Available online at www.sciencedirect.com
ScienceDirect
Acta Psychologica 125 (2007) 361–389
acta psychologica
www.elsevier.com/locate/actpsy

Of gnomes and leprechauns: The recruitment of recent and categorical contexts in social judgment

Jonathan C. Pettibone [a,*], Douglas H. Wedell [b]

[a] *Department of Psychology, Southern Illinois University Edwardsville, Edwardsville, IL 62026-1121, United States*
[b] *University of South Carolina, Columbia, SC 29208, United States*

Received 10 August 2006; received in revised form 25 October 2006; accepted 26 October 2006
Available online 2 January 2007

Abstract

Participants rated schematic faces from two categories, gnomes and leprechauns, on feature widths and pleasantness of facial configuration. Three target faces shared critical facial features across categories while two contextual faces extended the range for that category to include either very wide or very narrow features. In Experiment 1a, results indicated contextual effects on judgments of target faces when they were rated in separate categorical blocks [Wedell, D. H., & Pettibone, J. C. (1999). Preference and the contextual basis of ideals in judgment and choice. *Journal of Experimental Psychology: General*, *128*, 346–361], but not when faces from both categories were rated together in Experiment 1b. Two additional experiments explored this failure to use categorical information. Categorical context was found to produce contrast effects on descriptive ratings of feature width when participants were forced to rely on name cues rather than actual faces in Experiment 2. In Experiment 3, both contrast effects on descriptive ratings and assimilation effects on ideals for pleasantness were found when the names for the faces were learned separately for each category. These results identify constraints on the nature of category-based stereotyping effects on judgment while isolating the influence of recent and categorical context.

PsycINFO classification: 2340

Keywords: Decision making; Judgment; Classification; Stereotyping; Learning

* Corresponding author. Tel.: +1 618 650 3346; fax: +1 618 650 5087.
E-mail address: jpettib@siue.edu (J.C. Pettibone).

0001-6918/$ - see front matter
doi:10.1016/j.actpsy.2006.10.004

1. Introduction

Two important influences on social judgments of persons, behaviors or events are the recent set of persons, behaviors or events one has encountered and the information represented in categories relevant to the target being judged. The first of these influences has been documented in the literature on context effects in judgment (for reviews, see Eiser, 1990; Mussweiler, 2003; Schwarz & Bless, 1992; Wedell & Parducci, 2000) and the second has been documented in the literature on stereotyping (for reviews, see Hilton & von Hippel, 1996; Macrae & Bodenhausen, 2000). These two types of judgment phenomena have been studied largely in isolation of each other, although some researchers have attempted to examine the relationship between them (e.g., Manis, Biernat, & Nelson, 1991). Others have pointed out how these influences are difficult to separate within commonly used social judgment paradigms (Wedell & Pettibone, 1999). The purpose of the research reported here was to determine how these two sources of evaluative influence might operate on two different types of judgment, dominance judgments and ideal-point judgments (Coombs, 1964), and to understand the conditions under which each may be used.

Dominance judgments are perhaps the more common type of evaluation examined in the social literature and reflect the situation in which values on the response scale are monotonically related to values on the underlying dimension. Thus, for example, ratings of a person's aggressiveness increase with increase in the number of aggressive behaviors, or ratings of a person's likableness increase with number of positive behaviors. Ideal-point judgments, on the other hand, typically follow a nonmonotonic relationship between responses and the underlying stimulus attribute. The specific form of ideal-point judgments tends to be a single-peaked function. These types of judgments are most often found in attitudinal endorsements (Eagly & Chaiken, 1993). For example, a politically moderate individual will disagree with attitudinal positions at the liberal or conservative ends of the spectrum and endorse those in the middle. The stimulus value corresponding to the peak of the endorsement function is referred to as the ideal point and represents the individual's own attitude (Thurstone & Chave, 1929). Item characteristic curves for Likert-type scales typically follow a single peaked function and thereby are consistent with an ideal-point judgment process (Roberts & Laughlin, 1996; Roberts, Laughlin, & Wedell, 1999). Ideal-point functions are also characteristic of many preference domains, in which there can be too much or too little of an attribute, such as the preferred amount of sugar in one's tea or coffee.

1.1. Context effects

The studies reported here build on research concerning how context effects operate differently for dominance and ideal-point judgments (Riskey, Parducci, & Beauchamp, 1979; Wedell & Pettibone, 1999). Context effects have been explored most extensively for dominance judgments and typically can be classified into one of two varieties: contrast or assimilation. Contrast is said to occur when judgments of the target stimuli are displaced away from the values of contextual stimuli. For example, after judging extremely aggressive persons, a moderately aggressive person will tend to be judged as low in aggressiveness (Martin, 1986). Assimilation is said to occur when judgments of the target stimuli are displaced toward values of the contextual stimuli. For example, after unscrambling words indicative of high degrees of aggression, an ambiguous description of a moderately aggressive individual will tend to be judged as high in aggressiveness (Srull & Wyer, 1979). Note that both assimilation and contrast effects tend to push judgments in only one direction when the context consists of extreme values on the underlying dimension.

Contextual effects for ideal-point judgments have received less attention and are often more complex in nature than the simple assimilation and contrast observed for dominance judgments. This is because rather than judgments being displaced toward or away from contextual values, the value defining the ideal point may be displaced toward or away from contextual values. When context has been manipulated for judgments of the same stimuli on dominance and ideal point scales, the typical finding has been a contrast effect for dominance judgments accompanied by assimilation of ideals on the corresponding ideal point scale (Riskey et al., 1979; Wedell & Pettibone, 1999; Wedell, Santoyo, & Pettibone, 2005).

Fig. 1 illustrates these typically obtained context effects. The top panel of Fig. 1 illustrates the contrast effect that occurs when the same set of target stimuli are presented together with one of two different contextual sets of stimuli. For example, Wedell and Pettibone (1999) manipulated the widths of the gaps between the eyes of schematic faces and had participants rate how wide the gap appeared to them, a dominance judgment. They found that when contextual faces had narrowly spaced eyes, moderate target faces were judged to have wide spacing between the eyes; conversely, when the contextual faces had widely spaced eyes, the target faces were judged to have narrow spacing, a contrast effect. Participants were also asked to rate the pleasantness of the facial configuration, an ideal-point judgment task. As expected in such domains, the participants typically judged moderate levels to be most pleasant, with very narrowly spaced or very widely spaced eyes judged less pleasant. The peak or ideal point describing the pleasantness function shifted toward the values of contextual stimuli, an assimilation of ideals illustrated in the bottom panel of Fig. 1. Thus, in the narrow eye gap context, the preferred face had narrower features than in the wide eye gap context. As shown in Fig. 1, this shift in ideals produces a strong cross-over interaction of the perceived pleasantness of faces, reflecting contextually induced preference reversals.

This pattern of results for dominance and ideal-point judgments has been demonstrated across multiple domains. Wedell et al. (2005) demonstrated similar context effects for judgments related to body image. Participants were shown target bodies of moderate widths in two different contexts, one in which body shapes were mostly narrow and one in which they were mostly wide. Similar to the Wedell and Pettibone (1999) study, dominance judgments of body width demonstrated a contrast effect, while ideal-point judgments of pleasantness demonstrated an assimilation of ideals. These results were found with both simple silhouettes and with more realistic, computer generated figures. These effects have also been shown to occur for judgments pleasantness of drinks varying in sweetness (Riskey et al., 1979), judgments of preferred tempo of musical pieces (Holbrook & Anand, 1990), and judgments of the appropriateness of bringing different food items for a dinner party (Cooke, Janiszewski, Cunha, Nasco, & de Wilde, 2004).

1.2. Recruitment of categorical and recent contexts

There is abundant evidence that manipulation of context alters perceptions of a target. There is also evidence that these contextual effects may be moderated by perceived category or group membership. For example, several researchers have shown that factors governing whether stimuli are perceived as belonging to the same or different categories will

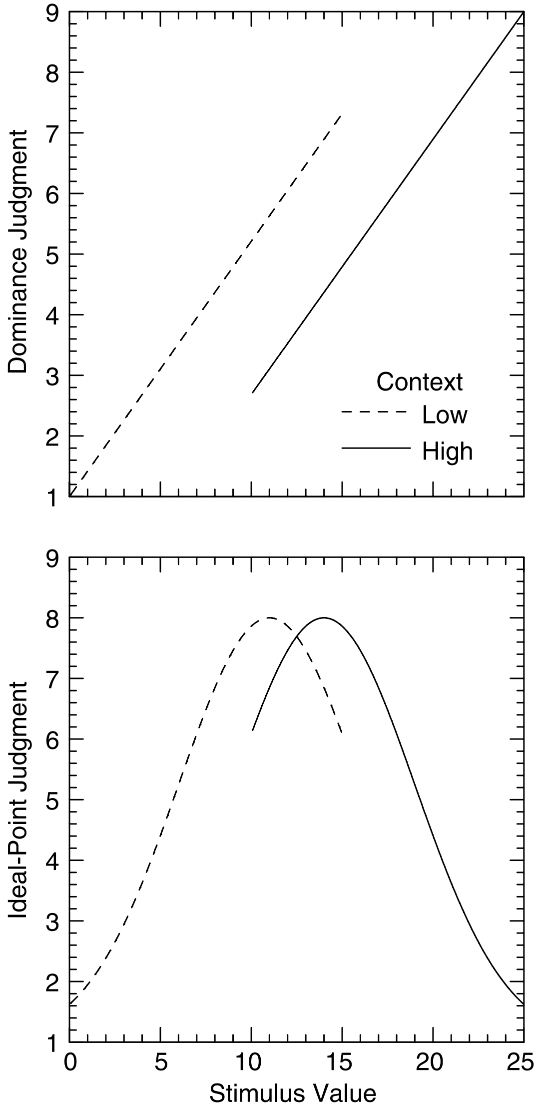

Fig. 1. Typical results of a contextual manipulation for ideal point and corresponding dominance judgment domains. Values for judgments are represented on the *Y*-axis, and values corresponding to characteristics of the stimuli are represented on the *X*-axis. The solid line represents the expected pattern of judgments of stimuli for participants in a context that extends the range of stimulus values upward. The dashed line represents the pattern of results for judgments of stimuli when the context extends the range of stimulus values downward. The top panel shows contrast effects for dominance judgments, reflecting the higher ratings of targets when the range is extended downward. The bottom panel shows assimilation of ideals producing cross-over interactions for ideal-point based preference judgments for the same contextual manipulation.

influence the strength and nature (assimilation versus contrast) of contextual effects (Huttenlocher, Hedges, & Duncan, 1991; Marks, 1992; Mussweiler, 2003; Stapel & Koomen, 1997; Stapel & Winkielman, 1998; Wedell, 1995; Zellner, Rohm, Bassetti, & Parker, 2003). Other researchers have sought to explain stereotyping effects as the result of including category based information in the context used for judgment (Manis, Nelson, & Shedler, 1988), indicating that sometimes these effects lead to assimilation and at other times to contrast.

In examining these effects, we assumed a fairly simple view of how categorical information might influence judgment. We hypothesized that the context for evaluating a stimulus derives from two sources: the recent set of experiences and experiences associated with the categories to which the stimulus is a member. These two routes are depicted schematically in Fig. 2. As elaborated by Kahneman and Miller (1986), we assume that presentation of a target stimulates a process for recruiting a contextual set with which to evaluate the target. One source for constructing this set is the recent set of experiences residing to working memory. In a typical judgment experiment, this might consist of information from the last 8–10 judgment trials (Parducci & Wedell, 1986). In Fig. 2, the dashed arrow from the target to recent experiences indicates that not all recent experiences are recruited into the contextual set, but rather these are selected on the basis of relevance to the target. Selective recruitment is necessary to explain why recent stimuli that are considered different in class or irrelevant to the current task have little impact on judgments (Brown, 1953; Zellner et al., 2003). In Fig. 2, the dashed right arrow indicates the categorical route to contextual recruitment. Here the target stimulus activates related categories in long term memory that provide information constituting the categorical context. One possibility is that once a category is activated, a distribution of category exemplars is brought to mind and used to influence judgment (Smith & Zárate, 1992). Alternatively, categorical information may consist of prototypical information or feature sets (Sherif & Hovland, 1961). In either case, the categorical recruitment process changes the activated context. Recruitment of categorical and recent context may interact such that the recent context highlights a specific categorical associate more than another (Klauer, Ehrenberg, & Wegener, 2003) or differences in activated categories may lead to bias in recruitment of recent experiences falling into the same category (Zellner et al., 2003). Regardless of these possible interactive complications, the recent and categorical contexts are conceptually distinct sources for influencing social evaluations.

To illustrate these two routes, consider how one might judge the height of a 170 cm individual. If the recent context consisted of mostly shorter persons, one would tend to

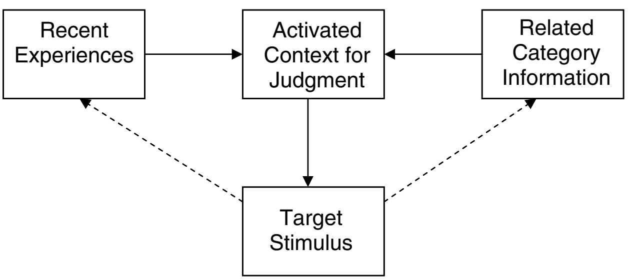

Fig. 2. A simple model for combining recent and categorical contexts. Presentation of the target stimulus recruits recent relevant experiences from working memory into the context for judgment. Similarly, it recruits through shared categorical links stored information from memory into the context. The target is then evaluated based on information in the activated context that contains information from both sources of context.

judge the target "tall," but if the recent context consisted of mostly taller persons, one would tend to judge the target "short" (Manis et al., 1991). This type of contrast effect is well accounted for by Parducci's (1965, 1995) range–frequency theory of judgment, in which ratings reflect the relative location of the stimulus value in the distribution of contextual values. Range–frequency theory describes how judgments are made relative to a context, but it does not describe how that context is recruited. Clearly, categorical markers might strongly influence the context retrieved for judgment. For example, now consider if the person being judged were a 170 cm woman. Based on the activation of the category, "woman," a distribution of heights for previously encountered women could be recruited. Because this distribution would likely be centered below 170 cm, one would tend to judge the target "tall." On the other hand, if the person being judged were a 170 cm man, a different distribution of exemplars would be recruited from the category, "man." As that distribution would likely be above 170 cm, one would tend to judge the target "short."

Although a contrast effect is predicted in the above example, Biernat and colleagues (Biernat, Kobrynowicz, & Weber, 2003; Biernat & Manis, 1994; Biernat, Manis, & Nelson, 1991) have shown that the type of effect produced from recruitment of stereotype information depends in part on the type of judgment. For more objective judgments, such as estimates of heights in feet and inches, one might expect an underestimate of the 170 cm woman's height and an overestimate of the 170 cm man's height, an assimilation effect. More generally, the solid arrow from the contextual set to the target shown in Fig. 2 is meant to depict a variety of contextual judgment processes that could be applied, including assimilation and contrast for dominance judgments and assimilation and contrast of ideals for ideal-point judgments.

1.3. Overview of experiments

Our research builds on these ideas in three ways. First, consistent with the discussion of different types of judgments, we aimed to determine whether category membership may produce effects on both dominance and ideal-point judgments parallel to those described for manipulations of recent context when the recent context was held constant. Second, we examined how changes in the type of information provided by cues used to elicit judgments impact categorical context effects. Third, we sought to determine if these effects were moderated by how information about category members was learned.

To do this, we adapted the schematic faces used by Wedell and Pettibone (1999) for study of categorical context because these materials provided robust effects of context for both dominance and ideal-point judgments. We created two groups to which faces could belong: gnomes and leprechauns. As shown in Fig. 3, gnomes were defined by their red[1] clothing, pointy hats, angular shoulders, and facial hair. Leprechauns were defined by their green clothing, square hats, rounded shoulders, and lack of facial hair. Fig. 3 also describes the distributions of faces making up the two groups. In all experiments reported here except Experiment 1a, distribution of feature widths differed with group, so that in a particular experimental condition, gnomes might have mostly narrow features and leprechauns might have mostly wide features. In addition to the two unique contextual faces,

[1] For interpretation of color in Fig. 3, the reader is referred to the web version of this article.

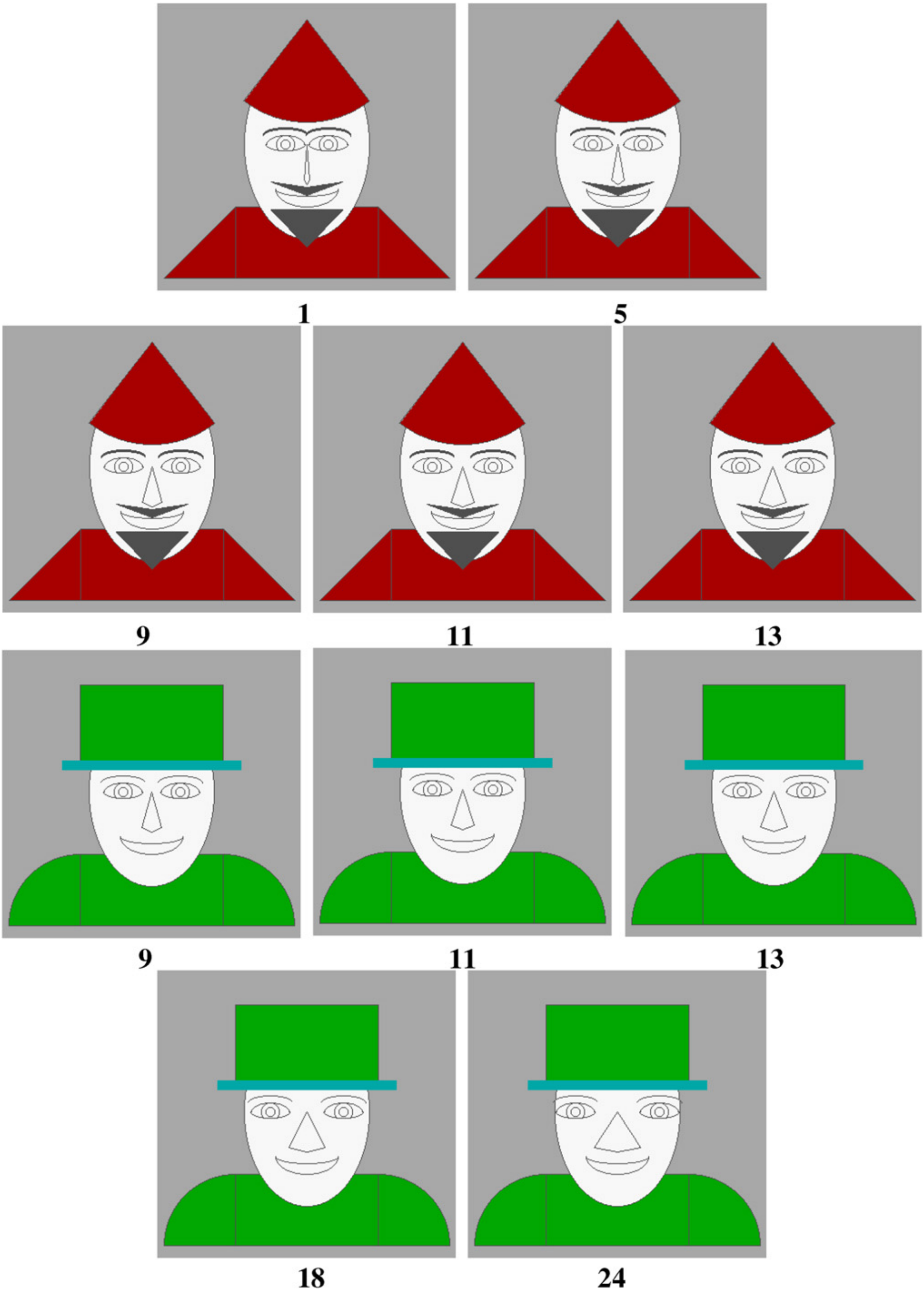

Fig. 3. Examples of a typical set of gnomes and leprechauns used in Experiments 1–3. Target faces 9, 11 and 13 had the same increasing values of nose width and eye gap for the two groups. The critical features for the two groups differ for the two contextual faces, with narrow features (1 and 5) for one group and wide features (18 and 24) for the other group.

each group shared three target faces of moderate widths. Both nose width and eye gap were manipulated as a correlated dimension, so that an increase in one was linked to the same increase in the other feature. All participants saw both groups of faces so that categorical context was manipulated within subject. In all experiments, dominance judgments of feature width and ideal-point ratings of pleasantness were made in separate trials. Although we could have asked participants to make other ideal-point related judgments, such as judgments of typicality (e.g., 'how typical is this facial configuration'), we used the ideal point domain of pleasantness of facial configuration because it clearly relates to affective reactions that drive preferences and hence relates to socially relevant behavior.

Because participants had to associate names with individual faces in several conditions of the experiments reported here, we reduced the number of target and contextual faces greatly from the experiments described by Wedell and Pettibone (1999), in which there were 14 contextual and 7 target faces. As shown in Fig. 3, there were just 7 facial configurations altogether (3 target and 4 contextual).

Experiments 1a and 1b examined whether recent or categorical recruitment occurs for dominance and ideal-point judgments when all relevant target information is available through the use of face cues. In Experiments 2 and 3, a learning phase was implemented so that the retrieval cue for judgment could be either the face itself or the name associated with the face. In Experiment 2 learning of name information was conducted in an integrated learning context, so that participants had to learn the names associated with faces for both categories simultaneously. In Experiment 3 learning was conducted in segregated learning contexts, with participants learning the names associated with faces from one category at a time.

For all experiments, we focused on how context effects operated in dominance and ideal point domains by using width ratings and pleasantness ratings, respectively. With the exception of Experiment 1a, all experiments held the recent context constant at judgment so as to isolate the effects of categorical context. Thus, the pattern of results across the various experimental conditions should provide insight into the application of categorical information in social judgment. Although there are a variety of potential factors we could have manipulated, we focused upon these because they are common features of every day tasks in which stereotypes are often used. That is, we sometimes learn about members of different groups together or separately, and we may be asked to judge an individual based on only a name cue or based on our direct observation of relevant features. Further, we felt that a good starting point for understanding categorical recruitment of context would engender holding recent context constant. Our initial manipulation of recent context is simply to demonstrate the nature of these effects in comparison to manipulations of categorical context.

1.4. Overview of hypotheses

In our current experimental paradigm, we predicted that the stereotype effects of the categorical context when they occurred would parallel the effects of recent context: contrast effects for dominance judgments and assimilation of ideals for pleasantness judgments. Partly, this was because we felt the individuating information that was provided obviated the need for using the stereotypes as information. Furthermore, as our dependent variable was in the form of relative rating scales, contrast resulting from the stereotypes invoking different standards of judgment was more likely (Biernat et al., 1991). No prior work has examined how the ideal-point judgments would be affected by categorical context, but the theoretical explanations for these effects in manipulations of recent context suggest that a parallel effect would occur for manipulations of categorical context. These hypotheses are summarized as follows:

H1: Dominance judgments will show the same type of contrast effects for categorical as recent contexts.

H2: Ideal-point judgments will show the same type of assimilation of ideals for categorical as recent contexts.

A second issue we wanted to explore was related to the cognitive mechanisms that would lead to context effects for ideal and dominance judgments. Specifically, do both types of context effects arise from similar or different processes? Wedell and Pettibone (1999) proposed that assimilation of ideals is consistent with two theoretical mechanisms. The first of these they called the judgment-mediated model, according to which the ideal is tied to a particular value on a subjective scale. The contrast effects on the underlying subjective judgment scale change the stimulus value corresponding to the ideal in an assimilative fashion. To see this, consider the values judged "5" on the dominance rating scale of the top panel of Fig. 1. These values correspond roughly to the peaks of the ideal-point judgment function in the lower panel for the low and high contexts, respectively. The judgment mediated model suggests that the assimilative shift of the ideal point is due to the contrastive shift of standards on the underlying dimension. In support of this relationship, Wedell and Pettibone (1999) observed a moderate negative correlation between contrast effects on width ratings and the inferred location of the ideals from the pleasantness ratings ($r = -.59$). However, because this correlation was only moderate and context was manipulated between subjects, Wedell and Pettibone argued that the data are also consistent with an alternative explanation they called the prototype-mediated model. According to this model, the category prototype is influenced by the current set of stimuli and shifts towards them as a moving average. As such, there is an assimilative shift of the ideal. A key element of this explanation is that as a separate mechanism it does not imply a close correspondence between contextual contrast on dominance judgments and assimilation of ideals on pleasantness judgments. The within-subjects manipulation of context in the current set of experiments allows for a more rigorous test between these two explanations. Recently, Wedell et al. (2005) demonstrated that some individuals can show strong contrast effects while showing no corresponding shift in ideals, providing support for independent mechanisms. We based our research hypothesis on this finding and formulated the following hypothesis:

H3: Given that contrast on dominance judgments arises from different processes than assimilation of ideals, it will be possible to demonstrate context effects on dominance judgments without context effects on ideal-point judgments.

Lastly, a potentially important moderator of whether categorical context effects will occur is the type of judgment cue. If the judgment cue includes the relevant information for the object to be judged, then it may be quite easy for judges to avoid using categorical information to evaluate the individual. In the experiments presented here, participants rated the widths of facial features and the pleasantness of the facial configuration. Categories were quite obvious, as gnomes and leprechauns differed strikingly in clothing and defining features. In the face-cue condition, the face or the gnome was presented directly for judgment. Because all the relevant information is available in this condition, there seems little need to refer to categorical information and so we predicted that categorical context effects would be minimal. Note, however, that categorical context effects have been found in perceptual experiments of this type (Marks, 1992), so that we cannot discount them completely.

In many social judgment situations, one can argue that the relevant information is not immediately available but must be retrieved from memory. Thus, if asked to judge the honesty or aggressiveness of an acquaintance, one would need to retrieve the relevant information from memory. Research by Wedell (1996) provides evidence that contextual effects are more likely to occur when stimuli are evaluated from memory. In a series of experiments, participants rated the dissimilarity of pairs of dot patterns or pairs of squares drawn from skewed distributions. When stimuli were simultaneously present on the screen, disordinal contextual effects on similarity relations were minimal. However, when the presentation of the members of the pair was delayed by three seconds, strong disordinal contextual effects were found. These results were consistent with the idea that context is used to encode stimuli in memory so that retrieval of stimulus information will likely reflect the initial encoding context. In the set of experiments described here, the use of name cues rather than face cues required the retrieval of relevant stimulus information from memory. Given the much stronger effects of context found by Wedell (1996) when information had to be retrieved from memory, we formulated the following hypothesis:

H4: Effects of categorical context effects should be greater in the name cue condition in which relevant target information is available only by retrieving it from memory.

2. Experiments 1a and 1b

In order to test specific hypotheses about the two sources of contextual recruitment for both dominance and ideal-point judgment domains, it is necessary to develop a paradigm in which recent and categorical context are not confounded. Studies like those described by Wedell and Pettibone (1999) confound these two contextual sources because participants were exposed to a single category of stimuli (e.g., schematic faces) and a single recent contextual set. Thus, in Experiments 1a and 1b, the participants were exposed to either multiple recent contexts or multiple categorical contexts. The pattern of contextual effects across these two conditions can provide a basis for inferring the relative contributions of recent and categorical contexts to judgment.

Given this change in the contextual manipulation, as well as the within-subject nature of the manipulation instead of the between-subjects manipulations reported previously, we conducted Experiment 1a to determine if strong within-subject contextual effects from our manipulation were possible. In Experiment 1a, all faces were from a single category, but the narrow and wide distributions were evaluated in separate blocks of trials. Thus, for example, participants might rate widths of features and pleasantness of facial configuration for leprechauns drawn from the narrow set (i.e., the top five facial configurations in Fig. 3 but with leprechaun rather than gnome features) and then make similar ratings of leprechauns drawn from the wide set (i.e., the bottom five facial configurations of Fig. 3) in a second block of trials. The categorical context did not vary across trial blocks because faces were all from the same category (e.g., leprechauns). However, assuming that the recent context consists of the last 8–10 trials, then the recent context did vary across these trial blocks. If participants rated faces based on the recent context, then the pattern of results shown in Fig. 1 would be expected for width and pleasantness ratings, respectively, unless strong transfer effects from the initial context are observed.

Having established strong effects of context in Experiment 1a using the same category and different recent contexts, Experiment 1b was constructed to isolate effects of categorical context. In Experiment 1b, narrow or wide distributions corresponded to either leprechauns or gnomes, respectively, so that categorical context varied in the same way as recent context varied in Experiment 1a. Unlike Experiment 1a, faces from the two distributions were interspersed within the same rating block so that the recent context was held constant and consisted of the full range of faces. Context effects occurring in Experiment 1b could be unambiguously attributed to recruitment and application of different categorical contexts, as the recent context was equated for both groups.

2.1. Method

2.1.1. Participants and design

The 49 participants in Experiment 1a and 80 participants in Experiment 1b were students enrolled in psychology courses at the University of South Carolina, who received course credit for their voluntary participation. In Experiment 1a, separate recent contexts were created by presenting two distributions of faces in separate blocks of trials. All faces from both blocks belonged to the same category. Thus, between-subjects variables for Experiment 1a included judgment task order (one of four possible orders), distribution order (narrow feature distribution presented first or wide feature distribution presented first), and category (all gnomes or all leprechauns). Within-subjects variables included target (one of three faces with moderate features) and context (narrow or wide features). Experiment 1b was similar to Experiment 1a, except that all stimuli were presented in a single block using different categories for the two distributions. Thus, between-subjects variables included judgment task order along with a new variable, category-distribution match, indicating which feature distribution was assigned which group (e.g., gnomes or leprechauns with the narrow or wide distribution). Within-subjects variables included target level and categorical context (narrow or wide features), indicating to which contextual distribution the stimuli belonged.

2.1.2. Materials

Computer generated schematic faces, dubbed either ‘gnomes’ or ‘leprechauns’, were used as judgment stimuli. These faces were similar to those developed by Wedell and Pettibone (1999), save for the addition of features necessary for categorization, and differed by nose width and eye gap. Gnomes were defined by red clothing, pointy hats, angular shoulders, and facial hair. Leprechauns were defined by green clothing, square hats, rounded shoulders, and no facial hair. Feature widths are described by scale values that are linearly related to actual pixel widths. For nose width, a scale value of 1 represents a width of six pixels, increasing by two pixels for each increase in scale value. For eye gap, a scale value of 1 represents a gap of four pixels, increasing by two pixels with each increase in scale value. Each category (Gnomes and Leprechauns) included the same three target faces with moderate feature widths (scale values of 9, 11, and 13). Across categories, these faces differed only in their clothing and facial hair. The category assigned to the wide feature distribution contained two contextual faces with wide features (scale values of 18 and 24), while the category assigned to the negative feature distribution contained two contextual faces with narrow features (scale values of 1 and 5). The spacing of stimuli reflects model fits of Wedell and Pettibone (1999) indicating that subjective valuation of widths followed a power function on actual pixel width, with a power coefficient close to .80. Thus contextual stimuli in the wide context were spaced at greater pixel intervals than those in the narrow context. Eye gap and nose width manipulations were perfectly correlated. Thus, a narrow face with a scale value of 1 had both the narrowest nose and the narrowest eye gap.

2.1.3. Procedure

In both Experiments, participants were shown 10 faces and asked to make three different types of judgments about each face. Participants made dominance ratings of width for nose width and eye gap, as well as a single overall preference rating of "pleasantness to view." The order of these ratings was counterbalanced such that half of the participants started in the pleasantness judgment task. All faces were judged in a single task before proceeding to the next task. All ratings were made on a 1–9 scale, with 1 representing "very narrow" and 9 representing "very wide" for width judgments. For pleasantness judgments, 1 represented "very unpleasant" and 9 represented "very pleasant." Faces in each block were presented twice each in block randomized fashion for each rating task. Prior to initial ratings, a preview of the faces for that block was presented. During the preview, faces were presented one at a time until all faces had been viewed.

In Experiment 1a, the faces were presented in separate sessions of five faces each, with each session including the three target faces and one of the contextual sets of faces. Between sessions, participants were given a 10 second break in which they were told to rest their eyes and that a new set of faces would be presented shortly. All faces were either gnomes or leprechauns. In Experiment 1b, all 10 faces were presented in a same rating session, with half of the faces belonging to the gnome category, and half belonging to the leprechaun category.

2.2. Results

In each experiment, participants who failed to rate the widest features wider than the narrowest features were eliminated, due to misuse of the rating scale. This resulted in the elimination of one participant in Experiment 1a and three participants in Experiment 1b. All analyses were conducted on the remaining participants' data ($n = 48$ in Experiment 1a and $n = 77$ in Experiment 1b).

2.2.1. Experiment 1a

Fig. 4 presents the results for Experiment 1a, with the top left panel representing the width rating combined across feature (nose and eye gap) and the bottom left panel representing rating of pleasantness of facial configuration. Both types of judgments show very strong effects in the recent context, as documented by the difference in ratings assigned to

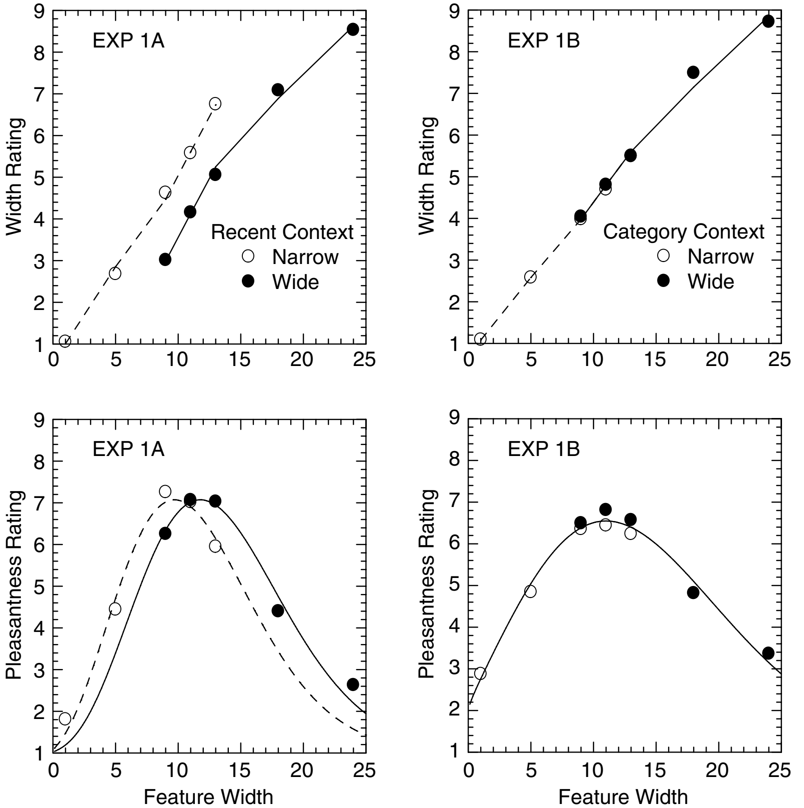

Fig. 4. Results for Experiments 1a and 1b. Large contrast effects on width ratings and assimilation of ideals on pleasantness ratings result from the within-subject manipulation of recent context in Experiment 1a (left panels). No effects on either type of judgment resulted from a similar within-subject manipulation of categorical context in Experiment 1b (right panels). Solid and dotted lines show model fits for each context.

the common target faces across contexts. The pattern of effects was consistent with prior research in which the dominance judgments of width show strong contrast effects and the ideal-point judgments of pleasantness show strong assimilative shifts of the ideals (Wedell & Pettibone, 1999).

The results for width ratings were analyzed using a repeated measures multivariate analysis of variance (MANOVA), with judgment order (4 orders) and distribution order (2 orders) as between-subjects variables, target level (3 faces) and contextual set (high or low) as within-subject variables, and the dependent variable being the mean of width ratings across features. The large contrast effect for width ratings was reflected in the main effect of context, $F(1,40) = 192.6$, $p < .001$. Ratings of the three target faces averaged 1.58 categories higher in the narrow context than in the wide context. Note that although there were other significant effects, none of these interacted with the context factor. Furthermore, the same basic contrast effect was found in all eight between-subjects cells. Thus, the within-subject manipulation of context across judgment blocks with relatively few stimuli produced robust contrast effects.

A parallel repeated-measures MANOVA was conducted on the mean pleasantness ratings. The critical effect that reflected the predicted assimilative shift in ideal-points was the Context × Target interaction, which was significant, $F(2,80) = 45.5$, $p < .001$, and in the predicted direction. In the narrow context, the peak of the pleasantness function was close to face 9, whereas in the wide context it was close to face 11. This shift in ideals produced a large preference reversal. In the narrow context the pleasantness rating of face 9 ($M = 7.26$) was significantly higher than the rating of face 13 ($M = 5.95$), $t(47) = 5.1$, $p < .05$, but in the wide context the pleasantness rating of face 9 ($M = 6.26$) was significantly lower than the rating of face 13 ($M = 7.03$), $t(47) = -3.4$, $p < .05$. As a second test of the ideal point shift, ideal points were inferred for each participant by determining which of the three targets in each contextual set the participant rated highest (with ties being averaged). The parallel repeated measures analysis of variance (ANOVA) conducted on these inferred ideal-points revealed a significant effect of context $F(1,40) = 30.1$, $p < .001$, with the mean ideal in the narrow context inferred to be 10.10 and the mean ideal in the wide context 11.56.

2.2.2. Experiment 1b

Fig. 4 presents the results for Experiment 1b, with the top right panel representing the width rating combined across feature (nose and eye gap) and the bottom right panel representing rating of pleasantness of facial configuration. In this experiment, the recent context was held constant and the categorical context was manipulated. As shown in Fig. 4, neither type of judgment showed any hint of an effect of the categorical context, even though the extent of the manipulation was comparable to that of Experiment 1b.

A repeated measures MANOVA parallel to that run for Experiment 1a was conducted on the width ratings of Experiment 1b. The key finding was that there emerged no indication of a contrast effect resulting from the categorical context manipulation for the combined width ratings, $F(1,69) = 1.5$, $p > .10$. Ratings of the three target faces averaged 0.05 categories lower in the narrow context than in the wide context. Note that although there were other significant effects, none of these interacted with the context factor.

The pleasantness ratings were analyzed in the same way as described for Experiment 1a. The critical effect testing for shifting ideals was the Context × Target interaction and was not significant, $F(2,138) < 1$, $p > .10$. In both narrow and wide category groups, the peak of the pleasantness function was close to face 11. In a more direct test, an ANOVA was conducted on inferred ideal points and again revealed no significant effect of categorical context $F(1,69) < 1$, $p > .10$, with the mean ideal in the narrow context inferred to be 10.80 and the mean ideal in the wide context 11.00.

2.2.3. Additional analyses for Experiments 1a and 1b

A set of additional analyses examined how well the width ratings could be explained by application of Parducci's (1995) range–frequency theory. Parallel to analyses of Wedell and Pettibone (1999), nonlinear regression with a least squares loss function was used to fit the following range–frequency equation to the data:

$$C_{ik} = 1 + 8[w(\phi_i^p - \phi_{\mathrm{MIN},k}^p)/(\phi_{\mathrm{MAX},k}^p - \phi_{\mathrm{MIN},k}^p) + (1-w)F_{ik}], \tag{1}$$

where C_{ik} is the category rating of stimulus i in context k, w is the range–weighting value, p is a power exponent for scaling purposes, ϕ_i is the physical scale value of stimulus i, $\phi_{\mathrm{MIN},k}^p$ and $\phi_{\mathrm{MAX},k}^p$ are physical values corresponding to the subjective minimum and maximum values brought to mind by the judge, and F_{ik} is the frequency value of stimulus i in context k (i.e., the proportion of stimuli ranked below it). Note that two different sets of frequency values (F_{ik}'s) are available: one based on categorical context and the other based on recent context. Details for fitting models to the data for all experiments are given in Appendix A. Our purpose for including the fits here is to provide a quantitative description of the various effects using model parameters rather than provide a test of how well different models are able to capture the pattern of data.

As can be seen by the close adherence of the data points to the theoretical lines of predictions in the top panels of Fig. 4, the range–frequency model is able to provide a reasonable account of the dominance judgments (width ratings). To simplify data fitting, range values were assumed to be constant across contexts so that all contextual effects were due to differences in ranks or frequency values. In both Experiments 1a and 1b, the fits shown are based on frequency values corresponding to the recent contexts. Predictions based on using frequency values from the categorical contexts provided inadequate fits to the data, as these would predict no context effects in 1a and context effects in 1b (the opposite of what occurred).

Also parallel to the analyses of Wedell and Pettibone (1999), nonlinear regression with a least squares loss function was used to fit the Gaussian ideal point model to the data using the following equation:

$$C_{ik} = 1 + b[\exp(-c(\phi_i^p - \mathrm{Ideal}_k^p)^2)], \tag{2}$$

where Ideal_k is the physical value corresponding to the ideal point for context k, c determines the width of the preference function and b determines its height. These parameters may be related to the basic features of participants' attitudes as follows: Ideal corresponds to attitude location, c corresponds to the width of the latitude of acceptability, and b corresponds to the strength of endorsement. As shown in the bottom panels of Fig. 4, different ideal points were needed to account for the shift in the pleasantness ratings with manipulation of recent context, corresponding to a contextually induced attitude shift. However, when only categorical context was manipulated, the data were adequately described by a single function indicating no attitude shift (bottom right panel). The data fits thus indicate that for Experiments 1a and 1b, ideal points were determined by the recent context and not the categorical context. The latitude of acceptability and strength of endorsement (parameters c and b of Eq. (2) were unaffected by context).

2.3. Discussion

The results of Experiment 1a established that the contextual manipulation was sufficient to produce strong contrast effects on dominance based width judgments and strong assimilative shifts of ideals on ideal-point based pleasantness judgments in a within-subjects design. In Experiment 1a, the recent context was manipulated by having participants judge the narrow and wide contextual sets in separate sessions during the experiment. The categorical context was held constant by keeping the group type the same (leprechauns or gnomes in both sets of judgments). Thus, we can unambiguously attribute the obtained effects to the influence of recent rather than categorical context. These effects were as large as or larger than those obtained by Wedell and Pettibone (1999) in a between-subjects manipulation using many more stimuli.

Having established large within subject effects of manipulating recent context, we held recent context constant and manipulated categorical context in Experiment 1b. Although the size of the manipulation was matched with Experiment 1a and the number of participants was substantially increased, there was no hint of any categorical context effects. When judging a target face, participants appeared to recruit the recent context rather than using its group membership to recruit the corresponding categorical context. Thus, faces were compared to the full set of faces recently experienced and not just to faces from a specific group of gnomes or leprechauns. Given that Experiment 1b had power greater than .95 to detect effects one quarter of the size found in Experiment 1a, the null results of categorical context imply that participants can clearly ignore categorical context in judgment situations in which individuating information is provided and there is no need to remember stimulus values.

3. Experiment 2

Experiment 2 was designed to explore the lack of contextual effects arising from the differences in the categorical context presented in Experiment 1b. One potential difference between the task used here and typical judgment tasks involving stereotypes may be the fact that participants were presented with the actual face to be judged, reducing the role of memory in judgment. For example, if asked to judge the leadership ability of the current President, one would first need to retrieve relevant information from memory, because the retrieval cue, "current President," provides none of this directly.

To investigate if the lack of categorical context effects in Experiment 1b was due to the relatively light memory demands imposed by the manipulation, participants in Experiment 2 were asked to make their judgments based either upon a face cue (low memory demand) or a name cue (high memory demand). This experiment required the addition of a learning task in which names were associated with the different leprechauns and gnomes so that later judgments could be based either on the face cue or a name cue. The face cue condition provides a control to determine whether any effects that arise in Experiment 2 are from merely the addition of the learning task or if they depend on retrieval of stimulus information from memory. If these effects are dependent on retrieval of individuating information from memory, then no such effects should be observed in the face cue conditions, as all relevant information is present at the time of judgment.

Two basic hypotheses about categorical context effects in the name cuing situation seem reasonable. First, information in memory at retrieval might be considered degraded so that categories are needed to resolve actual stimulus values. This use of categorical information would produce assimilation effects on dominance based judgments of width and apparent contrast effects on ideal-point judgments of pleasantness. However, another possibility is that categorical context may influence the encoding of information in memory to produce the opposite effects. As Wedell (1996) demonstrated, range–frequency valuation may be used to encode stimuli when they have to be held in memory for even short periods of time. Such encoding would emphasize stimulus differences rather than similarities within categories and hence produce contrast effects on dominance judgments of widths (Mussweiler, 2003). If these shifts in value mediate ideal-point judgments of pleasantness, then one should also expect corresponding assimilation of ideals (as shown in Experiment 1a).

3.1. Method

3.1.1. Participants and design

The participants were 139 undergraduates drawn from the University of South Carolina participant pool. Within-subject variables included categorical context (narrow or wide features), category (gnomes or leprechauns) and target (one of three moderate faces). Between-subjects variables included judgment cue (face or name only) and judgment task order (preference or width ratings first). The category variable was a blocking variable that assigned the gnomes or leprechauns features to either the wide or narrow contextual distributions.

3.1.2. Materials and apparatus

The leprechaun and gnome faces were identical to those used in Experiments 1a, 1b, and 2, with the addition of providing a name for each face. Names within a group were assigned to the faces in alphabetical order, such that the name for the narrowest face in a category was at the beginning of the alphabet. For example, the narrowest face from the narrow width distribution had a name that started with the letter 'A.' Names for leprechauns all ended in the letter "y," while names for gnomes are all used with a single syllable. Lastly, all names for a particular category were presented on either the left or right side of the screen during the learning phase, providing a spatial organization for the categories. For leprechauns, the names used were Crafty, Floppy, Jumpy, Lucky, and Pappy, while gnomes were named Al, Bob, Gus, Jim, and Ken. These procedures were used to facilitate learning of the names and faces. Participants, however, were not informed of these rules but rather were exposed to them through experience.

3.1.3. Procedure

Before beginning the experiment, the participants were shown a preview of the faces they were about to view, paired with the correct names. Faces remained on the screen for 5 s, were separated by a short delay, and were shown in a randomized order. The learning phase consisted of a matching task where a single face was shown on the screen, along with a list of leprechaun names on the left and gnome names on the right. For each face, the participants were told to select the correct name. There were a total of 30 possible blocks of trials, each containing the ten faces from both categories. After each block participants were given accuracy feedback. After the tenth block, participants were judged to have learned the names sufficiently if they made 20 matches in a row. If this criterion was met, participants were allowed to leave the learning task and continue to the judgment phase of the experiment.

The procedures in the judgment phase were the same as used in Experiment 1b in the face cue condition, with ratings based on presentation of the face on the screen. In the name cue condition, the face was absent from the screen so only the name corresponding to the face was presented as a judgment cue.

3.2. Results

In order to provide a clear test of the influence of categorical context, it was critical that participants were able to demonstrate that they had learned the group members' features by being able to distinguish among the faces based on their name cues. To this end, data from participants were dropped if their mean width ratings for the features of the two most extreme faces within a category (face 1 versus face 13 for the narrow distribution and face 9 versus face 24 for the wide distribution) did not differ by at least two rating categories in the appropriate direction.[2] This resulted in the removal of 52 participants from the analysis, 2 from the face cue condition and 50 from the name cue condition. Of the 87 participants who remained, 53 were allowed to end the learning task early due to their successful matching of twenty names and faces in a row. Overall, in the name cue condition, participants saw an average of 24.5 (SD = 6.0) learning blocks while participants in the face cue condition saw an average of 24.0 (SD = 6.5) learning blocks. There were no outliers in either condition.

3.2.1. Dominance based ratings of width

The top panels of Fig. 5 present the mean ratings of width for the target and contextual faces from the two judgment cue conditions. Because the eye gap and nose width manipulations were parallel for each face, these ratings were combined into an overall rating of feature width. A repeated-measures MANOVA was conducted on the width ratings, with judgment order (four orders), judgment cue (face or name), and group distribution (leprechauns paired with narrow or wide distributions) as between-subject variables and target face (3 target levels) and categorical context (narrow or wide) as within subject variables. A strong main effect of categorical context was found in the overall analysis, $F(1,71) = 21.9$, $p < .001$, representing categorical contrast: target face features were rated significantly wider in the narrow feature context than in the wide context. However, this effect was moderated by judgment cue, as revealed in a significant two-way interaction, $F(1,71) = 18.4$, $p < .001$. Ratings for the target faces in the face cue condition were similar in both contexts, but corresponding ratings in the name cue condition showed large contrast effects. Simple effects analyses revealed that context was not significant for face cues ($F < 1$, $p > .10$) but was significant for name cues, $F(1,45) = 33.5$, $p < .001$. Thus, categorical contrast only occurred when width ratings were based on name cues.

The three-way interaction of context, target and judgment cue was also significant, $F(2,142) = 5.7$, $p < .01$. Simple effects analysis revealed a significant Context × Target interaction for participants in the name cue condition, $F(2,90) = 10.9$, $p < .001$, indicating that the effect of context differed across the three target faces. As shown in Fig. 5, the contrast effect was larger for faces 9 and 13 than for face 11. No other higher order interactions were found with the context variable.

3.2.2. Ideal-point based ratings of pleasantness

The bottom panels of Fig. 5 present the mean pleasantness ratings. Again, a repeated-measures MANOVA was used to analyze these ratings. Context effects on ideal points

[2] To address the concern that this procedure of excluding participants may bias the conclusions of this experiment, we repeated the analyses reported here with all 139 participants. Overall, the pattern of results for the critical effects was very similar, with the major differences between analyses being effect size. For ratings of width, both the main effect of context ($F = 5.7$) and the interaction between context and judgment cue ($F = 4.2$) were still statistically significant at the .05 level. Further, the direction of both effects was identical. As with the reduced set of data, no effects of categorical context on the target faces were found with the full data set for ideal points. We focus on the subset of participants whose judgments show a clear evidence of remembering at least gross distinctions of feature widths for different faces, as otherwise the comparison between the name condition and the face condition would likely be obscured by the many errant data points in the name condition.

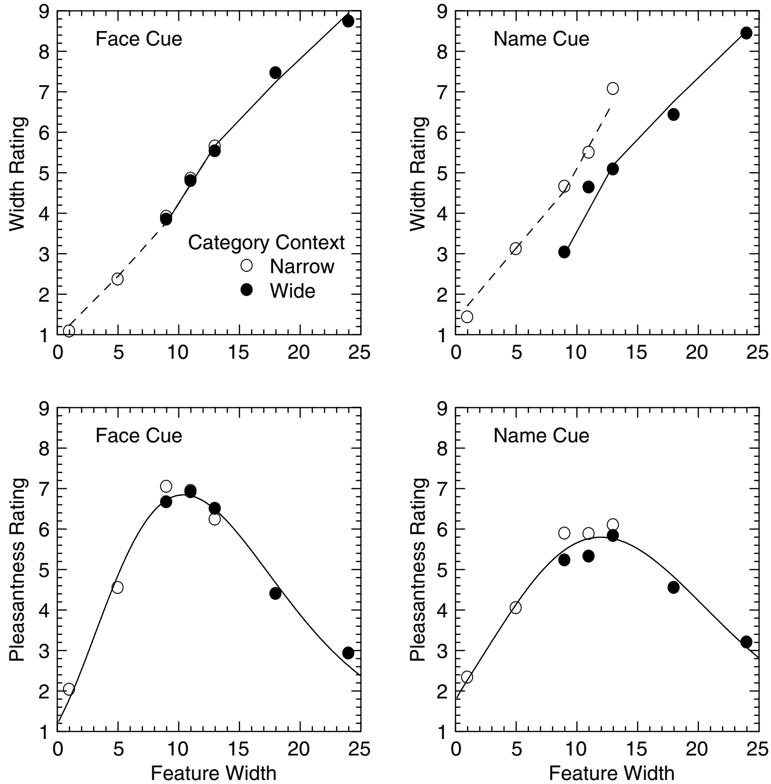

Fig. 5. Mean ratings from Experiment 2 segregated by context (narrow or wide), rated dimension (width or pleasantness), and judgment cue (name or face). The lack of categorical context effects for ratings based on face cues (left panels) replicates results from Experiment 1b. Strong categorical contrast effects were found for width ratings from name cues (top right panel); however, there were no corresponding shifts of ideals for pleasantness ratings from name cues (bottom right panel). Solid and dotted lines show model fits for each context.

would be demonstrated by an interaction involving context and target face, indicating that the location of the peak of the preference function differs between contextual distributions. This interaction was not significant ($F = 2$, $p > .10$), suggesting no overall effect of context on ideal points. Also of interest is the Context × Target × Judgment Cue interaction, which would indicate differential context effects in the when judgments are made with the face or the name cue. The lack of a three-way interaction ($F < 1$, $p > .10$) indicated that the lack of shift in preference ideals generalized across cue condition. No other higher order interactions involving context and target were found.

A significant main effect of judgment cue, $F(1, 71) = 24.7$, $p < .001$, indicated that, overall, ratings of the targets were higher in the face cue condition ($M = 6.72$) than in the name cue condition ($M = 5.73$). There was also a significant interaction of target and judgment cue condition found, $F(2, 142) = 4.9$, $p < .01$. Simple effects analyses revealed that the effect of target was only significant in the face cue condition, $F(2, 52) = 3.5$, $p < .05$. Thus in the face cue condition, targets differed in pleasantness but in the name cue condition they did not.

As an additional test of possible ideal point shifts, ideal points were inferred from the target ratings as described earlier. The main effect of categorical context was not significant, $F(1,71) < 1$, $p > .10$, providing no evidence for an overall shift of ideals with categorical context.

3.2.3. Modeling the data

The face cue data were modeled in the same way as in Experiment 1b. Width ratings were well described by a range–frequency model (Equation 1) in which the recent (global) context was used to generate frequency values. The Gaussian ideal point model of Eq. (2) fit the data well with a single ideal for both categories. Comparison of parameter values in the appendix shows that very similar values were obtained in Experiments 1b and 3, so that the learning phase had minimal impact on ratings.

The name cue width rating data were also reasonably well-modeled by the range–frequency Eq. (1), except that frequency values were based on rank within categorical context. This led to strong contextual effects so that the same target was rated much higher in the narrow than wide context. These results are consistent with the use of rank within the category as a way to code the values of the facial features, with these ranks then being used in later dominance based evaluations on this dimension. The Gaussian ideal-point model of Eq. (2) was used to model the pleasantness ratings based on name cues. Like the face cue condition, only a single ideal point was needed to adequately describe the pattern of data.

3.3. Discussion

In Experiment 2, judgments in the face cue condition demonstrated no significant effects of categorical context, effectively replicating Experiment 1b, in which there was no prior learning phase. However, when names served as judgment cues rather than faces so that individuating information had to be retrieved from memory, a significant effect of categorical context was found for descriptive ratings, similar to the contrast effects of Experiment 1a. Despite very strong contrast effects on the dominance based width judgments with the name cues, there was no corresponding assimilative shift of ideals in this condition. This pattern of results has several important implications.

First, the combination of the absence of contextual effects in the face cue conditions plus the presence of a contrast effect on descriptive judgments in the name cue condition indicates that the presence of the learning task alone was not sufficient to cause the observed context effects. Second, the significant context effect with descriptive ratings in the name cue condition implies that categorical context is more likely to be integrated into judgment when participants are forced to make judgments from memory. This conclusion was supported by the good fit to the data of a range–frequency based model that integrated the effects of the categorical context in the frequency component. It may well be that categorical context is used during learning to enhance differences among category members, perhaps encoding them in terms of their rank within the category on the given dimension. Such encoding alone is not sufficient to alter judgments, as shown by the lack of contextual effects for judgments of the face stimuli. However, in the absence of directly available individuating information, the category can be used to retrieve this information from memory. The results of Experiment 2 clearly show a tendency to enhance within group differences in this case, producing a categorical contrast effect.

A third important result was the absence of context effects for ideal-points in the name-cue condition. Although making judgments based on name cues was sufficient to produce context effects for dominance based width judgments, this was not the case for ideal-point based pleasantness judgments. Apparently, participants did not use the categorical context in evaluating pleasantness with either face or name cues, implying that the judgment process in dominance and ideal point domains may differ with regard to recruitment and integration of contextual information. This pattern of results is problematic for the judgment-mediated model. If the contextually altered values of width had served as input for the ideal-point process, then there should have been large assimilative shifts in ideal points for the pleasantness ratings. Experiment 3 explores predictions from one interpretation of the distinction between contextual effects on dominance and ideal-point judgments.

4. Experiment 3

The results of Experiment 2 showed that dominance and ideal-point judgments differ in the utilization of categorical context. The very large categorical contrast effects on the width judgments were consistent with a strategy in which facial features were remembered by their rank within the group, providing a basis for correctly identifying each gnome and leprechaun. The lack of effects on ideal points implies that these categorically encoded features did not serve as a basis for judging the pleasantness of the facial configuration. It is possible that pleasantness of faces may have been automatically encoded at the time of learning, relative to the context available at that time (Bargh, Chaiken, Govender, & Pratto, 1992; Fazio, Sanbonmatsu, Powell, & Kardes, 1986). Given the results of Experiment 1b, this encoding would have been based on the global recent context and a single ideal point would apply to both groups.

The results of Experiment 1a, however, clearly show that if these faces are considered in separate recent contexts for each group, then different ideals would be generated for the two groups. If pleasantness is automatically encoded, then segregating the learning contexts by group should lead to the establishment of separate ideal points for each group during the encoding of the pleasantness of group member faces. This implies that after segregated exposure to the different groups, the pleasantness ratings based on name cues should reflect the differences in encoding contexts and thus show the assimilative shift of ideals for each category. In Experiment 3, the participants learned to identify the faces for each group in separate learning sessions rather than learning them in a single integrated session. Based on the results of Experiment 2, we would again expect to find very large contrastive effects of categorical context on the width ratings from name cues. We further hypothesized that significant assimilative shifts in ideals would emerge in the segregated learning context. If no effects of categorical context are found for pleasantness ratings in the name cue condition, then it would appear that pleasantness judgments may be based on a single ideal point tied to the recent context available at retrieval rather than to any differences in encoding contexts.

4.1. Method

Ninety-one participants were selected from the undergraduate participant pool at the University of South Carolina. The design for Experiment 3 was identical to that of Experiment 2 with one key exception. Because the learning task for Experiment 3 consisted of two parts, with each contextual distribution learned in a separate section, it was necessary to add an additional blocking variable to control for the order of categorical learning.

The faces used in Experiment 3 were identical to those used in Experiment 2. Procedure for the judgment phase was also identical to Experiment 2. In the learning phase, however, participants learned the two sets of names separately. Thus, participants were given thirty blocks of trials for each category, with five faces in each block. The participants were allowed to continue to the next phase if they correctly matched 20 names and faces in a row after the first 10 trials. After completing the learning phase for one category, the participants were given a short break and told that they would now learn the names for a different group of faces. The participants began the judgment phase after learning the names and faces from both categories.

4.2. Results

The participants were again dropped if their mean descriptive ratings for the features of the two most extreme faces within the narrow and the wide context were not at least two categories apart.[3] This resulted in the removal of 23 subjects from the analysis, one from the face cue condition and 22 from the name cue condition. Of the 68 participants who were retained for this analysis, 40 were allowed to end both learning sessions prior to finishing all trials due to meeting the learning criterion. Overall, in the name cue condition, participants saw an average of 22.6 (SD = 5.5) learning blocks per category (gnomes and leprechauns) while participants in the face cue condition saw an average of 22.0 (SD = 5.2) learning blocks per category. There were no outliers in either condition.

4.2.1. Dominance based ratings of width

The top two panels of Fig. 6 present the mean ratings for width for the face and name cue conditions, combined across the nose and eye gap domains. A repeated measures MANOVA was conducted on the three target faces from both contextual conditions, with context (wide or narrow) as an additional repeated measures variable. Judgment cue (face or name), Task order (descriptive first or pleasantness first), learning order (gnomes first or leprechauns first), and feature assignment (wide gnomes/narrow leprechauns or wide leprechauns/narrow gnomes) were the between subjects variables included in the analysis.

A significant main effect of context, $F(1,52) = 55.6$, $p < .001$, indicates a significant contrast effect. However, as depicted in the comparison between the top panels of Fig. 6, there was a clear Context × Judgment Cue interaction for ratings of width, $F(1,52) = 32.66$, $p < .001$. A simple effects analysis revealed that there was no effect of context in the face cue condition ($F = 2.2$, $p > .10$), but there was a significant effect of context in the name cue condition, $F(1,22) = 59.0$, $p < .001$. As shown in Fig. 6, target faces were rated as wider in the narrow context, a difference in the mean ratings of 2.08. This pattern of results replicates that found in Experiment 3. There was also a significant three-way interaction of context, target, and judgment cue, $F(2,104) = 10.66$, $p < .001$, similar to that found in

[3] Again, to address the concern that excluding participants may bias the conclusions of this experiment, we repeated the analyses reported here with all 91 participants. For judgments of width, the pattern of key effects with the full set was similar to that of the reduced set. Both the main effect of context ($F = 15.9$) and the interaction of context and judgment cue ($F = 35.2$) were still significant and in the same direction. For ideal-point judgments with the full set of data, the interaction of context × target ($F = 3.3$) was still significant and was in the same direction. The three-way interaction of context × target × judgment cue was marginally significant ($F = 2.8$, $p = .068$), but was in the same direction as in the reduced set.

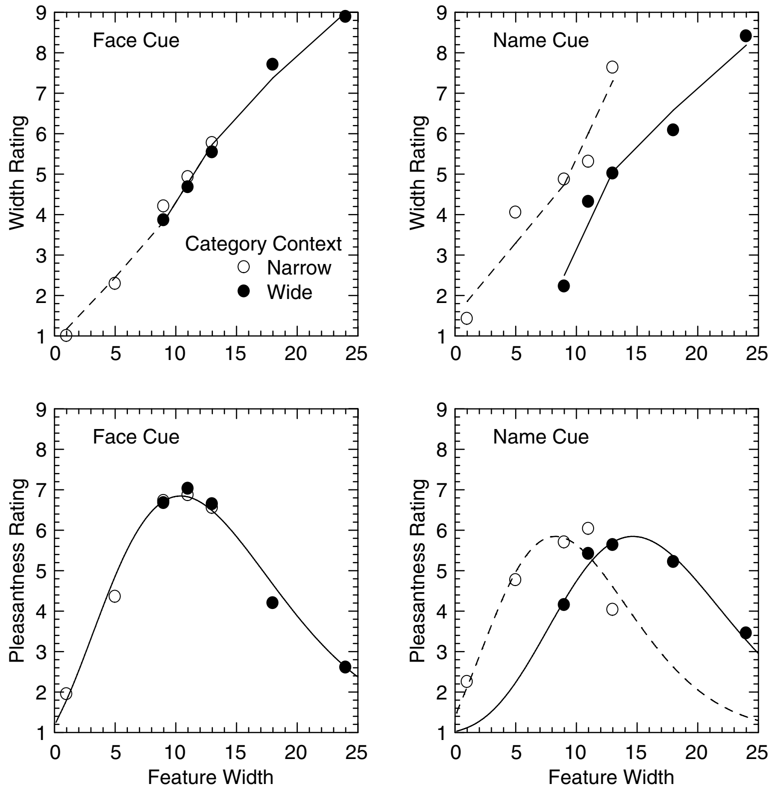

Fig. 6. Mean ratings from Experiment 3 segregated by context (narrow or wide), rated dimension (width or pleasantness), and judgment cue (name or face). The lack of categorical context effects for ratings based on face cues (left panels) replicates prior results. Under name cue conditions, there were large categorical context effects for both ratings of width and pleasantness. Solid and dotted lines show model fits for each context.

Experiment 3. A simple effects analysis revealed that the Context × Target interaction was not significant for the face cue condition ($F < 1$), but was significant for the name cue condition, $F(2,44) = 9.85$, $p < .001$. As in Experiment 3, this interaction reflected the fact that greater context effects were stronger for faces 9 and 13 than for face 11, with mean differences of 2.64, 2.62, and .99, respectively. There were no other higher order interactions involving the other variables with the context variable for the overall analysis.

4.2.2. Ideal-point based ratings of pleasantness

The bottom panels of Fig. 6 show the mean pleasantness ratings for all faces in both the face and name cue conditions. A repeated-measures MANOVA was used to analyze the results, with the same variables as in the analysis reported above. The significant Context × Target interaction, $F(2,51) = 8.58$, $p < .001$, indicated that the categorical context had an effect on ideals for pleasantness. A significant Context × Target × Judgment Cue interaction, $F(2,51) = 7.37$, $p < .001$, demonstrated that the context effect differed across the two judgment cue conditions.

A simple effects analysis revealed that the Context × Target interaction was not significant for the face cue condition ($F < 1$), but was significant for the name cue condition,

$F(2,51) = 8.58$, $p < .001$. As shown in the lower right panel of Fig. 6, ratings for the two contextual distributions in the name cue condition fell along single peaked functions, centered on different ideal points. The ideal face had narrower features in the narrow contextual group than in the wide contextual group. To illustrate this assimilation of ideals towards their categorical context, note that face 9 was rated as much more pleasant in the narrow context than face 13 ($M = 5.71$ vs. $M = 3.89$) and that face 13 was rated more pleasant than face 9 in the wide condition ($M = 5.45$ vs. $M = 4.26$). In contrast, no such shift of ideals occurred for judgments based on face cues. When participants made their judgments with the face on the computer screen, the ideal fell near face 11 for both contextual conditions. Only when participants made their judgments from name cues and the learning tasks were separated did context effects on ideal-point judgments occur.

A significant main effect of judgment cue, $F(1,52) = 22.4$, $p < .001$, indicated that ratings of pleasantness were higher in the face cue condition ($M = 6.78$) than in the name cue condition ($M = 5.19$). A main effect of target, $F(2,51) = 24$, $p < .001$, indicated that the target faces were judged differently. This main effect was moderated by a significant Target × Judgment Cue interaction, $F(2,51) = 6.62$, $p < .01$. A simple main effects analysis revealed that the main effect of target was significant in both the face cue condition, $F(2,29) = 6.37$, $p < .01$, and the name cue condition. Thus, the three-way interaction simply reflects a different ordering of targets across these cue conditions. There were no other significant higher order interactions involving target, context, and judgment cue.

4.2.3. Modeling the data

The same approach to modeling the data from the previous Experiments was used again to understand the data when the categories are learned separately. Equations used were the same as in Experiment 3, and the estimates from the best fitting models can be found in Fig. 6. For descriptive judgments in the face cue condition, the basic range–frequency model based on global frequency values fit the data well. This pattern of results is similar to that found for the same ratings in Experiment 3, further suggesting that the categorical context was not used when making judgments when the actual stimulus is present. In the name cue condition, the best fitting model was based on categorically determined frequency values. As in Experiment 3, these results point to the strong influence of the categorical context when judgments are made from only a name cue.

For pleasantness ratings in the face cue condition, the data were well modeled by a Gaussian ideal point function based on a single ideal point. However, the best fitting model for the name cue data required separate ideal points for narrow and wide groups. This pattern of results differs from those of Experiment 2 and implicate learning context as important in the determination of group ideals.

4.3. Discussion

Experiment 3 demonstrated the critical importance of the learning environment in categorical context effects for ideal-point based judgments when names serve as cues. When the group information was learned in an integrated fashion in Experiment 2, participants showed no effects of categorical context on pleasantness ratings. When information was learned in a segregated manner in Experiment 3, large categorical assimilation effects on ideals were observed for ratings of pleasantness based on name cues. This pattern was not observed for dominance based width ratings in which large categorical contrast was observed in the name cue conditions regardless of whether encoding context was integrated or segregated.

The overall pattern of effects implies that categorical context effects for ideal point and dominance based judgments are based on different processes. The contrast effects observed for width ratings are consistent with emphasis during encoding on cues that allow one to distinguish members within a group. The ranking process implicit in range–frequency theory is a good candidate for such processing. By focusing on rank information, one can learn the differences between members of each group. Ranks within groups would be helpful in learning members in either integrated or segregated learning environments, and so corresponding categorical context effects should occur for both situations. If the contrast effects on width ratings produce assimilation of ideals for pleasantness ratings as posited by the judgment mediated model, then these effects should have been observed on the pleasantness ratings in both Experiments 2 and 3. The lack of these effects in Experiment 2 provides evidence against the judgment mediated model.

An alternative mechanism for assimilation of ideals is that the ideal tends to move toward the mean of contextual values. Consistent with the idea that "what is average is good" is the finding that the average of faces is judged more attractive than the constituent faces (Langlois & Roggman, 1990). Just when are these ideals updated? Work on the processing of affective information would suggest that they are updated automatically (Bargh et al., 1992; Fazio et al., 1986), and without evidence of intent (Bargh, Chaiken, Raymond, & Hymes, 1996). This is consistent with the differences found between integrated and segregated learning. In integrated learning, the faces were evaluated relative to a common ideal, but in the segregated learning they were evaluated relative to different ideals consistent with different encoding contexts. Experiment 1b demonstrated that these shifts in context can quickly produce shifts in ideals, so we should expect implicit pleasantness ratings during segregated encoding to show these differences. When names cues are used, the affective reactions generated during encoding were retrieved and reflected in the pleasantness ratings. When face cues are available, the recent context provides the relevant contextual information and so there is no dependence on retrieved ranks or affective reactions. While this story is plausible, additional research should be conducted to test its implications compared with implications from other explanations.

5. General discussion

Previous research has demonstrated how people can recruit contextual information based on category membership (Goldstone, 1995; Marks, 1992; Parducci, Knobel, & Thomas, 1976). The current experiments expanded upon previous findings in several ways. First, the use of categorical context was demonstrated not only in descriptive judgments, but also in affectively driven ideal-point judgments of pleasantness. Second, several conditions for the use of categorical context in ideal-point judgments were explored. The use of categorical context was found to rely on both the representation of information (memory cues versus actual stimuli) and further upon the way that the information was learned. Ideals were found to shift towards the context when participants had to rely on a memory based representation of the stimuli and when participants learned the category information for each category separately in a segregated manner. In the following discussion, we will describe these contributions and suggest future areas of research.

5.1. Factors influencing categorical contextual recruitment

One clear result from the experiments presented here was the absence of categorical context effects for both descriptive and ideal-point judgments when participants did not have to rely on memory to make a judgment and when the recent context was held constant. When individuating information for each stimulus was easily available, participants appeared to use the unbiased recent context when making judgments. Reducing the information available at judgment by forcing participants to rely wholly on a memory-based representation produced strong contextual effects for descriptive judgments. Memory based retrieval of information was also a prerequisite for context effects on ideal-point judgments. This suggests that category use in social judgment is most likely when we must rely on category level information to supplement missing information about an individual. This pattern of data is consistent with spatially based models of category use (Huttenlocher et al., 1991) as well as with prior stereotype research (Fiske & Neuberg, 1990; Macrae & Bodenhausen, 2000). Both types of research have implicated a lack of individuating information or fine-grain memory as a reason for the use of categorical context, as does the research presented here.

Although the lack of individuating information was sufficient to cause participants to rely on the categorical context for descriptive judgments, this was not the case for ideal-point judgments. This dissociation between categorical context effects on dominance based judgments versus ideal-point based judgments suggests that these effects arise from different processes. One difference between most prior research on categorical context effects and the experiments presented in this paper is that we provided participants with artificially constructed social categories. Participants were forced to learn the categories during the course of the experiment, allowing us to explore how the different contexts in which social categories are learned may influence their usage. Although categories were clearly available for use in judgment across all experiments, participants appeared to use them for ideal-point judgments only when they had to learn the categories separately. When both categories were learned at the same time, ideal-point judgments of pleasantness were consistent with the use of an unbiased recent context. This result suggests that biased usage of stereotypes is enhanced by segregated exposure to members of the group.

Although we found a dissociation between dominance judgments of feature width and ideal-point judgments of pleasantness of faces in the integrated learning context, one may ask whether other ideal-point judgments would show a similar dissociation. For instance, rather than have participants rate pleasantness, we could have had them rate typicality. Would typicality judgments be made relative to the different group norms in the integrated learning context, or would they show the same pattern as pleasantness judgments and be made relative to the integrated context? Future research that contrasts these different ideal point scales may better distinguish whether the differences we obtained for pleasantness ratings from integrated and segregated learning contexts are distinctively linked to the affective nature of pleasantness judgments or are more generally linked to ideal-point processes.

5.2. Implications for models of stereotyping

Research has suggested that stereotypes are activated automatically, demonstrated by their pervasive effect on implicit judgment tasks. Devine (1989) demonstrated that after being subliminally exposed to a set of words, of which 80% were stereotypically associated to African–Americans, Caucasian subjects judged a race-unspecified male target to be more hostile than did a group exposed to a set of words with only 20% stereotype related content. Along similar lines, Gaertner and McLaughlin (1983) found that Caucasian subjects responded faster to Caucasian positive word pairs (white-smart) than to African–American positive pairs (black-smart). For both experiments, explicit measures of stereotype activation, i.e., self-report on racist beliefs, had no correlation to the implicitly demonstrated stereotype effects. These results suggest that though stereotype effects occur implicitly, people often do not demonstrate them in explicit tasks like the judgment tasks used in the current experiments.

In the current set of experiments, features that were diagnostic of the pleasantness of the face (width of nose, gap between the eyes) were common between all categories. Thus, it may have been fairly easy for participants to ignore the categorical information, if they were motivated to do so. Given that the pleasantness judgment task was a somewhat explicit measure of stereotype activation, and that the information vital to categorization may have been easily suppressed, the lack of contextual effects on ideal-points observed in Experiments 1 and 2 may have been due to suppression rather than the lack of established stereotypes. Future research to explore the use of categorical and recent context may benefit from the inclusion of an implicit stereotype activation task to explore this hypothesis.

Appendix A

Model fitting was based on Eqs. (1) and (2) and conducted using nonlinear regression and a least squares loss function. This appendix provides the parameter values for the model fits portrayed in Figs. 4–6. Table A.1 shows the fit of the range–frequency model to width ratings and Table A.2 shows the fit of the Gaussian ideal-point model. The tables include R^2 as a measure of proportion of systematic variance explained by the models.

The fits of the range frequency models to data shown in Table A.1 were generally good. A key point of differentiation was whether the frequency values were based on the full set of 10 faces or restricted to the five faces making up each group. The poorer fits of the range frequency model to the name cue data of Experiments 2 and 3 are due to the model's inability to account for the cue by Target interaction in these conditions.

The fits of the Gaussian Ideal Point models to data shown in Table A.2 were generally good. A key point of differentiation was whether different ideals were fit for the narrow and wide groups.

Table A.1
Model parameters for fit of range–frequency Eq. (1) to data

Experiment	R^2	Frequency values	w	p	S_{MIN}	S_{MAX}
1a (Faces)	.997	Restricted	.629	.752	0.983	26.350
1b (Faces)	.996	Full	.867	.868	0.796	24.568
2 (Faces)	.996	Full	.740	1.0	0.045	24.21
2 (Names)	.955	Restricted	.610	1.0	−3.530	27.110
3 (Faces)	.990	Full	.761	1.0	0.342	23.317
3 (Names)	.951	Restricted	.439	1.0	−9.643	34.102

Note: w = range weighting, Frequency values based on full set of faces or a set restricted to the group, p = power exponent and was set to 1.0 if it did not significantly differ from 1.0, S_{MIN} = subjective range minimum, S_{MAX} = subjective range maximum.

Table A.2
Model parameters for fit of Gaussian ideal point model of Eq. (2) to data

Experiment	R^2	b	c	p	Ideal_N	Ideal_W
1a (Faces)	.985	6.702	0.310	0.595	9.707	11.774
1b (Faces)	.982	5.547	0.056	0.713	10.993	10.993*
2 (Faces)	.984	5.845	0.108	0.549	10.393	10.393*
2 (Names)	.924	4.798	0.048	0.737	11.978	11.978*
3 (Faces)	.991	5.894	0.150	0.650	10.773	10.773*
3 (Names)	.924	4.846	0.180	0.638	8.266	14.662

Note: b = height of function, c = narrowness of function, p = power exponent, Ideal_N = ideal for narrow set, Ideal_W = ideal for wide set. *Indicates that the ideal for the narrow and wide sets were constrained to be equal.

References

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology, 62*(6), 893–912.

Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (1996). The automatic evaluation effect: unconditional automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology, 32*(1), 104–128.

Biernat, M., Kobrynowicz, D., & Weber, D. L. (2003). Stereotypes and shifting standards: some paradoxical effects of cognitive load. *Journal of Applied Social Psychology, 33*, 2060–2079.

Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology, 66*, 5–20.

Biernat, M., Manis, M., & Nelson, T. E. (1991). Stereotypes and standards of judgment. *Journal of Personality and Social Psychology, 60*, 485–499.

Brown, D. R. (1953). Stimulus-similarity and anchoring of subjective scales. *American Journal of Psychology, 66*, 199–214.

Cooke, A. D. J., Janiszewski, C., Cunha, M., Jr., Nasco, S. A., & de Wilde, E. (2004). Stimulus context and the formation of consumer ideals. *Journal of Consumer Research, 31*, 112–124.

Coombs, C. H. (1964). *A theory of data*. New York: Wiley.

Devine, P. G. (1989). Stereotypes and prejudice: their automatic and controlled components. *Journal of Personality and Social Psychology, 56*(1), 5–18.

Eagly, A., & Chaiken, S. (1993). *The psychology of attitudes*. Orlando, FL: Harcourt Brace Jovanovich.

Eiser, J. R. (1990). *Social judgment*. Pacific Grove, CA: Brooks-Cole.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*(2), 229–238.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.). *Advances in experimental social psychology* (Vol. 23, pp. 1–74). New York: Academic Press.

Gaertner, S. L., & McLaughlin, J. P. (1983). Racial stereotypes: associations and ascriptions of positive and negative characteristics. *Social Psychology Quarterly, 46*(1), 23–30.

Goldstone, R. L. (1995). Effects of categorization on color perception. *Psychological Science, 6*(5), 298–304.

Hilton, J. L., & von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology, 47*, 237–271.

Holbrook, M. B., & Anand, P. (1990). Effects of tempo and situational arousal on the listener's perceptual and affective responses to music. *Psychology of Music, 18*, 150–162.

Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: prototype effects in estimating spatial location. *Psychological Review, 98*, 352–376.

Kahneman, D., & Miller, D. T. (1986). Norm theory: comparing reality to its alternatives. *Psychological Review, 93*, 136–153.

Klauer, K. C., Ehrenberg, K., & Wegener, I. (2003). Crossed categorization and stereotyping: structural analyses, effects patters, and dissociative effects of context relevance. *Journal of Experimental Social Psychology, 39*, 332–354.

Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science, 1*(2), 115–121.

Macrae, C. N., & Bodenhausen, G. (2000). Social cognition: thinking categorically about others. *Annual Review of Psychology, 51*, 93–120.

Manis, M., Biernat, M., & Nelson, T. F. (1991). Comparison and expectancy processes in human judgment. *Journal of Personality and Social Psychology, 61*, 203–211.

Manis, M., Nelson, T. E., & Shedler, J. (1988). Stereotypes and social judgment: extremity, assimilation, and contrast. *Journal of Personality and Social Psychology, 55*, 28–36.

Marks, L. E. (1992). The contingency of perceptual processing: context modifies equal-loudness relations. *Psychological Science, 3*, 285–291.

Martin, L. L. (1986). Set/reset: use and disuse of concepts in impression formation. *Journal of Personality and Social Psychology, 55*, 28–36.

Mussweiler, T. (2003). Comparison processes in social judgment: mechanisms and consequences. *Psychological Review, 110*(3), 472–489.

Parducci, A. (1965). Category judgments: a range–frequency model. *Psychological Review, 72*, 407–418.

Parducci, A. (1995). *Happiness, pleasure and judgment: the contextual theory and its applications*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Parducci, A., Knobel, S., & Thomas, C. (1976). Independent contexts for category ratings: a range–frequency analysis. *Perception and Psychophysics, 20*(5), 360–366.

Parducci, A., & Wedell, D. H. (1986). The category effect with rating scales: number of categories, number of stimuli, and method of presentation. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 496–516.

Riskey, D. R., Parducci, A., & Beauchamp, G. K. (1979). Effects of context in judgments of sweetness and pleasantness. *Perception and Psychophysics, 26*, 171–176.

Roberts, J. S., & Laughlin, J. E. (1996). A unidimensional item response model for unfolding responses from a graded disagree–agree response scale. *Applied Psychological Measurement, 20*, 231–255.

Roberts, J. S., Laughlin, J. E., & Wedell, D. H. (1999). Validity issues in the Likert and Thurstone approaches to attitude measurement. *Educational and Psychological Measurement, 59*, 211–233.

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 & A. Tesser (Eds.), *The construction of social judgments* (pp. 217–245). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Sherif, M., & Hovland, C. I. (1961). *Social judgment: Asimilation and contrast effects in communication and attitude change*. Oxford, England: Yale University Press.

Smith, E. R., & Zárate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review, 99*, 3–21.

Srull, T. K., & Wyer, R. S. (1979). The role of category accessibility in the interpretation of information about persons: some determinants and implications. *Journal of Personality and Social Psychology, 37*, 1660–1672.

Stapel, D. A., & Koomen, W. (1997). Social categorization and perceptual judgment of size: When perception is social. *Journal of Personality and Social Psychology, 73*(6), 1177–1190.

Stapel, D. A., & Winkielman, P. (1998). Assimilation and contrast as a function of context-target similarity, distinctness, and dimensional relevance. *Personality and Social Psychology Bulletin, 24*, 634–646.

Thurstone, L. L., & Chave, E. J. (1929). *The measurement of attitude: A psychophysical method and some experiments with a scale for measuring attitude toward the church*. Chicago: University of Chicago Press.

Wedell, D. H. (1995). Contrast effects in paired comparisons: evidence for both stimulus and response based processes. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 1158–1173.

Wedell, D. H. (1996). A constructive-associative model of the contextual dependence of unidimensional similarity. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 634–661.

Wedell, D. H., & Parducci, A. (2000). Social comparisons: lessons from basic research in judgment. In J. Suls & L. Wheeler (Eds.), *Handbook of social comparison: Theory and research* (pp. 223–252). New York: Kluwer Academic.

Wedell, D. H., & Pettibone, J. C. (1999). Preference and the contextual basis of ideals in judgment and choice. *Journal of Experimental Psychology: General, 128*, 346–361.

Wedell, D. H., Santoyo, E. M., & Pettibone, J. C. (2005). The thick and the thin of it: contextual effects in body perception. *Basic and Applied Social Psychology, 27*(3), 213–227.

Zellner, D. A., Rohm, E. A., Bassetti, T. L., & Parker, S. (2003). Compared to what? Effects of categorization on hedonic contrast. *Psychonomic Bulletin and Review, 10*, 468–473.