

UC San Diego

UC San Diego Previously Published Works

Title

How Word Reversibility Impacts Judgment Confidence

Permalink

<https://escholarship.org/uc/item/4sd0m0nw>

Journal

Journal of the Association for Consumer Research

ISSN

2378-1815

Authors

Maimone, Giulia

Karmarkar, Uma R

Amir, On

Publication Date

2025-12-15

DOI

10.1086/740066

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NoDerivatives License, available at <https://creativecommons.org/licenses/by-nd/4.0/>

Peer reviewed

How Word Reversibility Impacts Judgment Confidence

Giulia Maimone, Uma R. Karmarkar, On Amir

Abstract:

From interpersonal conversations to commercial and political messaging, our world revolves around people making sense of communications. In the present research, we illustrate how specific lexical choices systematically shape recipients' confidence in their judgment of a message's truthfulness. Prior work has shown that words can differ in their "reversibility," that is, how easily their antonyms can be retrieved. We propose a novel theoretical framework that predicts when and how statements containing concepts with different levels of reversibility engage distinct psychological processes, which in turn differentially affect downstream judgments of confidence. We test this model across two experiments (N = 1,067) using truthfulness judgments arising from both participants' existing beliefs and experimentally manipulated beliefs.

Author Note: We thank Victor Ferreira, Rachel Gershon, and Thomas Urbach for helpful comments, and are grateful to Samantha Tieger and Project Kesher for their collaboration.

Open Access: All data, analyses code, and material for all studies are available at the following link: https://osf.io/kybhx/?view_only=b92c8b5f3e8c42ff8dbf8ef959174cec.

All experiments are preregistered on AsPredicted (Study 1: <https://aspredicted.org/q4jr-n8dz.pdf>, Study 2: <https://aspredicted.org/n23b-qttt.pdf>, Facebook Study: <https://aspredicted.org/4v23-q76b.pdf>). Note that all measures, manipulations, and exclusions of all experiments are disclosed. For all experiments, data collection was completed before any data analysis was carried out.

The contemporary media environment is crowded with persuasive messaging, from advertising to political communications. As a consequence, when consumers encounter an assertion from a firm or individual, they immediately evaluate whether they believe that statement to be true or false. However, these veracity judgments are not strictly binary. Instead, they are held with varying degrees of confidence, which in turn affect how likely people are to be persuaded and to act on the information in the future (e.g., Fazio and Zanna, 1978). In this research, we focus on how the lexical choices made in crafting a message interact with consumers' judgments about the statement's truthfulness. The combination of these factors can activate one of two distinct cognitive processes, and we show how this leads to differential judgment confidence.

Researchers and practitioners have illustrated a myriad of ways in which changing a single word in a message can impact attitudes and consequent behavior (e.g., Hussein & Tormala, 2024; Packard & Berger, 2021; Packard et al., 2018; see Kronrod, 2022 for review). Prior findings at the intersection of marketing, decision-making, and linguistics have largely focused on how this arises from consumers' sensitivity to semantic distinctions (see Berger & Packard, 2023 for review). In this paper, we extend this literature by integrating additional elements of linguistic theory, such as the polarity, or "reversibility," of key words, with consumer judgments.

To illustrate this, imagine a skin-care company selling a new sunblock. These products often have a characteristic scent. The firm wishes to highlight that their product has benefits for sensory-sensitive customers, using one of the following statements:

Statement 1: "The scent is not prominent."

Statement 2: "The scent is not intense."

While these negation-type ("not") statements are intended to express the same sentiment, we

argue they may have differential impact. Beyond any potential semantic differences, the words “prominent” and “intense” are qualitatively different in terms of reversibility, discussed in cognitive linguistics as *concept polarity* (e.g., Mayo, Schul, & Burnstein, 2004). “Prominent” is uni-polar, meaning it is difficult to reverse, or retrieve a word that expresses its opposite. As a result, people tend to negate it by adding the tag “not” to the original concept (i.e., “not prominent”). In contrast, “intense” is bi-polar, meaning that it is easy to retrieve a word that expresses its opposite, and people tend to negate it by using its antonym (i.e., “mild”). We show that if individuals believe the message is true, the bi-polar word engages a slower and more elaborate cognitive process that results in lower confidence in their judgment of truthfulness.

THEORETICAL BACKGROUND

Judgment Confidence and the Importance of Negation Language

Confidence has been defined as the subjective sense of conviction about one’s attitude or opinion (Festinger, 1950, 1954). Similarly, *judgment confidence* refers to the sense of conviction one has about a judgment. High levels of judgment confidence increase the judgment’s impact on a number of dimensions, from resistance to attitude or opinion change (e.g., Bassili, 1996; Tormala & Petty, 2002, 2004), to the likelihood to translate attitudes into behavior, that is, to act on the message (e.g., Fazio & Zanna, 1978; Fishbein & Ajzen, 1975; Tormala & Rucker, 2007). Notably, high confidence in one’s own judgments about a message can also increase the extent to which that message is persuasive (Petty, Briñol, & Tormala, 2002).

Judgment confidence is particularly relevant for advertising and other consumer

messaging, since evoking highly confident judgments can influence consumers' likelihood of purchasing a product or otherwise engaging with the firm. As a result, understanding how to craft a message that elicits high levels of judgment confidence has the potential for meaningful marketing benefits. Prior research related to this topic suggests that "negation" language increases engagement with advertising messages overall (Pezzuti & Leonhardt, 2023). However, the impact of negations can depend on the cognitive resources required for the specific form of language processing (e.g., Grant et al. 2004), underscoring the importance of incorporating further cognitive linguistic theory. To address this, our framework integrates consumer judgments with the psychology of negation processing, yielding a more sophisticated understanding of how negations can be used in ways that relate to judgment confidence.

Concept Polarity and Cognitive Processing

We first examine concept polarity as an important determinant of cognitive processes engaged by various types of messages. As described, concept polarity refers to the degree to which a concept is "reversible" (Mayo, Schul, & Burnstein, 2004). A concept is uni-polar if people find it difficult to retrieve its antonym and describe its opposite by negating the original concept instead (e.g., "prominent" and "not prominent"). This classification holds even if an antonym technically exists. From a cognitive perspective, concept polarity depends on the (in)accessibility of the antonym at the time of processing. In contrast, a concept is bi-polar if it is relatively easy for people to retrieve and use its antonym (e.g., "intense" and "mild").

Negations of uni-polar concepts engage an encoding process described as *Schema-Plus-Tag (SPT)*; Mayo, Schul, & Burnstein, 2004; consistent with Carpenter & Just, 1975; Gilbert,

1991; Wang et al., 2021; see also Kaup & Dudschig, 2020). For example, reading “The scent is not prominent” first triggers processing the statement’s affirmative form “The scent is prominent” (the Schema), and then encoding it with ‘not’ (the Tag) afterwards. In contrast, negations of bi-polar concepts engage a *Fusion* process (Mayo, Schul, & Bernstein, 2004; consistent with Kaup, Lüdtke, & Zwaan, 2006; Kaup & Zwaan, 2003). In this case, reading the easily reversible statement “The scent is not intense” triggers retrieval of the antonym and substitution in the statement, so that it is cohesively encoded as “The scent is mild”.

For negation statements, SPT promotes storing information as a negated concept, while Fusion promotes storing information as the opposite (affirmed) concept. We can then consider other types of statements that might occur in a persuasive message. All affirmations, like “The scent is intense/prominent,” are processed through SPT, regardless of concept polarity, because people can process the Schema directly. Explicit double negations are also processed via SPT regardless of concept polarity because a double negation (e.g., “not not intense”) becomes an affirmation, again leaving the concept to be processed as a Schema .

PROPOSED FRAMEWORK AND HYPOTHESES

Prior research has argued that all bi-polar negation claims engage the Fusion process (e.g., Mayo, Schul, & Bernstein, 2004). However, a key insight rooted in the belief formation literature is that people are not merely passive listeners. Indeed, consumers often hold relevant priors about the claim at hand and use them to judge whether messages are truthful and form beliefs. Existing work has shown that consumers often treat advertising claims as hypotheses to be tested against their own knowledge or expectations (Hoch & Ha, 1986). Returning to our sunblock example, when a shopper encounters the claim “The scent is not

prominent,” they are unlikely to passively accept it. Instead, they mentally simulate whether the claim fits with their prior experiences (e.g., whether the sunblock’s scent was prominent—or not—in the past) or with their broader understanding of how sunblocks are typically scented. In this sense, consumers actively engage in an implicit truth-testing process.

Thus, here we propose that a more complex framework is needed because veracity (true vs. false) judgments are fundamental to how consumers encode, represent, and respond to messages. As shown in Table 1, we propose that these veracity judgments interact with whether the claim is an affirmation or a negation to determine whether the message recipient will ultimately hold a belief that is an affirmation or a negation.

		Message (Claim)	
		Affirmation	Negation
Veracity Judgment	True	<i>Affirmation</i>	<i>Negation</i>
	False	<i>Negation</i>	<i>Affirmation</i>

Table 1. Matrix of Factors Determining Beliefs

To illustrate this directly, our framework proposes that if a claim is affirmative and judged as true, or negative and judged as false, it will result in an *Affirmation belief*:

- “The scent is prominent,” TRUE → *The scent is prominent.*
- “The scent is not prominent,” FALSE → *The scent is prominent.*

In contrast, if the claim is affirmative and judged as false or negative and judged as true, it will result in a *Negation belief*:

- “The scent is prominent,” FALSE → *The scent is not prominent*.
- “The scent is not prominent,” TRUE → *The scent is not prominent*.

By incorporating consumers’ veracity judgments, we shift the process determinant from negation claims to negation *beliefs*. While prior work suggests that all bi-polar negation claims lead to Fusion (e.g., Mayo et al., 2004), our framework proposes that a bi-polar negation claim is neither necessary nor sufficient to engage that process. More specifically, we propose the process engaged by a bi-polar negation claim depends on whether the individual’s *belief* is an affirmation or a negation. Bi-polar *Negation beliefs* engage Fusion, and bi-polar *Affirmation beliefs* engage SPT. Table 2 illustrates the full set of relationships for uni-polar and bi-polar claims as predicted by our framework. Notably, and consistent with prior theory, all uni-polar words engage SPT regardless of whether they are incorporated into affirmation or negation beliefs.

POLARITY	CLAIM	JUDGMENT	BELIEF	ENCODING PROCESS
UNI-POLAR Prominent	AFFIRMATION The scent is prominent	TRUE	<i>AFFIRMATION</i> <i>The scent is prominent</i>	SPT The scent is prominent
UNI-POLAR Prominent	AFFIRMATION The scent is prominent	FALSE	<i>NEGATION</i> <i>The scent is not prominent</i>	SPT The scent is prominent + not
UNI-POLAR Prominent	NEGATION The scent is not prominent	TRUE	<i>NEGATION</i> <i>The scent is not prominent</i>	SPT The scent is prominent + not
UNI-POLAR Prominent	NEGATION The scent is not prominent	FALSE	<i>AFFIRMATION</i> <i>The scent is prominent</i>	SPT The scent is prominent
BI-POLAR Intense	AFFIRMATION The scent is intense	TRUE	<i>AFFIRMATION</i> <i>The scent is intense</i>	SPT The scent is intense
BI-POLAR Intense	AFFIRMATION The scent is intense	FALSE	<i>NEGATION</i> <i>The scent is not intense</i>	FUSION The scent is mild
BI-POLAR Intense	NEGATION The scent is not intense	TRUE	<i>NEGATION</i> <i>The scent is not intense</i>	FUSION The scent is mild
BI-POLAR Intense	NEGATION The scent is not intense	FALSE	<i>AFFIRMATION</i> <i>The scent is intense</i>	SPT The scent is intense

Table 2. All combinations of Polarity, Claim Type, and Judgment to determine Process Type with practical examples. Belief Type is *italicized* to represent the interaction of Claim Type and Judgment, illustrated in Table 1.

By defining this framework, we are able to make novel predictions about message recipients' downstream judgment confidence about the veracity of the message. In contrast to SPT, Fusion requires the additional cognitive function of actively retrieving a word's antonym and substituting it into the schema being encoded. Thus, we argue that Fusion involves more elaborate and resource-intensive processing than SPT, which likely results in slower processing (Gough, 1965). Since longer response times have been associated with lower judgment confidence in problem-solving tasks (e.g., Ackerman & Zalmanov, 2012), we predict that statements engaging the more elaborate Fusion process will lead message recipients to hold lower confidence in their judgments about the truthfulness of a statement

(see Figure 1).

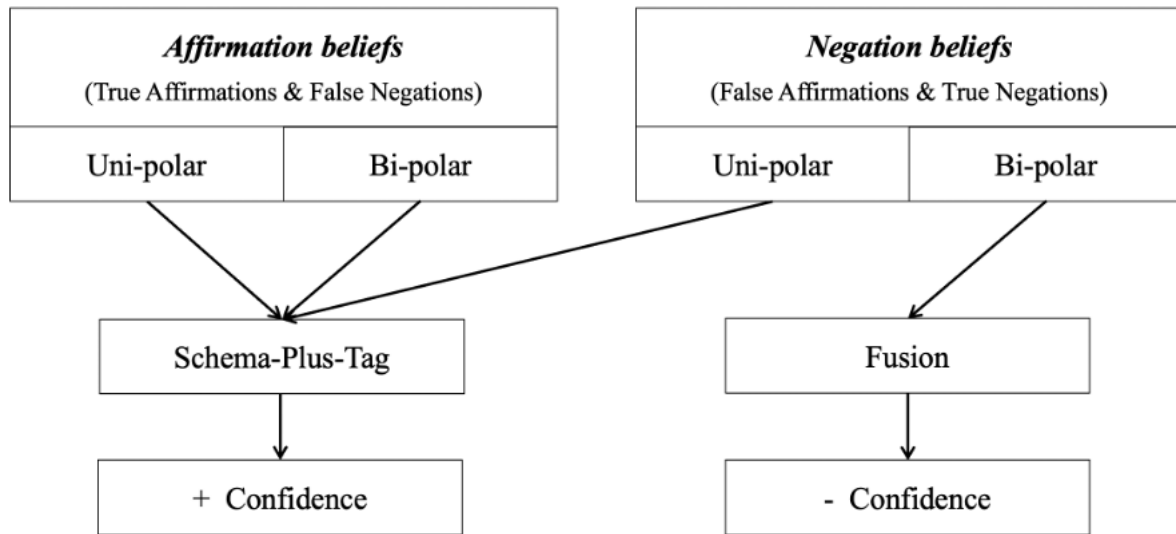


Figure 1. Belief-driven framework

In two preregistered experiments, we test our theory predicting the engagement of Fusion vs. SPT encodings and their differential impact on downstream judgment confidence. Study 1 tests a complete version of the theory using statements with uni-polar and bi-polar words, and measures participants' spontaneous veracity judgments to examine the downstream confidence effects across all message types. Given this evidence, Study 2 causally manipulates veracity judgments for bi-polar statements to isolate the contrast between SPT and Fusion, and replicates our predicted effects. All data, analysis code and experimental material are available on https://osf.io/kybhx/?view_only=b92c8b5f3e8c42ff8dbf8ef959174cec.

STUDY 1

Study 1 provides a full theory-driven test of our complete belief-centered framework and

its predictions about judgment confidence in a relatively neutral evaluative setting, which offers an appropriate platform for both conceptual and practical applications. As discussed, three linguistic elements are relevant for determining process type: *concept polarity* (uni-polar vs. bi-polar), *claim type* (affirmation vs. negation), and *veracity judgment* (true vs. false). This perspective yields a 2x2x2 design. However, our predictions related to judgment confidence are specific to the differences arising from engaging SPT versus Fusion. As illustrated in Figure 1 above, only bi-polar *Negation beliefs* engage Fusion. Bi-polar *Affirmation beliefs* and all uni-polar claims engage Schema-Plus-Tag. Therefore, while we analyze judgment confidence on the full 2x2x2 design for theoretical completeness, our (preregistered) focal analysis is the binary, process-driven comparison of interest.

Methods

Stimuli for this experiment were selected via two pre-tests. Pre-test 1 was used to create a list of words clearly perceived either as bi-polar or uni-polar. The critical factor in distinguishing uni- versus bi-polar words is the degree to which an antonym can be easily retrieved. Thus, participants were asked to generate antonyms for a range of words and to rate the difficulty of doing so. Uni-polar words were defined as those for which less than 20% of respondents provided a correct antonym, and for which the mean z-scored difficulty of doing so was greater than 0.25. Bi-polar words were defined as those for which more than 80% of respondents provided a correct antonym, and for which the mean z-scored difficulty rating was lower than -0.25 (see Table 3). These cutoffs were chosen to offer reasonable evidence that all selected stimuli would be reliably perceived by participants as uni-/bi-polar.

However, because the cutoff levels are arbitrary criteria, to ensure robustness, a different method is used for identifying a bi-polar concept in Study 2.

Bi-polar	Uni-polar
Before	Caloric
More	Deaf
Heavy	Precious
Negative	Territorial
Short	Charismatic
Northern	Insulating
Noisy	Eccentric
	Genetic

Table 3. Bi- and uni-polar words selected based on Pre-test 1 of Study 1.

Since our framework integrates concept polarity with veracity judgments, it was essential to use stimuli for which participants were likely to already hold prior beliefs. To do this, Pre-test 2 asked participants if they had “very basic knowledge” on a range of topics (Questions; *Q*) and asked them to rate their confidence in the veracity of paired Statements (*S*; Table 4). We regressed the judgment confidence ratings for each Statement on whether participants reported knowledge on the topic in the Question (yes/no). From the Question/Statement pairs showing a positive relationship, we then selected 12 Statements that represented the full scope of polarity and claim types (i.e., 3 bi-polar affirmations, 3 bi-polar negations, 3 uni-polar affirmations, and 3 uni-polar negations) with the fewest “completely uncertain” judgments (see Table 4). Details regarding both pre-tests are provided in the Web Appendix.

	Affirmations	Negations
Bi-polar	<i>Q</i>: Do you have very basic knowledge in American history? <i>S</i>: John F. Kennedy was elected before Ronald Reagan	<i>Q</i>: Do you have very basic general knowledge? <i>S</i>: Turtles can not live more than 100 years
	<i>Q</i>: Do you have very basic knowledge in chemistry? <i>S</i>: The charge of electrons is negative	<i>Q</i>: Do you have very basic knowledge in geography? <i>S</i>: Italy is not in the Northern hemisphere
	<i>Q</i>: Do you have very basic knowledge in natural sciences? <i>S</i>: One liter of oil is heavier than one liter of water	<i>Q</i>: Do you have very basic knowledge in acoustics? <i>S</i>: Explosions in outer space are not noisy
Uni-polar	<i>Q</i>: Do you have very basic knowledge about animal development? <i>S</i>: Dogs are deaf when they are born	<i>Q</i>: Do you have very basic knowledge about food properties? <i>S</i>: Avocado is not a caloric fruit
	<i>Q</i>: Do you have very basic knowledge in art history? <i>S</i>: Vincent Van Gogh had an eccentric personality	<i>Q</i>: Do you have very basic knowledge in material science? <i>S</i>: Glass is not an insulating material
	<i>Q</i>: Do you have very basic knowledge in genetics? <i>S</i>: Down Syndrome is a genetic disorder	<i>Q</i>: Do you have very basic knowledge in genetics? <i>S</i>: Traumas are not genetic conditions

Table 4. Study 1 stimuli.

This experiment is preregistered on AsPredicted (<https://aspredicted.org/q4jr-n8dz.pdf>). Participants (N = 270) were recruited through a behavioral laboratory at the Rady School of Management at the University of California, San Diego. While demographic information was not collected, the subject pool was drawn from the university's undergraduate population, and participation in the lab was not restricted by any demographic variables. Sample size was pre-set as one week of data collection in the lab. Two hundred and seventy undergraduates participated and were compensated with course credit. To incentivize accuracy and speed, participants were informed that an algorithm would identify the five participants who

correctly judged the veracity of the greatest number of Statement trials with the best-calibrated confidence ratings in the shortest time (as measured by fastest response times). These five participants were awarded a \$10 Amazon.com gift certificate.

This experiment employed a repeated-measures design—each participant made twelve judgments. Data were collected using DirectRT software (Empirisoft), allowing for measurement of behavioral responses and millisecond-level response time (RT). The full design, including participant instructions, is available in the Web Appendix. All participants first viewed instructions explaining that they were going to be shown Question trials, and that for each question they should answer by pressing either the key “Y”—for ‘yes’—or the key “N”—for ‘no’.

Following this, participants were instructed that they would complete multiple Statement trials, presented one at a time. Participants judged each statement as true or false by pressing the “T” or “F” key, respectively. Response time was measured as the duration from the appearance of the statement on the screen to the participants’ key press. Participants completed the task at their own pace with no time restrictions. After each Statement trial, participants rated their confidence in their judgment on a scale from 1 [completely uncertain about my answer] to 9 [completely certain about my answer]. All participants judged two practice trials (not included in the analysis), and then the 12 pre-tested statements, presented in randomized order.

Results and Discussion

Our framework centers on belief-based veracity judgments to determine engagement in one of the two cognitive processes of interest. To aid in confirming that participants held relevant prior beliefs, participants were asked Questions about whether they had basic topic-

level knowledge related to the specific Statements they would be judging (Table 4). As supported by Pre-test 2, these Question trials allowed us to identify whether participants were likely to have prior beliefs about the truthfulness of the Statement trials without disclosing the statements' specific content.¹ Following our preregistration plan, if a participant answered “no” to a Question, we excluded their responses for the paired Statement. Two participants answered “no” on all Question trials, resulting in their complete exclusion from the analysis. For the remaining participants (N = 268) the average number of observations retained was 7.81 (SD = 2.75).

To represent the full 2x2x2 set of linguistic factors in our theory-based framework, we extended our analyses beyond the focal preregistered contrast. We conducted a mixed-effects model regressing confidence on the full design, with subjects and Statement trials as random effects. According to our framework, only bi-polar *Negation beliefs* engage the Fusion process. All other combinations result in SPT. Thus, for bi-polar statements we should observe an interaction of claim type and veracity judgment such that affirmations judged as false and negations judged as true would be processed through Fusion leading to lower confidence, compared to affirmations judged as true and negations judged as false that would be processed through SPT. At the same time, we should observe no such interaction for uni-polar statements (all leading to SPT).

The results of this regression demonstrate the framework-predicted pattern. A significant three-way interaction ($b = -0.86$, $t = -2.00$, $p = .045$, 95% CI = [-1.70, -0.02]) indicates that the interaction of claim type and veracity judgment differs significantly between bi-polar and uni-polar Statements (see the Web Appendix for the full regression table). To unpack this, as shown in Figure 2, we found a significant crossover interaction between claim type and

¹ This test is specific to identify the likelihood of having prior beliefs, it does not measure, or make assumptions, about whether the individual's beliefs are accurate.

veracity judgment for bi-polar Statements ($b = 1.03$, $t = 3.73$, $p < .001$, 95% CI = [0.49, 1.57]). Specifically, bi-polar affirmations judged as false and bi-polar negations judged as true (Fusion cells) indeed led to lower confidence than their SPT counterparts. For uni-polar Statements, we observed no such interaction between claim type and veracity judgment ($b = 0.17$, $t = 0.54$, $p = .593$, 95% CI = [-0.46, 0.81]).

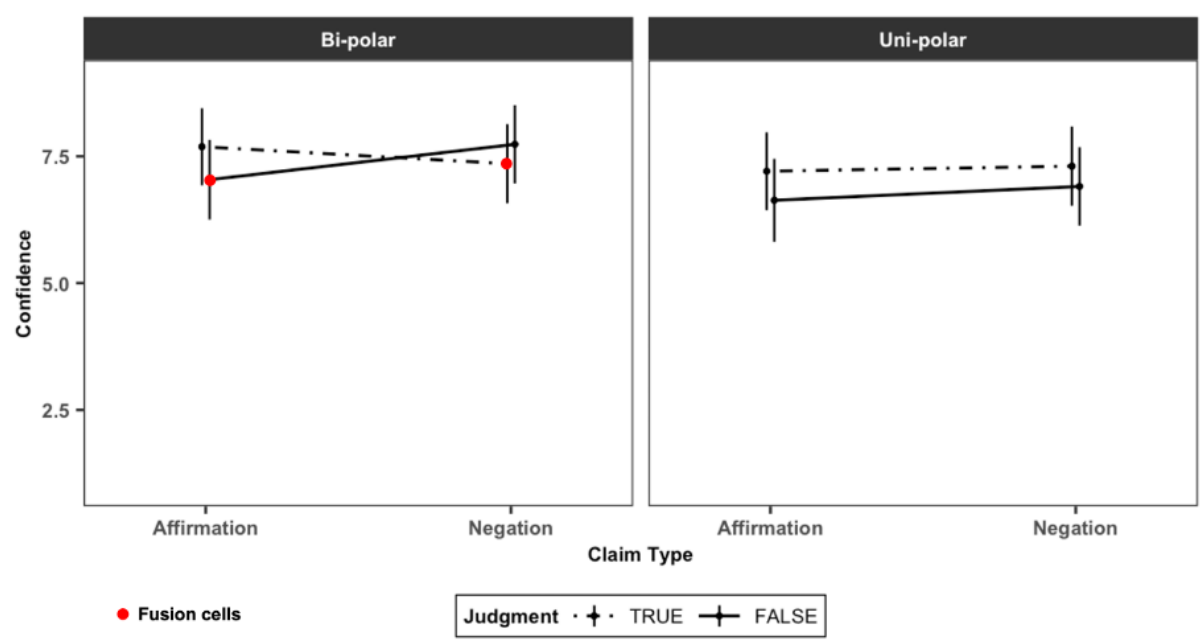


Figure 2. Confidence results by Concept Polarity (Bi-polar vs. Uni-polar), Claim Type (Affirmation vs. Negation), and Veracity Judgment (True vs. False) in the 2x2x2 mixed-effects model in Study 1.

These findings support our predicted dynamics of concept polarity, claim type, and veracity judgments on confidence.

While the full 2x2x2 design was necessary to demonstrate the full scope of the theoretical principles, we can collapse this down into the process-type comparison of interest between SPT and Fusion to test the implications for judgment confidence. The hypothesis that the Fusion process is more elaborate than SPT yields a prediction that statements

engaging Fusion would take longer for participants to process (Gough, 1965). Thus, as a manipulation check, we ran a linear mixed-effects model regressing judgment RT (i.e., response time) on a binary variable indicating the process type anticipated from the participants' combination of concept polarity, claim type, and veracity judgment (1 = Fusion, 0 = SPT), with random effects for subjects and Statements. This model confirmed that RTs were longer for Statements classified as engaging Fusion than for those classified as engaging SPT ($b = 655.6$, $t(1619.36) = 2.63$, $p = .009$, 95% CI = [166.5, 1144.7]).² This finding indicates that the claims our framework categorized as leading to Fusion were indeed slower (and likely more effortful) to process than those leading to Schema-Plus-Tag.

The key hypothesis tested in this experiment is whether statements that our framework predicted to engage Fusion would lead to lower judgment confidence than those engaging SPT. We ran a linear mixed-effects model regressing judgment confidence on process type, with random effects for subjects and Statement trials, which revealed that, as predicted, Statements categorized as engaging Fusion generated lower confidence than those engaging SPT ($b = -0.49$, $t(1715.02) = -3.54$, $p < .001$, 95% CI = [-0.76, -0.22]). This result offers evidence that judgment confidence is linked to the encoding processes evoked by the interactions of concept polarity, claim type, and veracity judgments arising from prior beliefs.

The higher confidence levels observed for SPT over Fusion appear independent from other statement characteristics like length. In fact, all results presented above replicate when controlling for the number of characters in each statement in each model ($b_{RT} = 618.8$, $t(1707.35) = 2.47$, $p = .013$, 95% CI = [128.0, 1109.6]; $b_{Confidence} = -0.50$, $t(1846.92) = -3.58$, $p < .001$, 95% CI = [-0.77, -0.22]), and number of characters itself does not have a significant

² This random-effects-model is slightly different from the preregistered one, which yields the same qualitative effect ($b = 1066$, $t(1435.3) = 4.63$, $p < .001$, 95% CI = [614.3, 1517.6], $r^2 = 0.27$). We judged the model in the main text as more appropriate because the Statement trials' order was already randomized by design.

impact on either RT or confidence (see Figure 3).

R Tjudg				Confidence			
Predictors	Estimates	CI	p	Predictors	Estimates	CI	p
(Intercept)	1328.52	-2088.62 – 4745.65	0.446	(Intercept)	6.75	4.37 – 9.12	<0.001
ProcessType [Fusion]	618.80	128.00 – 1109.59	0.013	ProcessType [Fusion]	-0.50	-0.77 – -0.22	<0.001
CharactersN	78.55	-8.91 – 166.01	0.078	CharactersN	0.02	-0.04 – 0.08	0.589
Random Effects				Random Effects			
σ^2	11472486.89			σ^2	3.44		
τ_{00} Subj	3234989.55			τ_{00} Subj	1.22		
τ_{00} Question	794061.90			τ_{00} Question	0.40		
ICC	0.26			ICC	0.32		
N Subj	268			N Subj	268		
N Question	12			N Question	12		
Observations	2093			Observations	2093		
Marginal R^2 / Conditional R^2	0.024 / 0.278			Marginal R^2 / Conditional R^2	0.006 / 0.324		

Figure 3. Regression tables for the mixed-effects models regressing RT and Confidence Process Type, with Character Number as fixed effect and Subjects and Statement trials as random effects in Study 1.

Prior literature might suggest that our observed effects could have been driven by semantic differences inherent to our specific lexical stimuli (see Berger & Packard, 2023 for review). The design of Study 1 helps to address this concern. As noted, the uni-polar words appear only in the SPT condition. However, because the study was designed to rely on flexible belief-driven judgments, bi-polar statements could be categorized as SPT for some participants and as Fusion for others, depending on how their veracity judgment interacted with claim type (see Table 2 and Figure 1). Beyond the preregistered analysis, we can thus examine only the six bi-polar Statement trials (so that *all* stimuli in the analysis appear in both conditions). Conducting the same linear mixed-effects models on this subset of the data yields the same set of findings: Fusion led to longer RTs ($b = 655.1$, $t(1063.27) = 2.09$, $p = .037$, 95% CI = [41.1, 1269.0]) and lower judgment confidence ($b = -0.51$, $t(1017.09) = -$

3.63, $p < .001$, 95% CI = [-0.78, -0.23]) than SPT. This result underscores the key contribution of the current framework: that the same exact statements can lead to different downstream judgments depending on the encoding process engaged.

In sum, this study provides evidence for the process-driven confidence dynamics predicted by our framework. Our conservative design enabled a rigorous test of the full set of the interactions possible from each element derived from the interactions of linguistic and belief theory. We further observed consistent results when restricting the dataset to only bi-polar statements, allowing us to compare Fusion vs. SPT using the exact same stimuli. The study's full design and analysis specifically tests our hypothesis that *existing prior beliefs* shape processing categorization, necessitating (preregistered) exclusions of participants who indicated no prior knowledge of the topics. Despite this, an exploratory version of our primary analyses without such exclusions is able to replicate a similar pattern of results.³ However, some limitations remain. First, eliciting individuals' own beliefs means that one of the three factors determining the encoding processes (i.e., veracity judgment) is non-randomized. Therefore, the SPT/Fusion classifications were potentially inferred based on partly endogenous factors. Second, we have a range of concepts and statements that vary in more ways than just polarity and number of characters.

To ensure that these elements were not driving our results, we conducted Study 2. This second experiment addresses the endogeneity concerns by manipulating veracity judgments. In addition, it tests a single concept, thereby removing concerns associated with comparing varied words or statements. Because uni-polar concepts can only lead to SPT, while bi-polar concepts can lead to both SPT and Fusion (see Table 2 and Figure 1), Study 2 uses a single bi-polar concept, focusing on one half of our framework.

³ Statement trials categorized as leading to Fusion led to longer RTs ($b = 444.5$, $t(2782.04) = 2.38$, $p = .017$) and marginally lower confidence ($b = -0.2$, $t(2761.05) = -1.71$, $p = .088$).

STUDY 2

Following the more neutral evaluative setting of Study 1, which employed a variety of concepts, Study 2 uses a single concept (and related statement) to experimentally control for any semantic, familiarity, or grammatical differences across conditions. Because only bi-polar concepts can lead to both SPT and Fusion, we employed a bi-polar concept. Moreover, whereas the encoding processes in Study 1 were partly elicited through participants' own judgments, in Study 2 participants were randomly assigned to make specific veracity judgments. Practically, the two linguistic processes arise from a fully randomized 2 (Claim Type: Affirmation vs. Negation) x 2 (Veracity Judgment: True vs. False) design that is implemented between-participants.

Methods

Stimuli were selected for this experiment through a pre-test that examined five potentially bi-polar concepts, each presented in both their affirmative and negative forms (see Web Appendix for full details). Participants were asked to rewrite the affirmations in a way that conveyed the opposing meaning, and the negations in a way that conveyed the same meaning. We coded rewrites expressed as affirmation claim types as 'Fusion-style,' and rewrites written as negation claim types as 'SPT-style.' Chi-square tests comparing the frequency of Fusion- versus SPT-style rewrites for each statement revealed a significantly stronger Fusion approach (all $ps < .01$). This suggests that the full set of concepts/statements tested qualified as bi-polar. From this pool of options, we selected the statements using

“strong” (i.e., “Cyclamen is [not] a strong plant”) for the Study 2 stimuli, because we expected participants to hold weak or minimal pre-existing opinions (priors) about them.

This experiment is preregistered on AsPredicted (<https://aspredicted.org/n23b-qttt.pdf>). Seven hundred and ninety-nine Connect workers from the U.S. and Canada (49.4% female, 0.3% nonbinary, $M_{\text{Age}} = 39.8$ years, $SD_{\text{Age}} = 12.7$ years) completed the experiment in exchange for monetary compensation (\$0.30 for an expected 2-minute completion time). As preregistered, participants who failed to correctly answer an attention-check question were not allowed to start the study.

Participants were randomly assigned to one of four conditions defined by the 2x2 design. First, the claim type of the statement was either an affirmation “Cyclamen is a strong plant,” or a negation “Cyclamen is not a strong plant.” Second, we randomly assigned participants to a (true vs. false) judgment. To do so, participants received information that created a specific belief (or “prior”; see Table 5).

Starting the study, participants read information constituting our manipulation of the veracity judgment. Specifically, participants read “Cyclamen is a plant native to Southern Europe that can live up to 30 years. Cyclamens have a high [low] survival rate of 80% [20%].” These two versions of the information were designed to create beliefs about the plant’s hardiness, using numerical evidence separate from the bi-polar word “strong.” After this “prior formation stage,” participants viewed the affirmation or negation version of the statement “Cyclamen is [not] a strong plant” (see Table 5) and were asked whether they believed the statement to be true or false (1 = True, 0 = False). They then rated their confidence in their judgment on a 9-point scale (1 = Completely uncertain about my answer, 9 = Completely certain about my answer). Finally, they provided demographic information. Note that, to avoid confusion between our veracity-judgment manipulation and participants’

actual veracity judgments, in what follows we refer to our manipulation as “Manipulated Judgment.”

CONDITION	PRIOR	STATEMENT	PROCESS
AFFIRMATION TRUE	Cyclamen is a plant native to Southern Europe that can live up to 30 years. Cyclamens have a high survival rate of 80%.	Cyclamen is a strong plant	SPT
AFFIRMATION FALSE	Cyclamen is a plant native to Southern Europe that can live up to 30 years. Cyclamens have a low survival rate of 20%.	Cyclamen is a strong plant	FUSION
NEGATION TRUE	Cyclamen is a plant native to Southern Europe that can live up to 30 years. Cyclamens have a low survival rate of 20%.	Cyclamen is not a strong plant	FUSION
NEGATION FALSE	Cyclamen is a plant native to Southern Europe that can live up to 30 years. Cyclamens have a high survival rate of 80%.	Cyclamen is not a strong plant	SPT

Table 5. Study 2 design and stimuli.

Results and Discussion

Attesting to the efficacy of our manipulation of participants’ veracity judgments, a chi-square test revealed that the majority of participants (75.06%) who were assigned to the ‘false’ Manipulated Judgment condition judged the assigned statements as false ($\chi^2(1) = 100.75, p < .001$). A second test showed that the majority (79.40%) of participants assigned to the ‘true’ Manipulated Judgment condition judged the statement as true ($\chi^2(1) = 137.58, p < .001$).

For the concept “strong” (bi-polar), our theoretical framework predicts that the Claim Type/Manipulated Judgment combinations of affirmation/true and negation/false are bi-polar *Affirmation beliefs* and therefore engage the SPT process. In contrast, affirmation/false and negation/true constitute bi-polar *Negation beliefs* that engage the Fusion process. Thus, we can re-categorize the four conditions into two Process Types (SPT and Fusion; see Table 5). Our preregistered prediction was that the Fusion condition (i.e., the collapsed

affirmation/false and negation/true conditions) would lead to lower judgment confidence than the SPT condition (i.e., the collapsed affirmation/true and negation/false conditions).

Regressing judgment confidence ratings on a binary variable for Process Type (1 = Fusion; 0 = SPT) confirms this hypothesis, as the Fusion condition was associated with lower confidence ($M = 6.30$, $SD = 1.92$) than the SPT condition ($M = 6.99$, $SD = 1.97$; $b = -0.69$, $t(797) = 5.02$, $p < .001$, 95% CI = [-0.96, -0.42]).

We can extend the analysis beyond the preregistered regression to also examine the underlying 2x2 nature of the experimental design. Conducting a type 3 2x2 ANOVA revealed that there were no main effects of Claim Type ($F(795, 1) = 0.01$, $p = .924$) or of Manipulated Judgment ($F(795, 1) = 3.34$, $p = .068$), but a significant crossover interaction of these two factors ($F(795, 1) = 25.29$, $p < .001$). As our theory predicts, the interaction (illustrated in Figure 4) shows that the affirmation/false and negation/true (Fusion) conditions yielded lower confidence than the affirmation/true and negation/false (SPT) conditions.

Our analyses offer conservative tests of our hypotheses, since participants who judged the claim counter to their condition's manipulation would only add noise to the model. Consistent with our model's predictions, we also replicate the same pattern of results when using Claim Type interacting with Reported Judgment instead of Manipulated Judgment (see Web Appendix). By directly manipulating participants' priors through random assignment, these results provide evidence supporting the predictions of our theoretical framework, complementing the self-generated priors that drove veracity judgments in Study 1.

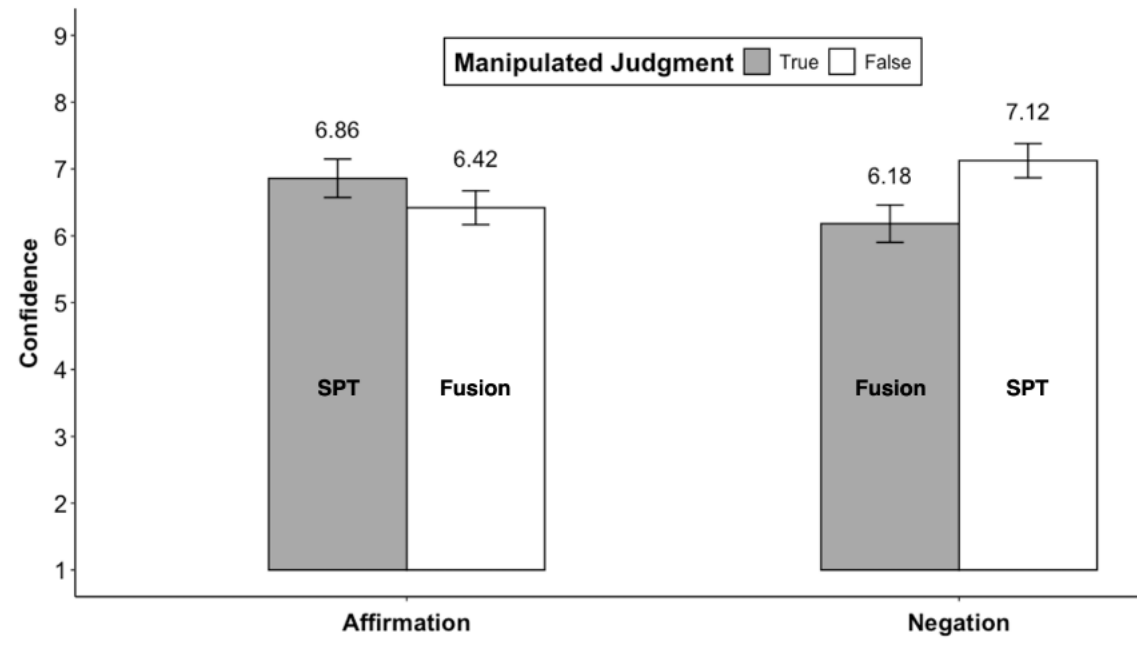


Figure 4. Confidence results by Claim Type (Affirmation vs. Negation) and Manipulated Judgment (True vs. False), given a bi-polar concept, in Study 2.

GENERAL DISCUSSION

Our results provide evidence for a novel conceptual framework that integrates consumer veracity judgments with the linguistic features of a statement (i.e., claim type and concept polarity) to determine the cognitive processes that ultimately influence judgment confidence.

Theoretical Contribution

From a theoretical standpoint, prior findings in cognitive linguistics have explored the interaction of concept polarity and claim type to determine SPT vs. Fusion encoding (e.g., Mayo, Schul, & Burnstein, 2004). Our research extends this work by incorporating consumers' veracity judgments to predict when statements will be encoded via one process or the other. The inclusion of individual priors, or beliefs, reveals two novel theoretical insights.

First, not all bi-polar negation claims are processed through Fusion—only those judged as true by the message recipient. Second, bi-polar negation claims are not the only statements processed through Fusion—affirmation claims judged as false are as well.

In this work, although we use RT as evidence suggesting that distinct processes are elicited, we do not directly investigate how those processes unfold. In psycholinguistics, SPT has been described as a “two-step procedure of validation of negation,” in which both the to-be-negated and the negated meanings are mentally represented (e.g., Carpenter & Just, 1975). Fusion has frequently been described as a single-step procedure in which only the negated meaning is represented (e.g., Kaup, Lüdtke & Zwaan, 2006; Kaup & Zwaan, 2003). This might lead us to expect the two-step process to be longer than the single-step one. However, more recent work taking advantage of temporally precise dynamic electroencephalography measurements has challenged this notion, finding sequential activity patterns suggesting that Fusion may also involve two steps (Dudschig & Kaup, 2018). The longer RT we observe for Fusion could be interpreted to support a multi-step process but could also be consistent with a more effortful single-step process. Thus, while our findings support a qualitative distinction in the nature of the mechanisms involved, the exact characterization of those mechanisms remains an important question for future research.

In work related to problem solving, RT has been directly related to judgment confidence, independently of judgment accuracy (Ackerman & Zalmanov, 2012). RT has also been found to be correlated with processing fluency—the experience of processing ease—which influences a variety of judgments (e.g., Reber & Schwarz, 1999; Reber, Schwarz, & Winkielman, 2004). This raises the question of whether the observed differences in RT between SPT and Fusion are capturing differences in fluency, and whether RT might mediate the effect of process type on judgment confidence. However, a mediation analysis failed to

show a relationship among process type, RT, and confidence (see Web Appendix). As a result, in this context, RT may reflect other cognitive elements, such as the additional semantic retrieval required for Fusion processes.

Practical Implications

Our theoretical framework not only clarifies how consumers process language at a cognitive level, but also offers implications for how real-world messages may shape behavior. First, by manipulating veracity judgments through priors on an unfamiliar topic in Study 2, nearly 80% of participants judged the veracity of the statements in alignment with the manipulation, representing a clear majority, but not the totality, of participants. This reflects a challenge inherent to real-world marketing communications: when launching a new product, firms often pair claims with supporting evidence but ultimately cannot ensure consumers believe them. Study 2 provides evidence that our theoretical framework has value for such naturalistic marketing contexts. Following its implications, for new products, marketers should favor the use of bi-polar affirmations (such as “The scent is intense” or “The scent is mild”), because most consumers would engage SPT by judging the affirmation as true, leading to higher judgment confidence. For those who judge the claim as false, it would engage Fusion, yielding a lower judgment confidence. This simple lexical choice could help the company maximize confidence in their desired message and minimize confidence among the doubters.

However, for an existing product, or line extension, consumers may already hold beliefs about the sunblock’s fragrance. If consumers beliefs are likely to agree with the firm’s desired message, our theory recommends phrasing it in a way that engages SPT to maximize confidence. However, if marketers fear consumers will be skeptical, potentially judging the

claim as false, they should use phrasing that engages Fusion to minimize consumer's confidence in their disbelief.

Process-dependent differences in judgment confidence could also lead to differences in other marketing-relevant outcomes, such as the persuasiveness of communications (e.g., Bassili, 1996; Fazio & Zanna, 1978; Fishbein & Ajzen, 1975; Tormala & Petty, 2002, 2004; Tormala & Rucker, 2007). To provide preliminary insight into such applications, we collaborated with the non-profit organization Project Keshet, which solicited potential financial donors through Facebook to support women during the 2022 military conflict in Ukraine. We designed ads intended to engage either Fusion or SPT processes, and compared their effectiveness in generating clicks in the “noisy” real-world setting of this social media platform. The ad engaging Fusion elicited a lower click through rate (4.07%) than the ad engaging SPT (4.74%; $p = .009$; $N=20,118$; see Web Appendix for details). Although the nature of the ad platform limits the extent to which causality can be inferred (Boegershausen et al., 2025, Braun & Schwartz, 2024), this difference illustrates how appropriately harnessing linguistic encoding processes may matter for real-world marketing outcomes. It also points to opportunities for future research investigating the effects of SPT- versus Fusion-engaging messages across contexts, platforms, and communication domains.

Limitations and Future Research

A key contribution of our conceptual framework is that SPT and Fusion systematically lead to different levels of judgment confidence depending, in part, on whether the underlying belief is an affirmation or a negation. However, questions remain regarding the degree to which these findings may generalize across languages. Negation markers vary widely across languages, with the main distinction being between double negation languages (i.e.,

languages where each negative marker yields an independent semantic negation) and negative concord languages (i.e., languages where two concurrent negative markers constitute a single semantic negation; e.g., Jespersen, 1917; Dudschig et al., 2021; Maldonado & Culbertson, 2021). English is a double negation language, so it seems reasonable to expect our framework to generalize to double negation languages with comparable concept polarity as English (such as German). Whether it would apply to negative concord languages remains an open question, highlighting the practical challenges in developing actionable communication insights that generalize across cultural boundaries and linguistic systems, even when findings arise from underlying psychological mechanisms.

While our findings suggest that structural features like word or statement length do not interfere with the downstream effects of encoding processes, it is possible that other message characteristics would do so. For example, perceptions of a message's source might interfere with our framework's validity, as various source characteristics can make statements more likely to be deeply processed versus handled via heuristics (e.g., central vs. peripheral route; Petty & Cacioppo, 1986), offering additional pathways and potential interactions to consider at the cognitive level.

Finally, it is worth considering the linguistic axiom that there are no true synonyms, because if different words exist, they must have different meanings (cf. Bloomfield, 1933; Bolinger, 1968; Haiman, 1980; Nida, 1958). This raises a central question for additional research regarding defining the threshold for an acceptable semantic difference between the bi- and uni-polar synonyms necessary to precisely elicit and differentiate the processes they engage.

In addition to expanding the scope of language research within consumer psychology, our findings also contribute to the concept-processing and psycholinguistic literatures.

Overall, the current work offers a novel cross-disciplinary framework for better understanding language encoding processes and their influence on judgments and behavioral outcomes.

REFERENCES

Ackerman, Rakefet, and Hagar Zalmanov (2012), “The Persistence of the Fluency—Confidence Association in Problem Solving,” *Psychonomic Bulletin & Review*, 19, 1187-1192.

Bassili, John N. (1996), “Meta-judgmental versus Operative Indexes of Psychological Attributes: The Case of Attitude Strength,” *Journal of Personality and Social Psychology*, 71(4), 637-653.

Berger, Jonah, and Grant Packard (2023), “Wisdom from Words: The Psychology of Consumer Language,” *Consumer Psychology Review*, 6(1), 1-16.

Bloomfield, Leonard (1933), *Language*, New York: Holt.

Boegershausen, J., Cornil, Y., Yi, S., and Hardisty, D. J. (2025). On the persistent mischaracterization of Google and Facebook A/B tests: How to conduct and report online platform studies. *International Journal of Research in Marketing*, 42(3), 886-903.

Bolinger, Dwight (1968), “Entailment and the meaning of structures,” *Glossa*, 2, 119-127.

Braun, M., and Schwartz, E. M. (2024), "Where A-B testing goes wrong: How divergent delivery affects what online experiments cannot (and can) tell you about how customers respond to advertising, *Journal of Marketing*, 87(2), 71-95.

Carpenter, Patricia A., and Marcel A. Just (1975), "Sentence Comprehension: A Psycholinguistic Processing Model of Verification," *Psychological Review*, 82(1), 45-73.

Dudschig, Carolin, and Barbara Kaup (2018), "How Does "Not Left" Become "Right"? Electrophysiological Evidence for a Dynamic Conflict-Bound Negation Processing Account," *Journal of Experimental Psychology: Human Perception and Performance*, 44(5), 716-728.

Dudschig, Carolin, Barbara Kaup, Mingya Liu, and Juliane Schwab (2021), "The Processing of Negation and Polarity: An Overview," *Journal of Psycholinguistics Research*, 50(6), 1199-1213.

Fazio, Russell H., and Mark P. Zanna (1978), "Attitudinal Qualities Relating to the Strength of the Attitude-Behavior Relationship," *Journal of Experimental Social Psychology*, 14(4), 398-408.

Festinger, Leon (1950), "Informal Social Communication," *Psychological Review*, 57, 271-282.

Festinger, Leon (1954), "A Theory of Social Comparison Processes," *Human Relations*, 7, 117-140.

Fishbein, Martin, and Icek Ajzen (1975), *Belief, Attitude, Intention, and Behavior*, Reading, MA: Addison-Wesley.

Gilbert, Daniel T. (1991), "How Mental Systems Believe," *American Psychologist*, 46(2), 107-119.

Gough, Philip B. (1965), "Grammatical Transformations and Speed of Understanding," *Journal of Verbal Learning and Verbal Behavior*, 4, 107-111.

Grant, Susan J., Malaviya, Prashant, & Sternthal, Brian. (2004). The influence of negation on product evaluations. *Journal of Consumer Research*, 31(3), 583-591.

Haiman, John (1980), "The Iconicity of Grammar: Isomorphism and Motivation," *Language*, 56, 515-540.

Hoch, Stephen J., and Young-Won Ha (1986), "Consumer learning: Advertising and the ambiguity of product experience," *Journal of Consumer Research*, 13, 221-233.

Hussein, Mohamed A., and Zakary L. Tormala (2024), "You versus We: How Pronoun Use Shapes Perceptions of Receptiveness," *Journal of Experimental Social Psychology*, 110, Article 104555

Jespersen, Otto (1917), "Negation in English and Other Languages," *AF Høst*, 1-151.

Kaup, Barbara, and Carolin Dudschig (2020), "Understanding Negation: Issues in the Processing of Negation," In Déprez, V. & Espinal, M.T. (Ed.), *The Oxford Handbook of Negation*, 635-655. Oxford, England: Oxford University Press.

Kaup, Barbara, Jana Lüdtke, and Rolf A. Zwaan (2006), "Processing Negated Sentences with Contradictory Predicates: Is a Door that is Not Open Mentally Closed?," *Journal of Pragmatics*, 38, 1033-1050.

Kaup, Barbara, and Rolf A. Zwaan (2003), "Effects of Negation and Situational Presence on the Accessibility of Text Information," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(3), 439-446.

Kronrod, Ann (2022), "Language Research in Marketing," *Foundations and Trends in Marketing*, 16(3), 308-421.

Maldonado, Mora, and Jennifer Culbertson (2021), "Nobody doesn't like negative concord," *Journal of Psycholinguistic Research*, 50, 1401-1416.

Mayo, Ruth, Yaacov Schul, and Eugene Burnstein (2004), "'I am not guilty' vs. 'I am innocent': Successful Negation May Depend on the Schema Used for its Encoding," *Journal of Experimental Social Psychology*, 40, 433-449.

Nida, Eugene A. (1958), "Analysis of Meaning and Dictionary Making," *International Journal of American Linguistics*, 24, 279-292.

Packard, Grant, and Jonah Berger (2021), "How Concrete Language Shapes Customer Satisfaction," *Journal of Consumer Research*, 47, 787-806.

Packard, Grant, Sarah G. Moore, and Brent McFerran (2018), "(I'm) Happy to Help (You): The Impact of Personal Pronoun Use in Customer-Firm Interactions," *Journal of Marketing Research*, 55, 541-555.

Petty, Richard E., Pablo Briñol, and Zakary L. Tormala (2002), "Thought Confidence as a Determinant of Persuasion: The Self-Validation Hypothesis," *Journal of Personality and Social Psychology*, 82(5), 722-741.

Petty, Richard E., and John T. Cacioppo (1986), "The Elaboration Likelihood Model of Persuasion," In Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*, 19, 123-205. San Diego, CA: Academic Press.

Pezzuti, Todd, & Leonhardt, James M. (2023). What's not to like? Negations in brand messages increase consumer engagement. *Journal of the Academy of Marketing Science*, 51(3), 675-694.

Reber, Rolf, and Norbert Schwarz (1999), "Effects of Perceptual Fluency on Judgments of

Truth,” *Consciousness and Cognition*, 8, 338-342.

Reber, Rolf, Norbert Schwarz, and Piotr Winkielman (2004), “Processing Fluency and Aesthetic Pleasure: Is Beauty in the Perceiver’s Processing Experience?,” *Personality and Social Psychology Review*, 8, 364-382.

Tormala, Zakary L., and Richard E. Petty (2002), “What Doesn’t Kill Me Makes Me Stronger: The Effects of Resisting Persuasion on Attitude Certainty,” *Journal of Personality and Social Psychology*, 83, 1298-1313.

Tormala, Zakary L., and Richard E. Petty (2004a), “Source Credibility and Attitude Certainty: A Metacognitive Analysis of Resistance to Persuasion,” *Journal of Consumer Psychology*, 14(4), 427-442.

Tormala, Zakary L., and Derek D. Rucker (2007), “Attitude Certainty: A Review of Past Findings and Emerging Perspectives,” *Social and Personality Psychology Compass*, 1(1), 469-492.

Wang, Shenshen, Chao Sun, Ye Tian, and Richard Breheny (2021), “Verifying Negative Sentences,” *Journal of Psycholinguistic Research*, 50, 1511-1534.