Contextual Contrast in Evaluative Judgments: A Test of Pre- Versus Postintegration Models of Contrast

Douglas H. Wedell

Two experiments examined at what stage contrast effects occur within the impression-formation process. University students rated the likability of persons described by 5 trait adjectives, with the distribution of trait adjectives manipulated between Ss. Theoretically, contextual stimuli may affect the valuation of component traits before integration, a preintegration model, or they may affect the valuation of the composite impression after the components have been integrated, a postintegration model. Context was manipulated so that the preintegration model predicted an interaction between context and target, but the postintegration model did not. The predicted interaction was replicated several times, supporting preintegration processing of contextual information. However, when subjects were divided into fast and slow judges (on the basis of judgment latencies), the pattern of results supported preintegration contrast for fast judges and postintegration contrast for slow judges. These results support the conclusion that contrast effects may operate at different levels in impression formation.

Social judgment is strongly dependent on context; the two major types of context effects are contrast and assimilation (Eiser, 1990). *Contrast* refers to the displacement of judgments away from the values of contextual stimuli. For example, Kenrick and Gutierres (1980) found that photographs of faces were rated much less attractive when preceded by exposure to a very attractive face. *Assimilation,* on the other hand, refers to the displacement of judgments toward the contextual standard. For example, Geiselman, Haight, and Kimata (1984) found that the judged attractiveness of a face was higher when it was presented simultaneously on a viewing screen with a more attractive face.

An important issue concerning research on context effects has been the question of the psychological significance of these effects. In part, this issue has been framed in terms of whether context effects have a perceptual basis or whether they simply reflect response conventions or biases (Krantz & Campbell, 1961). In this regard, the use of terms such as *perceptual contrast* has been rather unfortunate and confusing, especially because many of the verbal stimuli used in such experiments are undoubtedly processed by the perceptual system in the same manner regardless of the recently experienced context (Rayner & Pollatsek, 1989). The use of the term has been analogical in nature, reflecting a belief that contextual processes in social perception may resemble contextual processes guiding perceptual illusions such as the Ebbinghaus illusion or the Mueller–Leyer illusion. Following this analogy, the perceptual view of context effects is that they occur at an early stage in cognitive processing, require minimal resources, and are therefore beyond the subject's cognitive control. This view can be contrasted with the conception of context effects as response tendencies, which implies that people use strategic changes in responding to the same internal signal for communicational or other purposes.

The distinction between the perceptual and response bases of contextual processing is also problematic because it confuses two separate issues. The first of these concerns the stage at which contextual processing occurs. The second concerns the automaticity of these processes. Because encoding, integration, and response operations may each entail automatic and control processes, evidence that a process occurs at an early stage does not necessarily imply that it is automatic, or vice versa.

The focus of the present studies was to examine the stage at which contextual contrast occurs within the process of impression formation. The issue of the sequencing of processes has important implications for understanding social judgment. In particular, contextual effects that occur early are likely to have more general consequences than those that occur late. This is because the outputs of these early-occurring processes will serve as inputs for later occurring processes. This type of logic has been used successfully in a number of social judgment domains. Among these is the work on priming-based assimilation effects that has led to the conclusion that these occur at the encoding stage (before information integration) so that judgments along both specific descriptive and general evaluative attributes are altered (Bargh & Pietromonaco, 1982; Srull & Wyer, 1980; Wyer & Srull, 1986). However, it is important to note that several different processes may be responsible for contrast and assimilation effects so that conclusions concerning contextual processing in one study may not generalize to other judgment situations. Before presenting the judgment situation that was examined in the present set of studies, I first review experimental evidence concerning the locus of processes responsible for contextual contrast.

Evidence for Early and Late Occurrence of Contextual Contrast

A more extensive review of this literature is provided by Wedell (1990). The focus of this section is to describe evidence relevant to the question of when do contrast effects occur within

Correspondence concerning this article should be addressed to Douglas H. Wedell, Department of Psychology, University of South Carolina, Columbia, South Carolina 29208.

the sequence of information processing. One avenue of research that has been used to examine this question has been to determine how varying features of the response procedure affect contrast. The logic of these manipulations is to assume that if contrast effects are tied to features of the response process, they occur late in the sequence of information processing. For example, Krantz and Campbell (1961) argued that the reduction of contrast effects on length judgments observed when responses were expressed using the well-known scale of inches rather than an arbitrary rating scale provided a measure of the response component underlying these effects. Further support for a response-based interpretation of contrast is evidence that contrast effects may not generalize to other related scales (Upshaw, 1978) or to related behavioral indexes (Sherman, Ahlm, Berman, & Lynn, 1978). On the other hand, contrast effects have been found to generalize to open-ended written descriptions (Simpson & Ostrom, 1976), cross-modality matching procedures (Manis, 1967), and even physiological measures, such as skin conductance as a measure of anxiety (Krupat, 1974). Additional evidence that contrast effects may not be tied to the response stage of information processing is that contrast is typically observed even when no overt judgments of contextual stimuli are made (Kenrick & Gutierres, 1980; Smith, Diener, & Wedell, 1989; although see Pepitone & DiNubile, 1976, for an exception). Finally, one may question the logic of the argument that the dependency of context effects on response constraints implies that contextual processes occur at the response stage of processing. For example, Parducci and Wedell (1986) modeled the reduction of contrast effects with increase in the number of rating categories in terms of how the number of rating categories affects the stimulus distribution on which contextual processes operate.

Wedell (1990) has argued that the most powerful method for determining the loci of context effects is to study them in relation to other judgment, choice, or memory processes. Given that information must both be compared with contextual standards and used in these other cognitive operations, one may ask whether contrast occurs before or after these other processes. Several studies have examined the sequencing of contextual processing and information integration. For example, Wedell, Parducci, and Geiselman (1987) conducted a study in which subjects rated the attractiveness of photographs of faces presented in pairs. They observed both contrast and assimilation effects operating on the same judgment: Judgments were displaced away from the values of preceding faces (successive contrast) but toward the value of the other face in the pair (simultaneous assimilation). Modeling of the data led to the conclusion that the contrastive processes occurred before the assimilative processes.[1]

Contrast effects have also been examined in conjunction with comparative judgments. Such judgments may be modeled as a subtraction operation in which a stimulus is judged greater on an attribute dimension than another stimulus if the subtraction process yields a positive difference. Does contextual processing precede or follow the comparative process? A study by Manis, Nelson, and Shedler (1988) provided evidence for the priority of contrastive processes in comparative judgment. Subjects judged the psychopathology of statements attributed to patients from two hospitals. One hospital was characterized by predomi-

nantly high psychopathology statements and the other by predominantly low psychopathology statements. Comparative judgments of statements equated in psychopathology indicated that the psychopathology of each statement was first judged relative to the relevant contextual distribution (hospital) before being compared. Thus, the statements from the low psychopathology hospital were judged as exhibiting greater psychopathology than matched statements attributed to patients from the high psychopathology hospital.

Contextually induced preference reversals, resulting from the addition of a contextual decoy that is clearly inferior to or dominated by one of the targets, provide further evidence for the early occurrence of contextual effects on preference and evaluation (Huber, Payne, & Puto, 1982; Simonson & Tversky, 1992; Wedell, 1991). In addition, a much older literature on comparative judgment for psychophysical stimuli has supported the priority of contextual processing (e.g., Erlebacher & Sekuler, 1971; Helson, Michels, & Sturgeon, 1954).

An exception in this regard is a study by Mellers and Birnbaum (1982) in which judged differences in the lightness and darkness of dot patterns were found to be based on context-independent scale values. However, Wedell (1993) has demonstrated the priority of contextual processing for comparative judgments of dot patterns when a delay is introduced between the presentation of the members of each pair. Wedell's interpretation of the difference between delay and no-delay conditions was that when there is no delay, the perceptual difference in lightness and darkness can be computed directly and thus bypass contextual processing. However, in the delay condition, a constructed value of each stimulus is held in memory for later comparison, and this construction process is context dependent.

Consistent with this view of a context-dependent construction process is the fact that Mellers and Birnbaum (1982) found that comparative judgments in a cross-modality psychophysical matching task were based on context-dependent values (e.g., comparison of the lightness of a dot pattern to the size of a circle). In cross-modality matching, direct comparison is not possible because the stimuli are incommensurable and so context-dependent values must be constructed. Within the social judgment literature, equity judgments may be considered an example of cross-modality matching in that values on a salary dimension are compared with values on a performance dimension. Mellers (1983, 1986) has demonstrated that context-dependent values form the basis of comparison for equity judgments. Finally, Roberts and Wedell (1994) have recently demonstrated how similarity judgments of emotion words exhibit contextual dependencies. However, the magnitude of the context effects depended on the response mode, with single

---

[1] The logic for inferring the sequencing of contrast and assimilation processes used by Wedell et al. (1987) centered around the use of range-frequency theory to model contrast effects. If assimilation had occurred first, then the ranks of the moderately attractive target faces would have changed much more across contexts than the ranks of the high and low targets, producing a strong interaction effect. The lack of an interaction between target and context supported the occurrence of range-frequency processes prior to assimilation processes.

stimulus sorting tasks producing strong contextual dependencies and direct pairwise rating greatly attenuating these effects.

The literature reviewed above provides solid evidence that contrastive processes can operate either early, in the sense that they occur before other cognitive operations, or late, in the sense that they are tied to output processes. These mixed results suggest that one important goal of research in this area is to determine the conditions under which contrast effects may have high or low priority in the sequence of information processing.

When Does Contrast Occur in Judgments of Complex Stimuli?

The present set of experiments concerned the loci of contrast effects in impression formation of complex stimulus information. The type of complexity I am concerned with herein is the degree to which the stimulus consists of separable components that must be integrated to form a composite whole. From this viewpoint, simple social stimuli may be conceived as those that require minimal integration of components (e.g., a person represented by one trait or a salary represented by one number), and complex social stimuli may be conceived as those that require integration of component information to arrive at an overall impression (e.g., a person described by a series of traits or behaviors). Contrast effects have been demonstrated for both simple social stimuli, such as test scores (Wedell, Parducci, & Roman, 1989), tips (Smith et al., 1989), salaries (Mellers, 1983, 1986), and prison sentences (Higgins & Lurie, 1983), and for complex social stimuli, such as life event descriptions (Wedell & Parducci, 1988), abstracted clinical case histories (Wedell, Parducci, & Lane, 1990), and driver training films (Krupat, 1974).

One interpretation of the generality of contextual contrast across stimuli varying in complexity is that contrastive processes operate in the same manner on whatever is being judged, regardless of whether it is simple or complex. This interpretation is consistent with the view that contrast occurs late in the judgment process, at the point at which an integrated internal judgment is translated to a rating category. I refer to this view as the *postintegration model* of contrast.

An alternative view is that contrast occurs early in the judgment process, and thus it plays an important role in the formation of the internal representation. This view leads to a different model of how contrast operates on simple and complex stimuli. For the simplest type of stimuli, little if any information integration is necessary to form a coherent representation and so contrastive processes may be conceived as operating on that representation. On the other hand, more complex stimuli may be conceived as composed of several component elements that must be integrated together to form a coherent representation. The view that contrast occurs relatively early in the judgment process suggests that contrastive processes operate on the component elements before their integration rather than on the integrated composite stimulus representation itself. I refer to this view as the *preintegration model* of contrast.

Within the social judgment literature, the only study that has directly examined the question of pre- versus postintegration contrast is one conducted by Mellers and Birnbaum (1983). In their study, subjects rated the performance of students on the

basis of a combination of two test scores. Test scores were drawn from either positively or negatively skewed distributions, and subjects judged the entire set of pairs generated from the factorial combination of stimuli from each set. The data were fit with both pre- and postintegration models, and the results supported the postintegration model. Thus, subjects appeared to first integrate the scores together and then evaluate the composite stimulus relative to the contextual distribution of other composite stimuli.

There are two reasons why one may question generality of these results. The first concerns the nature of the stimuli. Test scores are clearly unidimensional, and their numerical form makes them particularly easy to integrate (i.e., average). On the other hand, typical complex social stimuli may be conceived as composed of multidimensional, nonnumerical components that are not so readily combined. For such stimuli, one may have to make an evaluation of each component on the relevant attribute scale before integration (the preintegration model). A second related problem with generalizing the Mellers and Birnbaum (1983) results is that many instructors explicitly tell students that the different test scores in their class are not individually graded and then combined, but rather the final grade is based on a composite score. Thus, subjects may have an explicit model that leads them to first combine scores and then evaluate them (postintegration contrast).

The present investigation was designed to test pre- versus postintegration models of contrast using stimulus materials of a more multidimensional nature. The stimuli were descriptions of hypothetical persons based on five trait adjectives, and the task was to rate how much one would like the person described. Although trait descriptions are not as rich and detailed as behavioral descriptions, they clearly are multidimensional, tapping dimensions as diverse as hostility, kindness, sociability, and intelligence. They have the added advantage of being consistently mapped onto evaluative dimensions (such as likability) to provide a rigorous test between models.

As with much of the literature reviewed earlier, the proposed test of the two models depends critically on the nonlinear form of contrast effects entailed by Parducci's (1965, 1983) range–frequency theory of judgment. Thus, before proceeding to experimental detail, I first describe range–frequency theory and how manipulation of the context may produce linear or nonlinear changes in category ratings. I then describe a design that makes differential predictions for preintegration and postintegration models of contrast. Finally, I report the results of two experiments, the first providing a qualitative test and the second a quantitative test of the proposed models.

Range–Frequency Theory

Range–frequency theory (Parducci, 1965, 1983) asserts that the judged value of a stimulus is determined by its location within the distribution of contextual stimuli brought to mind at the time of judgment. Two contextual principles determine the judged value of the stimulus. According to the range principle, judgments reflect the location of the stimulus relative to the most extreme values defining the relevant context. Thus, the range (R) value of stimulus i in context k describes the proportion of the contextual range that lies below it:

$$R_{ik} = (S_i - S_{min,k})/(S_{max,k} - S_{min,k}),\qquad(1)$$

where S_i is the context-independent scale value of stimulus i and $S_{min,k}$ and $S_{max,k}$ are the minimum and maximum values, respectively, defining the subjective context k. For example, a person whose likableness is halfway between the most likable and least likable persons brought to mind at the time of judgment will have a range value of .5. Because range values are a linear transformation of context-independent scale values, they will always be linearly related across different contexts.

According to the frequency principle, the location of the stimulus is described by its rank in the contextual set of stimuli. The frequency (F) value of stimulus i in context k describes the proportion of stimuli lying below it:

$$F_{ik} = (rank_{ik} - 1)/(N_k - 1),\qquad(2)$$

where $rank_{ik}$ is the rank of stimulus i, N_k is the rank of the highest valued stimulus, and 1 is the rank of the lowest valued stimulus in context k. For example, a person who is judged more likable than 90% of the persons in the contextual set will have a frequency value of .9. Because ranks are linearly related to scale values only when the distribution is uniform, the frequency principle will generally produce nonlinear effects of context (i.e., mean ratings of the same stimuli will be nonlinearly related across contexts).

The subjective judgment of stimulus i in context k is conceived as a compromise between the range and frequency principles. This subjective value is represented by a weighted average of range and frequency values:

$$J_{ik} = wR_{ik} + (1 - w)F_{ik}.\qquad(3)$$

Although the value of the weighting parameter, w, inferred from empirical data may approach its theoretical limits of 0 and 1 (Parducci & Wedell, 1986), it is typically closer to a value of .5, representing a roughly equal compromise between range and frequency principles (Parducci, 1983). Assuming $w = .5$, our example person with a range value of .5 and a frequency value of .9 will produce a subjective value of .7 on the internal judgment scale.

The function relating category ratings to subjective judgments will depend on the number of rating categories. However, when the number of rating categories is five or greater, category ratings are typically assumed to be linearly related to subjective values:

$$C_{ik} = bJ_{ik} + a,\qquad(4)$$

where C_{ik} is the mean category rating of stimulus i in context k. As a further theoretical constraint, Parducci (1983) has recommended equating the multiplicative constant, b, with the range of category ranks (e.g., 8 for a 9-point scale) and equating the additive constant, a, with the value of the lowest category. Accordingly, our example person with a judged value of .7 should receive a mean rating of $(8 \times .7) + 1 = 6.6$ on a 9-point rating scale.

The range–frequency model has provided good quantitative fits to category rating data across a wide range of judgmental attributes, including psychophysical judgments (Parducci, 1965; Parducci & Wedell, 1986), social judgments of simple stimuli (Mellers, 1983; Mellers & Birnbaum, 1983; Smith et al.,

1989; Wedell et al., 1989), and social judgments of more complex stimuli (Wedell et al., 1987; Wedell et al., 1990). In particular, the prediction of nonlinear effects of context as represented in the frequency principle has been used in critical tests to demonstrate the superiority of range–frequency theory to other theories of contextual contrast (Helson, 1964; Upshaw, 1969; Volkmann, 1951) that predict only linear effects of context (see Parducci, 1983, for a discussion of these tests).

The above presentation of range–frequency theory is consistent with the conception of contextual contrast as affecting subjective impressions, not just response tendencies. Indeed, Parducci's (1984) use of range–frequency principles as a theory of happiness only makes sense if one considers these contextual effects at the level of subjective impressions. However, range–frequency theory does not entail that contrast occurs at the subjective level. For example, Parducci's (1965) presentation of the theory in terms of principles guiding the use of response categories is consistent with an interpretation of contrastive processes operating at the response translation stage of judgment.

Experiment 1: A Qualitative Test of Pre-Versus Postintegration Models

Figure 1 presents a schematic representation of different ways in which context may affect judgments for composite stimuli. In this representation, X_{1jk} and X_{2jk} are the first and second stimulus components describing person j in context k, respectively. The top panel represents context-free processing of the information. According to this model, scale values, $S_{1..}$ and $S_{2..}$, are independent of person and context (denoted by the use of a period to sum over these subscripts). The integrated impression of the person, $I_{.j.}$, is assumed to follow a simple averaging process in accordance with Anderson's (1981) information integration theory:

$$I_{.j.} = (z_0 S_{0..} + \Sigma z S_{i..})/(z_0 + \Sigma z),\qquad(5)$$

where $S_{0..}$ is the scale value corresponding to an initial impression, z_0 is the weight of the initial impression, and z is the weight accorded to the scale values for the presented information.[2] This context-free integrated impression is then linearly transformed to a context-free category rating, $C_{.j.}$.

The second panel of Figure 1 represents the postintegration model. As with all of the models depicted in Figure 1, initial scale values are assumed to be independent of context and person. According to this model, an integrated impression that is independent of context is first formed, $I_{.j.}$, and it is this impression on which contextual processes operate to produce a context-dependent judgment, $J_{.jk}$. The inclusion of the person subscript j with the context subscript k indicates that the context is conceived in terms of the values of the other persons who have been rated rather than values of all the component stimuli. This is the model that was supported by the composite

---

[2] The assumption of a simple constant weight averaging rule is made here as only a first approximation to the pattern of data that is typically observed. A more precise characterization of the data is that the weight accorded different stimuli tends to depend on the rank of the stimulus within the set being judged, a configural weighting model (cf. Birnbaum, 1982).
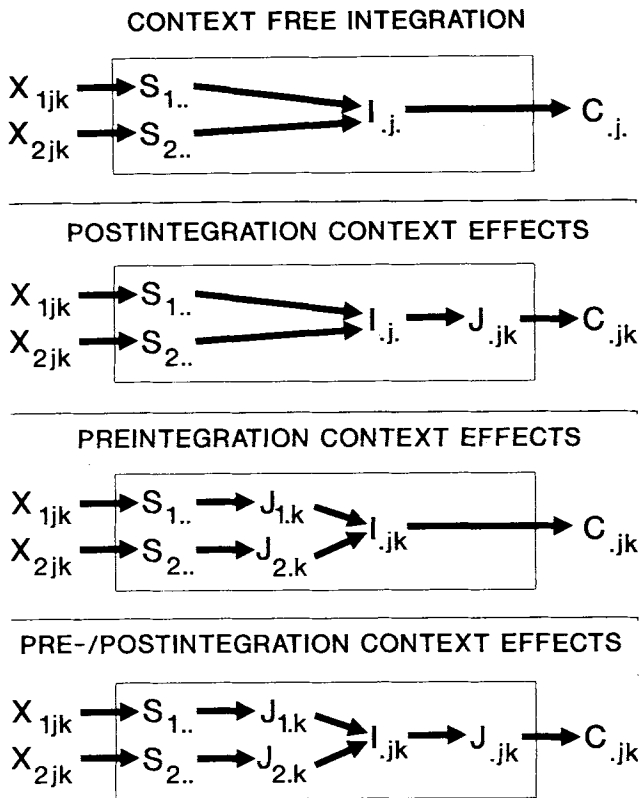
CONTEXT FREE INTEGRATION

POSTINTEGRATION CONTEXT EFFECTS

PREINTEGRATION CONTEXT EFFECTS

PRE-/POSTINTEGRATION CONTEXT EFFECTS

Figure 1. Four models of contextual processing. The symbols correspond to the physical stimulus (X), the initial scale value of the stimulus (S), the context-dependent judgment of the stimulus or stimulus set (J), the integrated impression of the stimulus set (I), and the category rating of the stimulus set (C). The first subscript corresponds to the stimulus, the second to the person with whom the stimulus is linked, and the third to the context in which the stimulus is embedded. A period is used to sum across a subscripted element so that the represented value does not depend on that element. For example, $S_{1..}$ indicates that the scale value of stimulus 1 does not depend on the person or context with which it is linked. The pre- and postintegration models differ in whether the context-dependent judgment occurs before or after stimulus integration.

score judgment study reported by Mellers and Birnbaum (1983).

The third panel represents the preintegration model. Here, context is portrayed as operating directly on the component scale values to produce context-dependent impressions of each component, $J_{1.k}$ and $J_{2.k}$. It is this context-dependent information that is then averaged to form a context-dependent integrated impression, $I_{.jk}$.

Finally, the bottom panel of Figure 1 presents a combination of pre- and postintegration models. Thus, it is possible that contextual effects operate at multiple levels, producing cascaded effects that build on (or cancel) one another. The models displayed in the top and bottom panels are presented for theoretical completeness. The primary focus of the present investigation was to test between preintegration and postintegration models.

The critical difference between these two models is that

context is represented by the distribution of component values in the preintegration model, but it is represented by the distribution of composite values in the postintegration model. Thus, a potential way to distinguish among these models is to manipulate the context so that the distribution of component values differs in important ways from the distribution of composite values. This was the strategy adopted in the present investigation.

Figure 2 schematically illustrates the experimental design that was used and how it leads to differential predictions for pre- and postintegration models. The stimuli were trait adjectives that were combined in sets of five adjectives to describe each hypothetical person. The task was then to rate how likable each person seemed. The top left panel of Figure 2 presents the contextual distribution of persons for low and high liking contexts. The histograms indicate that half of the persons were of very low (L) likability in the low context and that half the persons were of very high (H) likability in the high context. Three levels of target stimuli were common to the two contexts: moderately low (ML), moderate (M), and moderately high (MH) likability.

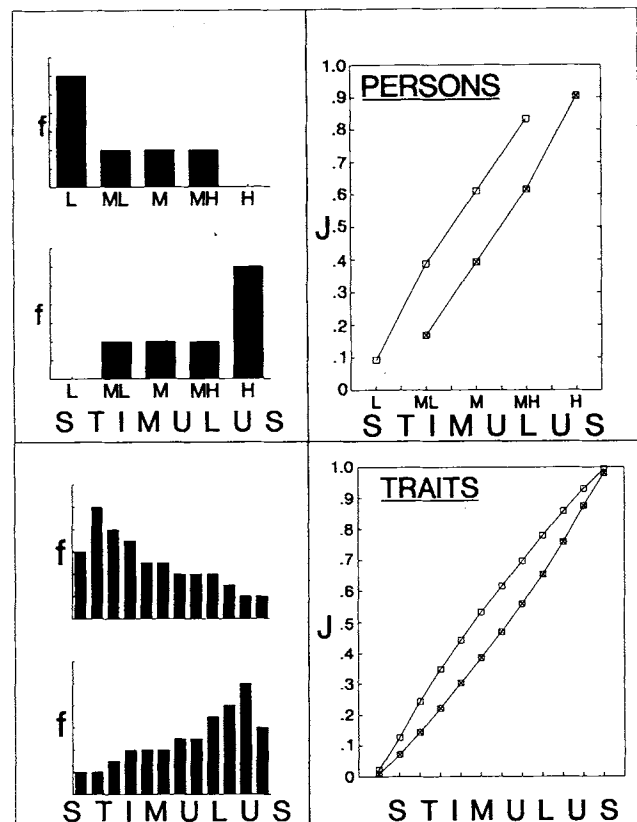The upper right panel presents the range–frequency gener-

Figure 2. Schematic illustration of the distribution of contextual stimuli for the postintegration model (the distribution of persons) and the preintegration model (the distribution of traits). Left panels show frequencies (f) and right panels show predicted judgments (J). (Levels of likability are as follows: L = low, ML = moderately low, M = moderate, MH = moderately high, and H = high.)

ated predictions, letting $w = .6$ and equating the range for low and high contexts. The important feature of these prediction functions is parallelism of the judgments of the target stimuli, indicating that these judgments are linearly related. The reason for this parallelism is that even though the frequency values for the target stimuli differ greatly across contexts, the difference is constant. This is because all of the contextual stimuli lie outside the range of target stimuli so that the relative ranks of these targets are the same across contexts even though the absolute ranks differ greatly. The upshot of this analysis is that the post-integration model, which is based on the distribution of persons, predicts no Target × Context interaction for common stimuli.

On the other hand, context effects for the preintegration model are based on the distribution of traits as depicted in the lower left panel of Figure 2. Like the distribution for persons, the trait distribution for the low context is strongly positively skewed and that for the high context is strongly negatively skewed. However, the person distribution differs in important ways from the trait distribution. First, the ranges for high and low contexts are now equated. Second, the frequency values across most of the trait range are not linearly related. This can be seen in the predictions generated by the range–frequency model for trait ratings, again assuming $w = .6$ (lower right panel). If these context-dependent trait values are then integrated to form the overall impression of the stimulus person, the nonlinearity of the functions should then apply to judgment at the person level as well. The upshot of this analysis is that the preintegration model predicts a Target × Context interaction for judgments of target persons, reflecting the differential curvature of these functions at the trait level. Specifically, if the target stimuli are equally spaced on the dimension of judgment, the Linear × Quadratic component of the interaction should be significant.

Several features of the design of Experiment 1 were developed to enhance the power of the statistical test. First, the trait descriptions were presented on a single sheet of paper rather than in succession. This type of simultaneous presentation has been found to increase the effects of context for social judgments (Wedell & Parducci, 1988). Second, the set of person descriptions constituting the contextual set were rated first, again to enhance the effects of context. Third, subjects used either 5-, 9-, or 100-point rating scales to make their ratings. A broad range of rating scales was used because it was unclear which would provide the most powerful test. In many situations, context effects tend to be enhanced when the number of rating categories is small (Parducci & Wedell, 1986; Wedell & Parducci, 1988; Wedell et al., 1990) so that the use of only five categories might lead to greater context effects and thereby create a more powerful test. On the other hand, use of just a few rating categories calls into question the assumption of response linearity and may obscure or create the nonlinear effects of context being studied.

Method

Subjects and design. Subjects were 225 undergraduates at the University of California, Los Angeles, who received psychology course credit for their participation. A 2 × 3 × 3 factorial design was used with

context (high or low) and number of categories (5, 9, or 100) manipulated between subjects and target likableness (ML, M, and MH) manipulated within subjects. Subjects were randomly assigned to participate in the different between-subjects conditions. Two adjective sets were constructed for each target level. For one set in each level, the adjectives were presented in ascending order, and for the other set, the adjectives were presented in descending order. The ordering of adjectives for each set was reversed for half of the subjects. These order manipulations were conducted to measure possible primacy or recency effects and determine their relation to the overall context effects. However, because the order manipulation did not produce significant primacy or recency, it was dropped from the analyses discussed below.

Stimuli. Stimuli were selected from Anderson's (1968) norming study, which lists mean likability ratings for 555 trait adjectives rated on a scale ranging from 0 to 6. Each adjective set consisted of 5 trait adjectives. In the high context condition, the trait adjectives ranged in likableness from 4.11 to 5.73. In the low context condition, the trait adjectives ranged in likableness from 0.27 to 1.99. The adjectives that made up the target sets ranged from 0.40 to 5.39. Originally, the two sets of adjectives comprising the MH target level were Set A = (*relaxed, attentive, shy, restless,* and *unconventional*) and Set B = (*intelligent, frank, aggressive, cautious,* and *materialistic*); and the two sets of adjectives comprising the M target level were Set C = (*trustworthy, excitable, emotional, persistent,* and *stubborn*) and Set D = (*curious, impulsive, unpredictable, bold,* and *inconsistent*). These sets were matched in value based on Anderson's (1968) norming data. However, the data from both Experiments 1 and 2 indicated that Set C belonged to the MH group and Set B belonged to the M group. Thus, for all data reported these groupings were used. The single adjective conditions of Experiment 2 provide ratings for these adjectives that confirm the validity of the groupings that were used.[3] Finally, the two sets of adjectives comprising the ML target level were Set E = (*perfectionistic, lonesome, rebellious, angry,* and *cruel*) and Set F = (*quiet, bashful, dependent, oversensitive,* and *dishonest*).

Procedure. Subjects participated in the experiment, which was embedded within a series of other judgment tasks, in groups of 8 to 12. Instructions, which were primed at the top of a sheet of paper, asked subjects to rate how much they would like or dislike the persons described by the set of trait adjectives. The endpoints of the scale were labeled *dislike* and *like,* and the midpoint of the scale was labeled *ok.* In addition to these verbal labels, numerical labels were displayed, evenly spaced along a line with tick marks corresponding to each number. For 5- and 9-point scales, all numbers were printed. For the 100-point scale, the 11 numbers that were printed ranged from -50 to 50 by increments of 10. Subjects were told to read the descriptions in the order in which they appeared on the page and to mark their rating on the blank next to each description.

The hypothetical persons being described were labeled *A* through *L.* Each consisted of a set of five trait adjectives, separated by commas. The first 6 descriptions consisted of a random ordering of the descriptions that made up the contextual set. The last 6 descriptions consisted of a random ordering of the descriptions that made up the target set.

Results

Figure 3 presents the mean rating functions for the different rating scales. Mean ratings of the target stimuli were much

---

[3] On the basis of the single trait ratings of Experiment 2, the lower value for Set B was due mostly to much lower ratings of the traits *frank* and *materialistic* than reported in Anderson (1968), and the higher value for Set C was due mostly to much higher ratings of *excitable* and *emotional* than reported in Anderson (1968).
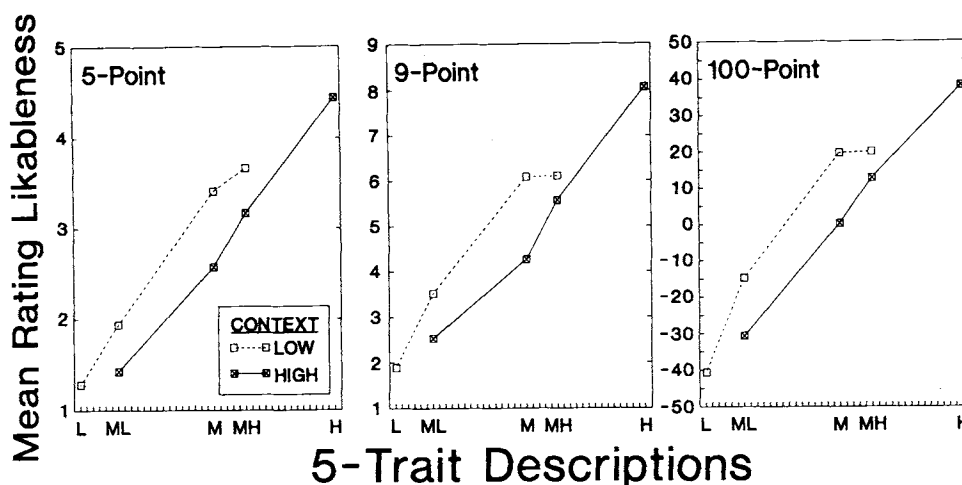
Figure 3. Likableness ratings for 5-, 9-, and 100-point scales of Experiment 1. The differential curvature of functions corresponding to the context effects on target stimuli provides support for the preintegration model. (Levels of likability are as follows: L = low, ML = moderately low, M = moderate, MH = moderately high, and H = high.)

higher in the low likability context than in the high likability context, indicating strong contextual contrast. The critical comparison that distinguishes the preintegration model from the postintegration model is whether the rating functions for the target stimuli are parallel across contexts. The rating functions for all scales exhibited the type of differential curvature predicted by the preintegration model. Thus, the form of the rating functions generally appears to support the preintegration model.

For purposes of statistical analysis, each subject's ratings were linearly converted to a scale ranging from 0 to 100 by subtracting the value of the lowest rating category and multiplying by $100 \times (1/(C - 1))$, where C = the number of rating categories in the scale. The ratings were then submitted to a $2 \times 3 \times 3$ repeated measures multivariate analysis of variance (MANOVA). The main effect of context was highly significant, $F(1, 213) = 67.0$, $p < .001$, and reflected the strong effects of contextual contrast. The main effect of target level was naturally significant as well, $F(2, 212) = 455.5$, $p < .001$, indicating differences in likableness ratings for the ML, M, and MH targets. The critical comparison was the interaction between target level and context, which was significant, $F(2, 212) = 11.1$, $p < .001$, supporting the preintegration model. The preintegration model specifically predicts that for evenly spaced targets, the interaction should be reflected in an interaction contrast of context with the quadratic component of target. Using the marginal means to represent the spacing of target stimuli, this Linear $\times$ Quadratic component of the interaction was highly significant, $F(1, 213) = 22.1$, $p < .001$, accounting for virtually all of the interaction sums of squares.[4] Separate planned contrasts of the Linear $\times$ Quadratic component of the Context $\times$ Target interaction effect were conducted for each number of categories and all were significant ($p < .05$).

Discussion

The qualitative test of predictions from pre- and postintegration models of contrast supported the view that contextual

contrast operated on the component information before integration. The differential curvature of the rating functions for high and low contexts was not predicted by the postintegration model, which equated the context with the set of previously judged persons rather than the set of previously encountered traits. The postintegration model predicted linear effects of context because the frequency values for the target persons were related by an additive constant across contexts.

On the other hand, the preintegration model proposed that traits were evaluated before integration, so that the context became the set of previously encountered traits. Differences in frequency values of the same trait across contexts would then produce nonlinear effects of context. These differences in frequency values at the component level led to the prediction that the rating function in the low context would be negatively accelerated relative to the rating function in the high context. This prediction was replicated three times.

Experiment 2: A Quantitative Test

Although the results of Experiment 1 supported the operation of contrastive processes prior to integration, they did not address whether the preintegration model provides a sufficient account of these effects. For example, the preintegration model may not account for the magnitude of context effects observed, so that additional postintegration contrast might also have been operating (see the model depicted in the bottom panel of Figure 1). The issue of whether the preintegration model was sufficient to explain the observed contrast effects required a study in which the data are quantitatively fit by the preintegration model. This was the strategy adopted in Experiment 2.

Subjects in Experiment 2 rated the same 12 sets of adjectives used in Experiment 1; however, the adjectives were presented

---

[4] All polynomial contrasts reported in this article were based on using the marginal means as estimates of the stimulus spacing.

successively on a computer screen, one every 2 s. After the five trait sets were rated on a 9-point scale, the 60 adjectives making up the 12 hypothetical persons were each rated on the same 9-point scale. Equations 1–5 were then used to fit the preintegration model to the data. An additional set of subjects rated only the single trait descriptions to test for any transfer effects from rating the five trait sets first.

The methodology of Experiment 2 also provided an opportunity to record judgment latencies. These latencies were hypothesized to be diagnostic in distinguishing those subjects who process contextual information before integration from those who process contextual information after integration. Specifically, the preintegration model implies that contextual processing occurs with the presentation of each component trait, so that no additional contextual processing is required after the presentation of the last trait describing a person. On the other hand, the postintegration model requires waiting until all component information has been integrated before making the contextual comparison of how likable this person is relative to the other persons one has rated. Thus, postintegration contextual processing should be associated with longer judgment latencies. If subjects differ in the stage at which they process contextual information, then short-latency subjects should exhibit the Context × Target Level interaction predicted by the preintegration model and long-latency subjects should exhibit the linear (or parallel) effects of context on the different targets predicted by the postintegration model.

Method

Subjects and design. Subjects were 272 undergraduates at the University of South Carolina who received psychology course credit for their participation. Of the full set of subjects, 133 rated only the single trait descriptions and 139 rated both the five trait and the single trait descriptions. The experimental design was essentially a replication of the 9-point conditions of Experiment 1, with some procedural changes. Subjects were randomly assigned to the high or low context condition. The adjective sets and their presentation orders were the same as those used in Experiment 1. Again, because the order manipulation did not produce significant primacy or recency, it was dropped from the analyses discussed below.

Stimuli and apparatus. The stimuli were the same as those used in Experiment 1. All instructions and stimuli were displayed on IBM Model PS2/50Z computers with color monitors. Subjects were tested in groups of up to 5, each on a separate microprocessor.

Procedure. The experimental task was the first of several judgment tasks in which the subjects participated. Instructions for the first phase of the experiment were the same as those in Experiment 1. An experimental trial consisted of the presentation of the five adjectives comprising the hypothetical person. Adjectives were presented in the middle of the screen for 2 s and then erased, with an interstimulus interval of 0.5 s. After the final adjective in each series was erased, the subject was prompted with a rating scale and asked to enter a likableness judgment using numbers from 1 to 9. Judgment latencies were measured from the onset of the rating scale prompt until the number key was pressed.

After the 12 hypothetical persons were rated, a second set of instructions appeared on the screen. These told the subjects that in the second phase of the experiment, they were to use the 9-point scale to rate the likableness of hypothetical persons described by a single trait adjective. Single adjectives were presented one at a time, for 2 s each. After the adjective was erased, the subject was prompted with a rating scale and asked to enter a likableness judgment. The 60 adjectives were presented

in the same order as they were during the first phase of the experiment, with the 30 contextual adjectives preceding the 30 target adjectives. The 133 subjects who rated only the 60 trait adjectives participated only in this second phase.

Results

Figure 4 presents the mean ratings for the five trait adjective descriptions and for the single trait adjectives that made up the target sets. The contrast effects on the five trait sets were strong, although somewhat smaller than that observed in the simultaneous presentation conditions of Experiment 1. The differential curvature of the rating functions is again consistent with the preintegration model, providing a further replication of the effects described in Experiment 1. The single trait ratings also clearly showed the expected contrast effects, which were nearly identical for subjects who only rated the single trait descriptions and those who made prior ratings of the five trait descriptions.

Statistical analyses. A 2 × 3 repeated measures MANOVA was conducted on the five trait rating data. The results demonstrated large and significant effects of context, $F(1, 137) = 35.5$, $p < .001$, and of target, $F(2, 136) = 396.9$, $p < .001$. Again, the major focus was on the Context × Trait interaction, which was significant, $F(2, 136) = 3.6$, $p < .05$. Supportive of the preintegration model, this effect was largely contained in the Linear × Quadratic component, $F(1, 137) = 7.2$, $p < .01$.

A 2 × 30 repeated measures MANOVA was also conducted on the ratings of the 30 traits comprising the target sets. The major purpose of this analysis was to establish the significant effects of context at the component level. The main effect of context was significant, $F(1, 137) = 26.0$, $p < .001$, reflecting the strong tendency to rate the same trait higher in the low context than in the high context condition. The significant Context × Target interaction, $F(29, 109) = 2.5$, $p < .001$, indicated that the effects of context were greater for some traits than for others. As predicted by range–frequency theory, a large part of this interaction was located in the quadratic component, $F(1, 137) = 19.7$, $p < .001$.

Fit of the preintegration model. The strategy used to fit the preintegration model was to (A) first fit the range–frequency model as described in Equations 1–4 to the single trait ratings, (B) use the ratings of the 30 contextual traits in each condition to estimate the z_0, z, and S_0 parameters of Equation 5, and (C) use the range–frequency values in conjunction with the integration parameters of Equation 5 to predict the ratings for the five trait target sets. This method was predictive in the sense that none of the fitted parameters used to estimate the mean ratings of the five trait target stimuli were based on those mean ratings. Instead, these parameters were based on the ratings of the component stimuli and the ratings of the five trait contextual stimuli. Parameter estimates for the five-trait condition were based on only those data from subjects who rated both single-trait and five-trait descriptions.

The fit of the range–frequency model was based on two key assumptions. First, the constants relating internal judgments to mean ratings (Equation 4) were set a priori to $b = 8$ and $a = 1$ (in accordance with Parducci, 1983). Second, the subjective values defining the minimum and maximum range values (S_{min} and S_{max} of Equation 1) were assumed to be equal for high and
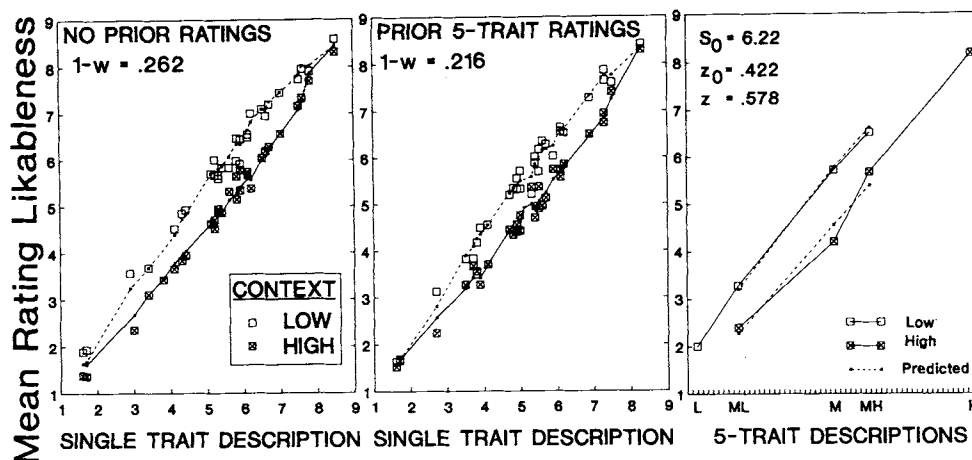
Figure 4. Ratings of single and five-trait descriptions from Experiment 2. Range–frequency theory provides a good fit to the single trait ratings. These theoretical values are used to predict context effects for the five-trait judgments in accordance with the preintegration model. (Levels of likability are as follows: L = low, ML = moderately low, M = moderate, MH = moderately high, and H = high. S_0 = scale value of the initial stimulus; z = weight; w = weighting parameter.)

low contexts. These two assumptions allow one to solve for the weighting of frequency values using the following relationship:

$$(1 - w) = (J_{i,low} - J_{i,high})/(F_{i,low} - F_{i,high}),$$ (6)

where $J_{i,low}$ and $J_{i,high}$ are linear transformations (from Equation 4) of the mean ratings of each stimulus in its respective context and the frequency values are calculated by substituting the ranks of the mean ratings into Equation 3. A single estimate of $1 - w$ was calculated by dividing the sum of the differences in the judgments by the sum of the differences in frequency values. The value of $1 - w$ was quite similar for those who only rated the single trait descriptions (.262) and those who made prior ratings of the five trait descriptions. Observed judgments, frequency values, and w were then used to estimate range values by substituting into the following equation:

$$R_{ik} = (J_{ik} - (1 - w)F_{ik})/w.$$ (7)

Because range values are assumed to be equal for high and low contexts, a single range value for each trait was calculated by averaging the range values estimated from the two contexts. Model predictions were then generated by substituting frequency values, range values, and w into Equation 3 and then using the linear transformation of Equation 4. The fit of the range–frequency model (shown in Figure 4) was quite good, capturing the differences between ratings in high and low contexts with a single fitted parameter, w.

The simple constant weight averaging model of Equation 5 was assumed to characterize the integration of the component trait values into an impression of the five trait sets. This model requires three constants to be fit to the data. The value of S_0 characterizes an initial impression or starting value in the integration process, with z_0 representing its weight and z representing the weight given each subsequent piece of information. The 30 high and 30 low contextual traits making up the six contextual persons in each condition were used to estimate

these parameters. An iterative estimation procedure solved for parameter values S_0, z_0, and z, under the constraint that these values were the same for single and five trait judgments.

The range–frequency predicted trait ratings were substituted into Equation 5 as estimates of $I_{.j}$, along with the values of S_0, z_0, and z obtained from the analysis of contextual stimuli, to solve for the context-dependent scale value of each trait adjective. These context-dependent scale values were then substituted back into Equation 5, along with values of S_0, z_0, and z, to predict the five trait ratings for the target stimuli. As shown in Figure 4, the preintegration model predicts the magnitude of the context effects on the five trait composites, confirming that the preintegration contextual processing is sufficient to account for the data with no need for additional postintegration contextual processing. These are predictions rather than fitted values in the sense that no information from the mean ratings of the composite target stimuli was used to estimate model parameters. These predictions do underestimate to some degree the observed curvilinearity of the scales, a point that is taken up in the next section.

Segregating subjects by judgment latencies. The analyses thus far have treated subjects as a homogeneous group. However, subjects might have differed in whether they processed the contextual information before integration or after integration. Because the preintegration and postintegration models differ in the time course of context effects, subjects were divided into two groups on the basis of a median split of judgment latencies for the five trait judgments. A 2 × 2 × 3 repeated measures MANOVA was conducted on the target ratings, with judgment speed (fast or slow) as an additional between-subjects factor. If longer latencies for the slow subjects were due to postintegration contextual processing, then a three-way interaction should be obtained, with the fast subjects exhibiting a Target × Context interaction and the slow subjects exhibiting parallel effects of context on each target. The three-way interaction was statisti-
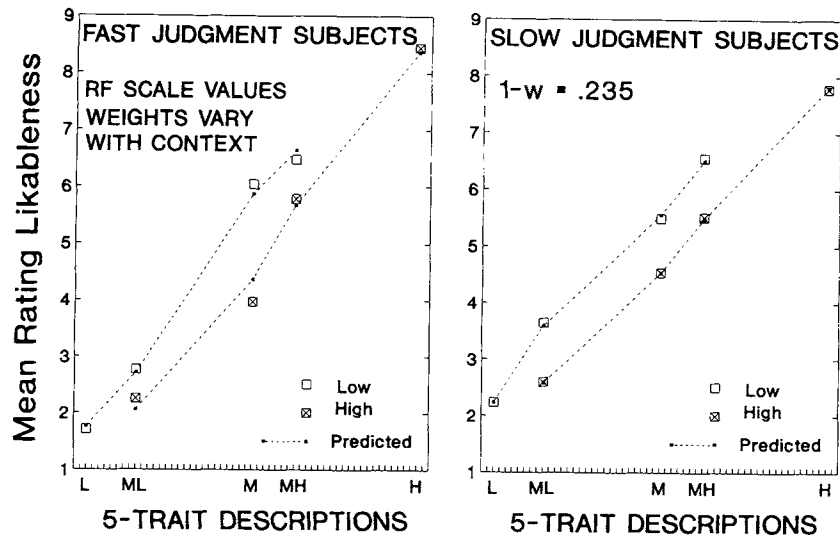
Figure 5. The pattern of five-trait ratings differs for fast- and slow-judgment subjects. Fast subjects exhibit the pattern predicted by the preintegration model, and slow subjects exhibit the pattern predicted by the postintegration model. (Levels of likability are as follows: L = low, ML = moderately low, M = moderate, MH = moderately high, and H = high. w = weighting parameter.)

cally significant, $F(2, 134) = 5.4$, $p < .01$, and was of the predicted form.

Figure 5 shows the rating functions for fast and slow subjects separately. The fast judgment subjects showed a strong interaction of context and target level, but the slow subjects showed virtually no differential effects of context on judgments of the different targets. These conclusions were supported by separate MANOVAs run for the fast and slow subjects; the interaction was significant for fast subjects, $F(2, 67) = 9.221$, $p < .001$, but not for slow subjects, $F(2, 66) = .04$, $p > .90$.

The assertion that the differences in judgment functions for fast and slow subjects reflected preintegration versus postintegration contextual processing is supported by two additional analyses that tested alternative explanations. The first alternative explanation was simply that slow subjects showed parallel judgment functions because they gave little or no weight to frequency values. This interpretation implies that context effects for trait ratings should be smaller for the slow subjects because the manipulation at the trait level was due to unequal frequencies rather than unequal ranges. However, context effects on trait ratings did not significantly differ for fast and slow subjects ($p > .25$), and indeed the inferred frequency weighting was slightly greater for the slow subjects.

A second class of interpretations is that the greater latencies did not reflect an additional postintegration contextual process, but rather they reflected some other more general judgment process. This class of interpretations predicts that the same interaction should be obtained if latencies from the single trait judgment task are used. In contrast, the pre- versus postintegration interpretation places the locus of the effect on processing additional information in the five trait condition. Thus, latencies from the single trait judgments should not predict the three-way interaction, but the difference in latencies between single trait and five trait tasks should predict the interaction.

Subsequent MANOVAs supported the processing specificity prediction of the pre- versus postintegration model. The three-way interaction was not significant when subjects were split on the basis of their single trait judgment latencies, $F(2, 134) = .5$, $p > .50$, but the three-way interaction was significant when subjects were split on the basis of the difference between five trait and single trait judgment latencies, $F(2, 134) = 5.1$, $p < .01$.

Fits to the fast and slow subject data. The functions presented in Figure 5 represent fits of preintegration and postintegration models to the fast and slow subject data, respectively. The postintegration model was fit to the slow judgment subjects using just three parameters, w, S_{min}, and S_{max}. Frequency values were generated a priori from the experimental design (top left panel of Figure 1). The spacing of the stimuli was determined by the mean of range values of the component adjectives generated from the range–frequency fit to the single trait data. Context effects were modeled in terms of differences in frequency values for the five trait stimuli, and so a single range function was used to fit the data. The range–frequency predictions fall nearly on top of the data points. It is interesting to note that the fitted value of $1 - w$ is very similar to that for the single trait judgment data. This is consistent with the idea that these subjects applied the same range–frequency judgment rule to single trait descriptions (using the single traits as the relevant context) and to the five trait descriptions (using the five trait composites as the relevant context).

The fit to the data of the fast judgment subjects should be viewed with some caution because it was based on six fitted parameters for eight stimuli. The greater number of parameters was needed because a constant weight averaging rule did not adequately predict the differential curvature across the two contexts. The predictions of Figure 5 were based on range–frequency scale values from the single trait data, but weights ac-

corded to the higher and lower ranked traits varied with context (i.e., a contextually dependent configural cue model). In the low context condition, the lower ranked traits within each set of five adjectives were accorded relatively more weight, and in the high context condition, the higher ranked traits within each set were accorded relatively more weight. This weighting scheme operated in the opposite direction of contrast, pulling judgments down in the low context condition and up in the high context condition. Because ML and MH target sets were slightly more heterogeneous than the M sets, they were pinched together more by this weighting scheme. Although this fit was post hoc in nature, it demonstrates that the data for the fast judgment subjects can be adequately explained by assuming that contextual processes operate during the integration process rather than after integration.

General Discussion

As discussed in the introduction, there is ample support in the social judgment literature for interpreting contrast as occurring relatively early in the judgment process and also for occurring relatively late, perhaps in the response selection process. The focus of the current set of experiments concerned the stage at which contrast occurs in impression formation of complex or composite stimuli. Prior research has supported a postintegration interpretation of contrast in that comparison to contextual standards only occurred after the components were first integrated together (Mellers & Birnbaum, 1983). However, the materials used in that research (test scores) were unidimensional, numerical, and easily added together. The trait descriptions used in Experiments 1 and 2 were multidimensional and nonnumerical so that integration might require an initial context-dependent evaluation. Overall, the results provide support that under these conditions, contextual contrast appears to occur before integration. However, these conclusions must be tempered by the analyses based on judgment latencies, which implied that fast judgment subjects engaged in contextual processing at an earlier stage than slow judgment subjects.

The Time Course of Judgment

The time course data in Experiment 2 provided converging evidence for the processing distinction between pre- and postintegration context. In preintegration contextual processing, each component is evaluated and integrated on-line so that the judgmental response is available almost immediately after the presentation of the last piece of information. On the other hand, postintegration contextual processing requires one to postpone contextual comparisons until the components have been integrated, resulting in additional processing time after the last component has been presented. The individual differences in processing time observed may be related to the concept of need for cognitive closure described by Kruglanski and his colleagues (Kruglanski, Peri, & Zakai, 1991). This concept suggests that subjects may differ in their need to quickly achieve a firm judgment on an issue (closure) or their need to suspend judgmental commitment until more information is available (avoidance of closure), depending on the domain of judgment, evaluative consequences, and self-perceptions. From this perspective, the long latency subjects may have had a higher need for the avoidance of closure and therefore waited until all information was available before making an evaluative comparison to relevant judgmental standards. If need for closure mediated differences between long and short latency subjects, then manipulations designed to affect need for closure, such as time pressure, initial confidence level, and costs of errors (Heaton & Kruglanski, 1991; Kruglanski et al., 1991), should also affect the time course of contextual processing. The mixed results concerning the loci of contextual contrast in the sequence of information processing may then be partially due to differences in situational factors that lead subjects to adopt different strategies with regard to quickly forming an evaluation or using a more deliberate strategy.

Using Range–Frequency Theory to Understand Social Judgments

In critical tests (Birnbaum, 1974; Parducci, 1963; Parducci & Perrett, 1971), range–frequency theory has repeatedly proven superior to other theories of contrast. The key difference has been range–frequency theory's ability to account for nonlinear effects of context through the role of the frequency principle. Thus, it is somewhat surprising that the social judgment literature continues to describe contrast in terms of theories that do not measure up to range–frequency theory, for example, Helson's (1964) adaptation-level theory, Sherif and Hovland's (1961) social judgment theory, and Upshaw's (1969) variable perspective theory.

Perhaps the adherence to older theories of contrast is due to the fact that these may provide a reasonable first-order approximation to contrast effects. For example, all of the above theories predict the general direction of the context effects shown in Figures 3 and 4, (i.e., higher favorability judgments of the target stimuli in the low likability context). However, none of these theories predict the observed nonlinearities of scale. From an applied perspective, such subtle differences may seem irrelevant. However, from a theoretical perspective, it is precisely this nonlinearity of scale that has served to develop important theoretical distinctions. The critical test between pre- and postintegration models reported in this article rested squarely on the assumption of nonlinear contextual effects operating at the component level. Similarly, several of the tests of sequencing of contextual processing in relation to other operations described in the introduction depend critically on the nonlinear effects of context described by range–frequency theory (Mellers & Birnbaum, 1982, 1983; Wedell, 1993; Wedell et al., 1987).

At an intuitive level, the frequency principle underlying these nonlinearities is consistent with the basic principle that social comparisons guide evaluations of oneself and others. How a person ranks relative to others is clearly an important factor in how people evaluate that person's test score (Mellers & Birnbaum, 1982; Wedell et al., 1989), income (Mellers, 1983, 1986; Smith et al., 1989), and attractiveness (Wedell et al., 1989), to name just a few dimensions. The principle of using ranks to value stimuli is the frequency principle, and it appears to have wide generality.

The usefulness of the range–frequency conception of judgment does not mean that it is the only viable theory of contextual contrast. Instead, range–frequency theory falls into

a class of theories that conceive of contrast as a change in judgmental standards. Schwarz and Bless (1992) have argued that there are processes guiding contrast effects other than changes in standards. For example, Martin and his colleagues (Martin, 1986; Martin, Seta, & Crelia, 1990) have explored the cognitive processes involved in one type of contrastive process, which they refer to as a reset or category inhibition process. The reset process appears to demand more cognitive resources than associative-based assimilation processes. It would be of interest to determine the degree to which the change-of-standard type contrast described by range–frequency theory differs from the reset process. Indeed, one possibility is that the reset process may be modeled in terms of the type of postintegration range–frequency process attributed to the slow judgment subjects of Experiment 2, who clearly expended more time and cognitive resources in making their judgments.

References

Anderson, N. H. (1968). Likableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology, 9,* 272–279.

Anderson, N. H. (1981). *Foundations of information integration theory.* San Diego, CA: Academic Press.

Bargh, J. A., & Pietromonaco, P. (1982). Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation. *Journal of Personality and Social Psychology, 43,* 437–449.

Birnbaum, M. H. (1974). Using contextual effects to derive psychophysical scales. *Perception & Psychophysics, 15,* 89–96.

Birnbaum, M. H. (1982). Controversies in psychological measurement. In B. Wegener (Ed.), *Social attitudes and psychophysical measurements* (pp. 401–485). Hillsdale, NJ: Erlbaum.

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

Erlebacher, A., & Sekuler, R. (1971). Response frequency equalization: A bias model for psychophysics. *Perception & Psychophysics, 8,* 315–320.

Geiselman, R. E., Haight, N. A., & Kimata, L. G. (1984). Context effects on the perceived physical attractiveness of faces. *Journal of Experimental Social Psychology, 20,* 409–424.

Heaton, A. W., & Kruglanski, A. W. (1991). Person perception by introverts and extraverts under time pressure: Effects of need for closure. *Personality and Social Psychology Bulletin, 17,* 161–165.

Helson, H. (1964). *Adaptation-level theory.* New York: Harper & Row.

Helson, H., Michels, W. C., & Sturgeon, A. (1954). The use of comparative rating scales for the evaluation of psychophysical data. *American Journal of Psychology, 67,* 321–326.

Higgins, E. T., & Lurie, L. (1983). Context, categorization and recall: The 'change-of-standard' effect. *Cognitive Psychology, 15,* 525–547.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and similarity hypothesis. *Journal of Consumer Research, 9,* 90–98.

Kenrick, D. T., & Gutierres, S. E. (1980). Contrast effects and judgments of physical attractiveness: When beauty becomes a social problem. *Journal of Personality and Social Psychology, 38,* 131–140.

Krantz, D. L., & Campbell, D. T. (1961). Separating perceptual and linguistic effects of context shifts upon absolute judgments. *Journal of Experimental Psychology, 62,* 35–42.

Kruglanski, A. W., Peri, N., & Zakai, D. (1991). Interactive effects of need for closure and initial confidence on social information seeking. *Social Cognition, 9,* 127–148.

Krupat, E. (1974). Context as a determinant of perceived threat. *Journal of Personality and Social Psychology, 29,* 731–736.

Manis, M. (1967). Context effects in communication. *Journal of Personality and Social Psychology, 5,* 326–334.

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.

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

Martin, L. L., Seta, J. J., & Crelia, R. A. (1990). Assimilation and contrast as a function of people's willingness and ability to expend effort in forming an impression. *Journal of Personality and Social Psychology, 59,* 27–37.

Mellers, B. A. (1983). Equity judgment: A revision of Aristotelian views. *Journal of Experimental Psychology: General, 111,* 242–270.

Mellers, B. A. (1986). "Fair" allocations of salaries and taxes. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 80–91.

Mellers, B. A., & Birnbaum, M. H. (1982). Loci of contextual effects in judgment. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 582–601.

Mellers, B. A., & Birnbaum, M. H. (1983). Contextual effects in social judgment. *Journal of Experimental Social Psychology, 19,* 157–171.

Parducci, A. (1963). Range-frequency compromise in judgment. *Psychological Monographs, 77*(92, Whole No. 565).

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

Parducci, A. (1983). Category ratings and the relational character of judgment. In H. G. Geissler, H. F. J. M. Buffort, E. L. J. Leeuwenberg, & V. Sarris (Eds.), *Modern issues in perception* (pp. 89–105). Berlin, Germany: VEB Deutsher Verlag der Wissenschaften.

Parducci, A. (1984). Value judgments: Toward a relational theory of happiness. In J. R. Eiser (Ed.) *Attitudinal judgment* (pp. 3–21). New York: Springer-Verlag.

Parducci, A., & Perrett, L. F. (1971). Category rating scales: Effects of relative spacing and frequency. *Journal of Experimental Psychology Monographs, 89,* 427–452.

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.

Pepitone, A., & DiNubile, M. (1976). Contrast effects in judgments of crime severity and the punishment of criminal violators. *Journal of Personality and Social Psychology, 33,* 448–459.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs, NJ: Prentice Hall.

Roberts, J. S., & Wedell, D. H. (1994). Context effects on similarity judgments of multidimensional stimuli: Inferring the structure of the emotion space. *Journal of Experimental Social Psychology, 30,* 1–38.

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: Erlbaum.

Sherif, M., & Hovland, C. I. (1961). *Social judgment: Assimilation and contrast effects in communication.* New Haven, CT: Yale University Press.

Sherman, S. J., Ahlm, K., Berman, L., & Lynn, S. (1978). Contrast effects and their relationship to subsequent behavior. *Journal of Experimental Social Psychology, 14,* 340–350.

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research, 29,* 281–295.

Simpson, D. D., & Ostrom, T. M. (1976). Contrast effects in impression formation. *Journal of Personality and Social Psychology, 34,* 625–629.

Smith, R. H., Diener, E., & Wedell, D. H. (1989). Intrapersonal and

social comparison determinants of happiness: A range-frequency analysis. *Journal of Personality and Social Psychology, 56,* 317–325.

Srull, T. K., & Wyer, R. S., Jr. (1980). Category accessibility and social perception: Some implications for the study of person memory and interpersonal judgments. *Journal of Personality and Social Psychology, 38,* 841–856.

Upshaw, H. S. (1969). The personal reference scale: An approach to social judgment. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 4, pp. 315–371). San Diego, CA: Academic Press.

Upshaw, H. S. (1978). Social influence on attitudes and on anchoring of congeneric attitude scales. *Journal of Experimental Social Psychology, 14,* 327–333.

Volkmann, J. (1951). Scales of judgment and their implications for social psychology. In J. H. Roherer & M. Sherif (Eds.), *Social psychology at the crossroads* (pp. 279–294). New York: Harper & Row.

Wedell, D. H. (1990). Methods for determining the locus of context effects in judgment. In J. P. Caverni, J. M. Fabre, & M. Gonzalez (Eds.), *Cognitive biases* (pp. 285–302). New York: Elsevier Science.

Wedell, D. H. (1991). Distinguishing among models of contextually induced preference reversals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 767–778.

Wedell, D. H. (1993). *A constructive-associative model of the contextual dependence of unidimensional similarity.* Manuscript submitted for publication.

Wedell, D. H., & Parducci, A. (1988). The category effect in social judgment: Experimental ratings of happiness. *Journal of Personality and Social Psychology, 58,* 319–329.

Wedell, D. H., Parducci, A., & Geiselman, R. E. (1987). A formal analysis of ratings of physical attractiveness: Successive contrast and simultaneous assimilation. *Journal of Experimental Social Psychology, 23,* 230–249.

Wedell, D. H., Parducci, A., & Lane, M. (1990). Reducing the dependence of clinical judgment on the immediate context: Effects of number of categories and type of anchors. *Journal of Personality and Social Psychology, 58,* 319–329.

Wedell, D. H., Parducci, A., & Roman, D. (1989). Student perceptions of fair grading: A range-frequency analysis. *American Journal of Psychology, 102,* 233–248.

Wyer, R. S., Jr., & Srull, T. K. (1986). Human cognition in its social context. *Psychological Review, 93,* 322–359.