

Inclusion and Exclusion: Implications for Stereotypic Judgments of Groups and Individuals

Jane Isabel Nussbaum

Senior Thesis in Social Psychology

Supervised by Professor Sid Perloe
Second Reader: Professor Ben Le

In collaboration with:
Prachi Dave
Julia Grayer
Cham Santé

Haverford College
4/30/04

Table of Contents

Abstract	4
Introduction	4
Assimilation and Contrast.....	4
<i>Assimilation Effects</i>	4
<i>Contrast Effects</i>	5
Types of Judgment	5
Comparative Judgment Theory	7
<i>Comparative Judgment and the Contrast Effect</i>	7
<i>Helson and Brown: Early Research in Basic Judgment</i>	7
<i>The Range-Frequency Compromise</i>	8
<i>Norm Theory</i>	10
<i>Summarizing the Comparative Process</i>	11
Interpretive Judgment Theory.....	11
<i>Set/Reset Theory</i>	12
<i>Empirical Evidence for Set/Reset Theory</i>	13
<i>The Inclusion/Exclusion Model</i>	15
<i>Empirical Evidence for the Inclusion/Exclusion Model</i>	17
<i>The Interpretation/Comparison Model</i>	25
<i>The Selective Accessibility Mechanism</i>	27
<i>Empirical Support for the Selective Accessibility Mechanism</i>	29
<i>Summarizing the Interpretive Process</i>	30
Prior Theory and Continuing Research.....	30
<i>Summarizing the Judgment Research</i>	30
<i>Continuing Research</i>	31
<i>Bless et al.: Inclusion and Exclusion of Groups and Individuals</i>	32
<i>Assimilation and Contrast in Bless et al. Explained</i>	33
The Present Research.....	34
Hypothesis.....	35
Methods	35
Participants.....	35
Design.....	36
Procedure.....	36
Results	38
Testing The Composite.....	39
Reliability Testing.	41
Individual Target Ratings.....	41
<i>The Effects of Condition on Target Ratings</i>	42
<i>Ratings for Traditional</i>	42
<i>Ratings for Cautious</i>	43
<i>Ratings for Dependent</i>	43
<i>Ratings for Forgetful</i>	44

<i>Analyzing the Forgetful/Condition Interaction</i>	44
Insight Effects.....	45
The Effects of Order	46
Discussion	47
The Composite and Reliabilities Tests.....	48
<i>The Dependent Variable</i>	48
Individual Target Ratings: Possible Explanations for the Results.....	49
<i>Insight</i>	50
<i>The Dependent Variable</i>	50
<i>The Manipulation of the Dependent Variable</i>	50
<i>Overshadowed: The Order Interaction</i>	51
<i>Summing up the Possibilities</i>	51
Understanding the Effects of Order.....	52
<i>The Present Findings Within a Broad Context</i>	52
Conclusion: Recommendations for Further Research.....	55
References	57
Appendices	60
Appendix A: Consent Form.....	60
Appendix B: Participant Instructions.....	60
Appendix C: Video Description.....	60
Appendix D: Sample Questionnaire.....	61
Appendix E: Responses to Question Seven.....	63

Abstract:

This paper presents a broad overview of various models of the judgment process, in an effort to place the present research within a larger theoretical context. Particular attention is paid to theories proposed by Parducci, Kahneman and Miller, Martin, Schwarz and Bless, Stapel and Koomen, and Mussweiler. The present research aimed to extend the prior finding that categorization of a moderately atypical exemplar as either within or without a group affects subsequent evaluations of both the group and the exemplar, but in opposite ways (Bless, Schwarz, Bodenhausen and Thiel, 2001). In this prior study, both assimilation and contrast effects were found. The present research, a methodologically similar study to Bless et al., employed a new method of presenting exemplar information (i.e., through film clips), intended to increase the ecological validity of the study, allowing participants to gather exemplar information from both auditory and visual domains. As well, the present research utilized a stereotyped group (i.e., the elderly) not used in the previous research. The results of this research did not support the main hypothesis. While people did evaluate the exemplar and the group differently, such evaluative differences were not the effect of differential categorization of the exemplar. It is hypothesized that the lack of empirical support for the main effects may have been due to the fact that the manipulation of the dependent variable was weak; alternately, the measure of the dependent variable may not have been effective. One strong and surprising—although interesting—interaction did emerge from the data analyses: an interaction between the order in which the targets were evaluated, and the evaluations of the targets themselves. Specifically, when the exemplar was evaluated first, evaluations of the exemplar and group were contrasted away from each other. It is suggested that this finding may be able to be accounted for by several factors—such as distinctiveness and category width—presented in prior judgment theories.

Introduction:

Assimilation and Contrast

Recent research in social psychology has investigated the role of assimilation and contrast in stereotypic judgments of both groups and individuals. With respect to the evaluative process, prior research has shown that inclusion of an exemplar within a group category (i.e., inclusion) leads to assimilation effects, while exclusion of an exemplar from a group category (i.e., exclusion) leads to contrast effects.

Assimilation Effects. When an assimilation effect results from an evaluative process, this means that judgments of the target are biased towards information contained within the context. In instances of evaluation where the target is an individual categorized as a member of a stereotyped group, there is a reduction in the degree to which individuating information about that person is relied upon; instead, the evaluation

of the individual is assimilated towards the implications of the group stereotype. As well, evaluations of the group are assimilated towards those of the individual. Hypothetically speaking, if given individual, Betty, is categorized as a member of the group “comedians,” then subsequent relevant evaluations of Betty should reflect the implications of the group stereotype. Assuming that humorousness is trait central to the stereotype of comedians, Betty should be judged as more humorous than she would be had she been *excluded* from the group category.

Contrast Effects. Prior research has also shown that when a category member is excluded from a group category, a contrast effect results. This means that judgments of the target are biased away from contextual information. In instances of evaluation of an individual categorized as *not* a group member, there is a decrease in the degree to which the group stereotype is relied upon in making judgments about that person; therefore, the impact of individuating information about that person is increased, and, relative to the group, the individual should seem less stereotypical. As well, the group should be judged as more stereotypical in contrast to the individual (i.e., judgments of the group are biased away from those of the target.) Again, speaking hypothetically, if Betty is categorized as *not* a comedian, then we would expect that she would be evaluated as less humorous than she would be had she been included within the group category, as judgments about her are biased away from the implications of the group stereotype.

Types of Judgment

According to current theories of the evaluative process, two types of judgment exist: comparative judgment and interpretive judgment (Stapel and Koomen, 1998.)

Depending on which type of judgment is made, a person will utilize stimulus and contextual information in different ways. As well, the type of judgment made influences the emergence of assimilation and contrast effects.

In comparative judgment, contextual information is used as a standard of comparison against which a target stimulus is judged; as well, the context is selected prior to evaluation of the target. Because this process necessarily emphasizes the differences that exist between the target and standard, only a contrast effect can result. Furthermore, in a comparative judgment, the stimulus will only be evaluated in terms of one dimension. If, for example, that dimension is 'weight,' and the target is a feather, judged in comparison to a brick, then the feather may only be evaluated along this one dimension. It may be concluded that the feather is 'lighter than the brick,' 'heavier than the brick,' 'the same weight as the brick' and so on, but no judgments can be made along other dimensions, such as 'age' or 'color.'

In contrast, interpretive judgments do allow for a multidimensional analysis of the target. In an interpretive judgment, the target stimulus acts as a probe, in response to which a representation of contextual information is invoked; this representation may be prior knowledge and/or novel constructs relevant to the target. Depending on the kind of information that is called to mind, the context may either guide the interpretation of the stimulus or serve as a standard against which it is compared. Thus, an interpretive judgment may lead to the emergence of assimilation *or* contrast effects.

In the following sections of this paper, various theories of comparative and interpretive judgment will be presented. In order to understand why the present judgment

research is relevant, it is important to take a broad look at these various theories, examining how, over time, they have developed and built upon one another.

Comparative Judgment Theory

Comparative Judgment and the Contrast Effect. “Classic” theories of categorization, rooted in the study of psychophysics, focus on the comparative process involved in judgment making, suggesting that the evaluation of a given stimulus is the result of a comparison of the target stimulus against some standard. Comparative judgment processes can *only* result in the emergence of contrast, as the target is necessarily compared and contrasted with the context in instances of evaluation. Thus, theories of comparative judgment do not account for the emergence of assimilation effects.

Helson and Brown: Early Research in Basic Judgment. The earliest studies of the judgment process were carried out by psychophysicists who sought to measure sensation. As a result of such research, Helson (1964) constructed a theory of adaptation level (i.e., AL theory) that has greatly influenced contemporary explanations of the contrast effect. AL theory rests on the assumption that in all uni-dimensional judgments, the target stimulus will be compared with a standard, which is constructed by averaging the experiences of all previously presented stimuli within the same context. Unfortunately, this theory is limiting in that it assumes that a stimulus will be compared with a standard that is the result of pooling all prior stimuli. Helson did not recognize that a higher order cognitive operation takes place in the construction of standards used in the

evaluation process, that contextual information is actively—though not necessarily consciously—selected.

In 1953, a breakthrough study by Brown brought this issue to the forefront of judgement research. In Brown's study, participants were asked to lift a series of weights and then judge their heaviness. He inserted an anchor weight into the stimulus series; the anchor was either a weight or a tray (of equal weight) on which the weights had been resting. It was found that when the anchor was a weight, a strong contrast effect emerged, whereas when the anchor was the tray, the effect was significantly less powerful. Presumably, because the tray was not seen as similar to the weights, it did not form a strong standard against which they could be compared. Brown thus concluded that not all prior stimuli that vary with respect to the dimension of judgment will be used in constructing a standard. Rather, people actively select the standards against which they compare a target stimulus; they make rich judgments taking into account only that contextual information which is most relevant.

The Range-Frequency Compromise. Today, the accepted theory of comparative judgment is the Range-Frequency Compromise theory put forth by Parducci (1961). Unlike Helson, Parducci does not take as evidence for the pooling process the fact that the standard can be predicted fairly well by taking an average of the series of stimuli. In constructing his theory, Parducci noted that while in a symmetrical distribution of stimuli, the median, mean, and midpoint of the series all fall at the same point, the average of a skewed distribution of stimuli generally falls between the median value of the series and the point midway between the series' extremes. This suggests that the average value of a series of stimuli is *not* in fact a good predictor of the standard. Instead, Parducci

proposes that: “assigning a stimulus to a given category represents a ‘compromise’ between a ‘range principle’ and a ‘frequency principle’” (Eiser, 1990, p.19).

The range principle assumes that that people will construct a scale of judgment that has endpoints that coincide with the endpoints of the range of contextual stimuli that is recalled; furthermore, the range of this scale will be divided up into equal intervals in proportion to the number of response categories. Meanwhile, the frequency principle assumes that people’s use of categories also reflects their rank ordered position. In other words, the range principal suggests that ratings of stimuli that are ‘medium’ for a series should be predicted by the midpoint of the series, while the frequency principal suggests that ‘medium’ ratings should be predicted by the series median. Thus, the Range-Frequency Compromise is a means of explaining the phenomenon noted by Parducci; that is, that the ‘medium’ for a series generally falls somewhere between the median and the midpoint of a series of stimuli, and thus the prediction of a standard for any stimulus series should take both the median and midpoint into account. Finally, it is important to note that in instances where stimuli distributions are symmetrical, all three measures of central tendency will fall at the same point on the judgment scale.

Having constructed a theory wherein the value of a stimulus is dependent upon comparisons with prior stimuli, and all values are relative, Parducci went on to put forth a theory about what values are “relative to” (Parducci, 1984, p.1). He was interested in understanding people's judgments about the happiness of their lives. Just as it has been found that the value of a stimulus is judged relative to the value of prior stimuli, Parducci assumed that the happiness of an event can be judged in relation to the degree of happiness of other events. This theory demonstrates how Range-Frequency theory is

applicable beyond judgments of physical stimuli, to rich judgments of complex social stimuli—the kind of real world evaluations that people make each day.

Norm Theory. Taking up where Parducci left off, Kahneman and Miller (1986) constructed a theory, known as ‘Norm Theory,’ which accounts for such target judgments made within rich contexts. This theory takes into account the kinds of judgments made in situations where people must define for themselves the standard against which a stimulus will be compared, and assumes (in accordance with Brown’s theory) that people actively—though not explicitly—select contextual information which is most relevant to the evaluation at hand. Kahneman and Miller (1986) propose an activation process model that, in response to some stimulus, produces norms (i.e., contextual information based on “remembered and constructed representations”) against which the stimulus is compared (p. 136.)

Norm theory assumes that a stimulus event comprises a probe, which simultaneously brings into consciousness multiple imagined representations of similar stimuli. These representations can be memories of previously presented stimuli or novel constructs based on an individual’s knowledge and understanding of the stimulus category; regardless, the norm will reflect the particular category members that come to mind and will be an aggregate of this set. The stimulus is then compared to this norm, and, to the extent to which the stimulus is seen as similar to the norm, it will be judged as normal and unsurprising.

Among social judgment theories, this model is able to go beyond others in its ability to account for the specific contextual information that is recalled as the result of presentation of arbitrary stimuli. As well, this model begins to make some sense of how

inclusion or exclusion of a stimulus can occur with regards to a given category. However, Norm theory does not go so far as to predict the emergence of assimilation or contrast effects. What Norm theory *does* do is challenge “the conception of norms as pre-computed structures” and takes into account how evaluations of rich, multidimensional stimuli might be made (Kahneman and Miller, 1986, p. 150). This theory thus sets the stage for an understanding of models of interpretive judgment, which deal explicitly with the relationship between inclusion and exclusion, assimilation and contrast.

Summarizing the Comparative Process. This section has presented various theories of comparative judgment that have been proposed over the years. While they do differ to some extent, each of these theories rests on the assumption that the evaluation of a given stimulus is the result of a comparison of the target with some standard. Only contrast can emerge from such a judgment process. In the following section, theories of interpretive judgment will be presented, all of which are able to account for the emergence of both assimilation and contrast effects.

Interpretive Judgment Theory

Theories of interpretive judgment account for rich, multidimensional analyses of complex stimuli. In interpretive judgments, contextual information may serve either to guide the interpretation of a target stimulus or become a standard against which the target is compared. Thus, an interpretive judgment may lead to the emergence of assimilation *or* contrast effects. The models of interpretive judgement presented in this section all attempt to elaborate upon and explain these phenomena.

Set-Reset Theory. One of the first models of the interpretive judgment process developed to explain directional variation in impression formation (i.e., whether assimilation or contrast results) was Martin's (1986) set/reset model, which suggests that people are able to process information in ways that are flexible and adaptive, although this process is not necessarily a conscious activity.

This theory rests on the assumption that judgments about a given stimulus will be made based on multiple informational sources. When people must form impressions, they refer back to prior knowledge or memories, general category information, and so on. Yet, not every source of information can be utilized in a given judgment; instead, when faced with the task of categorizing some stimulus, people are selective in the types of information that they choose to take into account. According to Martin's theory, "individuals seek information that appears appropriate (or at least not inappropriate) to their current processing objectives" (Martin and Achee, 1992 p. 200). In other words, in order to make an accurate assessment of a stimulus, an individual must attempt to take into account only target-relevant information, and not information about his or her own personal mood state, opinions about other stimuli in a set, or primed stimuli.

Thus, Martin suggests that when irrelevant information is made especially distinct (i.e., blatantly primed) before a person must assess an ambiguous target, then the individual will "delete the irrelevant reaction" (i.e., actively suppress or ignore the primed information) while making a judgment (Martin and Achee, 1992, p. 205). This, Martin argues, should result in a contrast effect, such that the judgment of the target is contrasted away from the suppressed contextual information. Martin (1986) refers to the

“suppressed use of the contextually activated response and the generation of a context-distinct response for the target” as “resetting” (p. 495).

At the same time, Martin’s theory also states that in instances of evaluation where contextual information is allowed to come into play (i.e., when irrelevant information is not detected, the prime is not blatant) the evaluation of the target should be assimilated towards the implications of the context. Martin (1986) refers to this “use of textually activated response during the formation of the target impression” as “set” (495).

Empirical Evidence for Set/Reset Theory. Martin tested his theory in an experiment employing a judgment task, wherein all participants were given the same contextual information (i.e., were blatantly primed with the same stimuli) but the degree to which this information was made salient later on, and thus “preserved or terminated,” was manipulated (Martin, 1986, p. 495). In order to accomplish the latter task, Martin (1986) employed a “task interruption,” the rationale for which comes from prior research that “has demonstrated that individuals are more likely to continue thinking about a goal when that goal has not been met than when it has been met” (p. 495). In other words, Martin assumed that task interruption should lead participants to focus on the uncompleted task, and thus on contextual information primed within it, making it more difficult for participants in this condition to ignore such ‘irrelevant’ information while making a judgment.

With relation to the set/reset model, the task interruption condition represents ‘set,’ and it was hypothesized that when making a judgment, participants in this condition would assimilate the target towards the primed contextual information. On the other hand, because the prime was so blatant, Martin assumed that participants in the

completed task condition (i.e., the condition where the task was not interrupted) should 'reset' and suppress primed information, resulting in a contrast effect.

Specifically, participants in Martin's study all completed a task in which the prime, either a positive or negative personality trait, was embedded. In this task, all participants categorized eight phrases. However, one group of participants (the task-interruption condition) was told beforehand that they would be completing a series of 12 such categorizations; thus, they felt that their task had been cut short when they were asked to move on to a different task after completing just eight categorizations. Following the phrase categorization task, participants were asked to read a paragraph containing an ambiguous person description and were asked to form an impression of that person.

Martin (1986) hypothesized that interruption of the priming task should lead participants to assimilate their target impressions towards evaluations of the information that was primed; on the other hand, he predicted that target impressions would be contrasted away from those evaluations when the priming task was completed.

The results of the experiment were consistent with Martin's predictions. It was found that subjects in the task-interruption condition assimilated their impressions of the target person towards primed information (the personality trait,) such that participants evaluated the target more positively when positive personality traits were primed than when the primed traits were negative. At the same time, participants in the task-completion condition contrasted their evaluations of the exemplar away from the implications of the primed contextual information.

These findings suggest that interpretive approaches to information processing are correct in assuming that people are flexible and adaptive in their evaluations of different stimuli. Furthermore, these findings provide strong evidence for the applicability of the set/reset model; they show that both assimilation and contrast effects can be observed in impression formation, depending on whether or not an individual is prompted to take contextual information into account.

Martin (1986) stresses that the set/reset model is not “limited to explaining the difference between subtle and blatant priming....[It] assumes more generally that any processing objective that tells subjects not to use primed information will cause subjects to reset” (p.201). Looking at this model in another way, set/reset could be thought of a theory of social cognition that focuses on the role of ‘appropriateness’ in information processing. That is, according to set/reset theory, when a person makes a judgment about a target, he or she must assess whether or not it is appropriate to consider contextual information in the process. Depending on the situation, a person may or may not suppress such information, resulting in directional variation in impression formation.

The Inclusion/Exclusion Model. By the time that Schwarz and Bless (1992) introduced their inclusion/exclusion model of assimilation and contrast effects in social judgment, it was “no news” that context matters in instances of evaluation. Already, numerous studies and experiments had provided support for the contention that “the evaluation of a target stimulus may be either assimilated or contrasted to the context in which it is presented” (Schwarz and Bless, 1992, p.217). Moreover, a variety of different factors had already been shown to affect the directional value of judgments, many of which had “been conceptualized in independent, and apparently unrelated, theories”

(Schwarz and Bless, 1992, p, 217.) Thus, in designing their inclusion/exclusion model, Schwarz and Bless hoped to integrate many of these theories within a more complete, overarching framework.

The inclusion/exclusion model builds off of previous models, such as Norm theory, in suggesting that before a target evaluation can be made, a person will construct a complex representation relevant to the target, made up of an evoked set of various informational elements. Furthermore, recognizing that not *all* potentially relevant information will be retrieved in a given instance of judgment, Schwarz and Bless again echo Kahneman and Miller in suggesting that an individual will recall only that information which is most accessible at the moment. As well, they further specify that only those elements which possess most or all of a specific set of categorical features will be included within a representation. The most important point highlighted in Schwarz and Bless' theory is the contention that it is the way in which target information is *categorized* that will determine whether a contrast or assimilation effect will result.

Briefly, the inclusion/exclusion model rests on the assumption that *inclusion* of a target stimulus or exemplar within a category will result in an assimilation effect, whereas *exclusion* of a target stimulus from a category should result in contrast. This is very similar to Martin's Set/Reset theory. However, Schwarz and Bless extend their theory beyond Martin's in suggesting that there are *multiple* processes that underlie the emergence of *contrast* effects.

On the one hand, they agree with Martin in suggesting that contrast could be the result of subtraction or suppression of contextual information from a target representation (i.e., a subtraction effect). Schwarz and Bless suggest that there are a variety of reasons

that a person might deliberately exclude context information when making a target judgment. For example, they cite prior research (Lombardi, Higgins and Bargh, 1987; Strack, Schwarz, Bless, Küber and Wänke, 1990; Martin, 1986) suggesting that awareness of external, irrelevant influences on the decision making process may lead individuals to ‘delete’ such information at the time of judgment.

On the other hand, Schwarz and Bless also acknowledge that a contrast effect could be the result of a comparison process—essentially, the same basic judgment process described in the early models put forth by Parducci and others—wherein contextual information (in the form of prior stimuli, knowledge, a mental representation, etc.) serves as a standard against which a target stimulus is compared.

With respect to the process of subtraction, Schwarz and Bless further specify that contrast effects will *only* result in evaluations where the target stimulus is a specific exemplar, as subtraction will influence only those target judgments that are related directly to the context. Meanwhile, contrast effects due to a comparative process may extend to other target stimuli within a set, for which the standard may continue to be relevant.

Empirical Evidence for the Inclusion/Exclusion Model. Two studies conducted by Schwarz and Bless (1992) in which “the inclusion or exclusion of context information was directly manipulated” support the assumptions of the inclusion/exclusion model (p.222). Specifically, it was found that judgments of the same target produce different effects depending on whether or not the target is included within a category representation. The first study, focused on the manipulation of categorization,

investigated the impact of a specific target (a politician) on category evaluations (specifically, evaluations of the political party to which the politician belongs.)

Participants in this study were asked a number of questions relating to German politics. In one condition, participants were asked a question, which was intended to increase the likelihood that a specific, well regarded politician, President Richard von Weizäcker, would be included in participants' representations of the Christian Democratic Party later on in the experiment when they were asked to evaluate the group. Specifically, participants in this condition were asked to recall the party of which "Richard von Weizäcker has been a member for more than 20 years" (Schwarz and Bless, 1992, 223). It was hypothesized that evaluations of this party should be more positive among participants who were asked this question, as von Weizäcker was a well regarded politician and including him as a party member should have made the party look more positive overall. In the other condition, subjects were asked a question intended to elicit exclusion of von Weizäcker from the party; specifically, they were asked which office was held by von Weizäcker that "sets him aside from party politics" (Schwarz and Bless, 1992, 223). Participants were expected to answer that he was the president. In Germany, the presidency is a non-partisan political position, and it was assumed that highlighting this should force participants to exclude the president from representations of the Christian Democratic Party. It was hypothesized that exclusion of von Weizäcker from the party should result in a lower party evaluation overall.

The results of the study supported both of the above hypotheses, indicating that the cognitive accessibility of specific information used in answering a question was increased as a result of presentation of the question. Furthermore, this increased

accessibility of specific information meant that it was more likely to be recalled when participants were later asked a relevant question. Schwarz and Bless also suggest that the impact of such accessible information on the direction of a judgment was dependent upon its inclusion within participants' representations of the target category. In other words, including von Weizäcker within the Christian Democratic Party meant that he was included as part of the relevant information which was utilized when participants made an evaluation of the party, resulting in an assimilation effect. On the other hand, excluding von Weizäcker from the party resulted in contrast.

Schwarz and Bless allow that it is not immediately obvious whether a comparative or subtractive process was at work in creating the observed contrast result. It could be the case that information about von Weizäcker—including his party affiliation—was chronically accessible among a portion of the participants who were *not* prompted to include him within the Christian Democratic Party. If this is true, then the observed assimilation effect would be a reflection of the fact that the question intended to lead to inclusion of the president within the party *did* in fact lead to an increase in the number of participants who included him within their category representation, while the question intended to lead to exclusion of the president of the party had the opposite effect—it led to a decrease in the number of subjects who included him within the party. If this is the case, then the contrast effect should be due to subtraction; that is, it should be due to the fact that information about von Weizäcker was suppressed from the data considered when participants made their evaluations. However, it could also be the case that participants went beyond simply excluding von Weizäcker from the category

representation, to using him as a standard against which members of the Christian Democratic Party could be compared.

Schwarz and Bless contend that if the contrast effect reflects exclusion *only*, then the question of presidency should only affect evaluations of the Christian Democratic Party—the effects of this question should not extend to evaluations of other political parties. On the other hand, if von Weizäcker was not only excluded from the category, but also used as a standard against which comparisons were made, then the contrast effect that resulted should extend to members of *other* political parties as well.

In order to determine which process lead to the obtained contrast effect, Schwarz and Bless asked study participants the same questions as before but concluded the procedure by asking them to evaluate a different political party—the Social Democratic Party. They found that evaluations of this party were not affected by the questions, suggesting that the contrast effect obtained from evaluations of the Christian Democratic Party was the result of subtraction, as opposed to a change in comparison standard.

A second study by Bless and Schwarz, conducted to further test principles of the inclusion/exclusion model, focuses on ‘category width.’ They assert that the likelihood of a given stimulus being included within a category increases as the category grows wider. Similarly, the likelihood that a given stimulus will be excluded from a category increases as a category grows narrower. Bless and Schwarz (1992) explain that, “this study bears on this assumption and extends the inclusion/exclusion logic by demonstrating changes in the standard of comparison...as indicated by a generalization of contrast effects across different angles” (p.225).

Bless and Schwarz hypothesized that when participants were asked to make an evaluation of the trustworthiness of politicians in general, after having been primed with specific politicians who were involved with a scandal, an assimilation effect should result. That is, these participants should rate politicians as less trustworthy, due to the fact that the contextual information—namely the group of politicians involved in the scandal—would most likely be included in the participants’ temporary representations of the category, ‘politicians.’

On the other hand, if subjects were instead asked to evaluate a specific politician, John Doe (who was not involved in the scandal,) in terms of trustworthiness, it was expected that a contrast effect would result. Specifically, this politician should be evaluated as particularly trustworthy in comparison with other politicians. The reasoning behind this expected result rests on the assumption that the specific politician should make up his own category, such that the contextual information—the scandal-ridden politicians—should not be included in the temporary representations that participants make of John Doe. Yet, at the same time, this contextual information was expected to remain highly accessible in participants’ minds, allowing it to be activated when participants attempted to make a judgment about John Doe. If this was the case, then it was expected that information about the scandal-ridden politicians would be used as a standard against which John Doe could be compared and contrasted.

To test their hypotheses, Schwarz and Bless (1992) asked study participants to name at least two German politicians involved in a specific (recent) political scandal, either before or after they responded to the measure of the dependent variable. Depending on the condition to which a participant was assigned, the measure of the dependent

variable was either his or her evaluation of the trustworthiness of German politicians in general, or a trustworthiness evaluation of three *specific* politicians who were not involved in the scandal, but who nevertheless had been rated as especially *untrustworthy* in a pretest. In other words the target being judged was the dependent variable.

Schwarz and Bless found that trustworthiness evaluations of German politicians in general decreased when participants thought about the scandal before they responded on the dependent variable measure. Ostensibly, this finding of an assimilation effect reflects the fact that participants included the scandal-ridden politicians within their representation of the category ‘German politicians in general.’

On the other hand, it was found that ratings of the trustworthiness of the three specific politicians increased when people thought about the scandal before responding on the measure of the dependent variable. Presumably, this contrast effect was due to the fact that the contextual information was easily accessible to participants who rated the dependent variable first, and thus it served as a standard against which the three specific politicians could be compared.

In response to these findings, Bless and Schwarz (1992) note that, “this contrast effect cannot be accounted for on the basis of a mere subtraction process” (p.226). They explain that the scandal questions primed information which would never have been part of the representations that subjects made for the three specific politicians. Thus, the contrast effect that resulted must have been the result of the scandal-ridden politicians, recalled from the previous task, serving as a standard against which comparative judgments could be made. Bless and Schwarz (1992) go on to say that, “This information could only be used in constructing the standard, however, when it was not

perceived to bear on the respective target category in the first place” (p.226). In other words, the assimilation effects that resulted from the priming emerged because the category being evaluated was wide; thus, these politicians were able to be included in its representation. At the same time, the contrast effect that resulted from the task of naming the scandal-ridden politicians before evaluating the specific politicians reflects the fact that the categorical representations of these specific people were *narrow* and thus the primed politicians could not be included within them. Instead, they served as a standard for comparison.

Overall, these findings serve as evidence for the notion that directional variation in the judgment process can occur even with use of the very same contextual information; what dictates the direction of an evaluation is whether or not the target stimulus being judged causes a person to include or exclude contextual information within a representation.

Another contention put forth by Schwarz and Bless is the suggestion that a main determinant of whether or not information will be included within a category representation is the degree to which a stimulus is perceived to be representative of the category. They suggest that highly representative information is much more likely to be included within a representation than non-representative information. They cite several different studies (Strack, Schwarz and Gshneidinger, 1985; Strack, Schwarz and Nebel, 1987) as evidence for this phenomenon. Both of these studies focused on representativeness with respect to life events; they showed that the subjective perception of a life event as pertaining (or not) to a person’s current situation moderates use of accessible information about that person’s life. In other words, these studies showed that

the degree to which a given target (i.e. the event) is representative of a category (i.e. a person's current life situation) influences use of accessible information, leading to assimilation if the information is included into a temporary category representation, and contrast if it is excluded.

Schwarz and Bless also suggest that the order in which target evaluations are made may influence the emergence of assimilation and contrast effects. As evidence, they cite a 1983 study by Martin & Seta, in which the timing of a judgment was manipulated. In this study, participants rated two target individuals; in one condition, information was presented about both targets prior to the evaluations. In the other condition, participants rated each target separately, immediately after having received information about the particular target. It was found that judgments about the two targets were assimilated towards one another when the information was presented together and the targets were judged together, whereas contrast resulted when the targets were presented and judged separately.

In summary, the inclusion/exclusion model put forth by Schwarz and Bless (1992) "assumes that assimilation and contrast effects are a function of categorization processes" (p.237). Depending on whether a target is included or excluded within the contextual representation called to mind, assimilation or contrast will result. Furthermore, observed contrast effects will reflect the operation of one two processes: either a process of subtraction or comparison. The authors of the theory also indicate that there are a multitude of other variables that can effect whether assimilation or contrast will result in instances of evaluation

The Interpretation/Comparison Model. Like Schwarz and Bless before them, in developing their interpretation/comparison model of social judgment, Stapel and Koomen (1998) hoped to create a framework that takes into account the multiplicity of factors that they saw as relevant to the judgment process. They suggest that several factors play a key role in determining whether contextual information will be used as “an interpretation frame” for evaluating a target stimulus, which would lead to assimilation, or whether it will be used as “a comparison standard,” against which the target is judged, which would lead to contrast effects (Stapel and Koomen, 1998, p.240). They argue that these factors include: *target ambiguity, interpretation relevance, distinctiveness, and comparison relevance.*

In proposing *target ambiguity* as a factor which influences social judgment, Stapel and Koomen note that the degree to which a target is ambiguous will determine the extent to which an individual relies on contextual information when faced with the task of making an evaluative judgment. If a target is relatively ambiguous, then contextual information will be heavily relied upon in the judgment process, and an evaluation of the target will be assimilated towards the context. They cite several studies in support of this contention (Koomen, Stapel, Jansen and In't Veld, 1998; Stapel and Koomen, 1997).

Stapel and Koomen then go on to suggest that *interpretation relevance* also plays a role in social judgment. Citing prior research (Higgins et al, 1977; Srull and Wyer, 1979) as evidence, they assert that in order for a target evaluation to be assimilated towards evaluations of contextual information, the context must be applicable and relevant to the target. The less similar the target and the context are perceived to be, the

less likely it is that contextual information will be taken into account when a target evaluation is made.

With regards to *distinctiveness*, Stapel and Koomen suggest that the more a contextual representation is perceived to have this quality, the more likely it is that the context will become a standard against which a target stimulus is compared. Presumably, this is because traits belonging to a distinct contextual representation are seen as unique to that stimulus, and thus it is unlikely that they would extend to other stimuli. Thus, instead of being included in the category to which the distinct contextual stimulus belongs, a target stimulus would be compared and contrasted with it. This concept of ‘distinctiveness’ is the complement of the concept of ‘category width’ proposed in Schwarz and Bless’ Inclusion/ Exclusion Model. Stapel and Koomen support this contention with empirical evidence from a 1983 study by Martin and Seta; it would also seem that further evidence for the effects of distinctiveness in the judgment process come from the second study by Schwarz and Bless (1992), described previously.

Stapel and Koomen also maintain that *comparison relevance* is a key factor affecting social judgments. Evidence for their contention comes from prior findings by Helson (1964) and Brown (1953), the latter of which were discussed previously, which indicated that in order for comparison contrast to occur, a contextual representation must be perceived as relevant to the target stimulus. This makes sense given that it would be unlikely for a person to judge a target against a completely irrelevant standard. For example, one cannot judge whether a given apple is ripe or not in comparison with a ripe banana.

In summary, Stapel and Koomen (1998) present a theory of assimilation and contrast effects that serves to integrate the various findings within interpretive judgment theory, as well as extend previous theories. The interpretation/comparison model assumes that factors of target ambiguity, interpretation relevance, distinctiveness and comparison relevance all affect how contextual information will be utilized in the judgment process, which in turn leads to directional variation in evaluation. In doing so, this theory provides a framework that can account for a multiplicity of factors that lead to directional variation in social judgments.

The Selective Accessibility Mechanism. One of the more recent interpretive judgment models to be proposed is Mussweiler's (2003) Informational Perspective on Comparison Consequences. Mussweiler proposes that there are three main steps involved in the judgement process, *standard selection*, *target-standard comparison*, and *evaluation*. In the first step, a standard is chosen, in light of which the target will be evaluated. Once this contextual representation has been created, an individual then selects the particular features of the target and category upon which to base an evaluation. Finally, once the judge has identified the critical features upon which an evaluation will be based, the actual evaluation is carried out.

Mussweiler suggests that it is the second, comparative stage, which is most important, as it determines the outcome of the evaluative process. He asserts that the primary process underlying this particular stage, the "search for and activation of judgment-relevant knowledge during a comparison" is *the selective accessibility mechanism* (Mussweiler, 2003, p. 475). Specifically, before an individual can make an evaluation, he or she must gather information relevant to both the target stimulus and the

category representation, upon which judgments can be based; Mussweiler supposes that the most efficient means of obtaining this knowledge is through a process of hypothesis testing in which prior knowledge of the target is related to the present judgment task.

Specifically, an individual will test whether or not a target stimulus is similar to the category. Mussweiler assumes that the individual will determine which of these two hypotheses to test based on a “quick holistic assessment of the target and the standard” in which he or she assesses just a few features of the target and category in order to judge whether they are similar or dissimilar to one another. In line with prior theories, Mussweiler notes that extremity, relevance, and target ambiguity all influence this assessment. Depending on the results of this initial comparison, the individual will then precede to test the overall similarity or dissimilarity of the target to the context. It is important to note that because the memory search is biased, information consistent with either the similarity or dissimilarity hypothesis will become more accessible and is therefore more likely to be included in the target representation.

If the results of the initial holistic assessment indicate that the target and standard are similar, then the judge should precede to search for “standard-consistent knowledge—evidence indicating that the target’s standing...is indeed similar to that of the standard” (Mussweiler, 2003, p.476). An evaluation indicating that the target is similar to the context along a given dimension will result in an assimilation effect. This is because the search for hypothesis-consistent information produces a representation of the target that contains a disproportionate number of features that resemble the features in the context.

On the other hand, if the result of the initial assessment is an indication that the target and category are dissimilar, then the judge should precede to evaluate the degree of

dissimilarity between the two, by searching for information about the target that is inconsistent with the standard. An evaluation indicating that the target is dissimilar to the context on a given dimension will result in a contrast effect.

Empirical Support for the Selective Accessibility Mechanism. A variety of studies lend support for the different aspects of Mussweiler’s selective accessibility mechanism. Support for the active hypothesis testing mechanism comes from research indicating that “the nature of the tested hypothesis determines which knowledge is activated” and that the accessibility of that knowledge is “selectively increased” as a result of hypothesis-consistent testing (Mussweiler, 2003, p. 477). For example, in an anchoring study (Mussweiler and Strack, 1999b) the researchers were able to manipulate the type of hypothesis that participants tested across conditions by altering the wording used in the initial evaluative judgment. Whereas participants in one condition were asked to judge whether a target was larger than the value of the anchor, participants in the other condition were asked to determine if the target was *smaller* than this value (which remained constant across conditions). It was found that participants’ evaluations of the target were highly dependent on the nature of the wording of the evaluative questions, such that lower estimates were given when the participants were asked if the target was *smaller* than the anchor, and higher estimates were given among participants who were asked if the target was *larger* than the anchor value. This study serves as an indication that during the judgment process, people search actively for target information based on the influence of contextual information. Therefore, similarity testing will selectively increase the degree to which standard-consistent information about the target becomes

accessible, whereas dissimilarity testing will selectively increase the degree to which standard-inconsistent target information becomes accessible.

Mussweiler maintains that there is ample evidence in support of the operation of the selective accessibility mechanism in the evaluative process. In addition, he suggests that the basic processes of selective accessibility may be supplemented in a variety of ways by auxiliary processes. Overall, his model serves to integrate and extend prior interpretive judgment theory.

Summarizing the Interpretive Process. The theories of interpretive judgment presented in this section each account for the multidimensional analyses of complex stimuli that people make each day. Research into the interpretive judgment process has shown how inclusion of a target within a contextual representation leads to assimilation effects, while exclusion of the target from the context leads to the effect of contrast. While the majority of researchers agree that the most important factor mediating the direction of an evaluation is the way in which target information is categorized, these theories have also shown that there are a variety of other factors such as relevance, representativeness, distinctiveness, category accessibility, target ambiguity, and category width, which may also affect the direction in which a judgment is made.

Prior Theory and Continuing Research

Summarizing the Judgment Research. The previous sections of this paper were intended to provide the reader with a brief overview of the core aspects of judgment theory in order to place the present research within the context of a larger theoretical

background. The following chart briefly summarizes the most important points of the major theories that have been presented thus far:

Parducci	<i>Range/Frequency Compromise.</i> The accepted theory of comparative judgment; suggests that assigning a stimulus to a category represents a compromise between principles of range and frequency.
Kahneman and Miller	<i>Norm Theory.</i> A comparative judgment which suggests that a stimulus event acts as a probe in response to which multiple remembered and constructed representations come to mind; a 'norm,' the context against which the stimulus is evaluated, is an aggregate of these representations.
Martin	<i>Set/Reset Theory.</i> An interpretive judgment theory developed to explain directional variation in the evaluative process. In "set," relevant contextual information is allowed to come into play in an evaluation, leading to assimilation effects. In "reset," irrelevant or blatantly primed contextual information is suppressed during judgments, leading to contrast effects.
Schwarz and Bless	<i>Inclusion/Exclusion Model.</i> Interpretive judgment theory suggesting that it is the way in which contextual information is categorized that mediates directional variation in the judgment process. Inclusion of a target within the context leads to assimilation effects, exclusion leads to contrast effects. Contrast may result from a comparative or subtractive process.
Stapel and Koomen	<i>Interpretation/Comparison Model.</i> Interpretive judgment theory which takes into account the influence of multiple factors in determining whether contextual information will be used as an interpretation frame or comparison standard for evaluating a target stimulus. Factors include: appropriateness, extremity, target ambiguity, interpretation and comparison relevance.
Mussweiler	<i>Selective Accessibility Mechanism.</i> Interpretive judgment theory suggesting that the degree to which contextual knowledge affects evaluation is mediated by accessibility. Through the search for and activation of judgment-relevant knowledge during an evaluation, the comparative stage determines the outcome of the evaluative process,

Continuing Research. It is now generally accepted that a host of processes are at work in instances of evaluation, and that the emergence of assimilation and contrast effects are mediated by a variety of factors. Yet, there is still a great deal of interest among social psychologists in further exploring this theoretical realm, in an effort to produce more accurate and comprehensive models of social judgment processes. Much recent research has focused specifically on the task of exploring assimilation and contrast effects in the evaluation of groups and individuals. These phenomena have been evidenced in a variety of studies, most of which have examined the effects of inclusion

and exclusion on either group evaluations *or* on exemplar evaluations, but not *both* in the same study. Recent research has shown that making an evaluation of one target necessarily changes the evaluation of the other. Bless, Schwarz, Bodenhausen and Thiel (2001) conducted a study that addresses this phenomenon; that is, they had participants evaluate both the individual *and* the group, as opposed to just evaluating one or the other.

Bless et al.: Inclusion and Exclusion of Groups and Individuals. Bless et al. explored this problem using a negatively stereotyped group (i.e., the Roma and Sinti, a gypsy group that is widely discriminated against in Germany, where the study was conducted.) In this study, participants heard a tape recording in which a moderately atypical member of this group was favorably described. Then, they answered a question designed to either elicit inclusion or exclusion of the exemplar in or from the group. In the inclusion condition, participants were asked how well integrated the individual was into the category; given four options, they were expected to select “very well integrated” or “well integrated.” In the exclusion condition, participants were asked to select which category best described the exemplar from the choices provided. Participants were expected to choose “Rom, but an exception” (Bless et al., 2001, p. 390). A third group, the control condition, was asked a neutral question, unrelated to the target. The dependent variable being investigated was participants’ evaluations of the stereotypicality of both the individual and the group.

As predicted, Bless et al. (2001) found an assimilation effect when the atypical exemplar was included in the representation of the group. Furthermore, a contrast effect resulted when that same person was excluded from the group. Specifically, when participants were prompted to categorize the individual as a well integrated member of

the Roma and Sinti, the group was evaluated less stereotypically than it was by participants in the exclusion condition. As well, the exemplar was evaluated more stereotypically by participants in the inclusion condition than by those in the exclusion condition.

Assimilation and Contrast in Bless et al explained. Bless et al. (2001) assert that, “both of these assimilation effects are presumably based on changes in the representation formed of the respective target of judgment” (p. 391). In other words, assimilation effects occurred because the context (i.e., the question the participants answered) increased participants’ readiness to view the exemplar as a member of the group category. Specifically, when the individual was included in the group category, positive traits associated with that person influenced evaluations of the negatively stereotyped group. As well, negative traits associated with the group negatively influenced evaluations of the individual. Thus, evaluations of the exemplar and group were assimilated towards one another.

In the exclusion condition, positive traits associated with the exemplar were compared to and thus contrasted with negative traits ascribed to the group; as well, the negative traits associated with the group were contrasted with the positive traits of the exemplar. According to the authors of the study, these contrast effects resulted from a change in the participants’ standard of comparison. Bless et al. (2001) explain these effects, stating:

The exclusion of the exemplar from the group decreases the likelihood that group related information enters the representation of the exemplar. On the other hand, it increases the likelihood that the group serves as a standard of comparison, relative to which the exemplar seems all the less stereotypical.

(Bless et al., 2001, p. 391)

This prior research has shown that the categorization of a moderately atypical exemplar as either within or without a group affects subsequent evaluations of both the group and the exemplar, but in opposite ways.

The Present Research

In the research by Bless et al. (2001) information about the exemplar was presented orally; however, it is important to note that in day to day life, individuals make evaluative judgments based on more than just what they hear. Oral information is usually supplemented by visual information, and individuals use both when making evaluations about other people.

The present research attempts to extend the findings of Bless et al. (2001) through a methodologically similar study using a different stereotyped group and, more importantly, a presentation of exemplar and contextual information that allows participants to absorb information through both the auditory and visual modalities. Specifically, the present research makes use of film as a medium for presenting such information, because it should allow study participants to generate a rich representation of the exemplar based on a variety of factors from multiple modalities, including: appearance, speech, intonation, posture, interactions with others, and so on. Furthermore, in order to increase the validity of the previous findings, as well as our own, it is important to show that the processes underlying categorization and evaluation remain the same across social groups. Thus, the present research makes use of the elderly as the stereotyped group.

Hypothesis

In modeling our methodology on the research of Bless et al. (2001), we have chosen specifically to examine the effects of inclusion and exclusion on evaluations of both the group and the exemplar, while allowing available information about the exemplar and group to remain constant across conditions. We expect to repeat the previous findings; we anticipate that our results will provide support for the hypothesis that the categorization of a moderately atypical exemplar as either within or without a group will affect subsequent evaluations of both the group and the exemplar, but in opposite ways.

Method:

Participants

145 students (80 males and 65 females) from Haverford and Bryn Mawr Colleges served as participants in this study. 117 of these students were Caucasian, four were African American, ten were Asian, four were Hispanic, and nine classified themselves as ‘other.’¹ The average age of study participants was 20.5 years. Participants were recruited into the study over the phone or by e-mail; every fifth person in the Haverford student phone directory was asked to participate, as well as friends and teammates of the experimenters. Experimenters did not run participants that they knew well; rather, they ‘gave’ such participants to one of the other three experimenters running the study. As

¹ Two other participants chose not to indicate their race.

incentive for participation, all participants were entered into a lottery for one of three \$50.00 prizes.

Design

The present research follows the model of Bless et al. (2001). The study was a 3 (participant condition: inclusion, exclusion and control) x 2 (judgmental target: exemplar, group) x 2 (question order: exemplar questions followed by group or visa versa²) x 2 (participant gender: male, female) x 4 (experimenter: Prachi Dave, Julia Grayer, Jane Nussbaum, Cham Santé) mixed factorial design.

Procedure

The study focused on stereotypes of the elderly. Information about an atypical category exemplar (an elderly woman named Olive) was presented to participants through a series of film clips [taken from the movie “Never Too Late” (Walker, 1997)] documenting her everyday actions. The content of the film provided the following information [adapted from the methods of Bless et al. (2001)]: first, the exemplar was presented in a favorable light; second, the exemplar’s actions demonstrated that she is well integrated into the culture of her social category, the elderly; third, the film suggested that she is somewhat atypical and exceptional for the group. Participants viewed the film on a television set in the Psychology Lab either alone or in small groups. Males and females were run separately. Prior to viewing the film, participants were

² A second order variable—the order of trait presentation within each question—was dropped from the study, because the trait order within each target evaluation was not varied in control questionnaires.

asked to focus their attention on Olive. The instructions which were given to participants are included in Appendix B of this paper.

Following the film presentation, participants were provided with different questionnaires to fill out. A sample questionnaire can be found in Appendix D. Participants were first asked to report, on a scale of one to nine, how absorbed they were in the film. They then were asked to describe what was going on in the video. This question was included in order to increase the salience of the movie in participants' minds, as well as serve as an indication of the extent to which participants understood the clips. The third question that participants answered was the one in which the category manipulation was embedded.

In the first condition, designed to elicit inclusion of the exemplar within the group category, participants were asked how well integrated the target person was into the elderly population. Given four response alternatives, participants were expected to judge the target as "very well integrated" or "well integrated." In the second condition, designed to elicit exclusion of the exemplar from the group category, participants were asked to characterize the exemplar as one of four response alternatives: gambler; elderly but an exception; homeless; invalid. Given the contextual information of the film, participants were expected to indicate that the exemplar was a group member, but atypical. In a third condition, the control condition, designed to elicit no categorization of the exemplar, participants responded to a question pertaining to the film clip, but not specifically to the exemplar; they were asked to indicate which of four card games (Go Fish, Poker, Bridge, or Rummy) the characters had played in the movie.

Next, participants in all three conditions evaluated Olive and the group on a nine-point scale in terms of four of stereotype-relevant and six stereotype-irrelevant traits. The order in which participants evaluated Olive and the group was varied so that half of the subjects evaluated Olive first, and visa versa. The evaluations of the stereotype-relevant traits served as the measure of the dependent variable. The stereotype-relevant traits, selected during pre-testing and based on elderly traits used in a previous study (Bargh, Chen, and Burrows, 1996.,) were: traditional, cautious, dependent, and forgetful. The stereotype irrelevant traits were: articulate, arrogant, natural, musical, jealous, and secretive.³

Participants were then asked to indicate, through an open-ended question, what they thought the study was about. This question was intended to allow the experimenters to analyze the study results in light of the degree to which participants had insight into the purpose of the study. Finally, participants were asked to indicate their age, gender, race, school, and year of graduation.

Results:

We first analyzed participants' responses to the categorization task. The data from twenty-one participants (18 males and 3 females) were excluded from our analysis because they did not respond in such a way that they fulfilled our a prior criteria (i.e., selecting "very well integrated" or "well integrated" in the inclusion condition, and "elderly, but an exception" in the exclusion condition.) Thus, the analyses were run on the data of the 124 remaining participants (62 males, 62 females.) In all analyses n=44 for

³ "Secretive" was dropped from the analysis because it was left off a significant number of questionnaires.

the inclusion condition; n=41 for the exclusion condition, and n=39 for the control. Based on their responses to the first open-ended question, all participants had a clear understanding of what was going on the film clips that they watched.

It was expected that participants in the inclusion condition should evidence an assimilation effect, evaluating the exemplar more stereotypically than participants in the exclusion condition (the degree of stereotypicality being measured on a scale of 1 to 9, with higher scores reflecting more stereotypic judgments.) The ratings of control participants were expected to fall in between. We expected to find an opposite pattern for participants' ratings of the group; that is, we expected the group to be rated less stereotypically in the inclusion condition than in the exclusion condition, with the control condition results again falling in the middle. Thus, it was expected that the difference in mean ratings of the group and exemplar for each trait should be largest for the exclusion condition and smallest for the inclusion condition, with the difference in means for the control group falling in the middle.

Testing the Composite:

The first step of the data analysis was to create a single measure depicting how the exemplar and group were rated in terms of the stereotype-relevant traits (i.e., the degree to which they were seen as stereotypical). To do so, we created a composite variable, a mean trait rating variable with two levels for the dependent variable: one level for ratings of Olive, and one for ratings of the group.

In order to test the hypothesis that there should be a significant interaction between condition and ratings of the exemplar and group, the composite data were

analyzed under a 3 (participant condition: inclusion, exclusion and control) x 2 (judgmental target: exemplar, group) x 2 (question order: exemplar first, group first) x 2 (participant gender: male, female) x 4 (experimenter: Prachi Dave, Julia Grayer, Jane Nussbaum, Cham Santé) mixed factorial ANOVA. As no main effects were found as the result of variations in participant gender or experimenter, we eliminated these variables and performed a new analysis of the composite date, a 3 (condition: inclusion, exclusion, control) x 2 (target: exemplar, group) x 2 (order of DVs: group first vs. exemplar first) mixed ANOVA.

This analysis revealed a highly significant main effect of target on the average trait ratings ($F_{1,118}=279.67, p < .01.$) This indicates that participant evaluations of the group and exemplar differed significantly from one another. The ANOVA also revealed a trend in the data, a near-significant interaction between the composite variable and condition ($F_{2,118}=2.88, p=.06.$) This finding was in line with our hypothesis that participants should give different ratings of the group and individual based on their condition. However, as can be seen in Table 1, below, the form of the interaction was not consistent with the hypothesis; the smallest difference between mean ratings of the exemplar and group did not occur in the inclusion condition—it occurred in the control.

Table 1
Mean Composite Score by Condition and Target

	Exemplar	Group	Difference in Means
Inclusion	3.255	6.432	3.177
Exclusion	2.772	6.170	3.398
Control	3.129	5.513	2.384

Note: Higher scores indicate more stereotypic ratings.

Reliability Testing:

As the composite analysis revealed a trend in support of our hypothesis, we then proceeded to run reliabilities analyses in order to determine the validity of using a composite variable for our data analysis; in other words, we wanted to determine if the four stereotype-relevant traits ('traditional,' 'forgetful,' 'cautious' and 'dependent') were all equally relevant to the elderly stereotype and were measuring the same thing. We calculated separate reliabilities for ratings of the group and ratings of the exemplar. The reliability coefficient for Olive was $\alpha = .527$, while the coefficient for the group was $\alpha = .767$, indicating that the composite variable was a reliable measure for evaluations of the group, but not the exemplar. Due to the fact that the reliability coefficient was low for exemplar traits, we examined the ratings in terms of each trait individually in subsequent analyses.

Individual Target Ratings:

Each of the four stereotype-relevant traits was analyzed under a 3 (participant condition: inclusion, exclusion and control) x 2 (judgmental target: exemplar, group) x 2 (question order: exemplar questions followed by group or visa versa) x 2 (participant gender: male, female) x 4 (experimenter: Prachi Dave, Julia Grayer, Jane Nussbaum, Cham Santé) mixed factorial ANOVA. As predicted, no main effects or interactions were found in any of the ANOVAs as the result of variations in participant gender or experimenter. Therefore, we eliminated these variables from subsequent analyses, including a new set of ANOVAs for the four stereotype-relevant traits; these were 3

(participant condition inclusion, exclusion, control) x 2 (target: exemplar, group) x 2 (question order: group first vs. exemplar first) mixed factorial ANOVAs.

The results of the second set of analyses revealed that the exemplar was indeed evaluated differently from the group, as indicated by the highly significant p-values for each trait: ‘traditional’: $F_{1,118} = 252.721, p < .000$; ‘forgetful’: $F_{1,118} = 190.430, p < .000$; ‘cautious’: $F_{1,118} = 115.735, p < .000$; ‘dependent’: $F_{1,118} = 85.518, p < .000$. Specifically, the exemplar was rated less stereotypically than the group (for a table of mean target ratings by trait see Table 2), a finding that was independent of the order condition.

Table 2
Mean Target Ratings by Trait

	Exemplar	Group	Difference in Means
Traditional	2.936	6.508	3.527
Forgetful	2.411	5.653	3.242
Cautious	3.613	6.226	2.613
Dependent	3.307	5.831	2.524

Note: Higher scores indicate more stereotypic ratings.

The Effects of Condition on Target Ratings. Our main hypothesis predicted a differential effect of condition (i.e. the categorization task) for evaluations of the group versus the exemplar. In general, the data did not support this hypothesis. The mean target ratings for each of the four traits appear by condition in Tables 3-6, below, followed by brief elaboration of their meanings.

Ratings for Traditional.

Table 3
Mean Ratings for Traditional as a Function of Target, Condition

	Exemplar	Group	Difference in Means
Inclusion	3.029	6.854	3.825
Exclusion	2.896	6.567	3.671
Control	2.818	6.036	3.218

Note: Higher scores indicate more stereotypic ratings.

The above data do not support the main hypothesis; there was not a significant interaction between condition and the target ratings for ‘traditional,’ ($F_{2,118} = .650$, $p < .524$.) The difference in mean ratings for the group and exemplar did not vary by condition as predicted; instead, the difference was *largest* in the inclusion condition and *smallest* for the control, with the difference in means for the exclusion condition falling in between.

Ratings for Cautious.

Table 4
Mean Ratings for Cautious as a Function of Target, Condition

	Exemplar	Group	Difference in Means
Inclusion	3.646	6.600	2.954
Exclusion	3.455	6.304	2.849
Control	3.721	5.730	2.009

Note: Higher scores indicate more stereotypic ratings.

The data in Table 4 do not support the main hypothesis; there was not a significant interaction between condition and the target ratings for ‘cautious’: $F_{2,118} = 1.492$, $p < .229$. Participants’ judgments were not moving the direction predicted; the difference in means was *highest* for the inclusion condition, followed by exclusion and then the control.

Ratings for Dependent.

Table 5
Mean Ratings for Dependent as a Function of Target, Condition

	Exemplar	Group	Difference in Means
Inclusion	3.858	6.137	2.279
Exclusion	2.599	5.868	3.269
Control	3.391	5.488	2.097

Note: Higher scores indicate more stereotypic ratings.

While the means for ‘dependent’ did not reach significance ($F_{2,118} = 1.789$, $p < .172$), the data in the chart above indicate that participants’ judgments *were* moving in

the predicted direction. The difference in mean ratings of the exemplar and group was *lower* for the inclusion than for the exclusion condition, although the difference in means for the control condition was still the lowest (i.e., not as predicted.)

Ratings for Forgetful.

Table 6
Mean Ratings for Forgetful as a Function of Target, Condition

	Exemplar	Group	Difference in Means
Inclusion	2.488	6.138	3.650
Exclusion	2.139	5.942	3.803
Control	2.587	4.796	2.209

Note: Higher scores indicate more stereotypic ratings.

There was a significant interaction between condition and ratings for ‘forgetful,’ ($F_{2,118} = 4.593, p < .012$.) As predicted, the difference in means was *highest* for the exclusion condition. However, contrary to our hypothesis, the difference in means for the control *was lower* than for the inclusion condition.

Analyzing the Forgetful/Condition Interaction. In order to understand the interaction between condition and ‘forgetful,’ we next calculated the difference scores for that trait, based upon which a post-hoc analysis could later be run. We calculated the difference scores by subtracting the mean exemplar rating for ‘forgetful’ from the mean group rating. We then ran a one-way ANOVA using the difference variables for each score. As predicted, the analysis revealed a significant interaction between condition and the difference score for ‘forgetful’ ($F_{2,121} = 4.152, p < .018$.) The mean difference score ratings were 3.650, 3.803, and 2.209 for the inclusion condition, exclusion condition and control, respectively⁴.

⁴ The actual values involved in the test were slightly different due to rounding errors; they were 3.636, 3.781, and 2.231 for the inclusion condition, exclusion condition and control, respectively.

We then ran a Bonferroni post hoc analysis for the above interaction. The results of the analysis revealed that there was a significant difference in ratings for 'forgetful' between the control and the exclusion condition ($p < .031$), as well as a similar near-significant trend between the control and the inclusion condition ($p < .053$.) These findings indicate that the significant interaction initially found between condition and ratings of 'forgetful' does not support our hypothesis that participants should rate the targets differently in terms of the traits as a result of condition. What these findings show is that the significant interaction between forgetful and condition was not due to differences in the ratings of 'forgetful' between the inclusion and exclusion conditions, but rather between the control and the inclusion condition and the control and the exclusion condition only.

Insight Effects:

In an effort to better understand why the results in general did not support our hypothesis, we decided to analyze whether there might be any interaction between responses to the evaluative tasks and the degree to which participants had insight into the nature of the study. While we predicted that no such interaction would be revealed, we wanted to ensure that participants who better understood the nature of the study were not responding differently on the categorization task than participants who were unclear about the nature of the study. In order to assess the degree of participant insight, we made use of the responses to the open-ended question asking what participants thought the study was about. Each response was rated separately by two different experimenters, who were blind to participant condition. The responses were rated on a scale of one to

three. A rating of ‘one’ indicated that the participant had no insight into the nature of the study; a rating of ‘two’ indicated that a participant had moderate insight into the study; a rating of ‘three’ indicated that a participant had a high degree of insight into the study. Appendix E of this paper includes examples of these responses and the ratings they were given. The inter-rater reliability was calculated at 83%.

The two ratings were then averaged for each individual response. These average ratings were then collapsed into two groups; average ratings of one were placed into the first group (N=83), while all other average ratings were placed into a second group (N=41.) We created a new two-level variable, Insight, to describe these two groups. Thus, we were able to compare the results of participants who had some insight into the study versus participants who had none. These data were then analyzed against the difference scores for each stereotype-relevant trait in a 3 (condition: inclusion, exclusion, control) x 2 (insight: no insight, some insight) x 2 (target: exemplar, group) x 2 (order: exemplar first, group first) x mixed factorial ANOVA. The results of this analysis revealed no significant interaction between condition and degree of insight ($F_{2,112}=1.194$, $p<.307$.)

The Effects of Order:

Our data analysis revealed a particularly interesting, though unexpected finding: a significant interaction between order (i.e., whether the exemplar or group was evaluated first) and the mean ratings of Olive and the group for all four of the traits⁵: ‘forgetful’: $F_{1,118}= 7.275$, $p <.008$; ‘cautious’: $F_{1,118}= 5.402$, $p<.022$; ‘dependent’: $F_{1,118} = 14.128$,

⁵ It would have been useful to run a post-hoc test on these data, in order to determine for certain the origins the observed interaction; however, we were unable to run this analysis using SPSS.

$p < .000$; and ‘traditional’: $F_{1,118} = 3.941$, $p < .049$. Across the board, Olive was rated less stereotypically when participants rated her before the group. As well, the group was rated more stereotypically when participants rated the exemplar before the group. These means appear below in Table 7.

Table 7
Mean Relevant-Trait Ratings by Order, Target

Trait Rated	Order	Exemplar	Group
Forgetful	Exemplar First	2.300	6.150
	Group First	2.509	5.100
Cautious	Exemplar First	3.400	6.567
	Group First	3.814	5.856
Dependent	Exemplar First	2.950	6.517
	Group First	3.615	5.146
Traditional	Exemplar First	2.717	6.733
	Group First	3.113	6.238

Note: Higher scores reflect more stereotypic evaluations

Discussion

While the results of the data analyses did not support our main hypothesis, the analyses did yield some very interesting interactions and overall results; there is still a great deal of value in attempting to understand and explain our findings. The following section attempts to shed some light on how and why we obtained the results that we did; furthermore, it attempts to explain the implications of these data, and what they suggest for further research.

The Composite and Reliabilities Tests:

Assuming that all of the stereotype-relevant traits were equally relevant to the elderly stereotype and were measuring the same thing, it would have been ideal to analyze target ratings in terms of the composite variable. Unfortunately, these results were essentially rendered meaningless when we ran reliabilities analyses on the composite variable and found a very low reliability coefficient for ratings of the exemplar. Even though the reliability coefficient was much higher for the group ratings, the former finding indicated that the stereotype-relevant traits we had selected were not in fact equally relevant to the stereotype, and that they were not measuring the same thing.

The Dependent Variable. That exemplar ratings in terms of the stereotype-relevant traits did not correlate well is probably due, for the most part, to time constraints; that is, we did not have adequate time to run all the necessary pre-tests that would have allowed us to select relevant traits that were accurate measures of the stereotype. In short, the measure of the dependent variable was poorly chosen. More specifically, it is quite possible that the stereotype-relevant traits were not equally relevant to the elderly stereotype, and may not have been measuring the same thing. Perhaps in evaluating the targets in terms of these traits, participants really weren't evaluating the degree to which the targets matched the group stereotype across the board, but only in some cases.

A further possibility is that these trait evaluations weren't measuring the descriptive value of the targets (i.e., the degree to which the exemplar and group were stereotypical), but rather were evaluative measures. According to Brown (1986) every trait term has two different meanings: an evaluative meaning and a descriptive meaning. In order to understand this concept, it is helpful to consider two similar trait words: thrifty

and stingy. Describing a person using either of these traits would imply that the individual is careful when spending money (i.e., the descriptive value of the words is the same) but the word stingy has a much more negative connotation than does thrifty.

With respect to the present research, it is worthwhile to consider that perhaps when people were completing the judgment task, they were judging the targets in terms of the evaluative—rather than descriptive—meaning of the stereotype-relevant traits. In other words, the traits that we chose may have caused people to make a value judgment about the target. It is certainly possible that participants felt that to categorize a person or group as ‘traditional,’ ‘forgetful,’ ‘dependent,’ or ‘cautious’ was to make a negative evaluation of that person. Taking that into account, it is possible that in making use of those traits, what we ended up measuring was participants’ value judgments of the targets, the degree to which they perceive Olive versus the elderly in general in a positive or negative light, rather than the stereotypicality of the targets.

Given that the composite score for judgments of the exemplar ($\alpha = .527$) was lower than that for the group ($\alpha = .766$), and given that a higher rating would indicate a more negative target evaluation, it seems highly plausible that participants were in fact responding to the evaluative meaning of these traits and were simply more willing to evaluate the group, rather than the exemplar, in a negative way.

Individual Target Ratings: Possible Explanations for the Results

Our analysis of the effects of differential target categorization on evaluations of the group and the exemplar in terms of the stereotype-relevant traits did not support our main hypothesis. Even the significant findings that there was an effect of condition on

target ratings in terms of the trait 'forgetful' did not support the hypothesis, as the post-hoc analysis revealed that any differences in the way the targets were categorized were due to differences between the ratings of participants in the control and inclusion/exclusion conditions, but *not* between the ratings of people assigned to the inclusion and exclusion conditions alone.

Insight. Initially, we thought that these findings might have been linked to variation in the degree to which participants had insight into the purpose of the study. However, the results of our analysis including insight as a factor revealed no interaction between participant insight and target evaluations. However, there are several other possible explanations for these findings, outlined in the following paragraphs.

The Dependent Variable. As was mentioned previously, the first possible explanation for these findings is that the measure of the dependent variable was poorly chosen. That is, it is quite possible that the stereotype-relevant traits were not equally relevant to the elderly stereotype, and may not have been measuring the same thing. Alternately, it is also possible that our findings were due to the fact that the evaluative task was actually a judgment of the targets in terms of the evaluative, rather than the descriptive, meanings of the trait. In other words, this task was not an effective measure of the stereotypicality of the targets.

The Manipulation of the Independent Variable. Another possible explanation for why our results failed to support the main hypothesis is that it was the measure of the independent variable rather than the manipulation of the dependent variable, that was ineffective. Given that we did find a strong interaction between order and target evaluations, given that people evaluated Olive and the group very differently depending

on who was evaluated first, it is possible that the relevant traits were in-fact measuring the same thing and were decent measures of the dependent variable. If this is the case, then one explanation for the fact that we didn't find an effect of condition could be that we didn't effectively manipulate this variable.

Perhaps the embedded manipulation was too subtle, and didn't elicit inclusion or exclusion of the exemplar within or from the elderly category. Although this might be difficult to argue, given that a similar manipulation did produce the desired effects in the study by Bless et al. (2001) upon which the present research is based, one must take into account that the present research and the former employed very different methods of presenting exemplar information. Specifically, our method allowed for the presentation of much more complex stimuli, and it is possible that people evaluate such rich stimuli differently than they do other stimuli; thus, it could be argued that we should never have expected to get similar results to Bless et al.

Overshadowed: The Order Interaction. One final explanation for these findings is the possibility that while both the manipulation and measure of the dependent variable were effective, the effects of differential exemplar categorization on target ratings were washed out by the interaction between the order in which the targets were rated and the ratings themselves. That is, while participant evaluations of the targets might have been effected by the categorization task, perhaps such effects were overshadowed and rendered null by the even stronger effects of order on the way in which targets were evaluated.

Summing up the Possibilities. While all of these are possible explanations for the findings that there were no significant main effects of differential target categorization on target evaluations in terms of the stereotype relevant traits, it is important to realize that

these are all merely post-hoc hypotheses. There are a variety of different factors, including the effectiveness of the measure of the DV, the strength of the manipulation of the DV, and the strength of the order/target rating interaction, which could have led to our obtaining the results that we did. However, it is virtually impossible to select one as the 'right' explanation.

Understanding the Effects of Order:

Though unexpected, the finding that there was a significant interaction between order and evaluations of the target is highly interesting, as it can likely be accounted for by the majority of theories of interpretive judgment presented in this paper. Specifically, when the group was evaluated prior to the exemplar, ratings of Olive and the group were assimilated towards one another (i.e., Olive was rated more stereotypically and the group was rated less stereotypically than in the other order condition). The opposite finding, that ratings of the exemplar and group were contrasted away from each other (i.e., Olive was rated less stereotypically and the group was rated more stereotypically than in the previous order condition), was found when the exemplar was evaluated first. This section of the paper attempts to understand these findings within the framework of prior judgment theory.

The Present Findings Within a Broad Context: It is quite possible that these findings could be explained as the result of factors such as category width, distinctiveness, and initial hypothesis testing, as highlighted in the prior theories of Schwarz and Bless (1992), Stapel and Koomen (1998) and Mussweiler (2003).

As was described earlier, Schwarz and Bless contend that one factor effecting directional variation in the judgment process is category width (i.e., the degree to which a category is able to encompass multiple stimuli.) They contend that the likelihood that a given stimulus will be included within a category increases as the category grows wider. Likewise, the likelihood that a given stimulus will be excluded from a category increases as the category grows narrower. They go on to assert, based on prior research by Wryer & Srull (1989), that when an evaluation is made of a specific exemplar, that exemplar will make up a relatively narrow category of its own, and, in instances of evaluation, judgements of that exemplar should be contrasted away from judgments of other stimuli.

This same idea is also articulated in the Stapel and Koomen's (1998) theory, which suggests that distinctiveness is an important factor affecting the evaluative process. They contend that the more a stimulus is seen to be distinct, the more likely it is that it will become a standard against which other stimuli are compared.

Not surprisingly, both of the above sets of theorists cite a 1983 study by Martin and Seta as evidence that assimilation effects are likely to emerge when stimuli are perceived to be part of a single unit, while contrast effects result when stimuli are perceived as distinct entities. A further conclusion of this study is that the timing of a judgment can strongly influence the "perceived relatedness" of different stimuli, which in turn should effect the direction in which subsequent judgments are made (Schwarz and Bless, 1992, p. 234).

With regards to the present findings, it is possible that in evaluating the exemplar first, participants came to see Olive as a very distinct category of her own, in a way that participants in the other order condition did not. This would make sense given that

participants who evaluated Olive first had a good deal of time to think about Olive specifically (both in the categorization task and the evaluative task) before rating the group, so that by the time they had to complete this second evaluative task, Olive stood out as a distinct personality, becoming a standard against which the elderly in general were compared.

On the other hand, when the group was evaluated first, participants weren't really asked to ruminate much about the exemplar; rather, they were forced almost immediately to construct a broad category representation of the group, the elderly, to be used in the evaluative task. Because participants were forced to construct such a broad representation, it is quite possible that they ended up including the exemplar within that category in the subsequent evaluation of Olive. This would explain the assimilation effects that resulted in this order condition.

Thus far in this discussion, Mussweiler's theory has not been highlighted. He theorizes that before a person makes an evaluation of a stimulus through a series of similarity or dissimilarity tests, he or she will determine which types of tests should be made through a quick holistic assessment of the stimuli. Mussweiler notes that factors such as target ambiguity, distinctiveness, extremity and so on all influence this initial assessment. Looking at the present research with respect to this theory, it could be argued that the same factors illuminated earlier (i.e., distinctiveness/category width) effected participants' initial assessments of whichever target they were judging first. If the exemplar was judged first, than probably her distinctiveness would come into play in this initial judgment, and thus participants in this condition proceeded to test for dissimilarity, leading to contrast effects. However, if the group was judged first, then

probably the broad, ambiguous nature of the group category would influence this initial assessment, leading participants to test for similarity. This would explain the contrast effects that resulted in this order condition.

Overall, it seems safe to say that while the interaction between order and target evaluation was not something that was expected, it isn't actually all that surprising a finding, given that such effects can seemingly be accounted for in much of the interpretive judgment theory. While each theory would explain these phenomena using different terms, they essentially account for them in a similar way.

Conclusion: Recommendations for Further Research

The present research attempted to examine the effects of inclusion and exclusion of an atypical exemplar within or from a group on evaluations of both the group and the exemplar. It predicted that inclusion of the exemplar should lead to assimilation effects, while exclusion of the exemplar should lead to contrast.

While the results of our research do not provide support for the main hypothesis, it is still important that future research continue to examine the effects of inclusion and exclusion on stereotypical evaluations of groups and individuals. Future research should go beyond the work of Bless et al. (2001), attempting to extend their findings to different stereotyped groups, as well as accounting for judgments made based on rich contextual information presented in realistic ways. Thus, future research might make use of a similar model to the one presented in the present research.

Clearly, our results indicated that both of the targets examined (i.e. the elderly and Olive) were effective selections, as people were willing and able to make differential

judgments about their stereotypicality. However, future research must include a more accurate measure of this dependent variable. This could most likely be achieved through a more extensive process of pre-testing than was allowed for in the present research. Selection of traits that prove to be equal measures of the stereotype is vital to obtaining significant results. Perhaps with such a change in place, future research along this vein might prove more 'successful' than the present research. Such continuing research is important as it allows for a more accurate understanding of the phenomena that social judgment theory attempts to understand and address.

References:

- Bargh, J., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71, 230-244.
- Bless, H., Schwarz, N., Bodenhausen, G.V., & Thiel, L. (2001). Personalized versus generalized benefits of stereotype disconfirmation: Trade-offs in the evaluation of atypical exemplars and their social groups. *Journal of Experimental Social Psychology*, 37, 386-397.
- Brown, R. (1986). *Social psychology, the second edition*. New York, NY: The Free Press.
- Eiser, J. R. (1990). *Social Judgment*. Pacific Grove, Calif.: Brooks/Cole Publishing Company
- Eiser, R. (1984). *Attitudinal judgment*. NY: Springer-Verlag.
- Helson, H. (1947). Adaptation-level as frame of reference for prediction of psychophysical data. *American Journal of Psychology*, 60, 1-29
- Herr, P. (1986). Consequences of priming: Judgment and behavior. *Journal of Personality and Social Psychology*, 51, 1106-1115. Cited in Moskowitz, G (1998), *Cognitive Social Psychology*, Mahwah, N.J, Lawrence Erlbaum Associates.
- Herr, P., Sherman, S., & Fazio, R. (1983). On the consequences of priming: Assimilation and contrast effects. *Journal of Experimental Social Psychology*, 19, 323-340. Cited in Martin, L. & Tesser, A. (1992), *The Construction of Social Judgments*, Hillsdale, NJ, Lawrence Erlbaum Associates.
- Higgins, E., Rholes, W., & Jones, C. (1977). Category accessibility and impression formation. *Journal-of-Experimental-Social-Psychology*, 13, 141-154. Cited in Stapel, D., and Koomen, W. (1998). Let's not forget the past when we go to the future: On our knowledge of accessibility. In G. Moskowitz (Ed.), *Cognitive Social Psychology*. Mahwah, NJ : Lawrence Erlbaum Associates.
- Kahneman, D., and Miller, D. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93, 136-153
- Koomen, W., Stapel, D., Jansen, S., & In't Veld, K. (1998). Priming and timing: A test of two perspectives. *European-Journal-of-Social-Psychology*, 28, 681-686. Cited in Stapel, D., and Koomen, W. (1998). Let's not forget the past when we go to the future: On our knowledge of accessibility. In G. Moskowitz (Ed.), *Cognitive Social Psychology*. Mahwah, NJ : Lawrence Erlbaum Associates.
- Lombardi, W., Higgins, E., & Bargh, J. (1987). The role of consciousness in priming effects on categorization: Assimilation and contrast as a function of awareness of the priming task. *Personality and Social Psychology Bulletin*, 13, 411-429.
- Martin, L. (1986). Set/reset: Use and disuse of concepts in impression formation. *Journal of Personality and Social Psychology*, 51, 493-504
- Martin, L., & Achee, J. (1992). Beyond accessibility: The role of processing objectives in judgment. In L.L. Martin and A. Tesser (Eds.), *The Construction of Social Judgments*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Martin, L., & Seta, J. (1983). Perceptions of unity and distinctiveness as determinants of attraction.

- Journal of Personality and Social Psychology*, 44, 755-764. Cited in L.L. Martin and A. Tesser (Eds.), *The Construction of Social Judgments*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Martin, L. & Tesser, A. (1992), *The construction of social judgments*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Moskowitz, G. (1998). *Cognitive social psychology*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, 110, 472-489
- Parducci, A. (1961). Supplementary report: The effects of the mean, midpoint, and median upon adaptation level in judgment. *Journal of Experimental Psychology*, 61, 261-262.
- Parducci, A. (1984). Value judgments: Towards a relational theory of happiness. In J. Richard Eiser (Ed.), *Attitudinal Judgment*. New York : Springer-Verlag.
- Schwarz, N & Bless, H. (1990). An inclusion/exclusion model of assimilation and contrast effects in social judgment. *Tagung experimentell arbeitender Psychologen*, Regensburg, FRG. Cited in Martin, L. & Tesser, A. (1992), *The Construction of Social Judgments*, Hillsdale, NJ, Lawrence Erlbaum Associates.
- Schwarz, N., & Bless, H. (1992a). Scandals and the public's trust in politicians: Assimilation and contrast effects. *Personality and Social Psychology Bulletin*, 18, 574-579.
- Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In L.L. Martin and A. Tesser (Eds.), *The Construction of Social Judgments*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Seta, J., Martin L., & Capehart, G. (1979). Effects of contrast and assimilation on the attitude similarity-attraction relationship. *Journal of Personality and Social Psychology*, 37, 462-467. Cited in L.L. Martin and A. Tesser (Eds.), *The Construction of Social Judgments*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Strull, T., & Wyer, R. (1979). Category accessibility and social perception: Some implications for the study of person memory and interpersonal judgments. *Journal-of-Personality-and-Social-Psychology*, 38, 841-856. Cited in Stapel, D., and Koomen, W. (1998). Let's not forget the past when we go to the future: On our knowledge of accessibility. In G. Moskowitz (Ed.), *Cognitive Social Psychology*. Mahwah, NJ : Lawrence Erlbaum Associates.
- Stapel, D., & Koomen, W. (1997). Social categorization and perceptual judgment of size: When perception is social. *Journal-of-Personality-and-Social-Psychology*, 73, 1177-1190. Cited in Stapel, D., and Koomen, W. (1998). Let's not forget the past when we go to the future: On our knowledge of accessibility. In G. Moskowitz (Ed.), *Cognitive Social Psychology*. Mahwah, NJ : Lawrence Erlbaum Associates.
- Stapel, D., & Koomen, W. (1997). Using primed exemplars during impression formation: Interpretation or comparison? *European-Journal-of-Social-Psychology*, 27, 357-367. Cited in Stapel, D., and Koomen, W. (1998). Let's not forget the past when we go to the future: On our knowledge of accessibility. In G. Moskowitz (Ed.), *Cognitive Social Psychology*. Mahwah, NJ : Lawrence Erlbaum Associates.
- Stapel, D., & Koomen, W. (1998). Let's not forget the past when we go to the future: On our knowledge of accessibility. In G. Moskowitz (Ed.), *Cognitive Social Psychology*. Mahwah, NJ : Lawrence Erlbaum Associates.

- Stapel, D., & Winkielman, P. (1998). Assimilation and contrast as a function of context-target similarity, distinctness, and dimensional relevance. *Personality and Social Psychology Bulletin*, 24, 634
- Strack, F., Schwarz, N., Bless, H., Küber, A., & Wänke, M. (1990). *Remember the priming episode!* Paper submitted for publication. Cited in Martin & A. Tesser (1992), *The Construction of Social Judgments*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Strack, F., Schwarz, S., & Gschneidinger, E. (1985). Happiness and reminiscing: The role of time perspective, affect, and mode of thinking. *Journal of Personality and Social Psychology*, 49, 1460-1469. Cited in Martin & A. Tesser (1992), *The Construction of Social Judgments*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Strack, F., Schwarz, S., & Nebel, A. (1987). Thinking about your life: Affective and evaluative consequences. Paper presented at the conference on "Ruminations, Self-Relevant Cognitions, and Stress." Memphis State University, Memphis, TN. Cited in Martin & A. Tesser (1992), *The Construction of Social Judgments*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Walker, G. (Director), & Wodoslawsky, S. & Berry, T. (Producer). (1997). *Never Too Late* [Motion picture]. USA: Warner Home Video.
- Wryer, R. & Srull, T. (1989). *Memory and cognition in its social context*. Hillsdale, NJ: Lawrence Erlbaum Associates. Cited in Martin & A. Tesser (1992), *The Construction of Social Judgments*, Hillsdale, NJ: Lawrence Erlbaum Associates.

Appendices:

Appendix A: Consent Form:

Thank you for participating in our study.

This study investigates how viewers interpret films that are meant to represent realistic people and scenarios. In this experiment, you will watch a series of film clips and then complete a questionnaire regarding the footage. The film is approximately twelve minutes long and the questionnaire should take 10-15 minutes to complete.

Before you begin, we ask that you read the following statement carefully and sign below if you agree to the terms of the study:

“I understand that my responses will remain anonymous, and that I can withdraw from the experiment at any point. I also understand that my name will be entered into a lottery for one of three fifty-dollar prizes and that at the end of the semester I will receive an email explaining the purpose of this study and its findings. Finally, I agree that I will not discuss any part of the study with anyone until I receive the email in May.”

Name: _____

Date: _____

Thank you,

Cham, Jane, Julia & Prachi

Appendix B: Participant Instructions:

Thank you very much for your participation in this study. Just a reminder--you will be entered into a lottery for one of three \$50 prizes; the winners will be announced in late April or early May, and we will contact you if you have won.

In this experiment, we're studying how people interpret video clips that resemble real life situations. Because it would take too long to show a video in its entirety, we will only be showing a series of clips from a single movie.

While this video is being shown, I will leave the room. Please get comfortable and try to enjoy this movie as you would any other, but do not talk while the film is running. The movie will run for about 15 minutes. When it is done, I will come back and give you a short questionnaire about the film. Your responses on this questionnaire will remain anonymous. If you do not feel comfortable answering a question, you may leave it blank.

Participants will be watching video excerpts under different conditions. In this condition, we ask that you focus your attention on the main character, Olive, the elderly woman washing her face during the opening scenes of the video.

Ok, that's it. Thanks again, and enjoy the movie!

Appendix C: Video Description:

Film clips were taken from the movie Never Too Late (Walker, 1997). The clips featured a series of interactions between a group of elderly friends, two men and two women. The series of clips focused in particular on the actions of one of the women, Olive. In the selected scenes, Olive is shown both interacting with the other characters, as well as engaging in such atypical behavior as instant messaging and hacking into a computer.

Appendix D: Sample Questionnaire:

Questionnaire

Experimenter's Name: _____

Date: _____

Participant #: _____

Participant Gender: _____

Please answer the following questions:

1. Please indicate on the following scale, the degree to which you were in absorbed in the film.

1	2	3	4	5	6	7	8	9
not at all								highly
absorbed								absorbed

2. Briefly describe what was going on in these scenes:

3. **(Control)**

Which of the following card games did the characters play in the movie?
(Circle the one):

- Go Fish
- Poker
- Bridge
- Rummy

OR

(Inclusion Condition)

Which of the following *best* describes the main character, Olive?
(Circle the best choice):

- Gambler
- Elderly, but an exception
- Homeless
- Invalid

OR

(Exclusion Condition)

How well integrated into the elderly population is the main character, Olive?
(Circle the best choice):

- Very well integrated
- Well integrated
- Poorly integrated
- Not at all integrated

4. Please evaluate the degree to which the following traits characterize Olive:

Jealous:	1	2	3	4	5	6	7	8	9
	strongly							strongly	
	disagree							agree	

Musical:	1	2	3	4	5	6	7	8	9
	strongly							strongly	
	disagree							agree	
Articulate:	1	2	3	4	5	6	7	8	9
	strongly							strongly	
	disagree							agree	
Dependent:	1	2	3	4	5	6	7	8	9
	strongly							strongly	
	disagree							agree	
Natural:	1	2	3	4	5	6	7	8	9
	strongly							strongly	
	disagree							agree	
Secretive:	1	2	3	4	5	6	7	8	9
	strongly							strongly	
	disagree							agree	
Traditional:	1	2	3	4	5	6	7	8	9
	strongly							strongly	
	disagree							agree	
Cautious:	1	2	3	4	5	6	7	8	9
	strongly							strongly	
	disagree							agree	
Arrogant:	1	2	3	4	5	6	7	8	9
	strongly							strongly	
	disagree							agree	
Forgetful:	1	2	3	4	5	6	7	8	9
	strongly							strongly	
	disagree							agree	

5. Please evaluate the degree to which the following traits characterize the elderly:

Articulate:	1	2	3	4	5	6	7	8	9
	strongly							strongly	
	disagree							agree	
Arrogant:	1	2	3	4	5	6	7	8	9
	strongly							strongly	
	disagree							agree	
Natural:	1	2	3	4	5	6	7	8	9
	strongly							strongly	
	disagree							agree	
Traditional:	1	2	3	4	5	6	7	8	9
	strongly							strongly	
	disagree							agree	
Secretive:	1	2	3	4	5	6	7	8	9
	strongly							strongly	
	disagree							agree	
Forgetful:	1	2	3	4	5	6	7	8	9
	strongly							strongly	
	disagree							agree	
Jealous:	1	2	3	4	5	6	7	8	9
	strongly							strongly	
	disagree							agree	
Cautious:	1	2	3	4	5	6	7	8	9
	strongly							strongly	
	disagree							agree	

Musical: 1 2 3 4 5 6 7 8 9
 strongly disagree strongly agree

Dependent: 1 2 3 4 5 6 7 8 9
 strongly disagree strongly agree

6. Prior to this study, had you seen the film from which these clips were taken?
 Yes No

7. Briefly describe what you believe to be the purpose of this study:

Please provide the following personal information:

Age: _____ Gender: M F School: HC BMC Class: 04 05 06 07

Race: Caucasian
 Asian
 African American
 Native American
 Hispanic
 Other

**Please return this questionnaire to the experimenter once it is completed.
 Thank you for your participation!**

Appendix E: Responses to Question Seven

In question seven, participants were asked the following: “Briefly describe what you believe to be the purpose of the study.” Their responses were then evaluated in terms of the degree to which the participant had insight into the actual purpose of the study. The responses were rated on a scale of one to three, one indicating a low degree of insight, and three indicating a high degree of insight.

-The following responses were given ratings of “one” by the experimenters:

“To understand my reactions to the film and whether the depictions of characters in the film are realistic and believable.”

“Determine the difference between male and female emotional responses to films.”

-The following responses were given a rating of “two” by the experimenters:

“Not quite sure--maybe stereotypes regarding elderly people—in contrast to the way we actually feel about them. For example, examining Olive as an individual and not just an elderly person.”

“To determine how our opinions of the elderly (predetermined) influence how we view an elderly individual we meet.”

-The following responses were given ratings of “three” by the experimenters:

“To evaluate the reaction people have of Olive—a fairly non-traditional elderly woman—to the stereotypical elderly person.”

“To see how perceptions of a specific individual could influence our perceptions of a group.”

