Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Experimental Social Psychology 40 (2004) 433–449
Journal of Experimental Social Psychology
www.elsevier.com/locate/jesp

"I am not guilty" vs "I am innocent": Successful negation may depend on the schema used for its encoding☆

Ruth Mayo,[a,*] Yaacov Schul,[a] and Eugene Burnstein[b]

[a] *Department of Psychology, Hebrew University, Mount Scopus 91905, Jerusalem, Israel*
[b] *University of Michigan, Ann Arbor, MI, USA*

Received 30 March 2002; revised 16 July 2003
Available online 28 October 2003

Abstract

Negations (e.g., "Jim is not guilty") are part of our daily language and communication. Linguistic and non-linguistic negations can occur when receivers counter-argue what communicators are saying, when hypotheses are disconfirmed, or through negative cognitive responses and many other social interactive processes. Our study explores how negations are encoded by considering the predictions of two theoretical models. According to the fusion model, the core of a negated message and the negation marker are integrated into one meaningful unit. Thus, Jim in the example might be encoded within the schema "innocence." According to the schema-plus-tag model, a negated message is represented as a core supposition and a negation tag, allowing for dissociation of the two at a later point in time. We compare the two models by examining the nature of inferences that are facilitated by negations. Our results show that the existence of a schema that accommodates the meaning of the original negation is critical in determining how a negation will be encoded. When such a schema is not readily available, processing a negated message facilitates negation–*incongruent* associations, in line with predictions of the schema-plus-tag model. This model is also supported by analyses of respondents' memory. We discuss implications of these findings for the communication of negated information, for discounting theories, and for the assessment of the truth of incoming information.

Introduction

"John is simply not a romantic person," Mary muttered. Before you continue, please try to think of at least two examples of John's behaviors that may have led Mary to this conclusion. Were your examples similar to "He forgets her birthday/wedding anniversary," "He never brings her flowers," "He does not express his love in poetry/love letters/surprises/gifts"? All of these behaviors are actually negated romantic behaviors. In other words, stating that John is not romantic can make one think of romantic behaviors and then negate them. Is this a general effect? What are the boundary conditions? What are its implications?

There is no doubt that negation is part of our daily language and communication. It also comes as no surprise that readers have no difficulty understanding the previous sentence, which contains a negation, as well as this one (which contains two negations). In spite of the greater complexity attributed to the comprehension of negations (Carpenter & Just, 1975; Clark & Chase, 1972; Wason, 1963), people usually succeed in understanding the intended meaning of negations. Yet, we believe that a message which is phrased as a negation can under certain conditions activate associations that are incongruent with the message meaning, and might thus introduce communication errors that may actually lead to inferences opposite to the message's intended meaning. This paper explores whether and how this can happen.

The presence and function of negation in social interactions have puzzled philosophers, linguists, and psychologists throughout the centuries (Jordan, 1998).

☆ The research reported here was supported by grants from the US–Israel Binational Science Foundation (BSF) and the Israeli Foundation Trustees. We would like to thank Liran Rasinski, Yasmin Folder, and Orly Carmi for their help in conducting the experiments and Rachel Giora for comments on an earlier version of this paper.
* Corresponding author.
E-mail address: msmayo@mscc.huji.ac.il (R. Mayo).

0022-1031/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.jesp.2003.07.008

This paper examines whether there are processing and coding differences between messages phrased as negations (''Jim is not guilty'') and those phrased as affirmations (''Jim is innocent''). These might be central for understanding what happens in situations that induce receivers to use negation. For example, negations might be generated as one counter-argues against an untrustworthy source of communication while listening to it or reading it (Eagly, Wood, & Chaiken, 1978; Papageorgis, 1968; Petty & Cacioppo, 1977; Zuwerink & Devine, 2000). More generally, receivers are often informed that some knowledge they have is ill-founded, so that after having learned ''X'' in the past, they should now process ''not-X'' (Fiedler, Walther, Armbruster, Fay, & Naumann, 1996; Hornby, 1974; Johnson, 1988; Johnson, Hashtroudi, & Lindsay, 1993; Johnson & Raye, 1981; Koehler, 1991; Loftus, 1979; Loftus & Palmer, 1974; Ross, Lepper, & Hubbard, 1975; Schul, 1993; Schul & Burnstein, 1985; Schul & Manzury, 1990; Schul & Mayo, 1999; Sellars, 1954; Strack & Bless, 1994). Obviously, these situations involve diverse sets of causal factors that have an impact on how people process the given information. To study the role of negation, the current study simplifies the situations considerably by examining the kind of inferences people access at the time they process affirmations and negations that are given to them explicitly. In this endeavor we took advantage of our participants' native language. In Hebrew, a language without any semantic prefixes or suffixes, a negated concept is simply a concept with the addition of the word ''no'' preceding it. Hence it provides a clear state of literal negation, with little leeway in interpretation. Therefore, we believe that any effect found for this simple form of negation may explain in part why some messages are easier to negate than others, in what cases discounting succeeds or fails, and why some beliefs are easier to change than others.

Let us start by considering what comes to mind spontaneously when one reads the statement, ''Michael says that ever since he came to this country, he has not known peace of mind.'' Does one think spontaneously about anxiety, or about serenity? Consider now our initial example ''John is simply not a romantic person'' and recall if you thought about romantic or unromantic gestures. Put more generally, does the processing of messages expressed as negations facilitate associations that are congruent with the intended meaning of the negation (e.g., unromantic gestures), or associations congruent with what is being negated (e.g., romantic gestures). In the latter case, the ensemble of spontaneously activated associations might leave residual meanings that can lead receivers astray when they are trying to think about the original message. The associative structures activated in processing negations can be explored in the light of two theoretical models, which will be discussed in the following sections.

The schema-plus-tag model

Clark and Chase (1972) suggested that encoding negations requires more time because it entails processing an additional operator, the negation operator, which is separated or dissociated from the message's core concepts. For example, they assume that the negation ''B is not above A'' is represented by the propositions (Not (B above A)). Accordingly, to process the negated message one must first process the core supposition (i.e., ''B above A'') and then negate it. The critical point for our discussion is that the core supposition is processed as a cognitive unit, which is then marked with a negation tag. For this reason we term this the schema-plus-tag model. Under this model the core (''B above A'') could be dissociated from the negation tag at a later time. There are several lines of evidence consistent with this model:

Just and Carpenter (1976) tracked respondents' eye movements during a verification task. They showed that respondents tend to focus on the core supposition more than on its negation. For example, when participants were presented with a probe saying ''is not north,'' referring to the location of a plus sign, their eye movements were focused on the north position rather than the south one. While this might be consistent with the logical way to test the truth value of negated information, it nevertheless implies that the focus of attention is on the information that is being negated.

Fiedler et al. (1996) asked their participants a series of questions about the interior of an apartment they had seen on a videotape. Some of the questions referred to objects that were not present in the apartment. Fiedler et al. reported that while participants correctly denied seeing the absent objects when they were questioned immediately afterwards, they mistakenly recognized these objects as having appeared in the video when asked to recognize them after a 20-min distraction task. Importantly, the extent of their mistaken memory was significantly higher than the rate of false alarms for other objects that had not appeared in the video and were not mentioned in the questions. Thus, it seems that thinking about and providing a negative response (e.g., ''there was no hat rack in the apartment'') led to a false memory (e.g., there was a hat rack in the apartment). Fiedler et al. (1996) attributed the memory intrusions of absent objects to constructive memory, suggesting that people encode negated messages by adding a tag denoting negation to the core of the message. Finally, somewhat indirect support for the schema-plus-tag model emerges from studies showing that, when negation is used to deny a plausible misconception, comprehension is facilitated by prior consideration of the preconception (Johnson-Laird & Tridgell, 1972; Wason, 1963).

The schema-plus-tag model has two related consequences. First, the original negated message is assumed

to be represented as a core supposition and a negation tag, allowing for dissociation between the two at a later point in time. Second, the consideration of the core supposition activates the associations that are congruent with the core, but incongruent with the intended meaning of the negation as a whole. For example, upon comprehending the message "Tom is not guilty," one first thinks about the proposition "Tom is guilty" and activates associations that are congruent with the schema of "guilt." Only then does the receiver attach the negation marker. This allows for the correct understanding of the intended meaning of the negation in spite of the activation of the negation–incongruent associations.

The fusion model

There is an alternative way to model the processing of negations. One may fuse the negation operator with the core, thus transforming the negation into an affirmation. For example, the message "Tom is not guilty" might be transformed into its affirmative counterpart and coded as "Tom is innocent." In fusing, one spontaneously activates associations that are congruent with "innocence" (and "not guilty"), thus reinforcing the intended meaning of the message. The feasibility of changing negations into affirmations presupposes that there is an affirmation that captures the meaning of the negation (e.g., "innocent" means "not guilty"). This premise is offered by the literalism account of negation (Horn, 1989; Lyons, 1995).

The fusion model is consistent with a suggestion made by Gannon and Ostrom (1996), who argued that in processing uni-polar rating scales, the scale-point labeled "completely not X" is associated with a category which is the opposite of X. The study of Brewer and Lichtenstein (1975) also lends some support for this model, by showing that receivers recalled negations as affirmative assertions which preserve the original meaning (e.g., recalling "X is not warm" as "X is cold").

The research of MacDonald and Just (1989) provides a mechanism for the fusion model by suggesting that a negation operator inhibits the activation level of concepts that appear in the core supposition (see also Lea & Mulligan, 2002). According to the findings of MacDonald and Just, for example, processing the negation "Tom is not guilty" should result in inhibiting the concept "guilty." Functionally, this implies that receivers of such a negation are less likely to think of guilty-congruent concepts and, speculatively, more likely to think of guilty-incongruent concepts.

Comparing the two models

A basic distinction between the schema-plus-tag and the fusion model involves the nature of schema activated as one encodes a negation. The schema-plus-tag model assumes that negations such as "Tom is not guilty" are processed within the schema that refers to the core supposition (e.g., "guilt"). In contrast, the fusion model assumes that a negation–congruent schema (e.g., "innocence") is activated. This distinction is highly significant because not only do the two schemas have opposite meanings, they are also embedded in markedly different associative networks. As Gannon and Ostrom (1996) asserted "The category dishonest is not merely the inverse or negative of the category honest, it is a different knowledge structure. That is, honesty and dishonesty do not differ only in terms of degree or positivity: The two categories are conceptually distinct." (p. 338). Hence, the schema in which we process and code the negated message may be crucial, as in impression formation (Asch, 1946) or constructing measurement scales (Dholakia & Morwitz, 2002).

The differences between the two models might not be evident if we probe for the meaning of the message using direct questions. As we noted earlier, under normal conditions individuals have no trouble understanding the intended meaning of a negation. The models might be distinguishable if we probe the pattern of activation indirectly. To illustrate, imagine that shortly after encoding the negation, "Tom is not lazy," one has to process one of the two probes, SLOW or QUICK. Assume that "quick" is message-congruent (it is associated with being not-lazy) while "slow" is message-incongruent (it is associated with being lazy). Our study examines the pattern of activation when processing negation by comparing the activation of the message-congruent ("quick") and the message-incongruent ("slow") probe. According to the fusion model, the description "Tom is not lazy" is interpreted within a negation-congruent schema (e.g., "industriousness"). Therefore, the "quick" probe should be activated more than the "slow" probe. In contrast, according to the schema-plus-tag model the description is interpreted within a negation-incongruent schema (e.g., "laziness"). Therefore, the "slow" probe should be activated more than the "quick" probe.

Of course the two models also differ in their implications for the long-term impact of negations. According to the schema-plus-tag model, the negation operator could be detached from the core supposition. As a result, individuals may remember the opposite of the intended meaning. Such state of affairs is unlikely under the fusion model, as the negation is interpreted within a congruent schema. Predictions about memory for negations are tested in Experiment 2. After describing the results of two experiments, we elaborate on other implications of the two models for the communication of negated information, for discounting theories and for the truth value assigned to incoming information.

Experiment 1

In each trial of Experiment 1 participants were presented with a description of a person, expressed as an affirmation or a negation (e.g., "Tom [is/is not] a tidy person"), followed by a probe sentence (e.g., "Tom's clothes are folded neatly in his closet"). Participants were to determine whether the probe fits the description of the person. The question of interest is whether processing of congruent probes is facilitated relative to processing of incongruent probes. Our main interest concerns the cases when descriptions are expressed as a negation. According to the fusion model, processing of negation-congruent probes ought to be facilitated compared to the negation-incongruent probes, whereas, according to the schema-plus-tag model, processing of negation-incongruent probes ought to be facilitated compared to the congruent ones.

External vs semantic negation

Negation is signaled in the experiment in two different ways, which we term *semantic* and *external*. We refer to descriptions in which the negation operator is the word "no" (as in "Tom is not guilty") as *semantic negations*. Negations are, however, often signaled by informational cues that are neither linguistic nor part of the message itself. For example, receivers might infer that a message is incorrect by noting who the source of the message is. In the present study, the color red signaled that the message was incorrect (cf. Gilbert, Tafarodi, & Malone, 1993). We refer to such negations as *external negations*. Comparing the effects of external and semantic negations is potentially interesting because it may show that there is a distinct cognitive operation that results from linguistic markers such as "not" or "un," as well as from situational cues, such as the credibility of the source. Yet we do not have a priori predictions about whether the effects are similar to each other, as there are grounds for making opposite predictions about semantic and external negations on the basis of the fusion and schema-plus-tag models.

Method

Participants

Fifty-one students participated in the experiment proper. Twenty-four additional students were run in a baseline condition. Participants were paid the equivalent of US$3 for their participation.

Procedure

Participants were run individually under computer control. In each trial the participants were presented with a description of a person, followed by a behavioral probe. Their task was to determine whether the probe was congruent with the description. Participants responded by pressing one of three keys that corresponded to "congruent," "incongruent," and "impossible to tell." Participants were informed that descriptions would appear on either a black or a red background. They were told that they should consider descriptions with a red background incorrect. "For example, the description *Bill is happy* [printed on red] means that Bill is not happy." Finally, participants were informed that we were measuring the time they took to read the descriptions and respond to the probes, as well as the accuracy of their judgments. Therefore, they were encouraged to respond as quickly and accurately as they could.

After these instructions participants were given five practice trials with descriptions and probes different from those used in the experiment proper. Following the practice trials, participants were informed that one-third of the behavioral probes would be congruent with the descriptions, one-third would be incongruent, and for the rest it would be impossible to tell. This was done to reduce the use of guessing strategies. Finally, participants were told that we would give a bonus worth the equivalent of $8.00 to the two participants whose accuracy and speed combined to yield the best performance.

The experiment consisted of 79 trials. Each trial was signaled by the appearance of the string XXX for 200 ms. The string was replaced by a description (e.g., "Bill is lazy"), which remained on the screen until the participant pressed any key. Immediately following a response, the description was replaced by a behavioral probe. In addition, the three response options ("congruent," "incongruent," and "impossible to tell") appeared at the bottom of the screen. Participants indicated "impossible to tell" by pressing the space bar. Pressing the "z" key signaled "congruent" for half the participants and "incongruent" for the other half. The key "/" indicated incongruity for the first group and congruency for the second group. Once a participant indicated a response, the screen was blanked for 1 s.

The 79 trials consisted of 72 experimental trials and 7 filler trials. The experimental trials were separated into six blocks. Hence, there were 12 experimental trials in each block. The first block started with two filler trials, and each of the other five blocks started with one filler trial. Participants could rest for as long as they wanted between blocks.

Stimuli construction

Based on pretest we constructed 24 sets. Each set consisted of a description of a person using a trait term (e.g.,"Tom is a tidy person") and three behavioral probes, one which was congruent with the description and incongruent with its negation (e.g.,"Tom's clothes are folded neatly in his closet"), one which was incongruent with the description and congruent with its negation (e.g.,"Tom forgets where he left his car keys"),

and one that was unrelated either to the description or its negation (e.g.,"Tom likes to have long conversations on the phone"). One of these sets is reproduced in Table 1, and the list of descriptions appears in Table 2 in a free translation from the Hebrew.

The description and the three probes in each set did not share any content words except for the name of the person. Sixteen of the sets included behaviors that were phrased affirmatively. In the remaining eight sets one of the three behaviors was a negation. These eight sets were excluded from all analyses.

On a more technical level, the experiment used several counter-balancing procedures. Each of the 24 descriptions was expressed in each experimental session either as an affirmation, a semantic negation, an external negation, or a double negation (see Table 1 for an example). Hence, each participant saw a particular description phrased in only one way. Each description appeared three times, once in the first two blocks, once in the middle blocks, and once in the last two blocks. In one of the presentations it was followed by a congruent behavioral probe, in another by an incongruent probe, and in the third by an unrelated probe. Finally, each block included three affirmations, three semantic negations, three external negations, and three double negations. Under these constraints, the assignment of descriptions to the negation condition and to the blocks, as well as the within-block order, were varied randomly between participants.

Baseline condition

The time participants took to make their congruency judgment depended critically on the number of words in the probe. Therefore, it was important to adjust for baseline differences in the time needed to process the probes even prior to making any judgment. To assess these differences, 24 additional participants were run in a baseline condition. In each trial the participants in the baseline condition were presented with a behavioral probe which was followed by a description. As in the experiment proper, each description was expressed as an affirmation, a semantic negation, an external negation, or a double negation. Participants were to determine whether the description was congruent with the behavioral probe. Unlike participants in the experiment proper, participants in the baseline condition processed the probe *prior* to being shown the descriptions. Therefore, the congruency between the probe and the description could not have influenced the processing of the probe. Except for this procedural difference, the baseline condition and the experiment proper were identical.

Results

Data preparation

Six of the 75 participants were eliminated from the analyses because they failed to follow the instructions regarding the meaning of the red background. They reported this failure during the debriefing and, indeed, more than 25% of their congruency judgments were erroneous. Three of the six participants were in the experimental condition and three were in the baseline condition. One additional participant from the experimental condition was eliminated from the analyses for completely ignoring the external negation in making the congruency judgments. The response latencies of the remaining participants were standardized within each

Table 1
An example of stimuli used in Experiment 1

Character description	Behavioral probe	Congruency
Affirmation		
Tom is a tidy person	Tom's clothes are folded neatly in his closet	Congruent
Tom is a tidy person	Tom forgets where he left his car keys	Incongruent
Tom is a tidy person	Tom likes to have long conversations on the phone	Irrelevant
Semantic negation		
Tom is not a tidy person	Tom's clothes are folded neatly in his closet	Incongruent
Tom is not a tidy person	Tom forgets where he left his car keys	Congruent
Tom is not a tidy person	Tom likes to have long conversations on the phone	Irrelevant
External negation		
~~Tom is a tidy person~~	Tom's clothes are folded neatly in his closet	Incongruent
~~Tom is a tidy person~~	Tom forgets where he left his car keys	Congruent
~~Tom is a tidy person~~	Tom likes to have long conversations on the phone	Irrelevant
Double negation		
~~Tom is not a tidy person~~	Tom's clothes are folded neatly in his closet	Congruent
~~Tom is not a tidy person~~	Tom forgets where he left his car keys	Incongruent
~~Tom is not a tidy person~~	Tom likes to have long conversations on the phone	Irrelevant

Note. ~~Strikethrough~~ signals external negations. In the experiment such negations were displayed on a red background. All other information (affirmations, semantic negations, and all probes) were displayed on a black background.

Table 2
Descriptions used in Experiment 1

careful	nosy*
boring	responsible*
gossipy	honest
humble*	polite
egoistic*	forgiving*
strong-minded*	punctual
relaxed*	miserly
sociable	industrious
spontaneous	worrying
adventurous	cowardly*
curious	self-confident
clever	tidy

Note. Descriptions flagged by * were excluded from the analyses since one of the behavior probes associated with them was phrased as a negation.

participant, and latencies three standard deviations or more above the participant's own mean (fewer than 2% of the cases) were trimmed by being set to the participant's mean plus three standard deviations.

Processing of the descriptions

The time participants took to process the descriptions in the four encoding conditions appears in Fig. 1. A one-way repeated-measures ANOVA indicated that the four encoding conditions differed in processing time, $F(3,141) = 20.34,\ p < .01$. It may not be surprising that the affirmations were processed faster than the semantic (or double) negations (in both cases, $p < .01$), since the semantic negations included an additional word, namely, "no." However, comparing affirmations to external negations provides a clean test for the amount of time negation requires, because the verbal information was identical in the two conditions. The results provide an unambiguous demonstration that negations increase the complexity of processing, as indicated by a prolonged processing time, $F(1,47) = 33.74,\ p < .01$. Finally, the time needed for processing double negations was similar to that for semantic negations [1759 vs 1750 ms, $F(1,47) < 1$]. This makes it unlikely that our participants attempted to transform negations into affirmations by accessing meanings opposite to the core supposition. Such attempts would have led to an increase in the complexity of processing of double negations compared to semantic negations. This did not occur. It appears that when participants discovered that the negation marker appeared in red (i.e., when it was negated externally), they simply canceled both negations.

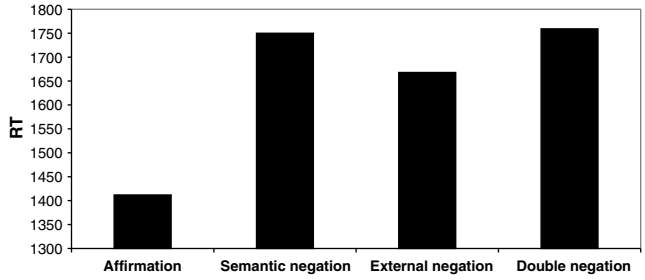

Fig. 1. Latency of description encoding.

Latency of the congruency judgment

Participants were shown three sets of behavioral probes. One set consisted of behaviors that were congruent with the affirmations (and the double negations) and incongruent with the negations' intended meanings. The second set consisted of behaviors that were incongruent with the affirmations and congruent with the negations. The third set consisted of behaviors irrelevant to the descriptions. Since our main analysis compares the time needed to process the first two sets, it was critical to equate them for ease of reading. We utilized the baseline condition for this purpose.

Baseline participants were shown the probes prior to viewing the descriptions. Therefore, their processing of the probes was not "contaminated" by congruency with the description. The three sets differed quite a bit from each other. Specifically, the first set took 125 ms longer to read than the second set, and the second set took 622 ms longer than the third set. The raw latencies in the analyses reported below were adjusted to reflect these differences. To do so, we computed the time baseline respondents took to read each of the probes. These were subtracted from the time experimental respondents took to decide whether or not the probes were congruent with the descriptions. Thus, the adjusted latencies correct for differences between the probes in ease of reading.

For each of the four types of descriptions we computed a *facilitation score* by subtracting the time needed to process the congruent probes from the time needed to process the incongruent ones (after adjusting for baseline differences). Facilitation scores are based only on trials in which the judgment responses were correct. Positive facilitation scores indicate that message-congruent probes gave rise to faster congruency judgments than message-incongruent probes.[1] Fig. 2 presents the adjusted latency of congruency judgments, as well as the means of the facilitation scores computed in the four types of descriptions.

The figure indicates that affirmations led to a strong facilitation effect, ($F(1,47) = 11.76,\ p < .01$), so that, in line with past research, congruent probes elicited faster responses than incongruent ones. The same effect appeared with the double negations; congruent probes gave rise to judgments that were faster by about 600 ms, ($F(1,47) = 24.59,\ p < .01$).

[1] We emphasize that we refer to facilitation in a relative sense, that is, as an advantage in processing congruent probes relative to incongruent ones. Our design does not allow for inferences about the magnitude of the absolute facilitation in processing the probes.

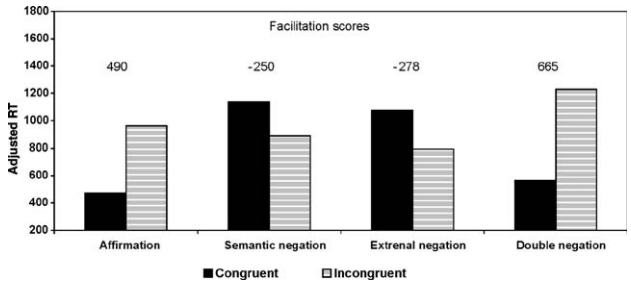

Fig. 2. Adjusted latency of judgments. *Note*. Adjusted judgment latencies are computed by subtracting reading time (baseline) from judgment time. Facilitation is computed as the latency of judgments involving incongruent probes minus the latency of judgments involving congruent probes.

Fig. 2 also shows that after descriptions were negated, either semantically or externally, judgments involving incongruent probes were significantly faster than those involving congruent ones, $F(1,47) = 8.29,\ p < .01$, in line with the predictions of the schema-plus-tag model. In order to compare the magnitude of facilitation in the different conditions, a series of two-way ANOVAs (type of probe [congruent vs incongruent] and type of description), comparing affirmations to the other three types of description, were performed. As Fig. 2 indicates, the extent of facilitation was similar in affirmations and double negations ($F(1,46) = 0.81,\ p = .37$), and also in semantic and external negation ($F(1,46) = 0.36,\ p = .55$). Importantly, facilitation differed significantly between responses to affirmations and semantic negations ($F(1,46) = 14.14,\ p < .01$), as well as between affirmations and external negations, ($F(1,46) = 21.23,\ p < .01$).

Accuracy of congruency judgment

For each participant we computed the response accuracy in each of the combinations created by the types of behavior (congruent, incongruent, and irrelevant) and the encoding of the description (affirmation, semantic negation, external negation, and double negation). Overall, judgments were highly accurate: 90% for congruent probes, 91% for incongruent probes, and 92% for irrelevant probes. The accuracy of the judgments involving congruent or incongruent probes depended on the type of message. Congruent probes were identified more accurately than incongruent ones when the descriptions were phrased as affirmations or double negations. Congruent probes were identified less accurately than incongruent ones when the descriptions were phrased as an external or a semantic negation. An analysis of planned contrasts shows that the difference in accuracy for congruent and incongruent behaviors was similar for affirmations and double negations, $F < 1$, as well as for semantic and external negation, $F < 1$. However, these two sets were significantly different from each other, $F(1,47) = 62.08,\ p < .01$.

The similarity of the patterns of accuracy and latency suggests that the difference between slow and fast congruency judgments does not reflect a speed-accuracy tradeoff. Differences in judgment latency would be less interesting if they reflected a tendency to invest more effort in some trials and less in others. Such a tendency would be indicated if fast judgments occurred when participants made many mistakes and slow judgments when mistakes were rare. This, however, did not happen.

Discussion

Experiment 1 has four main findings: First, processing negations is more complex than processing affirmations. This is indicated most clearly by comparing the time needed to process the affirmations and the external negations. Second, whereas affirmations (and double negations) gave rise to positive facilitation scores, negations gave rise to negative facilitation scores. Importantly, the facilitation measure was based only on trials in which participants gave correct judgments. Thus, even in cases where participants understood the intended meaning of the negation accurately, they processed negation-incongruent probes faster than negation-congruent probes. Third, we failed to find differences between semantic and external negations in the latency or the accuracy of the congruency judgments. Finally, it appears that the different pattern of activation for affirmations and negations does not reflect a speed-accuracy tradeoff, since the pattern of accuracy results was similar to that for judgment latency.

The judgment latency results are consistent with the predictions of the schema-plus-tag model. This model also receives support from the similarity of the affirmations to the double negations, where participants acted as if the presence of the red background served as a signal to cancel the negations altogether, rather than to access associations that are opposite to those accessed by the semantic negation. However, these results might also be explained by two other mechanisms, which we now discuss.

First, although the baseline condition provides a way to equate the congruent and incongruent sets of behavioral probes for ease of reading, it cannot equate the two sets with respect to their strength of association with the descriptions. Therefore, the judgment latency results could also be interpreted by capitalizing on this difference. To illustrate, it might be argued that crying is more diagnostic of sadness than is smiling, and therefore, that crying is more strongly associated with being sad or not being sad than is smiling. We believe that this interpretation is unlikely in light of the following analysis of the type of mistakes respondents made in their judgments. Respondents could make two types of

errors in responding to a probe: they could either indicate that the probe was not relevant, or they could choose the opposite response (e.g., respond "incongruent" to congruent behaviors). If one behavior (e.g., smiling) is more weakly associated with the trait than the other behavior (e.g., crying), then respondents should differ in their propensity to make the two kinds of errors. Specifically, in judging the probe that is more weakly associated with the trait they should make more mistakes involving selection of the "irrelevant" option than mistakes involving selection of the opposite option. This prediction was not borne out by the data. We shall discuss more evidence which is inconsistent with this interpretation after describing the findings of Experiment 2.

A second alternative explanation to the findings involves response processes. It could be that when processing negations participants were faster in responding "no" (i.e., "incongruent"), because the presence of a negation marker in the description speeded the "no" response option. Such an account attributes the effect to the response stage rather than the encoding of the descriptions. This alternative cannot be ruled out by the findings of Experiment 1.

Experiment 2 attempts to gain more insight into the encoding of negations and, at the same time, rule out the response-bias account, by studying two kinds of descriptions: bi-polar and uni-polar. A bi-polar description has a well-defined opposite construct which is easily accessible, whereas a uni-polar description does not. Since one can readily access an opposite construct when processing a negation of a bi-polar description, such processing is more likely to be consistent with the predictions of the fusion model, and therefore to depart from the pattern we observed in Experiment 1. Negations of uni-polar descriptions, in contrast, should show the pattern observed in Experiment 1, since they make it difficult to access the opposite schema. Moreover, the response-bias account is mute with respect to the nature of the description because its prediction is attributed to the correspondence between the presence of both a negation in the message and a negation as a response option. It therefore makes the same predictions for negations involving uni-polar and bi-polar descriptors. Hence, differences between bi-polar and uni-polar descriptions in their patterns of activation will suggest (1) that the impact of negated messages in our study involves the encoding stage and (2) that the two processing models—schema-plus-tag and fusion—co-exist and that their dominance depends on the existence of a readily available alternative schema during encoding. When receivers have a schema to accommodate the meaning of the negation as a whole, the negation should be encoded in line with the fusion model. However, the absence of such a schema should lead to processing according to the schema-plus-tag model.

Experiment 2 attempts to eliminate another potential mechanism that might have reinforced the encoding of negations in line with the predictions of the schema-plus-tag model. All participants in Experiment 1 were exposed to double negations. Perhaps the dominance of the schema-plus-tag model in Experiment 1 stems from the inclusion of this type of negation. To illustrate, assume that one is informed that "it is incorrect that Tom is not guilty" (i.e., "Tom is not guilty" presented in red). The use of the fusion model in such a case is confusing because it is unclear what the core supposition is and which associations should be activated. Thus, participants in Experiment 1 may have been forced to handle double negations by canceling both negations. This, in turn, may have undermined the use of the fusion model even in cases of single negations. To rule out this possibility, participants in Experiment 2 were not presented with double negations.

Experiment 2

The design and procedure of Experiment 2 are similar to those of Experiment 1, with several important differences. First, the descriptions used in Experiment 1 were not chosen systematically on the basis of the uni-polar/ bi-polar distinction. Therefore, some of them were uni-polar, some bi-polar, and the others neither clearly uni-polar nor clearly bi-polar. In Experiment 2 we selected descriptions that were clearly either uni-polar or bi-polar (see below). Second, whereas participants in Experiment 1 received descriptions expressed as affirmation, semantic, external, and double negation (a four-level within-participant factor), participants in Experiment 2 received descriptions expressed as affirmation and either semantic or external negation. Thus, although semantic and external negations did not give rise to significant differences in Experiment 1, they are included in Experiment 2, but as a between-subjects factor. Third, half the participants in Experiment 2 were instructed explicitly that we are using negation "to mean that the opposite of the negation is true." The remaining participants were not told how to interpret the negations. This allows us to compare the normal, unconstrained, understanding of negation to the more specific interpretation of negation as opposite. This is important because it is quite possible that in everyday communication both senders and receivers attach a special meaning to a characteristic of a person that is described using a negation. Such a meaning might be different from either the affirmative or the opposite formats of that characteristic (Colston, 1999). For example, by choosing to say that "Tom is not a stupid man" the sender may warn the receiver not to underestimate Tom, who is not the

smartest person but at the same time is not as stupid as one may think.[2]

After subjects received the descriptions and behavioral probes, as in Experiment 1, they were given a filler task to do for 5 min. They were then given a memory test. We wanted to see whether they would remember the negations correctly in form and meaning, whether they would transform them into affirmations so that they would remember their meaning correctly (but not their form), as the fusion model would predict, or whether they would remember their meaning incorrectly because they lost the negation marker (as the schema-plus-tag model would predict).

Method

Participants

Two hundred and twenty-four students participated in the experiment proper and an additional 32 participants participated in the baseline condition. Participants were assigned randomly to the eight experimental conditions. They were paid the equivalent of $3 for their participation.

Design

The experiment included three between-subjects factors: type of descriptions (uni-polar vs bi-polar); type of negation (semantic vs external); and meaning imposed on negation (opposite vs unconstrained).

Stimuli construction

Pretest participants were given a list of single-word descriptors denoting a trait characteristic (e.g., intelligent, lazy, and secure). For each descriptor they were instructed to write down the first word with an opposite meaning that came to mind. Each pretest participant saw a list of descriptors that did not include words which were opposites of each other. Based on the pretest results we selected 12 bi-polar pairs and 12 uni-polar descriptions. Bi-polar pairs of descriptions (e.g., stupid/smart) included descriptors that more than 80% of the pretest participants had given as opposites of each other. Uni-polar descriptors consisted of terms for which more than 80% of the participants had either failed to think of an opposite or had used a negation of descriptor (e.g., "not adventurous" for "adventurous") as a word with an opposite meaning. The list of descriptions appears in Table 3 in translation from the Hebrew.

For each uni-polar or bi-polar description we constructed three behavioral probes, one congruent with it and incongruent with its negation, one incongruent with the description and congruent with its negation, and a third that was unrelated to either the description or its negation. Since the bi-polar descriptions were actually opposite pairs on the same dimension (i.e., tidy/messy), the same behavioral probe was congruent with the affirmative description and incongruent with its negation. Moreover, it was also incongruent with the opposite term but congruent with its negation. For example, "Tom's clothes are folded neatly in his closet" is a congruent probe for both "tidy" and "not messy," while it is incongruent for "not tidy" and "messy." The correspondence between the probes and the descriptions was verified in a second pretest.

Table 3
Uni-polar and bi-polar descriptions used in Experiment 2

Uni-polar descriptions	Bi-polar descriptions
Creative	Industrious/lazy
Adventurous	Tidy/messy
Moral	Optimistic/pessimistic
Self-confident	Rich/poor
Talented	Warm/cold
Efficient	Arrogant/humble
Charismatic	Miserly/generous
Tolerant	Brave/cowardly
Responsible	Clever/stupid
Trustworthy	Strong/weak
Noisy	Autonomous/dependent
Romantic	Active/passive

Procedure

The procedure was identical to that of Experiment 1 except for the following changes. Half the participants were instructed that negation means that the opposite of the negated information is true. Specifically, these participants were told, "Descriptions in which the word 'not' precedes the characteristic are descriptions in which the opposite of the characteristic is correct. For example, the description 'Tom is not a loner' means that Tom is a sociable person." We call this condition *opposite meaning*. The remaining participants were not given specific instructions regarding the meaning of the negation: "Some of the information that is given to us regarding a specific person may be in an affirmative format (e.g., 'Tom is a loner') and some may be in a negative format (e.g., 'Tom is not a loner'). The experiment explores how people process information expressed in an affirmative or a negative format. Therefore, some of the descriptions will appear in a negative format— that is, the word 'no' will appear." We call this condition the *unconstrained meaning* condition.

The judgment phase consisted of 43 trials. On 36 of these trials participants were shown either uni-polar or bi-polar descriptions (12 descriptions × 3 probes) and on 7 trials participants were shown other descriptors, called

[2] Parenthetically, this may also explain why the participants in Experiment 1, who were not constrained in their interpretation of negation, did not show facilitation of negation-congruent associations. If this interpretation is correct, we should find different patterns of facilitation in the constrained and the unconstrained conditions.

fillers. The trials were separated into six blocks. The first block started with two filler trials and each of the other five blocks started with one filler trial. Participants could rest for as long as they wanted between blocks.

We employed several counter-balancing procedures, similar to those used in Experiment 1. Each participant saw each description expressed as either an affirmation or a negation. Half the participants saw semantic negations. The remaining saw external negations (negation signaled by a red background). Each of the descriptions appeared three times, once in the first two blocks, once in the middle blocks, and once in the last two blocks. It was followed by either a congruent probe, an incongruent probe, or an unrelated probe. Each of the blocks included three affirmations and three negations. Under these constraints, the assignment of the participants to type of negation condition (semantically or externally), type of description (uni-polar or bipolar), and negation meaning (opposite or unconstrained) as well as the within-block order were varied randomly among the participants.

Memory task

Following the judgment phase, subjects were given a 5-min filler task. They were given five different strings of letters and asked to generate as many words as they could from each string. Then they were presented with a surprise memory test. For each of the 12 experimental sets, participants were given the name of the protagonist together with the congruent and incongruent behavioral probes. They were asked to write down the description as it had been presented to them.

Baseline condition

Thirty-two participants were run in the baseline condition, which was identical in procedure to the experiment, except that probes were presented prior to the trait descriptions.

Results

Data preparation

As our main prediction involves judgment latency, we eliminated from the analyses the results of 18 participants who made more than three errors in classifying either the 12 congruent or the 12 incongruent probes following descriptions expressed as negations, and one participant who made more than three errors following descriptions expressed as affirmations. Two additional participants were eliminated because the post-experimental debriefing revealed that they had misunderstood the instructions. Thus, altogether, the responses of 21 of the 224 participants were eliminated from the analyses. As in Experiment 1, response latencies were standardized within participants and latencies that were three standard deviations or more above the participant's own mean (fewer than 2% of the cases) were trimmed by being set to the participant's mean plus 3 SD.

Processing of the descriptions

Participants responded faster to affirmations than to negations, $F(1,195) = 214.73,\ p < .01$. The difference between the time needed to comprehend affirmations and negations was less for external negations than for semantic negation, $F(1,195) = 4.98,\ p < .03$. Still, even when the negations and affirmations contained the same number of words (the case of external negations), affirmations were processed significantly faster, $F(1,96) = 86.77,\ p < .01$. There were no main effects or interactions involving the type of description (uni-polar vs bi-polar) or the meaning of the negation (opposite vs unconstrained).

Latency of the congruency judgment

As in Experiment 1, we computed facilitation scores by subtracting the judgment latency of the congruent probes from the judgment latency of the incongruent probes. This was done after we had adjusted for baseline differences between the different sets of probes for each type of description: bi-polar and uni-polar. Facilitation scores are based only on trials in which participants responded correctly to the congruency judgments.

Our main concern was comparing the relative facilitation produced by uni-polar and bi-polar descriptions. We hypothesized that since negations of uni-polar descriptions cannot be encoded within an opposite schema (since such a schema does not exist), they should give rise to negative facilitation scores. In contrast, as bi-polar negations can be encoded within the opposite schema, they should be associated with positive facilitation scores.

Fig. 3 displays the adjusted latency of congruency judgments, as well as the means of the facilitation scores computed as a function of the type of description (uni-polar vs bi-polar) and its formulation (affirmation vs negation). It shows that judgments involving the bi-polar descriptions showed positive facilitation for affir-

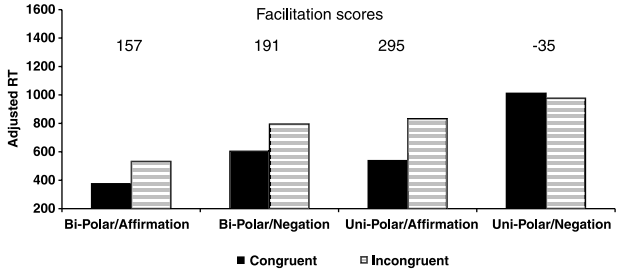

Fig. 3. Adjusted judgment latencies. *Note*. Adjusted judgment latencies are computed by subtracting reading time (baseline) from judgment time. Facilitation is computed as the latency of judgments involving incongruent probes minus the latency of judgments involving congruent probes.

mations as well as negations. In contrast, the pattern of facilitation of uni-polar descriptions was similar to that observed in Experiment 1, namely, positive facilitation for affirmations and negative facilitation for negations. To examine these patterns formally, the facilitation scores were analyzed in a four-way mixed-model ANOVA, with type of description (uni-polar vs bi-polar), type of negation (semantic vs external), and meaning of negation (opposite vs unconstrained) as between-participants factors and formulation of description (affirmation vs negation) as a within-participant factor.

Judgments of congruent probes following bi-polar descriptions were faster than judgments of incongruent probes (positive facilitation scores), $F(1, 105) = 15.96$, $p < .01$. Importantly, the magnitude of facilitation was similar for affirmations and negations, $F(1, 105) = 0.26$, $p = .60$, suggesting that receivers encoded the bi-polar descriptions by using the schema that was congruent with the intended meaning, even when the schema was incongruent with the core of the description. Moreover, the extent of facilitation did not differ between the unconstrained and the opposite meaning condition, $F(1, 105) = 0.73$, $p = .40$, suggesting tentatively that participants interpreted the negation as opposite whether or not they were explicitly instructed to do so.

In sharp contrast, the pattern of facilitation scores associated with the judgments of uni-polar descriptions revealed a significant difference between facilitation produced by descriptions expressed as affirmations and those expressed as negations, $F(1, 90) = 19.74$, $p < .01$, similar to the pattern observed in Experiment 1. The dissimilarity of the patterns associated with uni-polar and bi-polar descriptions is readily seen in a highly significant triple interaction, $F(1, 195) = 14.73$, $p < .01$. This triple interaction is inconsistent with the suggestion that the effect reflects a response-level facilitation, since this suggestion attributes the effect to the response stage and is therefore mute with respect to the type of description. Finally, the main effects and interactions involving type of negation (semantic vs external) and meaning of negation (opposite vs unconstrained) failed to reach significance.[3]

[3] We also recoded the descriptions used in Experiment 1 according to the classification criteria used to select descriptions in Experiment 2. It was found that of the 16 descriptions, four were uni-polar, three were bi-polar, and the remaining descriptors could not be classified to either category. The facilitation pattern found for these two categories was similar to the one reported for all 16 descriptions in Experiment 1. Thus, there was no significant difference between the uni-polar and bi-polar descriptions. This should be considered, however, in light of the small number of uni-polar and bi-polar traits, as well as the design differences between Experiments 1 and 2. In contrast to Experiment 2, where the type of description was a between-subject factor, Experiment 1 utilized a within-subject design, making it more likely that the processing the two types of descriptors influenced each other.

Accuracy of congruency judgment

Overall, judgments were highly accurate: 94% of the judgments involving congruent probes and 95% of the judgments involving either incongruent or unrelated probes were answered correctly. For each participant we computed the proportion of judgments in which he or she responded accurately in each combination of type of probe (congruent vs incongruent) and formulation of the description (affirmation vs negation). An analysis of the accuracy scores revealed that, as with judgment latencies, congruent probes were identified more accurately than incongruent probes when the descriptions were phrased as affirmations. In contrast, congruent probes were identified less accurately than incongruent ones when the descriptions were phrased as negations. This interaction between the type of expression (affirmative/negative) and type of probe (congruent/incongruent) was significant $F(1, 195) = 8.69$, $p < .01$. The only other significant effect was an interaction between the type of description (bi-polar/uni-polar) and type of negation (semantic vs external), $F(1, 195) = 5.36$, $p < .02$. It indicates that for external negations judgment accuracy associated with bi-polar and uni-polar descriptions was similar (92.8% vs 93.2%, $F(1, 96) = 0.07$), whereas for semantic negations the accuracy for the bi-polar descriptions was higher than that for the uni-polar descriptions (95% vs 92%, $F(1, 99) = 10.71$, $p < .01$). We return to this effect below, in the analysis of memory.

Memory of the descriptions

An analysis of the latency of judgments indicates that bi-polar and uni-polar negations elicited different reactions to the behavioral probes. The results are consistent with the suggestion that when a negation has a clear opposite (bi-polar negation), the opposite schema is activated and used to interpret the negation. In contrast, when there is no well-defined opposite (uni-polar negation), the negation is encoded within the schema of the core supposition together with a negation tag. The former is consistent with the fusion model, the latter with the schema-plus-tag model. If this interpretation of the finding is correct, then the mistakes people make in attempting to remember uni-polar and bi-polar negations should differ. Specifically, this analysis implies that, since people encode bi-polar negations with the aid of an opposite schema, they should remember it as such. Thus, for example, the description "Jim is not warm" might be encoded and remembered as "Jim is cold." Although this is not the exact form in which the original description appeared, such a memory preserves the original meaning of the description. Interestingly, receivers might also remember that the description involved a negation. In this case they might mistakenly remember "Jim is not cold."

Uni-polar negations, in contrast, are hypothesized to be encoded according to the schema-plus-tag model,

namely, as core plus negation marker. This allows for the dissociation of the negation marker from the core. As a result, the negation might be remembered mistakenly as the core. For example, the description "Jim is not responsible" might be remembered as "Jim is responsible." Thus, our analysis suggests that memory errors which reflect the loss of the negation tag would be more likely for the uni-polar negations than for bi-polar ones.

At the end of the experiment the participants were reminded of each original description through congruent and incongruent behavioral probes. They were asked to write down the original description. We coded their cued recall according to whether or not it preserved the original meaning and whether it was phrased as affirmation or negation. The recall protocols were coded by two judges, who were blind to how the original description had been phrased (affirmation or negation). Inter-judge agreement was high (97.4%), and differences were resolved by discussion with a third judge. Table 4 presents the proportion of cases in each of the four categories (meaning preserved/lost × recalled as affirmation/negation).

Since participants were given rich retrieval cues, their recall of descriptions that were phrased affirmatively was very good. When the original description was originally formulated as an affirmation, respondents recalled its meaning correctly in 91% of the cases when it was bi-polar and 93% when it was uni-polar (see Table 4). Not surprisingly, none of the uni-polar affirmations was remembered as a negation of the opposite, whereas 1% of the bi-polar affirmations were so remembered. This difference is consistent with the construction of the stimuli, since uni-polar descriptions did not have well-defined opposites.

The bottom part of Table 4 shows the recall of descriptions that had been formulated originally as negations. There are several noteworthy trends. First, using a gist criterion for accuracy, participants were more accurate in remembering the meaning of bi-polar negations (83% = 70% + 13%) than uni-polar negations (62%). Second, the memory errors were markedly different for the two types of negation. Whereas uni-polar negations led to loss of the negation marker (38%) but not errors of translation, bi-polar negations were associated with both types of errors (14% loss of the negation marker and 3% translation errors).

For each participant we computed the proportion of cases where a dissociation error (reporting the core supposition while losing the negation marker) was made. These proportions were analyzed in a three-way ANOVA, with type of description (uni-polar vs bi-polar), type of negation (semantic vs external), and meaning of negation (opposite vs unconstrained) as between-participants factors. The analysis revealed that the negation tag was more likely to be lost during recall for uni-polar negations than for bi-polar negations, $F(1, 199) = 43.73,\ p < .01$. The magnitude of this difference was, however, influenced by the type of negation, as indicated by an interaction between type of description and type of negation, $F(1, 199) = 15.04$, $p < .01$. The difference between bi-polar and uni-polar negations was stronger for semantic negations ($M_{\text{bipolar}} = 9.7\%$ vs $M_{\text{uni-polar}} = 45.3\%$) than for external negations ($M_{\text{bipolar}} = 19.1\%$ vs $M_{\text{uni-polar}} = 30\%$). None of the other effects or interactions reached an acceptable level of significance (all p's $> .15$).

Discussion

Experiment 2 suggests that the tendency of a negation to spontaneously activate incongruent associations more than congruent associations (the negative facilitation) depends on the type of information being negated. Whereas judgments involving negations of uni-polar descriptions showed negative facilitation, those involv-

Table 4
Memory performance (Experiment 2)

	Recalled as			
	Affirmation (meaning preserved)	Negation (meaning preserved)	Affirmation (meaning lost)	Negation (meaning lost)
	Affirmation			
	Warm	Not cold	Cold	Not warm
Examples of original description				
Bi-polar (e.g., "warm")	89% (571)	1% (9)	1% (9)	8% (51)
Uni-polar (e.g., "responsible")	93% (465)	0	0	6% (31)
	Negations			
	Cold (translation)	Not warm	Warm (**dissociation**)	Not cold
Examples of original description				
Bi-polar (e.g., "not warm")	13% (85)	70% (442)	14% (90)	3% (17)
Uni-polar (e.g., "not responsible")	0	62% (304)	38% (184)	0

ing bi-polar descriptions showed positive facilitation. This occurred even though the judgments involving affirmations of either uni-polar or bi-polar descriptions showed positive facilitation. We interpret this to mean that messages that were phrased affirmatively elicited associations congruent with them regardless of the type of description. However, the type of association structure activated by messages phrased as negations depended on the type of description—congruent associations were more strongly activated than incongruent associations following negated bipolar descriptions and incongruent associations were more strongly activated than congruent associations following negated uni-polar descriptions. We believe that the patterns of judgment latency and judgment accuracy are consistent with the suggestion that the negations of the uni-polar descriptions were encoded using the schema-plus-tag model, whereas the negations of the bi-polar descriptions were encoded using the fusion model. This distinction is further supported by the nature of the memory mistakes. In particular, negations involving uni-polar descriptions were more likely to lose the negation tag and hence the participants' memory of them lost the intended meaning.

Overall, semantic and external negations elicited similar effects in Experiments 1 and 2. Still, both accuracy and memory measures in Experiment 2 reveal that the differences between bi-polar and uni-polar descriptions were significantly greater in semantic negations than in external ones. This state of affairs prompts us to be cautious in proposing processing differences between semantic and external negations. One potential difference between external and semantic negations in the present study is that the former might require more effortful processing. That is, whereas receivers might have built-in structures for negation when the signal for it is semantic, they may have to use a more controlled process to translate the external signal into its meaning, especially since the external signal was novel—the color red. This post hoc analysis helps explain why the highest rate of dissociation of the negation marker from the description, as reflected in erroneous congruency judgments and faulty memory, occurred with the semantic negation of uni-polar descriptions. In this case respondents were least likely to elaborate on the description by considering its opposite. It should be noted, however, that the outcome of external negations might depend heavily on the nature of the signal. For example, whereas the color red may require special effort to be instantiated as a signal for negation, the face of an deceptive source might have built-in structure that can allow it to be utilized quickly (Cosmides & Toby, 1992).

The difference between semantic and external negations is further demonstrated by considering the processing of irony. Gibbs (1986) suggests that in certain circumstances an ironic statement (e.g.,"You are a big help") is read significantly faster than its non-ironic counterpart (e.g.,"You are not helping me"). Yet, as Giora (1995) noted, processing an ironic statement takes longer than processing the same statement when used in its literal sense. Hence, it seems that a message that is phrased affirmatively is easier to process than a message that is negated externally (irony in this case), which may suggest that understanding irony requires processing the literal meaning and then negating it. Moreover, it appears that, in the case of irony, processing the external negation is faster than processing semantic negation. However, irony is not irony if it is not immediately understood as such. Hence, it might be that with irony the alternative opposite schema clearly exists, while translating the irony into semantic negation reduces the accessibility of this alternative schema. It would be interesting to explore whether long-term memory for ironic statements corresponds to the intended meaning (the fusion model) or its opposite (as the schema-plus-tag model predicts), and whether memory mistakes differ as a function of the readiness with which the opposite comes to mind.

General discussion

Summary

Our study explored the consequences of processing negations. Experiment 1 showed that descriptions phrased as affirmations facilitated processing of behavioral probes that were congruent with their meaning as compared to probes that were incongruent. In contrast, descriptions phrased as negations facilitated processing of incongruent probes as compared to congruent ones. Experiment 2 demonstrated that this effect was limited to uni-polar descriptions. When negations had a well-defined and readily accessible opposite (i.e., bi-polar negations), then, as with affirmations, they facilitated the processing of congruent probes as compared to incongruent ones.

Moreover, the memory results in Experiment 2 indicated that, whereas memory for affirmations involving bi-polar traits was similar to that for affirmations involving uni-polar traits, memory for negations involving the two types of traits showed different types of mistakes. Mistakes of dissociation were more prevalent in the memory reports of the uni-polar traits (e.g., remembering "not responsible" as "responsible") than in the memory of the bi-polar traits (e.g., remembering "not smart" as "smart"; see Table 4). Dissociating the negation marker from the core trait is a type of mistake that would be predicted by the schema-plus-tag model but is not likely according to the fusion model. The fusion model predicts mistakes of expression instead,

such as translating a negated trait into its affirmative format. Indeed, this type of mistake appeared for the negations of bi-polar descriptions (e.g.,"not warm" was remembered as "not cold") but was absent in the case of uni-polar negations.

Taken together, the pattern of the judgment latency, memory results, and judgment mistakes are consistent with our suggestion that the phrasing of the description (affirmative vs negative) and the nature of the description (with or without a readily available opposite schema) determine the impact of a negated message. The overall pattern of data is inconsistent with two alternative mechanisms we discussed earlier, the one which attributed the effect to an artifact in the construction of the behavior probes, and the other which attributed it to response-level facilitation. The similarity of responses in the constrained and unconstrained conditions in Experiment 2 are inconsistent with a third alternative explanation that has to do with the interpretation of negations in the interaction. Consequently, we believe that our findings demonstrate that the existence of a schema that accommodates the meaning of the original negation is critical in determining how a negation will be encoded. When such schema is not readily available, processing a negated message facilitates negation–*incongruent* associations, in line with predictions of the schema-plus-tag model. When receivers do have such a schema, inferences that are congruent with the intended meaning are facilitated for negations and affirmations alike.

The associative structure activated in processing negation

Comparing our findings with those reported by MacDonald and Just (1989) may shed light on the activation produced by uni-polar negations. MacDonald and Just compared the activation level of a noun in an affirmative phrasing (e.g., bread) to its activation level when it was negated (e.g., no bread). They found that in the latter case, the activation was weaker. We compared the activation of probes that were congruent with a core trait to the activation of probes congruent with a negated trait. Taken together, the two studies suggest that, while a negation may decrease the activation level of a core compared to its activation without negation,[4] it still activates the core to a level higher than that of the concept with the meaning of the negated core trait (the opposite) when the negated information is of a uni-polar type. This conclusion is supported by MacDonald and Just's (1989) finding that there was no decrease in the response time to the associates of the negated noun (e.g., butter). Note that negating a noun ("no bread") involves an expression of the non-existence of an object, and hence, by definition, there is no opposite alternative. In this respect it is similar to uni-polar negation. Therefore, speculatively, comparing the incongruent associates of a negated core (e.g., "butter" in the case of "no bread") to its congruent associates (e.g., "hunger" in the case of "no bread") would have shown a higher activation level for the former.

This analysis has implications for the impact of communications phrased affirmatively or negatively. Affirmation tends to activate the core of the message with its associations (e.g., the assertion "John is innocent" activates associations of innocence during comprehension). It will be remembered consistently in the long term, regardless of the type of concept. When a concept is negated, however, activation varies as a function of the existence of an alternative opposite schema that can capture the meaning of the negation. If such a schema exists, then the effect of negation is similar to that of affirmation. If, however, there is no readily available schema with an opposite meaning, a boomerang effect might occur. In this case, during comprehension one may activate associations that are opposite to the intended meaning of the negation, so that in the long run receivers of the negated message information might remember the message as if it had not been negated (e.g., "John did not harass the secretary" will activate associations of harassment and John might be remembered as the one who did harass the secretary).

Next, we discuss several important theoretical as well as practical implications for understanding linguistic negations, as well as self-generated negations resulting from communication and life circumstances.

Prevalence of uni-polar negations

Let us start by noting that in constructing the stimulus material we found that it was appreciably easier to generate bi-polar than uni-polar descriptions. However, we believe that this asymmetry reflects our use of traits as descriptions. In the general case, we often negate characteristics that do not lie on a continuum and have no unique opposite. Thus, many statements which deny a particular act or event cannot be phrased affirmatively. For example, consider the statement "I did not have sexual relations with that woman." What could be an affirmative form of this statement? We believe that there is no well-defined opposite, and so such negations function like the uni-polar negations in our study and are encoded in line with the schema-plus-tag model. As a result, such statements may well activate associations opposite to their intended meaning and thus reinforce what they intend to deny.

Wegner, Wenzlaff, Kerker, and Beattie (1981) have shown that headlines phrased as questions (e.g., "Is Bob

[4] This effect was found in our study as well. Judgments of probes which were congruent with an affirmative description took less time than judgments of same probes when they were incongruent with a negated description; see Figs. 2 and 3.

Talbert linked with Mafia?") created impressions as negative as those induced by the affirmatively phrased headlines ("Bob Talbert linked with Mafia"). This finding could be attributed to two general mechanisms. On the one hand, readers might think that the question format is only a guise used to avoid libel charges, with the journalists writing what they actually believe, but presenting it in a question format. On the other hand, readers may spontaneously activate inferences associated with the core of the question, making this core appear more valid (see also Begg, Anas, & Farinacci (1992) for the relationship between frequency of exposure and perception of truth). Based on our findings, it is quite possible that even when a headline is phrased as an explicit negation (i.e., "Bob Talbert not linked with Mafia"), it would lead readers to associate Bob with the Mafia and, in the long run, to remember him as part of that organization. Indeed, Wegner et al. (1981) reported that, although the impact of negated headlines was not significantly different from the control headlines that were neutral assertions ("George Armsrtrong arrives in city"), it was also not significantly different from the impact of headlines aimed at providing incriminating evidence. This led Wegner et al. to conclude that "In certain extreme cases, the denial of criminality might itself prove incriminating" (p. 825).

Can people counter the biases involved in processing uni-polar negations? Comparing the processing of uni-polar and bi-polar negations suggests a way of protecting oneself against the potential bias of uni-polar negation. One has to think about an alternative schema, a counter-scenario, that can embody the negation. Rather than representing the uni-polar negation as ⟨not(X)⟩ one should look for a context in which not-X is positively characterized, namely, about what X is. This implies that in the course of a trial it would be less effective for the defense to argue that the evidence is invalid than to argue that it is part of a conspiracy against the defendant. "Not guilty" in the former case becomes "innocent" in the latter case.

This line of reasoning highlights the importance of political-correctness policies in combating stereotypes. A stereotype embodies a set of interconnected beliefs that are linked to a group-designation attribute. The interconnectedness, together with the causal relationships imposed on the beliefs (Anderson & Sedikides, 1991), provide a high degree of redundancy that makes the stereotype strongly immune to change (Schul & Zukier, 1999). Specifically, attempts to negate a particular belief within the stereotype are likely to encounter resistance because of the incongruence between the negated belief and the rest of the stereotypic structure. In this sense, even a bi-polar attribute becomes uni-polar because its counterpart can not emerge. This reasoning suggests that it may not be enough to just say "no" to stereotypes (Kawakami, Dovidio, Russin, Moll, & Hermsen, 2000). Rather, an alternative schema needs to be activated to replace the associative structure that embodies the stereotype. Recategorization may do just that.

Negation coding and message discounting

Our study found little difference between cases when a message contained the word "not" (semantic negation) and cases when the negation was signaled by a red background (external negation). It is interesting to consider other types of external negation, most importantly, negation due to processing messages from an unreliable source. Like messages with a red background, a message coming from a source known to be deceptive is likely to lose its immediate impact. Indeed, the literature on the "sleeper effect" (Cook, Gruder, Hennigan, & Flay, 1979; Mazursky & Schul, 1988; Pratkanis, Greenwald, Leippe, & Baumgardner, 1988) shows that a persuasive message attributed to an untrustworthy source is completely discounted in the immediate judgment condition. However, when the impact of such a message is not measured immediately, the message is dissociated from the source and discounting fails. We speculate that when the message coming from a deceptive source can be interpreted within a well-defined schema with an opposite meaning (as in the case of bi-polar negations), a sleeper effect would be less likely to occur. The effect is particularly likely when the message and the source are encoded in a schema-plus-tag fashion, allowing them to be dissociated at a later point in time (Mazursky & Schul, 1988).

On a more general level, our analysis suggests that manipulations which facilitate the generation of alternatives to the given message can protect receivers from the adverse effect of unreliable sources. The early research on immunization (McGuire, 1964), the more recent research into the effect of suspicion on the generation of inferences (Fein, McCloskey, & Tomlinson, 1997; Schul, Burnstein, & Bardi, 1996), and the demonstration that a meta-cognitive strategy can prevent the intrusions of false suggestions (Strack & Bless, 1994) are consistent with this suggestion.

A different implication of this analysis has to do with success in discounting. There is abundant evidence showing that integrative encoding impairs discounting. Discounting is more successful when receivers are prevented from elaborating on the message than when they are not prevented from doing so (Fleming & Arrowood, 1979; Schul & Burnstein, 1985), and is less successful when receivers are induced to encode the to-be-used and to-be-ignored messages integratively (Anderson, Lepper, & Ross, 1980; Schul & Mazursky, 1990). Our analysis suggests that this may depend on the way individuals interpret the instructions to discount a message. Specifically, it might be important to distinguish

between cases in which individuals are instructed to disregard or ignore a particular message, and cases in which the instructions to discount also provide information which directly refutes one or more of the original claims. The former case is analogous to uni-polar negations, the latter to bi-polar negations. In line with this analogy, Schul and Mazursky (1990) showed that whereas integrative encoding impairs successful discounting in the former case, it facilitates it in the latter case.

It is important to note that negation–incongruent associations were spontaneously activated in our study in spite of the fact that participants understood the negation well and used it appropriately in their judgments. Thus, our results are different from demonstrations that show failure to discount invalid messages or confusion of sources of information (Fiedler et al., 1996; Hornby, 1974; Johnson, 1988; Johnson et al., 1993; Johnson & Raye, 1981; Koehler, 1991; Loftus, 1979; Loftus & Palmer, 1974; Ross et al., 1975; Schul, 1993; Schul & Burnstein, 1985; Schul & Manzury, 1990; Schul & Mayo, 1999; Sellars, 1954; Strack & Bless, 1994). In these cases erroneous judgments reflect the use of inappropriate information (e.g., an invalid message, or a message from an irrelevant source). Our findings, in contrast, show that even in cases when recipients consider only the appropriate message information, they may err in the long run, though not in immediate judgments, as a result of the processing of negations.

The quest for truth

Gilbert (1991; Gilbert, Tafarodi & Malone, 1993) suggested that receivers initially tend to accept communications as true and that critical thinking and the appreciation that a message is false are achieved through a more controlled operation only at later stages of processing. Our study raises two questions regarding this suggestion: first, whether automatic acceptance is equally likely for affirmations and negations; second, whether the type of negation matters. Although we believe that the type of negation does matter, we are unsure about the direction of the effect. On the one hand, it could be that since bi-polar negations function like affirmations, in that they tend to activate inferences that support their intended meaning, they should be initially accepted. Uni-polar negations, in contrast, retain the negation tag. This makes it more likely that receivers will question the validity of the message in the early stages of processing, in line with the "Cartesian" rather than the "Spinozan" view (Gilbert, 1991). On the other hand, one may argue that a negation is an alternative way of communicating that the core is false. Accordingly, one has to understand the core prior to modifying it by the negation, in line with the schema-plus-tag model (Gilbert et al., 1993). However, given that bi-polar negations tend to be encoded as opposites, this analysis suggests that the effect of the "Spinozan" model of processing would be stronger for uni-polar negations. Such conflicting predictions are left for future research to test.

A caveat and final words on exploring negations

Before ending this discussion, a caveat is in order. As we noted earlier, our study was run in Hebrew, which has no prefixes like "dis-" or "un-," or suffixes to indicate the negation of a concept. It might be that this strengthens the tendency of Hebrew speakers to encode negations in line with the schema-plus-tag predictions. The existence of structures that allows the integration of an affirmative assertion and a negation marker may enable people to attach rich associative structures to negations. In such cases, people might define themselves and others according to what they are not rather than what they are. We believe that the diverse pattern of activation found for affirmations and negations which was moderated by the existence of an readily available opposite schema is by no means limited to the linguistic level of interpretation alone. Rather, as our very brief discussion suggests, the existence of alternative schema may influence how people cope with suspicion, discount invalid information, correct biases, and avoid stereotypes. Having an opposite schema is a no-nonsense answer to having to nullify a schema.

References

Anderson, C. A., Lepper, M. R., & Ross, L. (1980). Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology, 39*, 1037–1049.

Anderson, C. A., & Sedikides, C. (1991). Thinking about people: Contributions of typological alternative to associationistic and dimensional models of person perception. *Journal of Personality and Social Psychology, 60*, 203–217.

Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology, 41*, 258–290.

Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology General, 121*, 446–458.

Brewer, W. F., & Lichtenstein, E. H. (1975). Recall of logical and pragmatic implications in sentences with dichotomous and continues antonyms. *Memory & Cognition, 3*, 315–318.

Carpenter, P. A., & Just, M. A. (1975). Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review, 82*, 45–73.

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology, 3*, 472–517.

Colston, H. L. (1999). "Not good" is bad, but not bad is not good: An analysis of three accounts of negation asymmetry. *Discourse Processes, 28*(3), 237–256.

Cook, T. D., Gruder, C. L., Hennigan, K. M., & Flay, B. R. (1979). History of the sleeper effect: Some logical pitfalls in accepting the null hypothesis. *Psychological Bulletin, 86*, 662–679.

Cosmides, L., & Toby, J. (1992). Cognitive adaptation for social exchange. In J. E. Barkow, L. Cosmides, & J. Toby (Eds.), *The adapted mind evolutionary psychology and the generation of culture* (pp. 163–228). New York: Oxford University Press.

Dholakia, U. M., & Morwitz, V. G. (2002). The scope and persistence of mere-measurement effects: Evidence from a field study of customer satisfaction measurement. *Journal of Consumer Research, 29*, 159–167.

Eagly, A. H., Wood, W., & Chaiken, S. (1978). Causal inferences about communicators and their effect on opinion change. *JPSP, 36*, 424–435.

Fein, S., McCloskey, A. L., & Tomlinson, T. M. (1997). Can the jury disregard that information. The use of suspicion to reduce the prejudicial effects of pretrial and inadmissible testimony. *Personality and Social Psychology Bulletin, 23*, 1215–1226.

Fiedler, K., Walther, E., Armbruster, T., Fay, D., & Naumann, U. (1996). Do you really know what you have seen? Intrusion errors and presuppositions effects on constructive memory. *Journal of Experimental Social Psychology, 32*, 484–511.

Fleming, I., & Arrowood, J. (1979). Information processing and the perseverance of discredited self-perceptions. *Personality and Social Psychology Bulletin, 5*, 201–205.

Gannon, K. M., & Ostrom, T. M. (1996). How meaning is given to rating scales: The effects of response language on category activation. *Journal of Experimental Social Psychology, 32*, 337–360.

Gibbs, R. W., Jr. (1986). On the psycholinguistics of sarcasm. *Journal of Experimental Psychology: General, 115*, 3–15.

Gilbert, D. T. (1991). How mental systems think. *American Psychologist, 46*, 107–119.

Gilbert, D. T., Tafarodi, R. W., & Malone, P. S. (1993). You cannot believe everything you read. *Journal of Personality and Social Psychology, 65*, 221–233.

Giora, R. (1995). On irony and negation. *Discourse Processes, 19*, 239–264.

Horn, L. R. (1989). *A natural history of negation*. Chicago: The University of Chicago Press.

Hornby, P. A. (1974). Surface structure and presuppositions. *Journal of Verbal Learning and Verbal Behavior, 13*, 530–538.

Johnson, M. K. (1988). Discriminating the origin of information. In T. F. Ohmanns, & B. A. Mahler (Eds.), *Delusional beliefs* (pp. 34–65). New York: Academic Press.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*, 3–28.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*, 67–85.

Johnson-Laird, P. N., & Tridgell, J. M. (1972). When negation is easier than affirmation. *Quarterly Journal of Experimental Psychology, 24*, 87–91.

Jordan, M. P. (1998). The power of negation in English: Text, context, and relevance. *Journal of Pragmatics, 29*, 705–752.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8*, 441–480.

Kawakami, K., Dovidio, J. F., Russin, A., Moll, J., & Hermsen, S. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology, 78*, 871–888.

Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin, 110*, 499–519.

Lea, R. B., & Mulligan, E. J. (2002). The effect of negation on deductive inferences. *Journal of Experimental Psychology, 28*, 303–317.

Loftus, E. F. (1979). *Eyewitness testimony*. Cambridge: Harvard University Press.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13*, 585–589.

Lyons, J. (1995). *Linguistic semantics: An introduction*. Cambridge, England: Cambridge University Press.

MacDonald, M. C., & Just, M. A. (1989). Changes in activation levels with negation. *Journal of Experimental Psychology, 15*, 633–642.

Mazursky, D., & Schul, Y. (1988). The effects of advertisement encoding on the failure to discount information: Implications for the sleeper effect. *Journal of Consumer Research, 15*, 24–36.

McGuire, W. J. (1964). Inducing resistance to persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 1, pp. 192–229). New York: Academic Press.

Papageorgis, D. (1968). Warning and persuasion. *Psychological Bulletin, 70*, 271–282.

Petty, R. E., & Cacioppo, J. T. (1977). Forewarning, cognitive responding, and resistance to persuasion. *Journal of Personality and Social Psychology, 35*, 645–655.

Pratkanis, A. R., Greenwald, A. G., Leippe, M. R., & Baumgardner, M. H. (1988). In search of reliable persuasion effects: III. The sleeper effect is dead: Long live the sleeper effect. *Journal of Personality and Social Psychology, 54*, 203–218.

Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attribution processes in the debriefing paradigm. *Journal of Personality and Social Psychology, 32*, 880–892.

Schul, Y. (1993). When warning succeeds: The effect of warning on success of ignoring invalid information. *Journal of Experimental Social Psychology, 29*, 42–62.

Schul, Y., & Burnstein, E. (1985). When discounting fails: Conditions under which individuals use discredited information in making a judgment. *Journal of Personality and Social Psychology, 49*, 894–903.

Schul, Y., Burnstein, E., & Bardi, A. (1996). Dealing with deceptions that are difficult to detect: Encoding and judgment as a function of preparing to receive invalid information. *Journal of Experimental Social Psychology, 32*, 228–253.

Schul, Y., & Manzury, F. (1990). The effect of type of encoding and strength of discounting appeal on the success of ignoring an invalid testimony. *European Journal of Social Psychology, 28*, 337–349.

Schul, Y., & Mayo, R. (1999). Two sources are better than one: The effects of ignoring one message on using a different message from the same source. *Journal of Experimental Social Psychology, 35*, 327–345.

Schul, Y., & Mazursky, D. (1990). Conditions facilitating successful discounting in consumer decision making: Type of discounting cue, message encoding, and kind of judgment. *Journal of Consumer Research, 16*, 442–451.

Schul, Y., & Zukier, H. (1999). Why do stereotypes stick? In R. S. Wistrich (Ed.), *Demonizing the other: Antisemitism, racism, and xenophobia* (pp. 31–43). Harwood Academic Publishers.

Sellars, W. (1954). Presupposing. *The Philosophical Review, 63*, 197–215.

Strack, F., & Bless, H. (1994). Memory for nonoccurrences: Metacognitive and presuppositional strategies. *Journal of Memory and Language, 33*, 203–217.

Wason, P. C. (1963). The contexts of plausible denial. *Journal of Verbal Learning and Verbal Behavior, 4*, 7–11.

Wegner, D. M., Wenzlaff, R., Kerker, M. R., & Beattie, A. E. (1981). Incrimination through innuendo: Can media questions become public answers? *Journal of Personality and Social Psychology, 40*, 822–832.

Zuwerink, J. J., & Devine, P. D. (2000). Attitude importance, forewarning of message content, and resistance to persuasion. *Basic and Applied Social Psychology, 22*, 19–29.