



# A Cognitive Toolbox for Persuasion. Perceived Diagnosticity and Persuasibility Bias of Persuasive Features

Teresa Garcia-Marques and Filipe Loureiro

Ispa-Instituto Universitário William James Center of Research

## ABSTRACT

Persuasive features such as argument quality, source expertise, support, affect, consensus, source likability, and source attractiveness are linked to mental rules that vary in their reliability in supporting the persuasive process. We address these features as cognitive tools available to support judgments, investigating their self-other perceived diagnosticity and persuasibility, and using self-other differences as an index of perceived bias. Results show that participants explicitly consider argument quality, source expertise, and support (number of arguments) as diagnostic of the validity of a claim. Such diagnosticity predicted perceived persuasibility, and self-other differences linked non-diagnostic features to bias. We further discuss the relevance of these findings for a more comprehensive understanding of the cognitive toolbox that supports persuasive influences.

## Introduction

Persuasion plays a pervasive role in our everyday communications, subtly intertwining with the act of conveying information. Attitude researchers have documented various elements of a persuasive communication pertinent to the success of a persuasive attempt (e.g., Chaiken, 1980; Hovland et al., 1953; Petty & Cacioppo, 1981, 1986). Although the most crucial feature of a persuasive context is the quality of the arguments presented (message content), numerous other features (mostly peripheral to the message content) intervene in the process (for a review, see Petty & Briñol, 2012, 2015). These persuasive features often supplant the effects of argument quality, especially when conditions do not allow for high elaborative processing (see Petty, 2013). Importantly, these persuasive features may also be attended when people engage in the elaboration of message content, likely influencing thoughts about the validity of a claim (see Petty & Wegener, 1999).

While research has clearly elucidated when and how various persuasive features intervene in the persuasive process (e.g., Petty, 2013; Petty & Wegener, 2014), we lack information regarding their perception as “cognitive tools” that support our judgments and decisions—a hypothetical persuasive “cognitive toolbox”

(see Gigerenzer, 2020). It is likely that through lived experiences, individuals gain insight into the differential usefulness of these features as cues of message validity, their manipulative nature, and their fairness. This knowledge encompasses beliefs about their diagnosticity (i.e., how informative they are about the validity of a claim) and their influence (persuasibility) over themselves and others. Access to this information is not straightforward, as disparities between self-perception and perceptions of others likely arise when biases are anticipated (see Friestad & Wright, 1994). As such, this paper aims to unveil perceptions of these two properties of the persuasive toolbox both for the self and others.

## Perceived diagnosticity of persuasive features

Within a persuasion setting, uncertainty about the validity of a claim is likely anchored in a set of “learned rules” ranging from simpler to more complex, which are strategically utilized to maximize efficiency. These implicitly or explicitly learned rules comprise what Gigerenzer (2020) named “our cognitive toolbox.” Here, we focus on some of these rules closely associated with a persuasive context. The use of a cognitive tool does not imply a lack of rationality. Rules

are not inherently good or bad, and all are prone to bias and errors. Individuals likely take this reliability into account when using them. That is, tool use is likely sensitive to their perceived diagnostic value within the specific persuasive environment in which they are employed.

The perceived diagnosticity (epistemic value) of a persuasive tool refers to how much it is considered a valid guide for judgment (see Slovic & Lichtenstein, 1971). Perceived diagnosticity of a cue is known to impact the judgmental process, influencing levels of information search and confidence (e.g., Van Wallendael & Guignard, 1992). This impact has been demonstrated with regard to, at least, one of the tools proposed in our persuasive toolbox, consensus (Aaker & Maheswaran, 1997; Chen & Chaiken, 1999; Mondak, 1993; Pham, 1996). Aaker and Maheswaran (1997) hypothesized that “consensus” and “internal feelings” hold different diagnostic values for different cultures (collectivist vs. individualist, e.g., Martin et al., 1993; Petty et al., 1993) and showed that manipulating these features differently impacts individuals from different cultures, even in situations of high elaboration.

Although it is likely that not all elements of the persuasive toolbox are perceived similarly regarding their diagnostic value (see Chen & Chaiken, 1999; Darke et al., 1998), previous research has not yet clearly informed us about their relative value. What we know so far is that high diagnosticity is likely perceived in the “quality of the arguments”. This is because individuals rely on such argument quality when concerned with the implications of their judgments (high elaboration conditions; Petty & Cacioppo, 1981, 1986). Additionally, we can assume that individuals attribute some diagnostic value to other persuasive features, as they frequently rely on them as shortcuts or heuristic pathways to conclude about the validity of a claim or to add relevant insights into an elaborative process (e.g., Petty & Wegener, 1999).

We also find suggestions that the perceived diagnosticity of persuasive tools may relate to the extent to which they trigger expectancies related to message content. Persuasive features that generate expectations about the content of the message, such as expertise and the number of supporting arguments, are perceived as more diagnostic than features like attractiveness and likability, which are unrelated to the message content (see Bohner et al., 1995, 2002). Bohner et al. (1995, 2002) also highlight that positive and negative instances of a persuasive feature may differ in their perceived diagnosticity. Evidence outside the persuasive context suggests that negative information is perceived as more diagnostic compared to

positive information (Ahluwalia, 2002; Ito et al., 1998; Levin et al., 1998). However, the opposite seems to occur concerning persuasive features. Here, an asymmetry in perceived diagnosticity emerges, favoring positive features: while a high level of expertise is perceived as diagnostic, non-expertise is not (see Bohner et al., 2002). This asymmetry in perceived diagnosticity aligns with evidence showing an asymmetric influence of positive and negative instances of persuasive feature. Studies have demonstrated that messages from likable and attractive sources have a greater effect than messages from unlikable and unattractive sources (Kang & Herr, 2006; Koernig & Page, 2002; Messner et al., 2008; Reinhard et al., 2006).

### ***Perceived persuasibility and bias***

Individuals, cognizant of the persuasibility of specific features, perceive their influence as potentially undesirable for themselves (e.g., Albarracín & Mitchell, 2004). Understanding the perceived persuasibility of the cognitive toolbox is crucial in comprehending how individuals will actively resist or become more susceptible to these influences, especially if they are unaware of the influence (Brinöl, et al., 2004).

Individuals often believe they are immune to persuasive influences, especially when they possess knowledge about persuasion (e.g., Friestad & Wright, 1994; Rule et al., 1985) and are familiar with the strategies and tactics commonly employed in persuasion (Roskos-Ewoldsen, 1997). However, these beliefs do not lead individuals to perceive persuasive tools as lacking persuasibility. Despite asserting their resistance, individuals simultaneously view these tactics as effective for convincing others. This phenomenon, known as the third-person effect (for a review, see Perloff, 1996, and Sun et al.'s, 2008, and Eisend's 2017 meta-analyses), illustrates that individuals believe persuasive factors to be more persuasive for others than for themselves, assuming others are more susceptible to biases. The assumption that this difference between self-perception and perceptions of others is related to the level at which the influence is perceived to occur through bias is supported by two sets of empirical arguments. First, the third-person effect occurs mainly when the perceived influence is deemed undesirable (e.g., David & Johnson, 1998; Duck & Mullin, 1995; Gunther & Mundy, 1993; Sun et al., 2008). Second, when the influence at stake is perceived as socially desirable (as we assume it would be if perceived as diagnostic), the effect reverses: individuals report persuasive influences to have more impact on themselves than on others (e.g., Duck et al., 1995; Duck &

Mullin, 1995; Eveland & McLeod, 1999; Hoorens & Ruiter, 1996). These arguments support the assumption that the magnitude of the self-other difference indicates how much the influence at stake is perceived to be a bias and that it is likely related to perceived diagnosticity of the persuasive cue.

Our objective is simple. We aim to inquire about how individuals generally perceive the *diagnostic value* and *persuasibility* of some elements of their cognitive toolbox. This knowledge is relevant as it allows us to characterize how individuals socially differentiate the usefulness of their cognitive tools and how persuasive features differ in this regard. Importantly, social agreement does not imply a lack of individuality. Therefore, we also expect individuals to vary in their perceptions. This individual variability is relevant since differences in individuals' naive theories about persuasibility and the potential bias of different persuasive features have been shown to be important moderators of attitude change (Wegener et al., 2001; Wegener & Petty, 1997). Likely, at this individual level of analysis (contrasting with the persuasive tools level of analysis), perceived bias of an influence will also relate to the perceived diagnosticity of a specific persuasive feature.

In summary, we believe that both variabilities (the one defined at the tool level of analysis and the one defined at an individual level of analysis) and dimensions (diagnosticity and persuasibility/bias) are pertinent to our understanding of how individuals use persuasive cognitive tools and even intentionally and strategically choose between them to ease computational burden (Gilovich et al., 2002).

### Current studies

We rely on the available knowledge about the content of the persuasive toolbox to investigate how these cognitive tools vary in their perceptions of persuasibility, bias, and diagnosticity. We focus on persuasive features known (e.g., Chen & Chaiken, 1999; Petty & Briñol, 2012) to exert a role in a persuasive process such as Expertise, Support (number of arguments), Consensus, Attractiveness, Likability, and Affect. We expect to find evidence that these persuasive features vary in their perceived diagnosticity and persuasibility, and that these two dimensions are not independent. As Bohner et al. (1995, 2002) suggest, features more closely related to message content (e.g., expertise and support) are likely to be perceived as more diagnostic. We also aim to support the claim (see Bohner et al., 1995, 2002) of an asymmetry in perceived diagnosticity between positive and negative instances of the persuasive features.

Additionally, we address the interrelationship between the three perceived properties of the

persuasive tools at both the tool and individual level. By comparing different persuasive features, we expect to find evidence that those evaluated as having higher persuasibility for the self are also perceived to be more diagnostic. Self-other differences will signal which persuasive features are perceived to bias attitude judgments.

At the individual level of analysis, we investigate whether individuals who perceive a persuasive feature as having higher levels of diagnosticity will also report fewer self-other differences in its persuasibility. If this holds true, we will further address the hypothesis that perceived self-other differences in persuasibility can be linked to perceived self-other differences in diagnosticity.

We approach these aims in two studies involving distinct populations and operationalizations of main variables. Study 1 comprises two sets of data collected from Portuguese samples, evaluating either perceived diagnosticity or persuasibility of the set of persuasive features, for both self and others. Study 1a assessed only the perceived diagnosticity of the positive and negative persuasive features. Study 1b focused only on positive features and evaluated them in terms of perceived persuasibility for both self and others. Self-other differences in perceived persuasibility were used as an index for the perceived bias of the features' influence.

Study 2 conceptually replicates Study 1 with an American sample, simultaneously assessing (within participants) the two evaluative dimensions for both self and others. Study 2 allows us to replicate Study 1's results at the level of persuasive feature analysis and provides additional insights into the relationship between perceived diagnosticity and persuasibility at the individual level of analysis. This allows us to state that individuals perceive diagnostics related to perceived persuasibility, which supports to the statement that the persuasive features perceived by the group as more diagnostic are also the ones perceived by the group to be more persuasive.

All studies were conducted between 2016 and 2020, approved by Ispa's Ethical Committee, and all participants provided informed consent for participation. Sample sizes were determined by the availability of volunteers and/or ensuring sufficient power to detect effects with  $\alpha = .05$ ,  $1 - \beta = .80$  (see Faul et al., 2007), and a moderate effect size (based on a pilot study and Richard et al., 2003).

### Study 1

This study aims to describe how six persuasive features are generally perceived in terms of their diagnosticity and their persuasibility for both self and others. The assessment of these perceptions was conducted using two different samples of participants and

slight procedural differences. Study 1a examines both positive (e.g., expertise, likability) and negative (e.g., non-expertise, unlikability) instances of each of the six persuasive features. Study 1b focuses solely on the perceived persuasibility of positive instances of the six persuasive features, adding “argument-quality” to provide a clear referent to interpret participants’ responses (given that this message feature is the one individuals attend to when they are elaborating on the content of the message, and so, by default, it has been considered the one to which they attribute higher levels of diagnosticity). We expected higher perceived diagnosticity of this persuasive feature, as individuals are likely to pay attention to it when elaborating on the content of the message. In this study, persuasibility was assessed for both self and others in a counterbalanced order.

### Study 1a

A sample of 145 undergraduate students (73.87% females) aged between 18 and 46 years-old ( $M=22.1$ ,  $SD=5.12$ ) participated in this study in return for course credit, and evaluated the perceived diagnosticity of 6 (Persuasive features) x 2 (Instance: Positive vs. Negative).

### Pretests

We initially developed a diagnosticity measure for each persuasive feature using Likert’s methodology. A series of one-sentence statements regarding the diagnostic value of each persuasive feature was assessed by 10 judges to ensure face validity and uni-dimensionality. Subsequently, a sample of 20 university students rated their agreement with each statement on a seven-point scale, allowing for the study of psychometric properties.

From these pretests, we selected ten statements for each of the six persuasive variables (five focusing the positive instance of the persuasive features. e.g., “likable=valid” and five focusing the negative instance of the persuasive feature, e.g., “unlikeable=non-valid”) to be associated with a symmetrical agree-disagree continuum, constituting a measure of perceived diagnosticity (see [Supplemental Material](#)). Examples of these statements are: “Experts are closer to the truth” and “Non-experts don’t know what they talk about” (for *Expertise*); “The number of arguments gives strength to a position” and “We should not trust information not supported by many arguments” (for *Support*); “We can trust information shared by many people” and “An idea that few people agree with is generally incorrect” (for *Consensus*); “The opinion of

an attractive person is always preferable” and “We can’t trust the opinion of an ugly person” (for *Attractiveness*); “We can trust information shared by nice people” and “An unlikable person is never right” (for *Likeability*); and “We can trust a judgment that makes us feel good” and “Our negative feelings inform us about bad things” (for *Affect*).

Each of the six measures of perceived diagnosticity, when subjected to an exploratory factor analysis with an orthogonal rotation, showed a two-factor structure (separating positive and negative instances), explaining more than 64% of the variance and exhibiting reliable internal consistency (all Cronbach’s alphas  $> .70$ ).

### Procedure

Data were collected on a laboratory setting, having all instructions presented on a computer screen supported by E-prime 2.0 software. Each of the 60 items (5 exemplars x 6 Persuasive features x 2 Frames) was randomly presented in the middle of the screen in association with a 7-point scale (1-completely disagree; 7-completely agree).

### Results

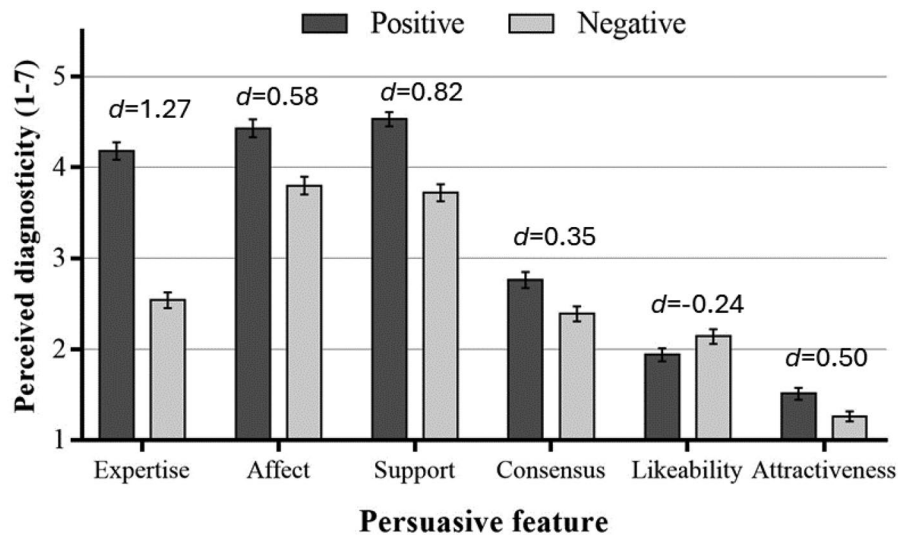
Participants’ responses to each of the five exemplars were averaged within the 6 (Persuasive features) x 2 (Positive vs Negative) design.

Results showed that the positive instance of all persuasive features were clearly perceived by participants as more diagnostic ( $M_p = 3.23$ ,  $SD=0.06$ ) than the negative instance ( $M_N = 2.64$ ,  $SD=0.06$ ;  $\eta^2_p = 0.55$ ). [Figure 1](#) shows that this occurs for most of the persuasive features, being the standardized difference between positive and negative instances smaller for Likeability and Attractiveness.

Participants also perceived clear differences in levels of diagnosticity within the set of presented persuasive features ( $\eta^2_p = 0.74$ ). As documented in [Figure 1](#), Expertise, Affect, and Support were perceived to be more diagnostic than Consensus, Likeability, and Attractiveness.

### Study 1b

Fifty Portuguese voluntary students (60% females) aged between 17 and 47 years ( $M=22.48$ ,  $SD=4.64$ ) participated in an online study conducted on Qualtrics. University students who voluntarily participated in this study subsequently distributed the study link to other participants through a snowball procedure. Participants were randomly allocated to two versions of a questionnaire defined by a mixed design: 7



**Figure 1.** Perceived diagnosticity of positive and negative instances of each persuasive feature.  
**Note.** Error bars are standard errors;  $d$  = Cohen  $d$  for self-other comparisons.

(Persuasive features)  $\times$  2 (Persuasive target: self vs. others)  $\times$  2 (Order: self-first vs. others-first), with the last factor between-participants to control for order effects.

### Procedure

Participants were contacted online and invited to participate in an opinion study. After accepting participation, they were informed that their task was to provide their opinion about features that should (or should not) be included in a petition to make it more effective. In the first task, they selected from a list of characteristics those they believed would be relevant in convincing themselves or others to sign the petition. In the second task, participants were asked to rate each of those features in terms of their perceived persuasibility/effectiveness for the self and others on a scale from 1 (not at all effective) to 5 (very effective). The order of self and others ratings was counterbalanced.

The features list was: Expertise (“*Author’s demonstration of knowledge about the petition’s cause*”), Support (“*The number of arguments presented on the petition*”), Argument-quality (“*The quality of the arguments presented on the petition*”), Consensus (“*High number of people reported to have already signed the petition*”), Attractiveness (“*The beauty/attractiveness of the petition’s author*”), Likeability (“*The friendliness expressed by the author of the petition*”) and Affect (“*What I feel while reading the petition*”).

Participants were lastly thanked for their participation and debriefed about the aims of the study.

## Results

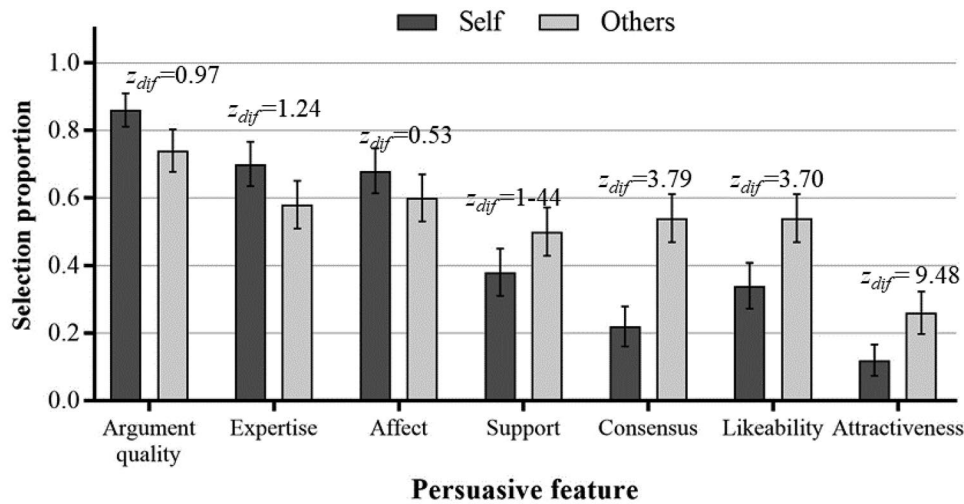
### Selection task

A mean of 3.76 ( $SD=1.39$ , Med = 4.00) persuasive features were selected by each participant to convince the self and of 4.26 ( $SD=1.35$ , Med = 4.00) were selected to convince others.

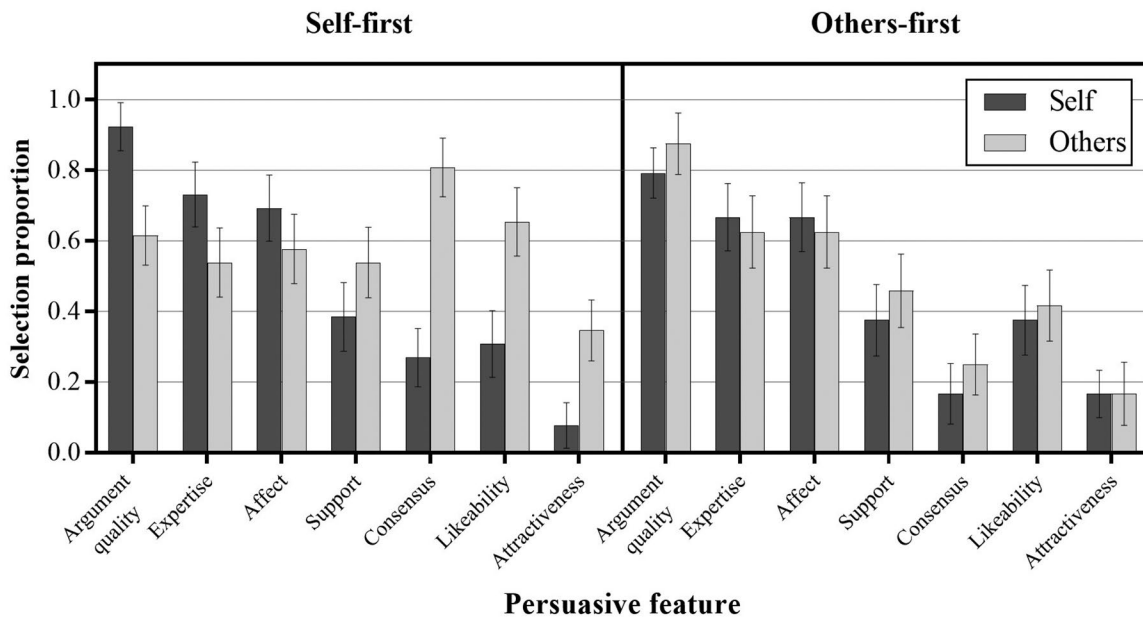
As expected, *Argument-quality* was the most frequently selected feature (80%), followed by “diagnostic features” such as *Expertise* and *Affect*, both of which were selected by 64% of participants. Non-diagnostic features, such as source *Attractiveness* (19%) and *Consensus* (38%), were selected by a smaller proportion of participants. Both *Likeability* and *Support* (number of arguments) were selected by slightly fewer than half (44%) of the participants (Figure 2).

For each persuasive feature, self-other differences were analyzed based on the frequency with which they were chosen to persuade oneself versus others. The analysis of *standardized proportion differences* (see Figure 2) suggests that three persuasive features were more frequently selected to convince others than to convince oneself: *Attractiveness*, *Consensus*, and *Likeability*. These were the same three persuasive features evaluated in Study 1 as non-diagnostic. Notably, those features characterized in Study 1a as “diagnostic features” were less prone to self-other differences: *Argument-quality*, *Expertise*, *Affect* and *Support*.

The order of measurement influenced the self-other differences found for each persuasive feature (see Figure 3). The pattern of self-other differences previously identified is more evident in the self-first condition,  $\chi^2(7) = 69.04$ , than in the others-first condition,  $\chi^2(7) = 1.89$  (see Figure 3). This is likely because the order of assessment



**Figure 2.** Selection proportion of the seven persuasive features for self vs. others.  
**Note.** Error bars are standard errors;  $Z_{dif}$  = Standardized Proportion Differences).



**Figure 3.** Selection proportion of the seven persuasive features for self and others, for two measurement orders (self-first and others-first).  
**Note.** Error bars are standard errors.

changes the selection of features assumed to be made by “others”,  $\chi^2(7) = 51.54$ , but not those made by the participant for the self,  $\chi^2(7) = 5.69$ .

### Perceived persuasibility ratings

Self-other differences were also examined regarding the perceived effectiveness of each persuasive feature, considering the order of assessment.

The self-others difference detected ( $d=0.26$ ) documented a third-person effect, showing that all

evaluated features were perceived as more effective in convincing others ( $M=3.70$ ,  $SD=0.38$ ) than the self ( $M=3.36$ ,  $SD=0.36$ ).

Differences in the perceived persuasibility of the set of persuasive features are clear ( $\eta_p^2 = 0.42$ ), suggesting variations in their perceived efficiency. As expected, *Argument-quality* was perceived as the most effective feature and *Attractiveness* as the least effective (see Table 1).

As expected, self-other differences were not the same for all persuasive features. Figure 4 clearly shows that

self-other differences are stronger for all the previously identified non-diagnostic cues, which are perceived as more effective in persuading others than oneself (all  $d$ 's  $< -0.67$ ). The difference is reduced for *Affect*. Figure 4 also documents that self-other differences can occur in a different direction. The perceived influence of both *Argument-quality* and *Expertise* showed the opposite tendency to those found for non-diagnostic features ( $d$ 's  $> 0.33$ ), as they were perceived as more effective in influencing the self than others.

There are also differences in the self-other assessment order regarding perceived persuasibility ( $\eta_p^2 = 0.15$ ), occurring because the ratings were more extreme in the self-first condition ( $M = 3.64$ ,  $SD = 0.43$ ) than in the others-first condition ( $M = 3.41$ ,  $SD = 0.45$ ). Changes promoted by order occurred mainly for others-ratings ( $d = 0.99$ ), with less evidence for the self-ratings ( $d = 0.35$ ). Others were perceived as more susceptible to all persuasive features after participants evaluated self-susceptibility. Figure 5 shows that the persuasive features factor qualified this order effect ( $\eta_p^2 = 0.08$ ), as the pattern of differences previously described is clearer when self is measured first. Only in the self-first condition does the pattern of data suggest higher perceived bias (indexed by self-other differences) for aggregated non-diagnostic than for aggregated diagnostic features ( $d = 2.45$ ).

## Discussion

In sum, as expected, results from both studies showed that persuasive features vary in their perceived

diagnosticity of the validity of a claim and in their persuasibility.

Regarding perceived diagnosticity: There are generally clear differences between positive and negative instances of persuasive features, suggesting that the positive ones are perceived as more diagnostic than negatively framed features (exceptions being *Likeability* and *Attractiveness*, which generally had lower diagnosticity, and *Affect*). Higher levels of diagnosticity were attributed to positively framed *Expertise* and *Support*, as well as to *Affect*, while lower levels were attributed to *Attractiveness*, *Likeability*, and *Consensus*.

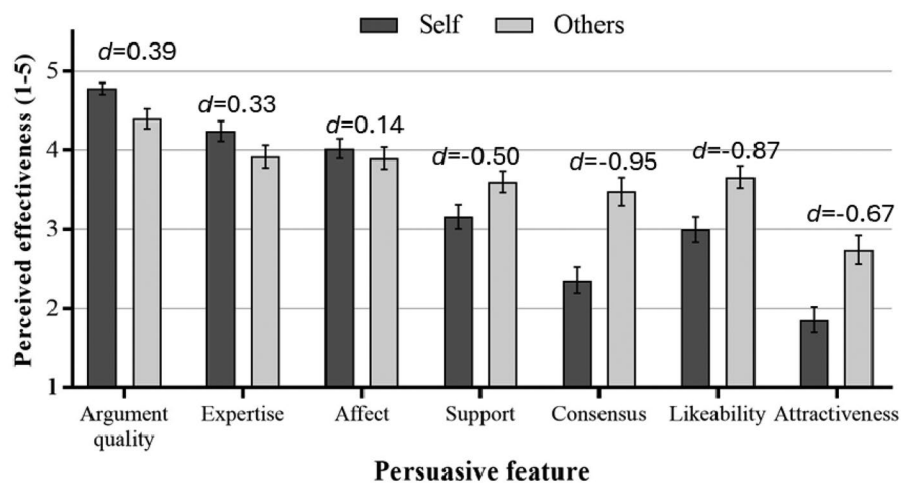
Regarding perceived persuasibility, *Argument quality* was perceived as the most effective feature, followed by features perceived as diagnostic, while non-diagnostic features (*Attractiveness*, *Likeability*, and *Consensus*) were perceived as less effective, with *Attractiveness* being the least effective.

Additionally, results corroborate expectations of third-person effect by suggesting self-other differences in perceived persuasibility of persuasive features. Using self-other differences to index the level of perceived bias of the persuasive feature, the data indicate that non-diagnostic features were also perceived as having an undesirable influence on the self. However, the self-other differences reverse for both *Argument quality* and *Expertise*, suggesting that participants perceived others to be biased for not considering the informative value of these features.

Importantly, the results indicate that participants perceive others, relative to themselves, as insensitive

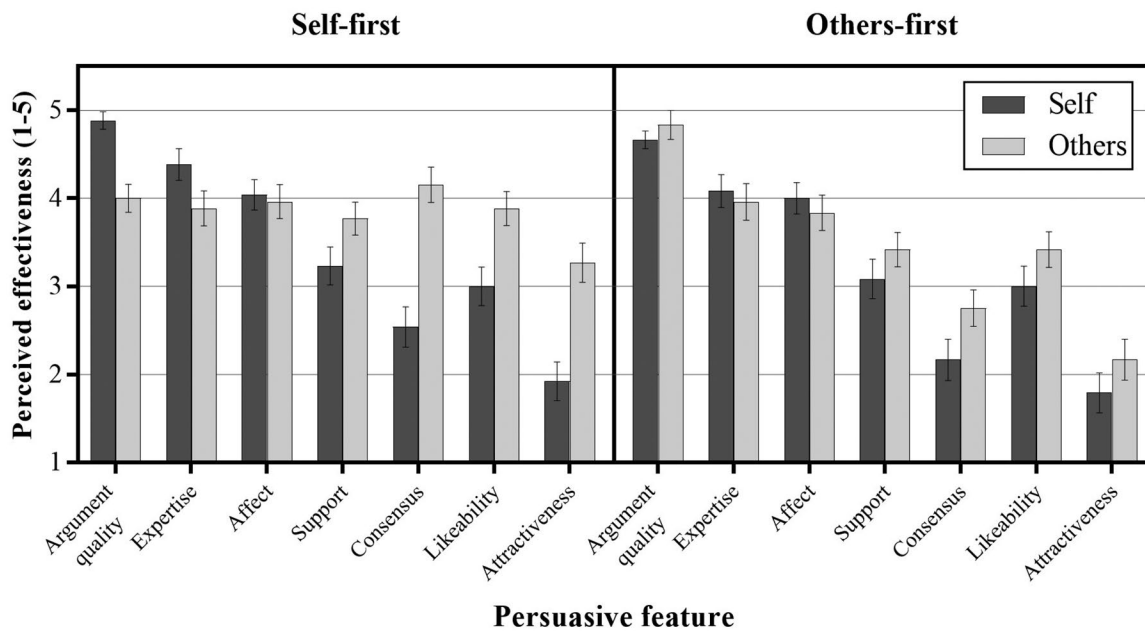
**Table 1.** Perceived persuasibility of the persuasive features (Study 1b).

	Arg. Quality	Expertise	Affect	Support	Consensus	Likeability	Attractiveness
Mean	4.59	4.08	3.96	3.38	2.92	3.33	2.30
Standard deviation	0.54	0.83	0.80	0.92	1.03	0.92	1.06



**Figure 4.** Perceived persuasibility of the persuasive features for self and others (Study1b).

**Note.** Error bars are standard errors;  $d$  = Cohen  $d$  for self-other comparisons.



**Figure 5.** Perceived persuasibility of persuasive features for self and others, for two measurement orders (self-first and others-first).

**Note.** Error bars are standard errors

to the varying diagnostic values of the persuasive features. As such, our data may contribute to the third-person effect by suggesting that perceived influence can relate to a lack of sensitivity to the diagnostic value of persuasive features. We will address this intra-individual relationship in Study 2.

In Study 2, we will also aim to overcome two unexpected results. First, the self-other differences order effect: expected self-other differences were more pronounced in the self-first condition than in the others-first condition. While references to such order-of-measurement effects are limited (Price & Tewksbury, 1996) within the studies of the third-person effect (see Perloff, 2002), the effects align with asymmetries found in self-other judgments of similarity (Srull & Gaelick, 1983). This asymmetry is suggested to occur because the self serves as a default habitual reference point, a phenomenon that future studies may also consider in the assessment of the third-person effect. Crucially, although the mentioned order effect did not compromise the informativeness of the obtained results, it imparts a different level of validity to the self-other index. Therefore, in Study 2, to mitigate order effects, we collect both measures simultaneously.

The second outcome we aim to overcome is that all diagnosticity measures provided scores below the midpoint of the scale. Although this may indicate that the diagnostic value of the persuasive toolbox is low, it may also have occurred due to how we assessed diagnosticity. The use of extreme statements as items

may have contributed to these results, supporting the need to replicate these results with a different measure, allowing the generalization of the conclusions. We will address this issue in Study 2.

## Study 2

Study 2 aggregates the two measures—perceived diagnosticity and persuasibility of the persuasive features—both for self and others, and uses a North American sample from M-Turk, allowing to replicate the previous results with a different population and with different measures. Importantly, by assessing both measures with the same participants, the study allows for testing, at an individual level of analysis, how perceived persuasibility is related to perceived diagnosticity, examining the hypothesis that perceived self-other differences in persuasibility may be rooted in corresponding self-other differences in perceived diagnosticity.

## Participants and procedure

A total of 124 M-Turk workers accessed our survey. Twenty-three participants were excluded for not matching age criteria. The final sample, aged between 17–44 years ( $M=30.01$ ,  $SD=6.85$ ), comprised 53 males and 48 females.

Participants were asked to evaluate the extent to which they believed that certain features of persuasive messages presented by advertisers, politicians, and

**Table 2.** Perceived diagnosticity of persuasive features (Study 2).

	Arg. Quality	Expertise	Affect	Support	Consensus	Likeability	Attractiveness
Mean	5.02	5.36	4.15	3.71	4.12	3.8	3.47
Standard deviation	1.38	1.43	1.29	1.59	1.45	1.53	1.62

attorneys could convince *them* and *other* people of a claim.

Seven persuasive features (positively framed) were individually presented through a short statement in the center of the screen on different pages. For all participants, the first feature was “*The good quality of the arguments used in the message*” (Argument-quality), to set the same standard of comparison for all participants. The other six persuasive features were then randomly presented: “*The high expertise of the source of the message*” (Expertise), “*The high number of arguments used in the message*” (Support), “*The fact that the message makes me feel good*” (Affect), “*The high number of people who agree with the message*” (Consensus), “*The likeability of the source of the message*” (Likeability), and “*The beauty/attractiveness of the source of the message*” (Attractiveness).

Within each screen, participants rated the extent to which they believed they and other people would be influenced by each feature. Below the question, two rating scales were presented, anchored in 1 (*I/Others would surely not be influenced by this feature*) to 7 (*I/Others would surely be influenced by this feature*). The two rating scales were presented on the same screen. After rating their own and others’ perceived susceptibility, the next persuasive feature was presented on a new screen.

Participants then completed a filler task that consisted of rearranging a set of colored bars by color gradient intensity using the drag/drop option to move bars (this task took 2-5 minutes).

After completing the filler task, a new screen informed participants that the same or different persuasive features would be presented again, but now their task was to indicate the extent to which they personally, and others, believed that a message containing the specific persuasive feature was likely to be more accurate than a message that did not contain that specific feature (perceived diagnosticity).

The *Argument-quality* statement was the first to be presented, randomly followed by the other features in the center of the screen. Below each feature, participants responded to two items. The first one stated, “For me”: 1-*This feature indicates nothing about the message being true*; to 7-*This feature indicates that the message is definitely true*. The second item stated, “For others” 1-*This feature indicates nothing about the message being true*; to 7-*This feature indicates that the message is definitely true*. The ratings regarding self and others were made on the same screen.

## Results

Ratings of perceived diagnosticity and perceived persuasibility of each persuasive feature were both analyzed in a within-participants design: 7 (Persuasive features) x 2 (Self vs. Others).

### Perceived diagnosticity

Result show perceived self-other differences ( $\eta^2_p = 0.38$ ), with participants attributing higher diagnostic value to all features for others ( $M = 4.63$ ,  $SD = 0.27$ ) compared to themselves ( $M = 3.83$ ,  $SD = 0.31$ ).

Furthermore, the results replicate Study 1a, by indicating differences in the perceived diagnosticity of each persuasive feature,  $\eta^2_p = 0.28$  (see Table 2). Both *Expertise* and *Argument-quality* were perceived to be more diagnostic than all other feature (all Cohen’s  $d$ ’s  $> .10$ ). Contrary to Study 1a, *Support* differed little from other persuasive features (all Cohen’s  $d$ ’s  $< .06$ ). *Affect* was the feature that most differed from diagnostic features (Cohen’s  $d > .13$ ), being here perceived as a “non-diagnostic feature”.

Self-other differences varied depending on the persuasive features ( $\eta^2_p = .38$ ), and their pattern clarifies the differences found in Study 1a. Figure 6 shows that self-other differences were evident for non-diagnostic features.

Additionally, differences in perceived diagnosticity between persuasive features were attenuated when considering others, with all pairwise comparisons between persuasive features having a median  $d = 0.04$ , compared to  $d = 0.10$  for self-ratings. For self-ratings, *Expertise* stands out as the most diagnostic feature, differing little from *Argument quality* ( $d = 0.04$ ). These two diagnostic features differed less from *Support* ( $d = 0.16$ ;  $d = 0.11$ ) than from all other features (Median  $d = 0.21$ ). Thus, these results generally replicated Study 1a, particularly when considering only positive instances of the persuasive features.

### Perceived persuasibility

There were clear differences in the perceived persuasibility of the different persuasive features ( $\eta^2_p = .16$ ), which replicates the results of Study 1a. This indicates that, across both self and others, *Argument-quality* and *Expertise* (see Table 3) were perceived to be more persuasive than the other features.

Once again, the results demonstrated general self-other differences ( $\eta^2_p = .30$ ), indicating evidence of the third-person effect, given that all persuasive features were perceived as more effective in convincing others ( $M=5.27$ ,  $SD=2.43$ ) than the self ( $M=4.73$ ,  $SD=1.95$ ).

Most importantly, we observed self-other differences dependent upon persuasive features ( $\eta^2_p = .30$ ; see Figure 7). Self-other differences emerged for *Affect*, *Attractiveness*, *Consensus* and *Likeability*, indicating that others were perceived as more susceptible to these influences than the self. Similarly to Study 1b, differences between self-perception and perceptions of others were inverted for *Argument Quality* and *Expertise* (and null for *Support*), as participants tended to perceive themselves as more responsive to these features compared to others.

Furthermore, differences in perceived persuasibility were attenuated for others; pairwise comparisons between persuasive features showed a median  $d = 0.02$  for others, compared to a median  $d = 0.10$  for self.

In summary, feature categorized as non-diagnostic in Study 1a (*Likability*, *Attractiveness*, and *Consensus*) and *Affect* were perceived as sources of bias, influencing others more than the self. Conversely, features perceived as diagnostic showed the opposite pattern.

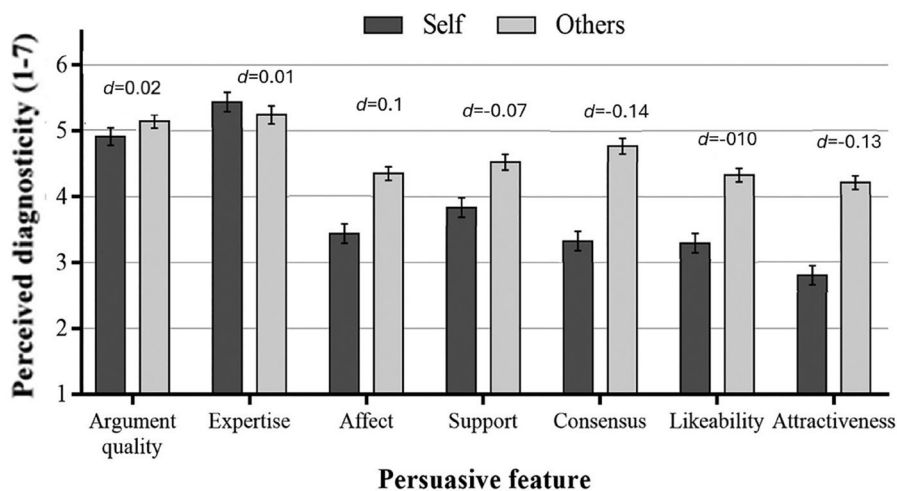
### Relationship between perceived diagnosticity and persuasibility

Table 4 summarizes the relationships between ratings of perceived diagnosticity and persuasibility for both

self and others. All reported correlations were positive, suggesting that judgments of perceived persuasibility of a feature are related to its perceived diagnosticity. Importantly, these relationships were stronger for diagnostic features compared to non-diagnostic features, and stronger for self-perceptions compared to perceptions of others.

We further addressed whether self-other differences in perceived diagnosticity accounted for the previously observed differences between perceived bias of diagnostic features (*Argument-quality*, *Expertise*, *Support*) and non-diagnostic features (*Likability*, *Attractiveness*, *Consensus*, and *Affect*), as measured by self-other differences in persuasibility. If so, we hypothesized that others' perceived lesser ability to recognize the level of diagnosticity compared to oneself would be associated with their greater bias.

Table 5 shows the correlations between these two indices, revealing no consistent pattern between diagnostic and non-diagnostic persuasive features. Interestingly, the perceived bias favoring the self in argument quality appears to be unrelated to the self's ability to evaluate the diagnosticity of this feature. For all other features (albeit to a lesser extent for likeability), greater discrepancies between self and others in assessing a feature's diagnostic value are positively associated with increased bias in persuasiveness.

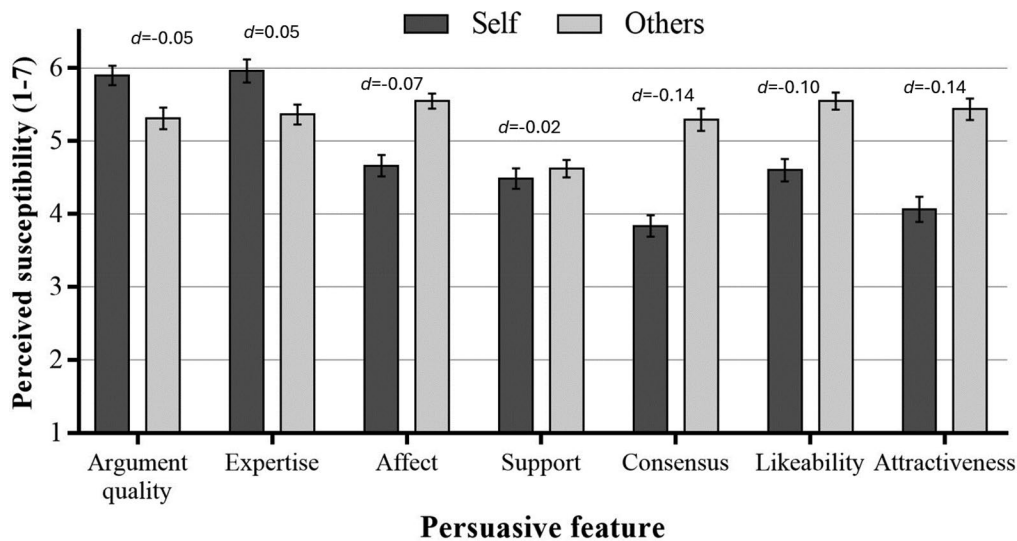


**Figure 6.** Perceived diagnosticity of each persuasive feature the self and others.

**Note.** Error bars are standard errors;  $d$  = Cohen  $d$  for self-other comparisons

**Table 3.** Perceived persuasibility of the persuasive features (Study 2).

	Arg. Quality	Expertise	Affect	Support	Consensus	Likeability	Attractiveness
Mean	5.52	5.57	4.51	5.11	4.55	5.03	4.70
Standard deviation	1.23	1.26	1.10	1.11	1.28	1.21	1.42



**Figure 7.** Perceived persuasibility of each persuasive feature for the self and other (Study 2).

**Note.** Error bars are standard errors;  $d$  = Cohen  $d$  for self-other comparisons

**Table 4.** Correlations between ratings of perceived persuasibility and diagnosticity for each persuasive feature.

	Argument Quality	Expertise	Affect	Support	Consensus	Likeability	Attractiveness	Mean
Self	.40	.46	.36	.40	.39	.38	.33	<b>.39</b>
Others	.39	.41	.21	.35	.25	.39	.32	<b>.33</b>

**Table 5.** Correlations between self-other differences in perceived persuasibility (Bias) and diagnosticity for each persuasive feature.

	Arg. Quality	Expertise	Affect	Support	Consensus	Likeability	Attractiveness
Correlation between bias and diagnosticity	.08	.43	.43	.43	.33	.20	.38

## Discussion

Results from Study 2, conducted with a different sample and methodology, replicate the primary findings of Study 1, indicating that individuals attribute higher diagnostic value to *Argument quality* and *Expertise* compared to *Likability*, *Attractiveness*, and *Consensus*. The findings regarding *Support* and *Affect* are less clear.

Moreover, the results align regarding the perceived persuasibility of each persuasive feature and the distribution of self-other differences in this regard. Thus, the findings suggest that features perceived as non-diagnostic are also perceived as sources of bias.

Additionally, this study provides evidence that participants perceive themselves as less susceptible to the influence of non-diagnostic features and more susceptible to the influence of diagnostic features compared to others.

In this study, we also focused on individual differences in how a feature's diagnosticity and persuasibility are perceived, addressing their relationship. The results demonstrate a positive relationship between diagnosticity and persuasibility, indicating that participants perceive others to be more influenced than

themselves by persuasive features, potentially due to a perceived lack of recognition of the features' diagnostic value. Notably, this belief contributes independently to social consensus, as it holds across persuasive features perceived to have varying levels of diagnosticity.

Furthermore, self-other differences in perceived diagnosticity exhibit a similar relationship with bias in persuasibility across persuasive features. Thus, an individual's perception of bias associated with a specific feature is linked both to the socially shared perception of that persuasive feature's characteristics and their personal belief that others underestimate its diagnostic value. An exception is observed for argument quality, which may indicate a unique status for this persuasive feature.

## General discussion

The main goal of this paper was to characterize the persuasive cognitive toolbox by examining a set of persuasive features known to influence people's responses to a persuasive attempt, focusing on perceived diagnosticity and persuasibility bias.

The studies conducted addressed different populations and proposed various methods for assessing persuasibility and perceived diagnosticity, yet their main results converge across these differences.

The results provide valuable insights into some of the tools in our cognitive toolbox. They suggest that positive instances of persuasive features such as likability, attractiveness, and consensus are generally perceived as non-diagnostic and sources of bias. However, features like expertise and support may be perceived to possess diagnostic value comparable to that attributed to argument quality. Additionally, individuals perceive these cognitive tools differently from others, both in terms of their perceived diagnosticity and persuasiveness. Features deemed more diagnostic by participants are also perceived as having a greater impact on their own attitudes compared to others'. Conversely, non-diagnostic features are perceived as having a stronger influence on others than on the self.

Moreover, these studies demonstrate that individual ratings of perceived diagnosticity correlate positively with perceived persuasiveness of the persuasive features. Participants who perceive a feature as more diagnostic also tend to perceive it as more persuasive. This relationship bears relevance to the third-person effect, specifically highlighting differences between self-perception and perceptions of others.

Below, we discuss how each of these sets of findings contributes theoretically to the existing literature.

### ***Persuasive features diagnosticity and persuasibility***

Here, we assume that several persuasive features, available in our persuasive toolbox, assist both rational and shortcut processes, in supporting attitude change. As anticipated, argument quality was perceived as highly diagnostic of claim validity, consistent with literature suggesting it instills judgmental confidence and promotes elaborate processing (Petty & Cacioppo, 1986; Chaiken, 1980). However, in contexts where efficiency outweighs optimization, people often rely on non-elaborative processes driven by salient contextual cues (e.g., Chaiken, 1980, 1987) that become accessible tools in our persuasive cognitive toolbox through learning. Our findings reveal not only varying levels of perceived diagnosticity across these features but also that some (e.g., expertise) achieved levels of perceived diagnosticity similar to argument quality. The observation that expertise is perceived as highly diagnostic aligns with evidence linking expertise to perceptions of reliability, goodwill, safety, qualification, and dynamism (Berlo et al., 1970). Future research

should further clarify whether similar perceptions extend to "Support" and other persuasive features not examined here (e.g. perceived credibility).

As anticipated, some persuasive features were perceived as non-diagnostic of claim validity. Participants not only attributed lower diagnosticity to attractiveness and likeability but also perceived them as sources of bias. These results are consistent with evidence suggesting that people tend to avoid exposure to such features (Wegener & Petty, 1998; Wilson & Brekke, 1994) and engage in bias correction processes when they are present (see assumptions of the Flexible Correction Model; Petty et al., 1998; Wegener & Petty, 1985, 1997). Some studies (e.g., Isbell & Wyer Jr, 1999; Wegener & Petty, 2013) suggest similar correction processes occur for affect as a persuasive feature, which was not consistently perceived as a source of bias in our studies. Future research could provide more direct evidence on how individuals' perceptions of diagnosticity related to perceived bias influence the likelihood of correcting their influence.

Surprisingly, although social consensus is often considered reflective of objective reality (Asch, 1952; Festinger, 1954; Mackie, 1987), our results indicate it is perceived as a non-diagnostic feature. However, we observed a stronger self-other difference in perceived diagnosticity regarding consensus, suggesting that others consider consensus a cue for validity. This apparent paradox may stem from consensus encompassing both informational and normative meanings (Deutsch & Gerard, 1955; Mackie, 1987), where different meanings may evoke distinct effects (Bohner et al., 2008). Thus, further data are needed to conclusively determine the diagnostic nature of consensus.

When characterizing the perceived diagnosticity of the persuasive toolbox, it is crucial to acknowledge the asymmetry between positive and negative instances of these tools. This asymmetry has received little attention in literature (e.g., van Leeuwen & Macrae, 2004) and generally favors positive instances, contrasting with negative instances that have been shown to hold greater relevance in other settings (see Baumeister et al., 2001; Skowronski & Carlston, 1987). This discrepancy may arise from the greater ambiguity inherent in negative instance information, which convey the absence of a feature (e.g., it is not an expert), contrasting with the presence in positive instances (e.g., it is an expert). Previous research has documented that ambiguous information is perceived as less diagnostic (Wyer, 1974).

The variability in perceived diagnosticity and persuasibility of different persuasive features likely contributes to their distinct roles in persuasion (Petty

et al., 1993; Petty & Cacioppo, 1986; Petty & Wegener, 1998; Tormala et al., 2009) in various ways. For instance, individuals who perceive themselves as less influenced by non-diagnostic features (and thus more resistant to them) are likely to have their attitude certainty or confidence (see Rucker & Petty, 2004; Tormala et al., 2006; Tormala et al., 2007) influenced by the perceived diagnosticity of a persuasive feature. Moreover, low confidence evaluations (Maheswaran et al., 1992) are likely to prompt the need for additional information to counteract associated costs (Chaiken, 1980; Van Wallendael & Guignard, 1992), indicating that cues perceived as diagnostic may deter engagement in this regard. Future research should investigate how the perceived diagnosticity of other persuasive features relates to attitude confidence, particularly given clear evidence that argument quality, the highest diagnostic tool in the cognitive toolbox, promotes greater certainty in attitude (Petty et al., 1995).

The persuasive features studied are also known to function as “arguments” under conditions of high elaboration (Petty & Cacioppo, 1986), influencing individuals’ thoughts (biasing elaboration of message arguments). Perceived diagnosticity of a feature is likely to determine the quantity and quality of thoughts generated in response to it. Future research could explore this hypothesis not only by examining features perceived as generally diagnostic but also by considering other contextual factors where additional features may emerge as ecologically diagnostic.

The tools within our persuasive cognitive toolbox are typically assumed to directly prompt agreement with a message, bypassing deliberative processes (e.g., Chaiken, 1980; Gigerenzer & Gaissmaier, 2011; Simon, 1963). However, the role of a tool’s diagnosticity in this direct influence remains unclear. Firstly, some studies suggest that perceived diagnosticity is primarily relevant for sustaining more elaborate thinking (see Miniard et al., 1991, 1992). Secondly, unless readily accessible, diagnosticity may exert minimal influence under conditions of low elaboration. While direct access to perceived diagnosticity is not guaranteed (see Diederich & Trueblood, 2018; Unkelbach et al., 2019), some authors (see Chaiken, 1980, 1987) have proposed that this might be the case. This is because access to heuristic confidence is assumed to either encourage or deter thoughtful processing. In this scenario, the perceived diagnosticity of a persuasive feature would be presumed to moderate levels of elaboration, particularly for individuals with a low need for cognition, who ‘do not like to think’ (Cacioppo & Petty, 1982). Access to a simpler yet diagnostic pathway would serve as a means to regulate

engagement in deeper processing (e.g., Garcia-Marques & Mackie, 2001; Thompson et al., 2013).

### ***Self-other differences***

Our findings provide evidence of a persuasive ‘bias blind spot’ (Pronin et al., 2002), where individuals appear unaware of their own susceptibility to the influence of non-diagnostic features (such as attractiveness, consensus, and likeability; e.g., Mello et al., 2020; Roskos-Ewoldsen et al., 2002), perceiving them as sources of bias primarily for others and not for themselves. These self-other differences associated with non-diagnostic features exhibit the expected pattern when bias is assumed (see Sun et al., 2008 meta-analysis), but we also observed a reversal with diagnostic features. Hence, these self-other differences are aligned with results observed in the third-person effect, where others are viewed as more susceptible to persuasion than oneself (see Paul et al., 2000), particularly when the persuasion is deemed undesirable (e.g., David & Johnson, 1998). This perception reverses when persuasion is perceived as socially desirable (e.g., Duck & Mullin, 1995).

Our data extend third-person effect findings by also highlighting self-other differences regarding perceived diagnosticity. Moreover, these differences appear related to self-other differences observed in perceived persuasibility. Both these self-other differences likely reflect a general tendency for individuals to evaluate themselves more favorably than others—a process underlying social comparison (Festinger, 1954). This tendency contributes to several self-serving bias effects, such as attributing internal causes to others’ failures more than one’s own failures and reversing this pattern with regard to success (Miller & Ross, 1975). It also aligns with better-than-average effect, where individuals tend to overestimate their own qualities compared to others’ (see Alicke et al., 1995). Therefore, our results align with previous research, suggesting that individuals are motivated to preserve self-integrity in the outcomes of persuasion attempts (see Grillo & Pizzutti, 2021).

### ***Limitations and future implications***

Several aspects of our studies warrant attention for their potential contributions to the results. One key aspect is the comparative context in which each persuasive feature was evaluated. While our experimental procedures ensured random evaluation order of features, the fixed evaluation of argument quality as the most extreme feature may have influenced subsequent

judgments. This influence should be considered in future research. Similarly, a comparable issue arises with the measurement of differences for self and others. Study 1 highlighted the impact of measurement order, prompting adjustments in Study 2 to mitigate these effects by making comparisons between self and others explicit to participants.

In our studies, we altered how we defined both persuasibility and perceived diagnosticity. Initially, we assessed diagnosticity using distinct items for each persuasive feature (small Likert scales). In Study 2, we standardized the evaluation of diagnosticity for all features using the same ratings. This standardized measure linked the presence of a cue with the perception of truth in the message. Importantly, findings were consistent across both approaches. However, future research should investigate the reliability of this measure in predicting truthfulness and explore the predictive power of perceived persuasibility to further our conclusions.

We also highlight that the way we define and present each of the persuasive characteristics to the participants will always be a source of variability. The presentation varied among the studies and could vary even more, from abstract forms to concrete cases. Future studies should consider this variability, as well as the potential influence of context on participants' evaluations, given that perceived diagnosticity is likely ecologically modulated.

At a theoretical level, it is also important to clarify the concept of the persuasive cognitive toolbox used in this paper. While inspired by Gigerenzer (2020), we do not adopt the same theoretical perspective. Instead, we use the concept in line with dual-route/process frameworks of persuasion models such as the Elaboration Likelihood Model (Petty & Cacioppo, 1986) and the Heuristic-Systematic Model (Chaiken, 1980). Our conceptualization of the persuasive cognitive toolbox focuses on facilitating processing under conditions of low elaboration, likely involving a global process that exhibits less strategic rationality and adaptiveness compared to Gigerenzer's proposals (2020; Gigerenzer et al., 2011). Individuals may not opt for what is "more rational for a given context". Instead, it is more probable that simple accessibility (e.g., Roskos-Ewoldsen et al., 2002) accounts for the selective use of a straightforward rule, and as stated above, only if it is easier to assess, may the diagnostic value of a persuasive feature be taken into consideration within a non-elaborative context.

Finally, we want to emphasize that our work is likely to help advance the field toward effective application in different directions. Reality is multifactorial, and within any specific persuasive setting, we should

expect individual differences, as people selectively attend to one or another factor. Acknowledging the individual's "mental toolbox" is likely of great importance. The toolbox analogy states the functionality of our belief system, and our results inform us about that functionality. By knowing that, metaphorically, individuals hold a "nail" in their hand, we can nudge their cognition by presenting them with a "hammer." An advertisement or other type of persuasive setting will gain a better understanding of the consequences of the type of cognitive resources they prime. A second direction anchors on the fact that by knowing about individuals' beliefs regarding the persuasive toolbox, we may better develop preventive strategies against undesirable influences (e.g., peer influence, addictive behavior, etc.). Future research should help to understand how we may "empower" individuals' toolboxes to develop attitude resistance abilities. A third direction focuses on the identified perceived self-other differences supporting the perception of the use of different tools in the persuasive toolbox. This uncovers potential biases in the decision-making process when developing prevention and marketing campaigns. When decision-makers contemplate what will influence others, they should prevent themselves from introspection.

## Conclusion

This paper introduces a novel perspective on persuasive features, framing them as cognitive tools that facilitate processing within persuasive contexts. The study delves into the perceived diagnosticity of these features and investigates how this variability connects directly to their perceived persuasibility over oneself and others. In doing so, it informs about the differences between various persuasive features, distinguishing those perceived as "diagnostic" from "non-diagnostic" cues. Importantly, the findings reveal that individuals perceive themselves as more sensitive to diagnostic features and less susceptible to non-diagnostic features compared to others. These insights hold implications for designing more effective interventions associated with attitude change, persuasion, or behavioral decision-making (e.g., prevention and marketing campaigns), by focusing intervention strategies on more diagnostic features rather than simply assuming others' higher sensitivity to sources of bias.

## Acknowledgements

We thank the editor and reviewers for their comments on an earlier version of this manuscript.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

This work is funded with national funds from FCT – Fundação para a Ciência e Tecnologia, I.P., in the context of the project UID/04810/2020.

## References

- Aaker, J., & Maheswaran, D. (1997). The effect of cultural orientation on persuasion. *Journal of Consumer Research*, 24(3), 315–328. <https://doi.org/10.1086/209513>
- Ahluwalia, R. (2002). How prevalent is the negativity effect in consumer environments? *Journal of Consumer Research*, 29(2), 270–279. <https://doi.org/10.1086/341576>
- Albarracín, D., & Mitchell, A. L. (2004). The role of defensive confidence in preference for proattitudinal information: How believing that one is strong can sometimes be a defensive weakness. *Personality & Social Psychology Bulletin*, 30(12), 1565–1584. <https://doi.org/10.1177/0146167204271180>
- Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., & Vredenburg, D. S. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology*, 68(5), 804–825. <https://doi.org/10.1037/0022-3514.68.5.804>
- Asch, S. (1952). Group forces in the modification and distortion of judgments. *Social psychology* (pp. 450–501). Prentice-Hall, Inc.
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, 5(4), 323–370. <https://doi.org/10.1037/1089-2680.5.4.323>
- Berlo, D., Lemert, J., & Mertz, R. (1970). Dimensions for Evaluating the Acceptability of Message Sources. *Public Opinion Quarterly*, 33(4), 563–576. <https://doi.org/10.1086/267745>
- Bohner, G., Dykema-Engblade, A., Tindale, R., & Meisenhelder, H. (2008). Framing of majority and minority source information in persuasion: When and how “consensus implies correctness. *Social Psychology*, 39(2), 108–116. <https://doi.org/10.1027/1864-9335.39.2.108>
- Bohner, G., Moskowitz, G. B., & Chaiken, S. (1995). The interplay of heuristic and systematic processing of social information. *European Review of Social Psychology*, 6(1), 33–68. <https://doi.org/10.1080/14792779443000003>
- Bohner, G., Ruder, M., & Erb, H. (2002). When expertise backfires: Contrast and assimilation effects in persuasion. *The British Journal of Social Psychology*, 41(Pt 4), 495–519. <https://doi.org/10.1348/014466602321149858>
- Briñol, P., Rucker, D. D., Tormala, Z. L., & Petty, R. E. (2004). Individual differences in resistance to persuasion: The role of beliefs and me. In J. T. Jost, M. R. Banaji, & D. A. Prentice (Eds.), *Perspectivism in Social Psychology: The yin and yang of scientific progress* (pp. 37–51). American Psychological Association. <https://doi.org/10.1037/10750-004>
- Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42(1), 116–131. <https://doi.org/10.1037/0022-3514.42.1.116>
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39(5), 752–766. <https://doi.org/10.1037/0022-3514.39.5.752>
- Chaiken, S. (1987). The heuristic model of persuasion. In M. P. Zanna, J. M. Olson, & C. P. Herman (Eds.), *Social influence: The Ontario symposium* (Vol. 5, pp. 3–39). Lawrence Erlbaum Associates, Inc.
- Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 73–96). The Guilford Press.
- Darke, P. R., Chaiken, S., Bohner, G., Einwiller, S., Erb, H.-P., & Hazlewood, J. D. (1998). Accuracy motivation, consensus information, and the law of large numbers: Effects on attitude judgment in the absence of argumentation. *Personality and Social Psychology Bulletin*, 24(11), 1205–1215. <https://doi.org/10.1177/01461672982411007>
- David, P., & Johnson, M. (1998). The role of self in third-person effects about body image. *Journal of Communication*, 48(4), 37–58. <https://doi.org/10.1111/j.1460-2466.1998.tb02769.x>
- Deutsch, M., & Gerard, H. (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal Psychology*, 51(3), 629–636. <https://doi.org/10.1037/h0046408>
- Diederich, A., & Trueblood, J. S. (2018). A dynamic dual process model of risky decision making. *Psychological Review*, 125(2), 270–292. <https://doi.org/10.1037/rev0000087>
- Duck, J., & Mullin, B. (1995). The perceived impact of the mass media: Reconsidering the third person effect. *European Journal of Social Psychology*, 25(1), 77–93. <https://doi.org/10.1002/ejsp.2420250107>
- Duck, J. M., Terry, D. J., & Hogg, M. A. (1995). The perceived influence of AIDS advertising: Third-person effects in the context of positive media content. *Basic and Applied Social Psychology*, 17(3), 305–325. [https://doi.org/10.1207/s15324834basp1703\\_2](https://doi.org/10.1207/s15324834basp1703_2)
- Eisend, M. (2017). The third-person effect in advertising: A meta-analysis. *Journal of Advertising*, 46(3), 377–394. <https://doi.org/10.1080/00913367.2017.1292481>
- Eveland, W., & McLeod, D. (1999). The effect of social desirability on perceived media impact: Implications for third-person perceptions. *International Journal of Public Opinion Research*, 11(4), 315–333. <https://doi.org/10.1093/ijpor/11.4.315>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7(2), 117–140. <https://doi.org/10.1177/001872675400700202>
- Friestad, M., & Wright, P. (1994). The persuasion knowledge model: How people cope with persuasion attempts. *Journal of Consumer Research*, 21(1), 1–31. <https://doi.org/10.1086/209380>
- García-Marques, T., & Mackie, D. M. (2001). The feeling of familiarity as a regulator of persuasive processing. *Social Cognition*, 19(1), 9–34. <https://doi.org/10.1521/soco.19.1.9.18959>

- Gigerenzer, G. (2020). What is bounded rationality? In *Routledge handbook of bounded rationality* (pp. 55–69). Routledge.
- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, 62(1), 451–482. <https://doi.org/10.1146/annurev-psych-120709-145346>
- Gigerenzer, G. E., Hertwig, R. E., & Pachur, T. E. (2011). *Heuristics: The foundations of adaptive behavior*. Oxford University Press.
- Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge University Press.
- Grillo, T. L. H., & Pizzutti, C. (2021). Recognizing and trusting persuasion agents: Attitudes bias trustworthiness judgments, but not persuasion detection. *Personality & Social Psychology Bulletin*, 47(5), 796–809. <https://doi.org/10.1177/0146167220946197>
- Gunther, A. C., & Mundy, P. (1993). Biased optimism and the third-person effect. *Journalism Quarterly*, 70(1), 58–67. <https://doi.org/10.1177/10776990930700010>
- Hoorens, V., & Ruiter, S. (1996). The optimal impact phenomenon: Beyond the third person effect. *European Journal of Social Psychology*, 26(4), 599–610. [https://doi.org/10.1002/\(SICI\)1099-0992\(199607\)26:4<599::AID-EJSP773>3.0.CO;2-7](https://doi.org/10.1002/(SICI)1099-0992(199607)26:4<599::AID-EJSP773>3.0.CO;2-7)
- Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion*. Yale University Press.
- Isbell, L. M., & Wyer, R. S. Jr. (1999). Correcting for mood-induced bias in the evaluation of political candidates: The roles of intrinsic and extrinsic motivation. *Personality and Social Psychology Bulletin*, 25(2), 237–249. <https://doi.org/10.1177/0146167299025002009>
- Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology*, 75(4), 887–900. <https://doi.org/10.1037/0022-3514.75.4.887>
- Kang, Y. S., & Herr, P. M. (2006). Beauty and the beholder: Toward an integrative model of communication source effects. *Journal of Consumer Research*, 33(1), 123–130. <https://doi.org/10.1086/504143>
- Koernig, S. K., & Page, A. L. (2002). What if your dentist looked like Tom Cruise? Applying the match-up hypothesis to a service encounter. *Psychology & Marketing*, 19(1), 91–110. <https://doi.org/10.1002/mar.1003>
- Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes*, 76(2), 149–188. <https://doi.org/10.1006/obhd.1998.2804>
- Mackie, D. (1987). Systematic and nonsystematic processing of majority and minority persuasive communications. *Journal of Personality and Social Psychology*, 53(1), 41–52. <https://doi.org/10.1037/0022-3514.53.1.41>
- Maheswaran, D., Mackie, D. M., & Chaiken, S. (1992). Brand name as a heuristic cue: The effects of task importance and expectancy confirmation on consumer judgments. *Journal of Consumer Psychology*, 1(4), 317–336. [https://doi.org/10.1016/S1057-7408\(08\)80058-7](https://doi.org/10.1016/S1057-7408(08)80058-7)
- Martin, L., Ward, D., Achee, J., & Wyer, R. (1993). Mood as input: People have to interpret the motivational implications of their moods. *Journal of Personality and Social Psychology*, 64(3), 317–326. <https://doi.org/10.1037/0022-3514.64.3.317>
- Mello, J., Garcia-Marques, T., Briñol, P., Cancela, A., & Petty, R. E. (2020). The influence of physical attractiveness on attitude confidence and resistance to change. *Journal of Experimental Social Psychology*, 90, 104018. <https://doi.org/10.1016/j.jesp.2020.104018>
- Messner, M., Reinhard, M.-A., & Sporer, S. L. (2008). Compliance through direct persuasive appeals: The moderating role of communicator's attractiveness in interpersonal persuasion. *Social Influence*, 3(2), 67–83. <https://doi.org/10.1080/15534510802045261>
- Miniard, P. W., Bhatla, S., Lord, K. R., Dickson, P. R., & Unnava, H. R. (1991). Picture-based persuasion processes and the moderating role of involvement. *Journal of Consumer Research*, 18(1), 92–107. <https://doi.org/10.1086/209244>
- Miniard, P. W., Sirdeshmukh, D., & Innis, D. E. (1992). Peripheral persuasion and brand choice. *Journal of Consumer Research*, 19(2), 226–239. <https://doi.org/10.1086/209298>
- Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin*, 82(2), 213–225. <https://doi.org/10.1037/h0076486>
- Mondak, J. J. (1993). Public opinion and heuristic processing of source cues. *Political Behavior*, 15(2), 167–192. <https://doi.org/10.1007/BF00993852>
- Paul, B., Salwen, M. B., & Dupagne, M. (2000). The third-person effect: A meta-analysis of the perceptual hypothesis. *Mass Communication and Society*, 3(1), 57–85. [https://doi.org/10.1207/S15327825MCS0301\\_04](https://doi.org/10.1207/S15327825MCS0301_04)
- Perloff, R. M. (2002). The third-person effect. In J. Bryant & D. Zillmann (Eds.), *Media effects: Advances in theory and research* (2nd ed., pp. 489–506). Lawrence Erlbaum Associates Publishers.
- Petty, R. E., Cacioppo, J. T., Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. Springer-Verlag New York Inc.
- Perloff, R. M. (1996). Perceptions and conceptions of political media impact: The third-person effect and beyond. In A. N. Crigler (Eds.), *The psychology of political communication*. Michigan University Press.
- Petty, R. E. (2013). Two routes to persuasion: State of the art. *International Perspectives on Psychological Science*, 2(1), 229–247.
- Petty, R. E., & Briñol, P. (2012). The elaboration likelihood model. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 224–245). Sage Publications Ltd. <https://doi.org/10.4135/9781446249215.n12>
- Petty, R. E., & Briñol, P. (2015). Emotion and persuasion: Cognitive and meta-cognitive processes impact attitudes. *Cognition & Emotion*, 29(1), 1–26. <https://doi.org/10.1080/02699931.2014.967183>
- Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. Springer Science & Business Media.
- Petty, R. E., & Cacioppo, J. (1981). *Attitudes and persuasion: Classic and contemporary approaches*. Wm. C. Brown.
- Petty, R. E., & Cacioppo, J. (1984). The effects of involvement on responses to argument quantity and quality:

- Central and peripheral routes to persuasion. *Journal of Personality and Social Psychology*, 46(1), 69–81. <https://doi.org/10.1037/0022-3514.46.1.69>
- Petty, R. E., Haugtvedt, C., & 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). Erlbaum.
- Petty, R. E., Schumann, D., Richman, S., & Strathman, A. (1993). Positive mood and persuasion: Different roles for affect under high- and low-elaboration conditions. *Journal of Personality and Social Psychology*, 64(1), 5–20. <https://doi.org/10.1037/0022-3514.64.1.5>
- Petty, R. E., & Wegener, D. T. (1999). The elaboration likelihood model: Current status and controversies. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 41–72). Guilford Press.
- Petty, R. E., & Wegener, D. T. (2014). Thought systems, argument quality, and persuasion. *Advances in Social Cognition*, 4, 147–161.
- Petty, R. E., & Wegener, D. (1998). Attitude change: Multiple roles for persuasion variables. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., pp. 323–390). McGraw-Hill.
- Petty, R. E., Wegener, D. T., & White, P. H. (1998). Flexible correction processes in social judgment: Implications for persuasion. *Social Cognition*, 16(1), 93–113. <https://doi.org/10.1521/soco.1998.16.1.93>
- Price, V., & Tewksbury, D. (1996). Measuring the third-person effect of news: The impact of question order, contrast and knowledge. *International Journal of Public Opinion Research*, 8(2), 120–141. <https://doi.org/10.1093/ijpor/8.2.120>
- Pham, M. (1996). Cue representation and selection effects of arousal on persuasion. *Journal of Consumer Research*, 22(4), 373–387. <https://doi.org/10.1086/209456>
- Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, 28(3), 369–381. <https://doi.org/10.1177/0146167202286008>
- Reinhard, M. A., Messner, M., & Sporer, S. L. (2006). Explicit persuasive intent and its impact on success at persuasion—The determining roles of attractiveness and likeableness. *Journal of Consumer Psychology*, 16(3), 249–259. [https://doi.org/10.1207/s15327663jcp1603\\_7](https://doi.org/10.1207/s15327663jcp1603_7)
- Richard, F. D., Bond, C. F., Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, 7(4), 331–363. <https://doi.org/10.1037/1089-2680.7.4.3>
- Roskos-Ewoldsen, D. (1997). Implicit theories of persuasion. *Human Communication Research*, 24(1), 31–63. <https://doi.org/10.1111/j.1468-2958.1997.tb00586.x>
- Roskos-Ewoldsen, D. R., Bichsel, J., & Hoffman, K. (2002). The influence of accessibility of source likability on persuasion. *Journal of Experimental Social Psychology*, 38(2), 137–143. <https://doi.org/10.1006/jesp.2001.1492>
- Rucker, D. D., & Petty, R. E. (2004). When Resistance Is Futile: Consequences of Failed Counterarguing for Attitude Certainty. *Journal of Personality and Social Psychology*, 86(2), 219–235; <https://doi.org/10.1037/0022-3514.86.2.219>
- Rule, B., Bisanz, G., & Kohn, M. (1985). Anatomy of a persuasion schema: Targets, goals, and strategies. *Journal of Personality and Social Psychology*, 48(5), 1127–1140. <https://doi.org/10.1037/0022-3514.48.5.1127>
- Simon, H. A. (1963). Experiments with a heuristic compiler. *Journal of the ACM*, 10(4), 493–506. <https://doi.org/10.1145/321186.321192>
- Simon, H. A. (1963). *The heuristic compiler*. Rand Corporation.
- Skowronski, J., & Carlston, D. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology*, 52(4), 689–699. <https://doi.org/10.1037/0022-3514.52.4.689>
- Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior and Human Performance*, 6(6), 649–744. [https://doi.org/10.1016/0030-5073\(71\)90033-X](https://doi.org/10.1016/0030-5073(71)90033-X)
- Sun, Y., Pan, Z., & Shen, L. (2008). Understanding the third-person perception: Evidence from a meta-analysis. *Journal of Communication*, 58(2), 280–300. <https://doi.org/10.1111/j.1460-2466.2008.00385.x>
- Srull, T. K., & Gaelick, L. (1983). General principles and individual differences in the self as a habitual reference point: An examination of self–other judgments of similarity. *Social Cognition*, 2(2), 108–121. <https://doi.org/10.1521/soco.1983.2.2.108>
- Thompson, V. A., Turner, J. A. P., Pennycook, G., Ball, L. J., Brack, H., Ophir, Y., & Ackerman, R. (2013). The role of answer fluency and perceptual fluency as metacognitive cues for initiating analytic thinking. *Cognition*, 128(2), 237–251. <https://doi.org/10.1016/j.cognition.2012.09.012>
- Tormala, Z. L., Briñol, P., & Petty, R. E. (2006). When credibility attacks: The reverse impact of source credibility on persuasion. *Journal of Experimental Social Psychology*, 42(5), 684–691. <https://doi.org/10.1016/j.jesp.2005.10.005>
- Tormala, Z. L., DeSensi, V. L., Clarkson, J. J., & Rucker, D. D. (2009). Beyond attitude consensus: The social context of persuasion and resistance. *Journal of Experimental Social Psychology*, 45(1), 149–154. <https://doi.org/10.1016/j.jesp.2008.07.020> [Mismatch]
- Wegener, D. T., & Petty, R. E. (1998). The naive scientist revisited: Naive theories and social judgment. *Social Cognition*, 16(1), 1–7. <https://doi.org/10.1521/soco.1998.16.1.1>
- Tormala, Z. L., DeSensi, V. L., & Petty, R. E. (2007). Resisting persuasion by illegitimate means: A metacognitive perspective on minority influence. *Personality & Social Psychology Bulletin*, 33(3), 354–367. <https://doi.org/10.1177/0146167206295004>
- Unkelbach, C., Koch, A., Silva, R. R., & Garcia-Marques, T. (2019). Truth by repetition: Explanations and implications. *Current Directions in Psychological Science*, 28(3), 247–253. <https://doi.org/10.1177/0963721419827854>
- van Leeuwen, M., & Macrae, C. (2004). Is beautiful always good? Implicit benefits of facial attractiveness. *Social Cognition*, 22(6), 637–649. <https://doi.org/10.1521/soco.22.6.637.54819>
- Van Wallendaal, L. R., & Guignard, Y. (1992). Diagnosticity, confidence, and the need for information. *Journal of Behavioral Decision Making*, 5(1), 25–37. <https://doi.org/10.1002/bdm.3960050105>
- Wegener, D. T., Dunn, M., & Tokusato, D. (2001). The flexible correction model: Phenomenology and the use

- of naive theories in avoiding or removing bias. In G. B. Moskowitz (Ed.), *Cognitive social psychology: The Princeton symposium on the legacy and future of social cognition* (pp. 277–290). Erlbaum.
- Wegener, D. T., & Petty, R. E. (1997). The flexible correction model: The role of naive theories of bias in bias correction. In *Advances in experimental social psychology* (Vol. 29, pp. 141–208). Academic Press.
- Wegener, D., & Petty, R. (1995). Flexible correction processes in social judgment: The role of naive theories in corrections for perceived bias. *Journal of Personality and Social Psychology*, 68(1), 36–51. <https://doi.org/10.1037/0022-3514.68.1.36>
- Wegener, D. T., & Petty, R. E. (2013). Understanding effects of mood through the elaboration likelihood and flexible correction models. In *Theories of mood and cognition* (pp. 177–210). Psychology Press.
- Wilson, T., & Brekke, N. (1994). Mental contamination and mental correction: unwanted influences on judgments and evaluations. *Psychological Bulletin*, 116(1), 117–142. <https://doi.org/10.1037//0033-2909.116.1.117>
- Wyer, R. (1974). *Cognitive organization and change: An information processing approach*. Lawrence Erlbaum.