise concerns. Rather than reflecting curious open-mindedness or the mere lack of an opinion, it represents a strong, a priori view of what is true, which also skews evaluations of other kinds of attitude-holders. Neutrality in the face of evidence that clearly supports one side over the other may prevent people from addressing major social issues. Given that only Study 1 assessed politically polarized topics specifically, however, future data applying attitude preferences to political polarization will be helpful in testing this possibility.

One interesting possibility is that these patterns stem from a moralized stance towards neutrality in which neutrality becomes a tribal identity commitment, like the group identities fostered through political identities (Clark et al., 2019; Graham et al., 2012; Sowell, 2002). Several preference for neutrality items assert neutrality's moral superiority (e.g., "striving to remain neutral on most topics is a virtuous stance"; "…taking a neutral stance is often an act of bravery"). Studies 1 and 3 assessed this moral conviction indirectly through its effects on perceptions of others. We had reasoned that since our neutrality and extremity preferring participants seemed to believe that holding neutral (extreme) attitudes is virtuous, this would bias the ways that they encountered new people and information. Indeed, perceived morality is a key component of impression formation (e.g., Brambilla et al., 2012, 2021) and so seeing other people as violating a moral principle ("thou shalt hold neutral attitudes") should and did greatly

impact perceptions of others. Accordingly, Study 1 revealed that preference for neutrality positively relates to a valuing of moderation as a general moral value, evidenced by a strong anti-extremist bias and a strong preference for neutral attitude holders (Drolet et al., 2021). This is interesting because moral attitudes are usually more extreme (e.g., Skitka et al., 2005), but people high in preference for neutrality judge others as immoral for not holding *neutral* stances.

Future connections of our constructs to the attitude extremity literature could be productive. For example, past work suggests that people sometimes process persuasive passages more carefully when they contain linguistic markers suggesting attitude extremity (e.g., "she detested him" versus "she didn't like him"; Craig & Blankenship, 2011, p. 291; also see Blankenship & Craig, 2011). This increased processing encompasses both more persuasive benefit of strong over weak arguments, and more behavioral intentions formed in connection with resulting attitudes. The most likely connection with our constructs is that preferences for neutrality (extremity) will simply prompt more rejection (acceptance) of messages that include more linguistic cues of extremity. This result would be almost a conceptual replication of the present Study 3. However, under other circumstances it may be that preferences for neutrality (extremity) will attenuate (amplify) the processing benefit cued by linguistic extremity markers. That is, our attitude preferences may lead people to be more cognitively engaged by messages whose linguistic markers match rather than mismatch their preferences. A broad literature on message matching effects (for a review, see Teeny et al., 2021) delineates conditions under which such matching of linguistic style to recipient attitudinal preference might produce directional bias effects versus processing effects. Relatedly, extreme (vs non extreme) attitude-holders have distinct linguistic characteristics (e.g., more confident, more "you" pronouns; Van

Swol et al., 2016); it would be interesting know if attitude preferences are responsive even to these more indirect cues of a source's attitude extremity.

We have generally focused on drawbacks of preference for neutrality, but future work could examine some positive aspects of this construct. In an increasingly polarized political environment (Gidron et al., 2019; Kubin & von Sikorski, 2021), the preference for neutrality may be helpful in moderating people's likelihood of polarizing. This may be particularly important when faced with meaning threats (e.g., mortality threat) because such threats motivate people to embrace their underlying ideologies (McGregor et al., 2010) as a means of buffering against these threats. This typically results in polarization (i.e., McGregor et al., 2013) but preference for neutrality might relate to depolarization under the same conditions (because for neutrality-preferring people, neutrality is their ideology). Furthermore, preference for neutrality might reduce the tendency for groups to split into polarized camps when debating (e.g., Bail et al., 2018) insofar as polarizing is antithetical to maintaining neutrality.

Indeed, the social consequences of preference for neutrality might be intriguing. Because neutrality preferring people are more ambivalent and less extreme, they might be able to moderate between polarized camps by representing a compromise position. Preference for neutrality may also moderate groupthink effects (Janis, 1972; Janis & Mann, 1977) by influencing the group's perception of information. A consistent neutrality motivation might help to reduce this tendency for groups to increasingly favor ideas without due criticism. Perhaps relatedly, neutrality preferring people might prefer neutral emotions (Gasper et al., 2019) and engage in mood regulation efforts (Larsen, 2000) to maintain emotional equanimity, which might affect group interactions by cooling off the hot emotions that commonly emerge in debates (Wollebæk et al., 2019).

5.2.2. Meta-attitudes

Ostrom (1989) argued that attitude theory and measurement are intertwined. In the present case, the fact that people vary meaningfully in their attitudes about their attitudes—and that such variance can be validly measured –signals something theoretically noteworthy about the nature of attitudes. Specifically, the present findings reveal an interesting form of attitudinal metacognition, in that preferences for extremity or neutrality entails evaluating one's evaluations (Petty, 2006; Petty et al., 2007). For instance, the preference for neutrality involves a sometimes strong liking of attitude positions that avoid liking and disliking. Furthermore, our findings reveal that people can be at least aware of what these metacognitive standards. Desiring an attitude which is other than what one currently holds is common (DeMarree & Rios, 2014), but research on these desires usually examines people's desires to like/dislike specific attitude objects (DeMarree et al., 2014, 2017; Vaughan-Johnston et al., 2023). In contrast, in the present work we highlight more global metacognitions about one's (and others') attitudes.

We propose that overarching preferences towards particular attitude positions (held across attitude objects in general) can be considered *meta-attitudes*, in that they characterize attitudes towards attitudinal positions and structures. For instance, preference for neutrality involves a very positive attitude towards neutral attitudes, and preference for extremity a very positive attitude towards extreme attitudes. Thus, these two constructs capture attitudes towards *attitude positions*. The interest in both sides is distinct because it suggests an attitude towards certain *attitude bases*; specifically, positive attitude towards attitudes formed from high information across multiple valences.

We think the notion of meta-attitudes is helpful to organize thinking about individual differences in the attitudes literature. Meta-attitudes have been studied at least indirectly. For

example, the need for affect (Maio & Esses, 2001) and need for cognition (Cacioppo et al., 1984; Petty et al., 2009) may be viewed as relatively direct measures of people's preferences for distinct types of *attitude content* (with important consequences for persuasion: Haddock et al., 2008; also see See et al., 2008, 2013). High (low) self-monitors are known to gravitate towards social-adjustive (utilitarian) information (Shavitt et al., 1992), and can thus be viewed as indirectly capturing individual difference variance in meta-attitudes for *attitude functions*. Furthermore, DeMarree and colleagues (2020) found that the desire to pursue high-certainty attitudes is stable across objects and scenarios, possibly suggesting individual differences in meta-attitudes concerning one type of *attitude strength*.

Studying meta-attitudes more explicitly, as in the present work, may be useful because it may highlight constructs that have not yet been studied. For example, some individuals tend to have more positive or negative attitudes towards objects in general (Hepler & Albarracín, 2013), but we are not aware of research concerning people's attitudes towards positive or negative attitudes. Yet such a construct would follow logically from a meta-attitudes perspective because attitude positivity and negativity are elements of attitude position, just like neutrality and extremity. People who prefer positivity/negativity might differ interestingly from those who just happen to have positive/negative attitudes (as examined by Hepler & Albarracín, 2013) in that only the former should strongly pursue new positive/negative information. Both constructs might relate to preferring other people with positive/negative attitudes, but for different reasons. For those who prefer positivity/negativity, this would reflect an evaluation of how well other people fulfill one's own standards; for those who simply have many positive/negative attitudes, it would reflect homophily (Ertug et al., 2022; McPherson et al., 2001).

One additional direction of note would involve combining meta-attitudinal properties. For example, an anonymous reviewer suggested that it would be useful to understand whether some people pursue neutrality or extremity for (e.g.) epistemic versus social reasons, or through more heuristic versus systematic elaboration. Our present preference for neutrality and extremity scales employ items that straddle attitude function dimensions, such as preference for neutrality items that reference value-based, social, epistemic, and utilitarian advantages of neutrality. However, future work that investigates how people prioritize and weight multiple of the above-noted dimensions (function, position, content, strength) could lead to a range of theoretical novelties. For example, does a preference for neutrality predicated in different functions, or combined with varying motivation to maximize attitude certainty, influence information processing differently? Relatedly, some people might be higher or lower in their commitment to generally neutral or extreme attitudes, just as people vary in their commitment to specific desired attitudes positions (DeMarree et al., 2017). Past research suggests that high versus low commitment to constructs can increase the predictive validity of those constructs (DeMarree et al., 2017; Petty et al., 1995; Shoots-Reinhard et al., 2014).

An exciting future research direction concerns the antecedents of preference for neutrality and meta-attitudes more broadly. We suspect that many variables may contribute to these beliefs arising within people. At the macroenvironmental level, journalists and other media personalities sometimes implicitly suggest that “the truth must lie in the middle” by giving equal speaking time even to epistemically unequal scientific sides (O’Connor & Weatherall, 2019; Oreskes & Conway, 2011), potentially encouraging a preference for neutrality. Microenvironmental factors like parenting styles also could contribute to attitude preferences. Past work has focused on parental influences on developing political reasoning and morality, often with a focus on roots of

political radicalization (Altemeyer, 1996; McClosky & Chong, 1985), so understanding how parenting may foster a comparatively strong enthusiasm for neutrality would be enlightening. Finally, past researchers have shown that attitude positions can be heritable (Olson et al., 2001); an intriguing future research direction would explore whether meta-attitudes are at least somewhat genetically mediated.

5.3. Limitations and Constraints on Generality

One possible objection to this set of findings is that we did not explicitly assess satisficing or disinterest in social topics, which could offer an alternative explanation for our results. For instance, some people might not really have cultivated attitudes towards most attitude objects, and might be described as holding inconsequential "nonattitudes" (see, e.g., Converse, 1974). These individuals might then default to neutral attitude responding as a form of satisficing, such as giving an uncommitted middle response as a default answer, potentially explaining the association of preference for neutrality with lower extremity and possibly higher ambivalence. The problem with this objection is that satisficing should not result in major consequences beyond selecting middlemost options. In other words, if some participants were simply satisficing they should not show any preference for an author or message that is arbitrarily labeled as "moderate" versus "extreme." Similarly, if people were simply holders of non-attitudes they should hold very weak and inconsequential attitudes and should be highly susceptible to persuasion regardless of message framing. Instead, we find that the characterization of others is especially consequential for those with a preference for neutrality (Study 1 and 3). We might also expect preference for neutrality to be negatively related to curiosity if they are simply disinterested in social topics, but instead Study 1 shows a near-zero association. This constellation of effects is therefore important because it suggests that these

meta-attitudes about global attitude positions are not simply a set of response patterns or the absence of an attitude. Indeed, we think that our meta-attitudinal constructs carry important implications for interpersonal dynamics and negotiation, information pursuit and processing, self-perception, and beyond.

Our samples were primarily composed of young adult women in Study 1 and 3 (Study 2 was only 58% women, with an average age of 38). Our conception of attitude preferences does not suggest any constraints on generalisability across these dimensions, and our data do not support any robust gender-based differences (and gender added as an interactive covariate was seldom significant and did not change any analyses meaningfully; see SOM-3). We did find that older participants in the only study with meaningful age variation (Study 2) were higher in interest in both sides, $r(99) = .26$, $p = .010$, and lower in the preference for extremity, $r(99) = -.24$, $p = .014$, but not different in preference for neutrality, $r(99) = .04$, $p = .691$. Future research, possibly with representative sampling, might better probe whether these demographic factors meaningfully change the psychological influences of our constructs.

Furthermore, our samples are overwhelmingly from Western cultures and our effects may require nuanced theoretical thinking to apply to non-Western populations. Preference for neutrality may initially sound like a philosophical orientation similar to dialectic or Zhongyong thinking, a common Confucian orientation to the world which places value on maintaining thoughts which are not extreme (Chiu, 2000; Spencer-Rodgers et al., 2009). However, an important difference is that "dialecticism predisposes people to more openly consider attitude-relevant information regardless of its valence" (Luttrell et al., 2022, p. 828), which if anything makes dialecticism more comparable to our interest in both sides construct. Given research suggesting that dialecticism is associated with a reduction of depressive symptoms and enhanced

subjective well-being for those who engage in this form of thinking (Yang et al., 2016), we look forward to future work examining how dialecticism is related to interest in both sides and preference for neutrality.

Another possible constraint is whether attitude preferences are equally applicable for controversial and non-controversial attitude objects. In principle, we think the attitudes preferences are relevant for both sorts of objects. In Study 1, although preference for neutrality was associated with less attitudinal extremity towards controversial objects and more political centrism, it was not related to extremity for non-controversial objects. In Study 2, preference for neutrality again related to attitude structures for a mostly controversial set of objects. However, in Study 3, attitude preferences related to attitudes towards lemphurs, which were not characterised as controversial. Thus, it is unclear if objects must be controversial for attitude preferences to capture variance in them. We interpret that the Study 1 discrepancy between controversial and non-controversial objects is attributable to a measurement difference—controversial objects by definition produce more variable attitudes, whereas non-controversial objects may have range restriction issues. Across the studies as a whole, it seems that attitude preferences have a broad utility that encompasses more controversial and less controversial issues.

5.4. Conclusion

Our present studies provide some interesting and often concerning observations about the preference for neutrality. For example, despite their elevated ambivalence about a range of topics, those who most strongly prefer neutrality appear low in intellectual humility; dislike other people with pronounced, non-neutral views, and see themselves as self-censoring to deal with disagreement; and reject information that is arbitrarily labeled as extreme versus moderate. This

constellation of findings represents, in our view, a potential hazard to the wellbeing of deliberative democracies, in which we might hope citizens will exchange opinions with a genuine curiosity about others' viewpoints, and engage with perspectives different from their own. At least in its strongest form, then, "motivated neutrality" can be concerning in many of the same ways that political extremity is concerning. We hope the present findings stimulate future research and discussion about how to best address the most problematic extremes of neutrality.

6. References

Abelson, R. P. (1995). Attitude extremity. In R. Petty & J. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 131–157). Erlbaum.

Altemeyer, B. (1996). *The authoritarian specter*. Harvard University Press.

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. F., ... & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences, 115*(37), 9216–9221.

Balietti, S., Getoor, L., Goldstein, D. G., & Watts, D. J. (2021). Reducing opinion polarization: Effects of exposure to similar people with differing political views. *Proceedings of the National Academy of Sciences*, *118*(52), e2112552118.

Binder, A. R., Dalrymple, K. E., Brossard, D., & Scheufele, D. A. (2009). The soul of a polarized democracy: Testing theoretical linkages between talk and attitude extremity during the 2004 presidential election. *Communication Research, 36*(3), 315–340.

Blankenship, K. L., & Craig, T. Y. (2011). Language use and persuasion: Multiple roles for linguistic styles. *Social and Personality Psychology Compass, 5*(4), 194–205.

Brambilla, M., Sacchi, S., Rusconi, P., Cherubini, P., & Yzerbyt, V. Y. (2012). You want to give a good impression? Be honest! Moral traits dominate group impression formation. *British Journal of Social Psychology, 51*(1), 149–166.

Brambilla, M., Sacchi, S., Rusconi, P., & Goodwin, G. P. (2021). Chapter Four - The primacy of morality in impression development: Theory, research, and future directions. In B. Gawronski (Ed.), *Advances in Experimental Social Psychology* (Vol. 64, pp. 187–262). Academic Press.

Briñol, P., & Petty, R. E. (2015). Elaboration and validation processes: Implications for media attitude change. *Media Psychology*, *18*(3), 267–291.

Brüggemann, M., & Engesser, S. (2017). Beyond false balance: How interpretive journalism shapes media coverage of climate change. *Global Environmental Change, 42*, 58–67.

Buder, J., Rabl, L., Feiks, M., Badermann, M., & Zurstiege, G. (2021). Does negatively toned language use on social media lead to attitude polarization?. *Computers in Human Behavior*, *116*, 106663.

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin, 115*(3), 401.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review, 1*, 3–25.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*(1), 116.

Clark, C. J., Liu, B. S., Winegard, B. M., & Ditto, P. H. (2019). Tribalism is human nature. *Current Directions in Psychological Science, 28*(6), 587–592.

Clarkson, J. J., Tormala, Z. L., & Leone, C. (2011). A self-validation perspective on the mere thought effect. *Journal of Experimental Social Psychology, 47*(2), 449–454.

Cohen, G. L. (2003). Party over policy: The dominating impact of group influence on political beliefs. *Journal of Personality and Social Psychology*, *85*(5), 808.

Conner, M., Sparks, P., Povey, R., James, R., Shepherd, R., & Armitage, C. J. (2002). Moderator effects of attitudinal ambivalence on attitude–behaviour relationships. *European Journal of Social Psychology*, *32*(5), 705–718.

Converse, P. E. (1974) Nonattitudes and American Public Opinion: Comment: The Status of Nonattitudes, *The American Political Science Review, 68*(2), pp. 650–660.

Costarelli, S., & Colloca, P. (2004). The effects of attitudinal ambivalence on pro-environmental behavioural intentions. *Journal of Environmental Psychology*, *24*(3), 279–288.

Craig, T. Y., & Blankenship, K. L. (2011). Language and persuasion: Linguistic extremity influences message processing and behavioral intentions. *Journal of Language and Social Psychology, 30*(3), 290–310.

Crites Jr, S. L., Fabrigar, L. R., & Petty, R. E. (1994). Measuring the affective and cognitive properties of attitudes: Conceptual and methodological issues. *Personality and Social Psychology bulletin, 20*(6), 619–634.

Dawson, E., Gilovich, T., & Regan, D. T. (2002). Motivated reasoning and performance on the Wason Selection Task. *Personality and Social Psychology Bulletin, 28*, 1379–1387.

DeMarree, K. G., Clark, C. J., Wheeler, S. C., Briñol, P., & Petty, R. E. (2017). On the pursuit of desired attitudes: Wanting a different attitude affects information processing and behavior. *Journal of Experimental Social Psychology*, *70*, 129–142.

DeMarree, K. G., Petty, R. E., Briñol, P., & Xia, J. (2020). Documenting individual differences in the propensity to hold attitudes with certainty. *Journal of Personality and Social Psychology*, *119*(6), 1239.

DeMarree, K. G., & Rios, K. (2014). Understanding the relationship between self-esteem and self-clarity: The role of desired self-esteem. *Journal of Experimental Social Psychology*, *50*, 202–209.

DeMarree, K. G., Wheeler, S. C., Briñol, P., & Petty, R. E. (2014). Wanting other attitudes: Actual–desired attitude discrepancies predict feelings of ambivalence and ambivalence consequences. *Journal of Experimental Social Psychology*, *53*, 5–18.

Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology, 63,* 568–584.

Dixon, G. N., & Clarke, C. E. (2013). Heightening uncertainty around certain science: Media coverage, false balance, and the autism-vaccine controversy. Science Communication, *35*(3), 358–382.

Drolet, A., Luce, M. F., Jiang, L., Rossi, B. C., & Hastie, R. (2021). The preference for moderation scale. *Journal of Consumer Research, 47*(6), 831–854.

Edwards, J. D., & Ostrom, T. M. (1971). Cognitive structure of neutral attitudes. *Journal of Experimental Social Psychology*, *7*(1), 36–47.

Edwards, K., & Smith, E. E. (1996). A disconfirmation bias in the evaluation of arguments. *Journal of Personality and Social Psychology, 71,* 5–24.

Ertug, G., Brennecke, J., Kovács, B., & Zou, T. (2022). What does homophily do? A review of the consequences of homophily. *Academy of Management Annals*, *16*(1), 38–69.

Fabrigar, L. R., & Petty, R. E. (1999). The role of the affective and cognitive bases of attitudes in susceptibility to affectively and cognitively based persuasion. *Personality and Social Psychology Bulletin, 25*(3), 363–381.

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression, Third edition*. Sage.

Gasper, K., Danube, C. L., & Hu, D. (2021). Making room for neutral affect: Evidence indicating that neutral affect is independent of and co-occurs with eight affective states. *Motivation and Emotion*, *45*, 103–121.

Gasper, K., Spencer, L. A., & Hu, D. (2019). Does neutral affect exist? How challenging three beliefs about neutral affect can advance affective research. *Frontiers in Psychology*, *10*, 2476.

Gidron, N., Adams, J., & Horne, W. (2019). Toward a comparative research agenda on affective polarization in mass publics. *APSA Comparative Politics Newsletter, 29*, 30–36.

Goldenberg, A., Abruzzo, J. M., Huang, Z., Schöne, J., Bailey, D., Willer, R., ... & Gross, J. J. (2023). Homophily and acrophily as drivers of political segregation. *Nature Human Behaviour*, *7*(2), 219–230.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*(5), 1029.

Gunther, A. (1988). Attitude extremity and trust in media. *Journalism Quarterly, 65*(2), 279–287.

Guyer, J. J., Fabrigar, L. R., Vaughan-Johnston, T. I., & Tang, C. (2018). The counterintuitive influence of vocal affect on the efficacy of affectively-based persuasive messages. *Journal of Experimental Social Psychology, 74*, 161–173.

Haddock, G., Foad, C., Windsor-Shellard, B., Dummel, S., & Adarves-Yorno, I. (2017). On the attitudinal consequences of being mindful: Links between mindfulness and attitudinal ambivalence. *Personality and Social Psychology Bulletin*, *43*(4), 439–452.

1217 Haddock, G., Maio, G. R., Arnold, K., & Huskinson, T. (2008). Should persuasion be affective
1218 or cognitive? The moderating effects of need for affect and need for cognition.
1219 *Personality and Social Psychology Bulletin, 34*(6), 769–778.
1220 Han, R., Proulx, T., van Harreveld, F., & Haddock, G. (2023). How people perceive
1221 dispositionally (non-) ambivalent others and why it matters. *Journal of Experimental*
1222 *Social Psychology*, *109*, 104518.
1223 Hanson, D. J. (1973). Dogmatism and attitude extremity. *The Journal of Social Psychology,*
1224 *89*(1), 155–156.
1225 Hayes, A. F., Glynn, C. J., & Shanahan, J. (2005). Willingness to self-censor: A construct and
1226 measurement tool for public opinion research. *International Journal of Public Opinion*
1227 *Research, 17*(3), 298–323.
1228 Hepler, J., & Albarracín, D. (2013). Attitudes without objects: Evidence for a dispositional
1229 attitude, its measurement, and its consequences. *Journal of Personality and Social*
1230 *Psychology, 104*(6), 1060.
1231 Higgins, E. T., Bond, R. N., Klein, R., & Strauman, T. (1986). Self-discrepancies and emotional
1232 vulnerability: how magnitude, accessibility, and type of discrepancy influence
1233 affect. *Journal of Personality and Social Psychology*, *51*(1), 5.
1234 Hu, D., & Gasper, K. (2022). Examining the link between neutral and ambivalent attitudes: Their
1235 association and their co-occurrence. *Social Cognition*, *40*(1), 1–28.
1236 Janis, I. L. (1972). *Victims of groupthink: A psychological study of policy decisions and fiascos.*
1237 Houghton Mifflin.
1238 Janis, I. L., & Mann, L. (1977). *Decision making: A psychological analysis of conflict, choice,*
1239 *and commitment*. Macmillan.

Jarvis, W. B. G., & Petty, R. E. (1996). The need to evaluate. *Journal of Personality and Social Psychology, 70*(1), 172.

Johns, R., & Kölln, A. K. (2020). Moderation and competence: How a party's ideological position shapes its valence reputation. *American Journal of Political Science*, *64*(3), 649–663.

Judd, C. M., & Johnson, J. T. (1981). Attitudes, polarization, and diagnosticity: Exploring the effect of affect. *Journal of Personality and Social Psychology, 41*(1), 26–36.

Kashdan, T. B., Stiksma, M. C., Disabato, D. J., McKnight, P. E., Bekier, J., Kaji, J., & Lazarus, R. (2018). The five-dimensional curiosity scale: Capturing the bandwidth of curiosity and identifying four unique subgroups of curious people. *Journal of Research in Personality, 73*, 130–149.

Knobloch-Westerwick, S., Mothes, C., & Polavin, N. (2020). Confirmation bias, ingroup bias, and negativity bias in selective exposure to political information. *Communication Research, 47*(1), 104–124.

Koehler, D. J. (2016). Can journalistic "false balance" distort public perception of consensus in expert opinion*?*. *Journal of Experimental Psychology: Applied, 22*(1), 24.

Krosnick, J. A., & Petty, R. E. (1995). Attitude strength: An overview. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 1–24). Lawrence Erlbaum Associates, Inc.

Krumrei-Mancuso, E. J., & Rouse, S. V. (2016). The development and validation of the Comprehensive Intellectual Humility Scale. *Journal of Personality Assessment, 98*, 209–221.

Kubin, E., & von Sikorski, C. (2021). The role of (social) media in political polarization: a systematic review. *Annals of the International Communication Association, 45*(3), 188–206.

Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psychology, 53,* 636–647.

Lammers, J., Koch, A., Conway, P., & Brandt, M. J. (2017). The political domain appears simpler to the politically extreme than to political moderates. *Social Psychological and Personality Science, 8*(6), 612–622.

Larsen, J. T., Hemenover, S. H., Norris, C. J., & Cacioppo, J. T. (2003). Turning adversity to advantage: On the virtues of the coactivation of positive and negative emotions. In L. G. Aspinwall & U. M. Staudinger (Eds.), *A psychology of human strengths: Fundamental questions and future directions for a positive psychology* (pp. 211–225). American Psychological Association. https://doi.org/10.1037/10566-015

Larsen, R. J. (2000). Toward a science of mood regulation. *Psychological Inquiry, 11*(3), 129–141.

Lennox, R. D., & Wolfe, R. N. (1984). Revision of the self-monitoring scale. *Journal of Personality and Social Psychology, 46*(6), 1349–1364.

Leone, C. (1989). Self-generated attitude change: Some effects of thought and dogmatism on attitude polarization. *Personality and Individual Differences*, *10*(12), 1243–1252.

Leone, C. (1996). Thought, objectivism, and opinion extremity: Individual differences in attitude polarization and attenuation. *Personality and Individual Differences*, *21*(3), 383–390.

1283 Luttrell, A., & Petty, R. E. (2021). Evaluations of self-focused versus other-focused arguments
1284 for social distancing: An extension of moral matching effects. *Social Psychological and*
1285 *Personality Science*, *12*(6), 946–954.
1286 Luttrell, A., Petty, R. E., Chang, J. H., & Togans, L. J. (2022). The role of dialecticism in
1287 objective and subjective attitudinal ambivalence. *British Journal of Social Psychology*,
1288 *61*(3), 826–841.
1289 Maio, G. R., & Esses, V. M. (2001). The need for affect: Individual differences in the motivation
1290 to approach or avoid emotions. *Journal of Personality, 69*(4), 583–614.
1291 McClosky, H., & Chong, D. (1985). Similarities and differences between left-wing and right-
1292 wing radicals. *British Journal of Political Science, 15*(3), 329-363.
1293 McGregor, I., Nash, K., Mann, N., & Phills, C. E. (2010). Anxious uncertainty and reactive
1294 approach motivation (RAM). *Journal of Personality and Social Psychology, 99*,133–147.
1295 McGregor, I., Prentice, M., & Nash, K. (2013). Anxiousuncertainty and reactive approach
1296 motivation (RAM) for religious, idealistic, and lifestyle extremes. *Journal of Social*
1297 *Issues, 69*, 537–563
1298 McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social
1299 networks. *Annual Review of Sociology*, *27*(1), 415–444.
1300 Miller, A. G., McHoskey, J. W., Bane, C. M., & Dowd, T. G. (1993). The attitude polarization
1301 phenomenon: Role of response measure, attitude extremity, and behavioral consequences
1302 of reported attitude change. *Journal of Personality and Social Psychology, 64*(4), 561–
1303 574.
1304 O'Connor, C., & Weatherall, J. O. (2019). *The misinformation age: How false beliefs spread*.
1305 Yale University Press.

O'Dea, C. J., Bueno, A. M. C., & Saucier, D. A. (2018). Social vigilantism and the extremity, superiority, and defense of attitudes toward climate change. *Personality and Individual Differences, 130*, 83–91.

Olson, J. M., Vernon, P. A., Harris, J. A., & Jang, K. L. (2001). The heritability of attitudes: a study of twins. *Journal of Personality and Social Psychology, 80*(6), 845–860.

Oreskes, N., & Conway, E. M. (2011). *Merchants of doubt: How a handful of scientists obscured the truth on issues from tobacco smoke to global warming*. Bloomsbury Publishing USA.

Ostrom, T. M. (1989). Interdependence of attitude theory and measurement. In A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *Attitude structure and function* (pp. 11–36). Lawrence Erlbaum.

Petty, R. E. (2006). A metacognitive model of attitudes. *Journal of Consumer Research*, *33*(1), 22–24.

Petty, R. E., Briñol, P., & DeMarree, K. G. (2007). The Meta–Cognitive Model (MCM) of attitudes: Implications for attitude measurement, change, and strength. *Social Cognition, 25*(5), 657–686.

Petty, R. E., Brinol, P., Loersch, C., & McCaslin, M. J. (2009). The need for cognition. In M.R. Leary & R.H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 318–329). Guilford.

Petty, R. E., Brinol, P., & Tormala, Z. L. (2002). Thought confidence as a determinant of persuasion: the self-validation hypothesis. *Journal of Personality and Social Psychology*, *82*(5), 722.

Petty, R. E., Haugtvedt, C. P., & Smith, S. M. (1995). Elaboration as a determinant of attitude strength: Creating attitudes that are persistent, resistant, and predictive of behavior. In R.

E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 93–130). Lawrence Erlbaum Associates, Inc.

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.

Pillaud, V., Cavazza, N., & Butera, F. (2013). The social value of being ambivalent: Self-presentational concerns in the expression of attitudinal ambivalence. *Personality and Social Psychology Bulletin*, *39*(9), 1139–1151.

Pillaud, V., Cavazza, N., & Butera, F. (2018). The social utility of ambivalence: Being ambivalent on controversial issues is recognized as competence. *Frontiers in Psychology*, *9*, 961.

Pittinsky, T. L., Rosenthal, S. A., & Montoya, R. M. (2011). Liking is not the opposite of disliking: The functional separability of positive and negative attitudes toward minority groups. *Cultural Diversity and Ethnic Minority Psychology, 17*(2), 134.

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology, 17*(3), 715–734.

Pretus, C., Servin-Barthet, C., Harris, E. A., Brady, W. J., Vilarroya, O., & Van Bavel, J. J. (2023). The role of political devotion in sharing partisan misinformation and resistance to fact-checking. *Journal of Experimental Psychology: General, 152*(11), 3116–3134. https://doi.org/10.1037/xge0001436

Price, E., Ottati, V., Wilson, C., & Kim, S. (2015). Open-minded cognition. *Personality and Social Psychology Bulletin, 41*(11), 1488–1504.

Priester, J. R., & Petty, R. E. (1996). The gradual threshold model of ambivalence: relating the positive and negative bases of attitudes to subjective ambivalence. *Journal of Personality and Social Psychology, 71*(3), 431.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reich, T., & Wheeler, S. C. (2016). The good and bad of ambivalence: Desiring ambivalence under outcome uncertainty. *Journal of Personality and Social Psychology*, *110*(4), 493.

Richard, F. D., Bond Jr, C. F., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, *7*(4), 331-363.

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.

Russell, J. A., Lewicka, M., & Niit, T. (1989). A cross-cultural study of a circumplex model of affect. *Journal of Personality and Social Psychology, 57*(5), 848.

See, Y. H. M., Petty, R. E., & Fabrigar, L. R. (2008). Affective and cognitive meta-bases of attitudes: Unique effects on information interest and persuasion. *Journal of personality and social psychology*, *94*(6), 938.

See, Y. H. M., Petty, R. E., & Fabrigar, L. R. (2013). Affective–cognitive meta-bases versus structural bases of attitudes predict processing interest versus efficiency. *Personality and Social Psychology Bulletin, 39*(8), 1111–1123.

Shavitt, S., Lowrey, T. M., & Han, S. P. (1992). Attitude functions in advertising: The interactive role of products and self-monitoring. *Journal of Consumer Psychology*, *1*(4), 337–364.

Shoots-Reinhard, B. L., Rucker, D. D., Petty, R. E., & Shakarchi, R. (2014). Not all contrast effects are created equal: extent of processing affects contrast strength. *Journal of Applied Social Psychology*, *44*(8), 523–535.

Siev, J. J., Philipp-Muller, A., Durso, G. R., & Wegener, D. T. (2024). Endorsing both sides, pleasing neither: Ambivalent individuals face unexpected social costs in political conflicts. *Journal of Experimental Social Psychology*, *114*, 104631.

Silver, I., & Shaw, A. (2022). When and why "staying out of it" backfires in moral and political disagreements. *Journal of Experimental Psychology: General, 151*(10), 2542–2561. https://doi.org/10.1037/xge0001201

Simonson, I., Sela, A., & Sood, S. (2017). Preference-construction habits: The case of extremeness aversion. *Journal of the Association for Consumer Research*, *2*(3), 322–332.

Skitka, L. J., Bauman, C. W., & Sargis, E. G. (2005). Moral conviction: Another contributor to attitude strength or something more?. *Journal of Personality and Social Psychology*, *88*(6), 895–917.

Sowell, T. (2002). *A conflict of visions: The ideological origins of political struggles*. Basic Books.

Spencer-Rodgers, J., Boucher, H. C., Peng, K., & Wang, L. (2009). Cultural differences in self-verification: The role of naïve dialecticism. *Journal of Experimental Social Psychology, 45*(4), 860–866.

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology, 89*(2), 342.

Stanovich, K. E., & Toplak, M. E. (2023). Actively open-minded thinking and its measurement. *Journal of Intelligence, 11*(2), 27.

Suedfeld, P., & Tetlock, P. (1977). Integrative complexity of communications in international crises. *Journal of Conflict Resolution*, *21*(1), 169–184.

Tannenbaum, P. H. (1956). Initial attitude toward source and concept as factors in attitude change through communication. *Public Opinion Quarterly*, *20*(2), 413–425.

Teeny, J. D., Siev, J. J., Briñol, P., & Petty, R. E. (2021). A review and conceptual framework for understanding personalized matching effects in persuasion. *Journal of Consumer Psychology*, *31*(2), 382–414.

Tesser, A., & Conlee, M. C. (1975). Some effects of time and thought on attitude polarization. *Journal of Personality and Social Psychology, 31*(2), 262.

Thompson, M.M., Zanna, M.P., & Griffin, D.W. (1995). Let's not be indifferent about (attitudinal) ambivalence. In R.E. Petty, & J.A. Krosnick (Eds.), *Attitude Strength: Antecedents and Consequences* (pp. 361–386). Erlbaum.

Toner, K., Leary, M. R., Asher, M. W., & Jongman-Sereno, K. P. (2013). Feeling superior is a bipartisan issue: Extremity (not direction) of political views predicts perceived belief superiority. *Psychological Science*, *24*(12), 2454–2462.

Toribio-Flórez, D., Van Harreveld, F., & Schneider, I. K. (2020). Ambivalence and interpersonal liking: The expression of ambivalence as social validation of attitudinal conflict. *Frontiers in Psychology*, *11*, 525301.

Van Boven, L., Judd, C. M., & Sherman, D. K. (2012). Political polarization projection: social projection of partisan attitude extremity and attitudinal processes. *Journal of Personality and Social Psychology*, *103*(1), 84.

1419 Van der Pligt, J., Ester, P., & Van der Linden, J. (1983). Attitude extremity, consensus and
1420 diagnosticity. *European Journal of Social Psychology, 13*(4), 437–439.
1421 Van Harreveld, F., Van der Pligt, J., & De Liver, Y. N. (2009). The agony of ambivalence and
1422 ways to resolve it: Introducing the MAID model. *Personality and Social Psychology*
1423 *Review*, *13*(1), 45–61.
1424 Van Swol, L. M., Prahl, A., Kolb, M. R., Lewis, E. A., & Carlson, C. (2016). The language of
1425 extremity: The language of extreme members and how the presence of extremity affects
1426 group discussion. *Journal of Language and Social Psychology, 35*(6), 603–627.
1427 Vaughan-Johnston, T. I., Fabrigar, L. R., Xia, J., DeMarree, K. G., & Clark, J. K. (2023).
1428 Desired attitudes guide actual attitude change. *Journal of Experimental Social*
1429 *Psychology*, *105*, 104437.
1430 Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive
1431 closure. *Journal of Personality and Social Psychology, 67*(6), 1049.
1432 Westfall, J., Van Boven, L., Chambers, J. R., & Judd, C. M. (2015). Perceiving political
1433 polarization in the United States: Party identity strength and attitude extremity exacerbate
1434 the perceived partisan divide. *Perspectives on Psychological Science*, *10*(2), 145–158.
1435 White, K. R., Kinney, D., Danek, R. H., Smith, B., & Harben, C. (2020). The Resistance to
1436 Change-Beliefs Scale: Validation of a new measure of conservative ideology. *Personality*
1437 *and Social Psychology Bulletin, 46*(1), 20–35.
1438 Wojciszke, B., Bazinska, R., & Jaworski, M. (1998). On the dominance of moral categories in
1439 impression formation. *Personality and Social Psychology Bulletin*, *24*(12), 1251–1263.

Wollebæk, D., Karlsen, R., Steen-Johnsen, K., & Enjolras, B. (2019). Anger, fear, and echo chambers: The emotional basis for online behavior. *Social Media+ Society, 5*(2), 2056305119829859.

Xu, M., & Petty, R. E. (2022). Two-sided messages promote openness for a variety of deeply entrenched attitudes. *Personality and Social Psychology Bulletin*, 01461672221128113.

Zmigrod, L., Rentfrow, P. J., & Robbins, T. W. (2020). The partisan mind: Is extreme political partisanship related to cognitive inflexibility? *Journal of Experimental Psychology: General, 149*(3), 407–418.