Affect Monitoring and the Primacy of Feelings in Judgment

MICHEL TUAN PHAM
JOEL B. COHEN
JOHN W. PRACEJUS
G. DAVID HUGHES*

Multidisciplinary evidence suggests that people often make evaluative judgments by monitoring their feelings toward the target. This article examines, in the context of moderately complex and consciously accessible stimuli, the judgmental properties of consciously monitored feelings. Results from four studies show that, compared to cold, reason-based assessments of the target, the conscious monitoring of feelings provides judgmental responses that are (*a*) potentially faster, (*b*) more stable and consistent across individuals, and importantly (*c*) more predictive of the number and valence of people's thoughts. These findings help explain why the monitoring of feelings is an often diagnostic pathway to evaluation in judgment and decision making.

The capacity to evaluate alternative states of the world and courses of action is among the most fundamental of human faculties, and the implications of doing so accurately and efficiently—or failing to do so—are profound. It is not a coincidence that the valuation of objects—whether products, issues, or people—has always been a major subject of inquiry in all social sciences.

Although many streams of research have characterized the process of evaluation as a cold, reasoned assessment and weighting of the component qualities of the target (e.g., Anderson 1981; Bettman, Luce, and Payne 1998; Fishbein and Ajzen 1975), there is increasing evidence that people also perform evaluations by monitoring their subjective affective responses (feelings and emotions) to the target (e.g., Damasio 1994; Pham 1998; Schwarz and Clore 1996; Wyer, Clore, and Isbell 1999). Studies on the "How do I feel about it?" heuristic (Schwarz and Clore 1988) have shown that people sometimes infer the direction of their preferences (liking vs. disliking) from the valence of their feelings toward the target (e.g., Gorn, Goldberg, and Basu 1993; Levine, Wyer, and Schwarz 1994; Pham 1998). People may additionally infer the strength of their preferences by monitoring the intensity of these feelings; that is, the level of arousal elicited by the target (e.g., Gorn, Pham, and Sin 2001).

According to the affect-as-information framework (Schwarz 1990; Schwarz and Clore 1996), people rely on their feelings because they perceive these feelings to contain valuable judgmental information. Evidence for this interpretation comes from studies showing that people's feelings cease to influence their judgments and decisions whenever people doubt that the feelings were elicited by the target itself (e.g., Ottati and Isbell 1996; Schwarz and Clore 1983) or when people consider their feelings toward the target to be immaterial for the judgment or decision to be made (e.g., Pham 1998; Raghunathan and Pham 1999)—contingencies that have been called "representativeness" and "relevance," respectively (Pham 1998).

Although affect-as-information research has generated important insights about the psychology of valuation (see Schwarz [2000]; Schwarz and Clore [1996]; Wyer et al. [1999] for recent reviews), it has left a number of fundamental questions about the informational properties of affective feelings largely unanswered. We contribute to the affect-as-information literature by examining three important questions regarding the monitoring of feelings toward supraliminal, focal objects: (1) Does the translation of stimulus-evoked feelings into evaluatively relevant responses

*Michel Tuan Pham is associate professor of business at the Graduate School of Business, Columbia University, 3022 Broadway, Uris 515, New York, NY 10027. Joel B. Cohen is Distinguished Service Professor of Marketing and Anthropology and director of the Center for Consumer Research at the University of Florida, P.O. Box 117150, Gainesville, FL 32611-7150. John W. Pracejus is assistant professor of marketing at the School of Business, University of Alberta, 3-40J Business Building, Edmonton, Alberta, Canada T6G 2R6. G. David Hughes retired as the Burlington Industries Professor of Business Administration at the Kenan-Flagler Business School, University of North Carolina, Chapel Hill, NC 27599. This research was supported by research grants from the Graduate School of Business of Columbia University and the Marketing Science Institute. The authors thank Dolores Albarracin, Chris Janiszewski, Kamel Jedidi, Alice Isen, Asim Ansari, the current editor, and the associate editor for their comments. Correspondence should be addressed to Michel Tuan Pham (tdp4@columbia.edu).

occur rapidly—as is sometimes assumed from research on less cognitively salient stimuli—or do intervening interpretational processes necessarily make resulting judgments slower than for reason-based assessments? (2) Are these feelings and their translation into valuation judgments particularly subjective and idiosyncratic—another widespread assumption (e.g., Cacioppo and Gardner 1999; Derbaix and Sjoberg 1994)—or is there substantial consistency across individuals, suggestive of considerable reliability? and (3) Are these feelings tangential to people's spontaneous thoughts about a target—as the feeling-versus-thinking dichotomy might suggest—or do people's immediate feelings guide (and hence predict) their overall thoughts about the target? These issues have important implications for understanding the role of feelings in evaluative judgments and for assessing the likelihood and consequences of relying on feelings in judgment and decision making.

This article describes a program of research in which over 670 respondents were exposed to a variety of stimuli that they had not evaluated previously and—unlike much prior research on affect and information processing—were moderately complex in semantic content and available to conscious assessment. Respondents' affective responses to the stimuli were measured using multiple instruments, then compared to evaluative responses obtained through a cold, nonaffective assessment of the same stimuli. The results show that, compared to a nonaffective, reason-based evaluative assessment of the stimuli, feelings provide judgmental responses that are (*a*) potentially faster (though possibly subject to further appraisal processes), (*b*) more consistent across individuals, and importantly (*c*) more predictive of the number of valenced thoughts. Implications for consumer research, the affect-as-information framework, and decision research are discussed.

INTEGRAL FEELINGS AS INFORMATION

Feeling-Based Evaluation

Numerous studies on the measurement of attitudes have shown that attitudes often have two separate components: one interpreted as a cognitive component—sometimes called "utilitarian" or "instrumental"—and the second as an affective component—sometimes called "hedonic" or "consummatory" (e.g., Batra and Athola 1990; Breckler and Wiggins 1989). These studies, however, do not specify the processes through which inputs underlying the affective component enter overall evaluation judgments (i.e., spontaneous associations vs. conscious inferential processes). Neither do these studies specify whether these affective inputs are actual feelings toward the target or merely affective beliefs about the target.

In contrast, research on affect as information does examine the processes through which affect enters judgments. In the affect-as-information framework, the evaluative inputs are not just affective/hedonic beliefs (e.g., "This car would be fun to drive"); they are actual feelings, that is, subjective experiences of affective states and responses with a somato-visceral component (e.g., experiencing sensory pleasure while driving a car; see Clore [1992]; Schwarz and Clore [1996]; Wyer et al. [1999]). Schwarz and Clore (1996) recently used the phrase "feelings-as-information" to emphasize that the judgment inputs are indeed feelings experienced at the time of evaluating the target. These feelings can be produced either (*a*) integrally, by a percept of the target (e.g., looking at a product) or a mental representation of the target (e.g., imagining using the product), or (*b*) incidentally, by a preexisting or contextually-induced mood that colors the experience. Unlike many previous studies on affect as information (e.g., Gorn et al. 1993; Ottati and Isbell 1996; Raghunathan and Pham 1999; Schwarz and Clore 1983), this research focuses on feelings that are produced integrally by the target.

Integral affective responses and feelings can be instantiated through three types of mechanisms (e.g., Buck 1985; Cohen and Areni 1991). Type I affect is based on the triggering of innate, sensory-motor programs that are essential to bioregulation. An example would be the affective response and feelings elicited by the intake of spoiled food. Type II affect is triggered by the mapping of stimulus features onto acquired schematic structures that have been previously associated, through conditioning, with particular emotional responses. An example would be the fear response triggered by suspense-generating features in movies. Type III affect is based on a controlled appraisal of the stimulus, which involves a subjective assessment of the stimulus' significance for well-being (Lazarus 1991). An example would be the guilt response that students may experience if they attribute an exam failure to their lack of effort (e.g., Weiner 1985). Although Type I and Type II affective responses are elicited very rapidly (e.g., Hermans, de Houwer, and Eelen 1994; LeDoux 1996), Type III affective responses often involve considerable cognitive mediation and should be decidedly slower (Cohen and Areni 1991).

In the affect-as-information view, feelings enter overall judgments, not through simple association, but through a controlled inferential process (e.g., Schwarz 2000; Wyer et al. 1999). A typical feeling-based inference is the "how do I feel about it" heuristic (Schwarz and Clore 1988), whereby people infer liking/disliking or satisfaction/dissatisfaction from the valence of their feelings (e.g., Gorn et al. 1993; Pham 1998; Schwarz and Clore 1983). Feelings have also been shown to guide inferences and judgments about causal responsibility (Keltner, Ellsworth, and Edwards 1993), the desirability of risk-reward trade-offs (Raghunathan and Pham 1999), proper problem-solving strategies (Soldat, Sinclair, and Mark 1997), and perceived risks (Lerner and Keltner 2000).

In summary, affect-as-information research shows that people often form overall evaluations based on their momentary feelings toward the target and appear to do so in an informed, deliberate manner. Much less is known about the judgmental implications of people's reliance on feelings compared to reason-based inputs (the traditional "cold" evaluation process). The key characteristics of feeling-based information have yet to be examined. To this point, affect-as-

information research has emphasized conditions that moderate the conversion of incidental, mood-produced feelings into judgments (e.g., Martin et al. 1997; Ottati and Isbell 1996), not the informational properties of integral feelings (e.g., Clore 1992; Wyer et al. 1999).

Integral Feelings versus Reason-Based Assessments

To examine the judgmental properties of consciously monitored, integral feelings toward the target, we separate ongoing evaluative processes into feeling-based and reason-based target assessments. In the latter process, people are assumed to consider the target's descriptive qualities and map these onto an evaluative continuum (e.g., good-bad, like-dislike). This process has been variously referred to as "descriptively-based evaluation" (Wyer et al. 1999, p. 30), "information-based judgment" (Strack 1992, p. 256), and "cognition-based attitude" (Edwards 1990, p. 203). For the sake of clarity, we shall adopt the phrase *reason-based assessment* since it seems to capture the essence of the information relied on during this evaluative process.

Evaluative responses based on reason-based assessments provide a meaningful benchmark for three reasons. First, consumer researchers have historically assumed that evaluations were based primarily on reason-based assessments of the target's characteristics (see Bettman et al. [1998] for a review). Second, models of attitude formation and social judgment often pit affect and feelings against "cognitive" (nonaffective) inputs (e.g., Edwards 1990; Epstein and Pacini 1999). Finally, reason-based assessments of the target are often regarded as having higher normative/evidentiary status than feelings, especially in the literature on judgment and decision making.

We focus here on stimuli that are representative of consumers' everyday experiences, namely, stimuli that have a moderately elaborate semantic content and are available for conscious assessment. These stimulus characteristics clearly set this research apart from experimental priming studies of affect and evaluation, which generally use simpler stimuli (e.g., polygons, drawings of common objects) presented in a subliminal or quasi-subliminal manner (e.g., Bargh et al. 1996; Murphy and Zajonc 1993; Winkielman, Zajonc, and Schwarz 1997). We consider only stimuli that have not been previously evaluated by the person, as prior evaluation can alter the judgment process (Lingle and Ostrom 1979).

The two types of assessments are compared on three dimensions of importance from a judgment and decision-making perspective. The first dimension is the speed with which the two types of assessments can be performed at a conscious level and translated into an overt response. This speed can be viewed as an indicator of the relative efficiency of affect-monitoring and reason-based assessment processes and may dictate their influence on subsequent thought generation (see Carlston and Smith 1996). The second dimension is the interpersonal consistency of the two types of assessments. The degree to which people experience commonality in feeling-derived and reason-based evaluations of a target can be regarded as an indicator of the reliability of the two types of processes. To the extent that either process produces subjective and idiosyncratic assessments, one should also observe increased variability over both people and occasions. The third dimension is the degree to which feeling and reason-based evaluations evoke and predict the number of valenced thoughts that are spontaneously generated by the target. This predictive ability speaks to the relative diagnosticity of the two types of responses for subsequent evaluative processes.

Relative Speed. It is generally assumed that feelings toward the target should be registered more rapidly than nonaffective assessments (e.g., Strack 1992; Wyer et al. 1999). Ever since Zajonc (1980, 1984) made his widely cited statements about "the primacy of affect," many researchers have taken this assumption for granted. However, empirical support for this assumption is not particularly strong. First, the evidence that Zajonc and his colleagues used to substantiate the primacy of affect has been challenged repeatedly (see Mandler, Nakamura, and Van Zandt 1987). Second, most supportive studies involve extremely brief exposures to rudimentary stimuli (e.g., Murphy and Zajonc 1993; Winkielman et al. 1997). A meta-analysis of over 200 studies indicates that mere exposure effects are largely limited to circumstances in which people were unaware of the presence of the affect-inducing stimuli (Bornstein 1989). Third, although some studies suggest that people are extremely fast at categorizing stimuli on a valence continuum, these elementary categorizations may not be representative of the processes underlying the conscious monitoring of feelings (Schwarz 2000; Wyer et al. 1999). Finally, some authors have suggested that emotional responses might be slower than certain nonemotional evaluative responses (Schmidt-Atzert 1988). Using a response-latency paradigm, Schmidt-Atzert (1988) observed that subjects who were shown affectively charged pictures took slightly longer to report their "emotional" responses to these pictures than to report their "evaluation" responses. (We shall return to this study later.) Clearly, the relative speed of affect is an unresolved and far more complex issue than was suggested by the early work of Zajonc and others.

We thus examine this issue for the kinds of stimuli that pervade consumers' lives—that is, moderately complex, consciously evaluated stimuli (e.g., products, movies, stores, advertisements). We predict that the conscious monitoring and recording of feelings toward such stimuli will generally (but not always) be faster than reason-based assessments of the same stimuli. This is because full-fledged affective responses may be triggered by innate sensory-motor programs (Type I affect) and schema matching (Type II affect), which operate very rapidly (e.g., LeDoux 1996). Only those affective responses that are based on a controlled appraisal of the stimulus (Type III affect) would be expected to be particularly slow, especially if the stimulus is motivationally ambiguous (Cohen and Areni 1991). Reason-based assessments should generally be slower because they typically

require multiple cognitive operations: the selection of a basis for stimulus assessment, the retrieval of evaluative standards from memory, comparisons against these standards, and the integration of the implications of these comparisons (e.g., Cohen 1990; Edwards 1990). One exception might, however, occur when the stimulus lends itself to a mapping into a previously formed evaluative category (see Cohen 1982).

Interpersonal Agreement. Interjudge agreement, or judgment consensus, is generally regarded as a desirable property (Kruglanski 1989), whereas interpersonal inconsistency is often equated with noise and instability. As illustrated by the old adage that "beauty is in the eye of the beholder" (see Etcoff 1999), it has been widely held that emotional feelings and affective judgments are highly idiosyncratic and unstable (e.g., Cacioppo and Gardner 1999; Derbaix and Sjoberg 1994; Medvec, Madey, and Gilovich 1995). To the extent that this is true, people should learn to put less faith in affect-based assessments because they are relatively unreliable. Reason-based assessments are generally thought to be more objective and reliable (e.g., Derbaix and Sjoberg 1994). Many applications of information integration theory, for instance, assume that the scale value of the stimulus information is constant across people and contexts (Anderson 1981).

We predict, however, that people are in fact more likely to agree on their integral feeling-based responses to everyday stimuli than they are to agree on their reason-based assessments of these stimuli. Affective responses involve processing structures that are phylogenetically and ontogenetically older (e.g., Damasio 1994; LeDoux 1996; Plutchik 1980). Though these structures may not have the flexibility that the cognitive system provides, they also lack its unpredictability (e.g., Epstein and Pacini 1999; Metcalfe and Mischel 1999). Although Type III feelings that are produced by controlled appraisal processes can be idiosyncratic (cf. Cohen and Areni 1991; Weiner 1985), feelings generated by the triggering of innate somatic structures (Type I) and by the matching of emotional schemata (Type II) should be fairly consistent across individuals. In comparison, reason-based assessments should generally be less consistent across individuals. Such assessments typically require a greater number of cognitive operations, which are all potential sources of interpersonal inconsistency. For instance, standards of comparison can be highly idiosyncratic (e.g., Biernat and Manis 1994), and the selection of judgmental criteria often depends on prior knowledge and preferences (e.g., Bettman et al. 1998). Note that our hypothesis about the greater interpersonal consistency of integral feelings does not extend to incidental feelings. By definition, incidental feelings such as those related to a preexisting mood should be contextually labile.

Relation to Thought Generation. Given the theoretical and empirical status of spontaneous thought generation in consumer research (as well as in other behavioral sciences), it is important to examine whether these thoughts are more closely linked to feeling-based responses or reason-based assessments. Cognitive-response models of attitudes would predict that spontaneous thoughts should be more closely related to reason-based assessments. This is because spontaneous thought generation is assumed to be the primary antecedent of the computational operations involved in reason-based assessments (e.g., Hastak and Olson 1989). Moreover, as the phrase *cognitive response* indicates, spontaneous thought generation and reason-based assessments are presumed to involve similar inputs (i.e., propositionally formatted beliefs). Inputs that are similar in nature presumably operate within a common attitudinal route (Eagly and Chaiken 1993).

We predict, however, that the number and the valence of spontaneous thoughts are in fact better predicted by feeling responses. This is because once feelings have been registered—which can occur very rapidly in the case of Type I and Type II affect—the initial affective response will prompt subsequent thought generation through both automatic and controlled processes. The initial affective response can automatically cue affect-congruent materials in memory (e.g., Blaney 1986; Isen et al. 1978). In addition, knowledge may be actively recruited to more fully assess the affect-eliciting stimulus and to transform the initial affective response into a motivationally relevant response (e.g., Cohen and Areni 1991; Damasio 1994; LeDoux 1996; Wyer et al. 1999).

By comparison, reason-based assessments may be only weakly related to spontaneous thought generation. Some reason-based assessments may simply "dead end" when there is no reason to explore them further. Moreover, cognitive operations performed with the explicit goal of reaching a reason-based assessment may bring to mind judgmental considerations (e.g., related judgments, standards, and norms) that are not representative of one's spontaneous thoughts about the target. The latter argument is consistent with studies showing that having people introspect about their reasons for liking or disliking certain options sometimes decreases the quality of their judgments and decisions (e.g., Wilson and Schooler 1991; Wilson et al. 1993). According to Wilson and his colleagues, this is because analyzing one's reasons increases the weight of judgmental criteria that are highly accessible and easy to verbalize—criteria that might be different from those one would normally use in spontaneous evaluations and choices. Although their reasoning is consistent with ours, Wilson and his colleagues did not typically test the relations between people's spontaneous thoughts and their reason-based and feeling-based assessments. Instead, their research focuses on the input-output relationship between reason-based processing and actual choice and overall evaluation. Furthermore, in their research feeling-based judgments are not examined explicitly (e.g., Wilson and Schooler 1991; Wilson et al. 1993).

We tested our hypotheses regarding relative speed, interpersonal agreement, and thought generation in a series of four studies, using multiple instruments to assess both types of responses, including real-time measures of each. The first two studies focused on responses to static stimuli (magazine

pictures). The last two studies focused on responses to dynamic stimuli (TV commercials).

STUDY 1

The first study compares feeling-based responses and reason-based assessments to pictures selected from general newsmagazines. Such pictures have several desirable characteristics for the purpose of this research. First, they cover a broad range of substantive domains (e.g., family, nature, news). Second, their semantic content is far richer than in the stimuli used in previous mere exposure and priming studies of affect elicitation. Finally, their exposure times can be controlled. One of the study's objectives was to verify that, in the context of the stimuli examined, feeling and reason-based evaluation responses could be empirically dissociated. This objective was addressed through a multitrait-multimethod (MTMM) analysis of the two types of responses. However, the study's primary objectives were to (*a*) compare the relative speed with which feeling and reason-based responses are registered and (*b*) compare the degree of interpersonal agreement (i.e., commonality) elicited by the two types of responses.

Method

A total of 161 undergraduates were exposed to a series of 35 magazine pictures. A pretest indicated that the selected pictures: (1) sampled a wide range of "pleasant-unpleasant" and "like-dislike" response levels; (2) exhibited relatively high homogeneity for both types of responses across subjects; and (3) could produce a dissociation between feeling and reason-based evaluative responses (e.g., some pictures were expected to produce negative feelings and be liked, and vice versa). Respondents were asked to report either the pleasantness of their feelings (henceforth feelings) or their overall reason-based assessments (henceforth assessments) of each picture. Two types of instruments were used to collect each type of response. Half of the respondents reported their responses (either pleasant/unpleasant feelings or like/dislike assessments) using paper-and-pencil (P&P) measures. The other half reported their responses (either feelings or assessments) using a continuous-measurement, dial-turning instrument (DTI). Half of the respondents (replication 1) saw the pictures in one random sequence; the other half (replication 2) saw the pictures in a different random sequence. The study thus involved eight groups of respondents in a 2 (feelings or assessments) × 2 (P&P or DTI) × 2 (replication 1 or 2) factorial design.

The validation of the two types of responses, and verification that they could be dissociated empirically, was based on an aggregate-level MTMM analysis of the average responses provided by the eight groups of respondents. The correlations among the eight sets of responses were computed across stimuli, with respondents serving as replicates (e.g., Holbrook and Batra 1987; Kahneman et al. 1993). The advantage of this approach is that it does not require the repeated measurement of respondents, which is often a source of spurious correlations. The comparison of the relative speed of feeling and assessment responses was based on individual-level analyses of the responses provided by the subjects in the DTI conditions. The comparison of the interpersonal agreement elicited by the two types of responses was based on analyses of the responses provided by subjects in all eight conditions.

Procedure. The study was conducted in small lab sessions. Subjects first received written instructions about the type of response they were to monitor and report (see below). They then received either a booklet with P&P measures or additional instructions on how to operate the DTI. Practice with the task and instructions were provided with three trial pictures. Subjects were then exposed to the 35 target pictures. The pictures were presented on a TV monitor, for 10 seconds each, followed by a 10-second blue screen. Subjects were instructed to record their responses during each of these 20-second windows, which was ample time.

Response Type and Focus. Subjects were prompted to focus on either their feelings or their assessments through a combination of task instructions and measurement instruments. Instructions in the feeling-monitoring condition explicitly asked subjects to concentrate on their feelings and ignore their cold assessment of the stimuli.

> What we are interested in is the overall pleasantness of your feelings in reaction to each picture. In other words, does each picture make you feel pleasant or "positive" (e.g., happy, joyful, pleased, proud) or does it make you feel unpleasant or "negative" (e.g., sad, angry, disgusted, scared)? It is important not to confuse the pleasantness of your feelings in reaction to the picture with a judgment or an evaluation of the picture. A picture may make you sad or uncomfortable—these are unpleasant feelings. But you may still judge the picture favorably because you find it interesting, unusual or because it is of high quality. Similarly, even though a picture may make you feel positive, you may still evaluate it unfavorably because you judge it uninteresting, ordinary, or because it is of poor quality . . . you should concentrate on how pleasant or unpleasant each picture makes you feel, regardless of whether you evaluate it as being a "good" (i.e., you like it) or a "bad" (i.e., you dislike it) picture.

Subjects in this condition received additional instructions on how to report their feelings using either the DTI instrument or the P&P affect scale. Practice on three trial pictures was given to help subjects internalize the proper response focus for the subsequent pictures.

Subjects in the reason-based assessment condition were explicitly asked to focus on their cold assessment of each picture and disregard their feelings:

> What we are interested in is your overall evaluation of each picture. By evaluation, we mean a judgment of whether you like (i.e., you think it is "good") or dislike (i.e., you think it is "bad") each picture. It is important not to confuse your evaluation of the picture with the pleasantness of your feel-

> ings in reaction to it. For instance, even though a picture generates unpleasant feelings (e.g., you feel sad or disgusted), you may still like it (i.e., find it to be a "good" picture) because you find it interesting, unusual or because it is of high quality. Similarly, even though a picture generates pleasant feelings, you may still dislike it (i.e., judge that it is "bad") because it is uninteresting, ordinary, or because it is of poor quality. . . . you should concentrate on how you evaluate each picture (i.e., like/dislike, good/bad) and ignore how it makes you feel.

Subjects in this condition also received additional instructions on how to report their assessments using the DTI instrument or the P&P assessment scale. Practice on three trial pictures was also given to help subjects internalize the proper response focus.

Measures

DTI Affect. Two groups of subjects (one per replication) reported the pleasantness of their feelings using an electronic dial whose position was recorded every 200 milliseconds on a 1 (negative affect) to 100 (positive affect) scale with 50 indicating neutrality. Subjects were instructed to adjust their dial until it correctly reflected their response to each picture, then leave it in that position until prompted to return to the midpoint in preparation for the next picture. Subjects' typical responses thus showed an ascending (or descending) trend leading to a plateau (or a valley), followed by a return to the midpoint. The analyses were based on these "plateau" or "valley" values.

DTI Assessment. Subjects in the two DTI assessment conditions were similarly asked to use the dial to report their liking/disliking of each picture. The usage instructions were the same as those in the DTI affect condition. Again these responses were summarized through their plateau or valley values.

Affect Scale. Subjects in the P&P affect conditions reported their feelings on a multi-item scale similar to ones used by Izard (1972) and Plutchik (1966). Subjects were asked to rate on a 1–5 scale (1 = not at all, 5 = very strongly) how strongly they felt each of 10 emotional experiences at the sight of each picture. The items were: "I had unpleasant feelings looking at the picture"; "The picture made me feel happy"; "I was disgusted by the picture"; "The picture made me feel good"; "I was fearful looking at the picture"; "The picture made me feel bad"; "The picture made me feel angry"; "The picture made me feel joyful"; "I had pleasant feelings looking at the picture"; and "The picture made me feel sad."

Picture Assessment Scale. Subjects in the P&P assessment conditions rated each picture on six seven-point semantic differential scales whose endpoints were counterbalanced: "The picture is good/bad"; "I dislike/like this picture"; "This picture is satisfactory/unsatisfactory"; "I am favorable/unfavorable to this picture"; "The picture is interesting/uninteresting"; and "This picture is valuable/worthless."

Results

Preliminary Analyses: Response Validation and Dissociation. To examine the aggregate correlations among the various measures, each P&P measure was first summarized by factor-analyzing the average responses (across subjects) to each measure's multiple items. A principal component analysis of the affect items revealed a single factor accounting for 84 and 86 percent of the variance in replications 1 and 2, respectively. The loadings clearly indicated that this factor captured the pleasantness of subjects' feeling responses to the pictures. A factor score for pleasantness was thus computed for each picture. A similar analysis of the assessment items showed that a single factor accounted for 68 and 74 percent of the variance in replications 1 and 2, respectively. The loadings indicated that this factor represented liking-disliking of the pictures. Additional analyses show that the second factor, in fact, captured affective responses. This suggests that the first factor (which was orthogonal to the second) was essentially affect-free, as intended. A factor score for reason-based liking was therefore computed for each picture based on the first factor.

Table 1 provides the across-pictures correlations among the different summary measures of feeling-based responses and reason-based assessment responses, aggregated across respondents within each replication. To save space, we report only a simple MTMM analysis of these correlations. Same-trait/same-method correlations across replications indicate that the two affect measures had very high reliability ($r = .96$ and .98), whereas the two assessment measures had slightly lower reliability ($r = .68$ and .79). The same-trait/different-method correlations indicate that the two affect measures also displayed high convergent validity (r's between .97 and .98). The convergent validity of the two assessment measures was more modest, although substantial (r's between .73 and .78). The convergence across methods is notable given that the measures came from independent groups of respondents using maximally different instruments. Finally, the remaining correlations (different-trait/same-method and different-trait/different-method) indicate that affect and assessment measures had good discriminant validity (r's between .09 and .47). Consistent with this MTMM analysis, a principal component analysis of the eight summary measures identified two distinct factors, accounting jointly for 91 percent of the variance. As expected, after a VARIMAX rotation, the four assessment measures loaded on one factor, whereas the four affect measures loaded on the other (see loadings in Table 1). When allowed to be correlated (using a PROMAX rotation), the two factors exhibited a correlation of $r = .26$. In summary, the aggregate analysis suggests that the combination of instructions and instruments used in this study was successful in generating feeling-based and assessment responses that were construct valid and empirically distinguishable.

TABLE 1

STUDY 1 (PICTURE STIMULI): CORRELATIONS AMONG MEASURES AND FACTOR LOADINGS

	Affect, DTI (R1)	Affect, DTI (R2)	Affect, P&P (R1)	Affect, P&P (R2)	Assessment, DTI (R1)	Assessment, DTI (R2)	Assessment, P&P (R 1)	Assessment, P&P (R2)	Factor 1 (59%)	Factor 2 (32%)
Affect, DTI (R1)									.97	.17
Affect, DTI (R2)	**.96*****								.98	.12
Affect, P&P (R1)	.97***	.97***							.99	.10
Affect, P&P (R2)	.97***	.98***	**.98*****						.99	.13
Assessment, DTI (R2)	*.46***	*.46***	*.43***	*.47***					.37	.83
Assessment, DTI (R2)	*.21*	*.16*	*.14*	*.17*	**.68*****				.05	.89
Assessment, P&P (R1)	*.17*	*.11*	*.09*	*.12*	.78***	.73***			.00	.93
Assessment, P&P (R2)	*.30*	*.25*	*.25*	*.25*	.73***	.77***	**.79*****		.15	.90

NOTE.—Bold type indicates stability coefficients, underscoring indicates convergent validity coefficients, and italics indicates discriminant validity coefficients.
*$p < .05$.
**$p < .01$.
***$p < .001$.

Response Latencies. The relative speed of monitoring and translating feelings and reason-based assessments into overt stimulus judgments was tested by examining the time that respondents in the DTI conditions took to adjust their dial to express their responses to each picture (i.e., the time elapsed between the onset of the picture and the moment the dial reached its plateau or valley value). These response latencies were submitted to a 2 (type of response) × 35 (pictures) mixed ANOVA. The analysis revealed a main effect of pictures ($F(34, 2{,}482) = 2.29$, $p < .001$), showing an expected substantial variation in response times across pictures. More important theoretically was a strong main effect of type of response ($F(1, 73) = 20.34$, $p < .0001$). Respondents took significantly less time to report their feelings about each picture ($M = 8.08$ seconds) than they did to report their assessments of each picture ($M = 9.94$ seconds). This effect was not qualified by an interaction with pictures ($F(34, 2{,}482) = 1.07$, $p = .36$), indicating that it was robust across stimuli. Furthermore, the difference in latencies between feeling-based and reason-based responses was the same for the first half of the pictures that subjects saw ($M_{\text{Feeling}} = 8.17$ vs. $M_{\text{Reason}} = 9.93$) as for the second half ($M_{\text{Feeling}} = 8.03$ vs. $M_{\text{Reason}} = 10.02$; $F < 1$). This suggests that this effect was not contingent on the amount of practice that respondents had. These findings support the prediction that the conscious monitoring of feelings toward meaningful stimuli is potentially faster than a reason-based assessment of the same stimuli.

TABLE 2

STUDY 1: REGRESSION ANALYSIS OF RESPONSE LATENCIES

Predictors	Estimate	*t*	*p*
Intercept	8.793	17.54	.001
Response level	.067	3.99	.001
Response level^2	−.00068	−4.54	.001
Response type (0 = feeling, 1 = assessment)	.5323	−.84	.398
Level × type	.06133	2.87	.004
Level^2 × type	−.000527	−2.63	.009

NOTE.—*t*-tests are two-tailed.

To explore the process underlying the greater speed of feeling responses, we investigated the relation between response latencies and extremity (positivity-negativity) of the responses. For evaluative responses, we predicted an inverted-U relationship between response time and response extremity. This is because consideration of multiple assessment dimensions should be time consuming, and moderate judgments are somewhat more likely to be based on an integration of multiple evaluative dimensions than are extreme judgments (e.g., Linville 1982). However, feeling response latencies should be less strongly related to their extremity, because feeling responses that arise from schema-matching and sensory-motor programs should be less sensitive to the dimensional complexity of the stimulus.

This prediction was tested in a random-coefficient regression, modeling response times as a quadratic function of the response levels. The regression was computed across subjects and pictures and included a random-effect term for each subject to account for dependencies due to the use of multiple observations from the same subject. The analysis included five predictors: (1) the response level (on a 1–100 scale), (2) its square value, (3) a dummy variable capturing the type of response (0 = affect, 1 = assessment), (4) the interaction between 1 and 3, and (5) the interaction between 2 and 3. As reported in Table 2, the analysis revealed that response times were an inverted-U function of the response levels: β_1 was significantly positive, whereas β_2 was significantly negative. Extreme response levels—positive or negative—were associated with lower latencies than were more moderate response levels. As predicted, significant interactions (β_3 and β_4) indicated a more pronounced inverted-U relationship between response level and response time for assessments than for feeling responses. As shown in Figure 1, assessments that were moderate—and presumably in-

FIGURE 1

STUDY 1: RESPONSE LATENCIES AS A FUNCTION OF TYPE OF RESPONSE AND EXTREMITY OF RESPONSE

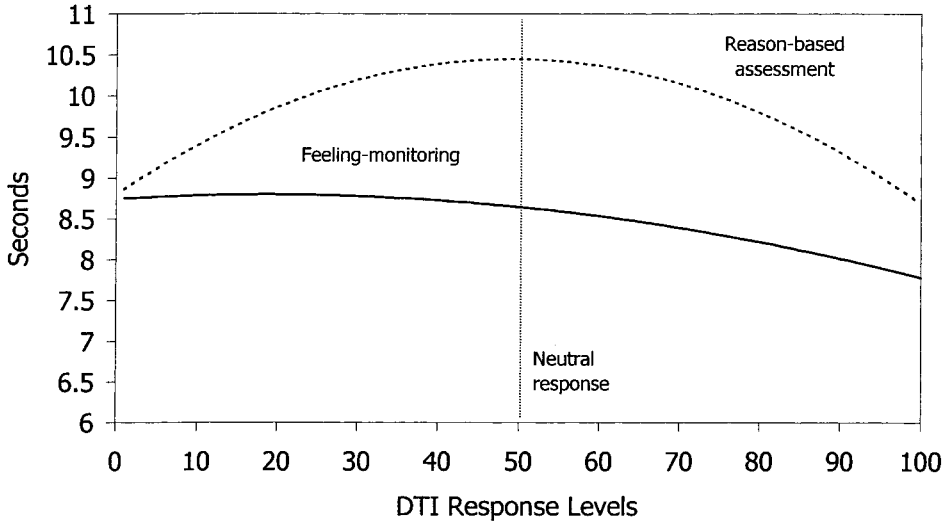

TABLE 3

STUDY 1: INTERPERSONAL AGREEMENT COEFFICIENTS

	Replication 1		Replication 2	
Instrument and item	*n*	α	*n*	α
Feeling DTI:				
Pleasantness	17	.97	18	.98
Feeling P&P:				
Angry	22	.93	17	.94
Disgusted		.96		.96
Fearful		.94		.90
Feel bad		.98		.97
Feel good		.98		.98
Sad		.98		.98
Happy		.97		.98
Joyful		.98		.98
Pleasant		.98		.98
Unpleasant		.98		.98
Assessment DTI:				
Liking	20	.83	20	.79
Assessment P&P:				
Like/dislike	24	.80	23	.83
Good/bad		.72		.76
Interesting/ uninteresting		.69		.71
Satisfactory/ unsatisfactory		.74		.77
Favorable/ unfavorable		.81		.83
Valuable/worthless		.74		.76

volved the integration of a greater number of stimulus dimensions or reasons—were particularly slower compared to feeling-based responses.

Interpersonal Agreement. Interpersonal agreement was assessed through a "transposed" coefficient alpha calculated by treating each picture as an observation and each respondent as a separate judge (see Holbrook and Batra 1987). This coefficient was calculated for each item within each of the eight between-subjects conditions. It indicates the degree of correlation among participants' responses. The results are summarized in Table 3. Overall, there was substantial interpersonal agreement in subjects' responses (average $\alpha = .84$), which justifies the aggregate-level analyses described earlier. More important, the agreement coefficients for the assessment measures, although substantial, were noticeably lower (average $\alpha = .77$) than those for the feeling measures (average $\alpha = .97$). Agreement on picture assessments was lower than agreement on feeling responses regardless of whether the measures were DTI-based or P&P-based. It is also noteworthy that the feeling responses exhibited a remarkably high degree of interpersonal agreement for both global feelings (overall pleasantness/unpleasantness) and more specific emotional responses, which presume greater cognitive differentiation (e.g., anger, sadness, pride). An ANOVA of the 36 agreement coefficients (transformed into Fisher *Z*'s) shows that the difference in agreement across types of response was highly significant ($F(1, 34) = 220.03$, $p < .0001$).[1]

[1]The greater number of items in the affect P&P scale was not a source of bias because (*a*) the analyses were performed item by item and (*b*) respondents, not the items, were the replicates. Although the number of

Discussion

The aggregate MTMM results indicate that the feeling and picture assessment responses observed in this study were construct valid. Confirming earlier intuition that people can like (dislike) stimuli that induce negative (positive) emotions (Buck 1985; Martin et al. 1997), we found the two types of responses to be empirically distinguishable. The modest correlation between the two types of responses observed in this study replicates earlier results (e.g., Breckler and Wiggins 1989). More important, the results support two of our predictions about the basic properties of feeling-based and reason-based evaluations.

First, the conscious monitoring of feelings toward the stimuli was found to be faster than reason-based assessments of the same stimuli. This finding substantiates a widely held assumption that so far had received only limited empirical support, especially in the context of meaningful, supraliminal stimuli (see Verplanken, Hofstee, and Janssen 1998 for one exception). We argue that this is because many everyday consumer stimuli trigger affective sensory-motor programs (Type I affect) and emotional schemata (Type II affect) that involve automatic associative processes and are very rapid. In contrast, reason-based assessments are often slower because they require a greater number of cognitive operations, especially when the multiple evaluative criteria need to be integrated.

Note that our methodology clearly overestimates the amount of time actually required to register privately the two types of evaluative responses. Once respondents had mentally registered their feelings or reason-based assessments, they had to translate them into an overt response by adjusting their dial. This physical translation was obviously time consuming. The total (overt) response latencies observed in this study were in the 8–10-second range, compared to the two-to-three-second range typically observed in studies where respondents expressed their responses by simply pressing a key (see Schmidt-Atzert 1988). Nevertheless, to the extent that our instrument (hence, its associated overestimation) was the same across conditions, the difference in latencies between feeling- and reason-based overt responses remains meaningful.

One could argue that reason-based assessments were comparatively slower because they might be represented mentally in a propositional format, whereas feelings might be represented in a sensory-motor format. As a result, the "nonverbal" DTI may have been less commensurable with the former than with the latter, thus requiring additional processing. One result seems inconsistent with this explanation. According to a format commensurability explanation, feeling responses should have been faster than assessment responses regardless of the extremity of these responses. As noted earlier, it was mostly moderate assessments that tended to take significantly longer than feeling responses (see Fig. 1). This finding is noteworthy because moderate responses required smaller physical movements (away from the starting neutral point) than more extreme responses, hence the former should (physically) have been reported more rapidly than the latter. The longer latencies observed for moderate assessments seem consistent with the view that reason-based evaluative responses tend to be slower because they reflect the integration of a greater number of judgmental dimensions (Edwards 1990).

One could also advance the view that reason-based assessments took longer because the process of excluding feelings from the evaluations was less natural than the reverse. However, if feeling-exclusion had indeed been more difficult than reason-exclusion in this task, the difference in latencies across conditions should have decreased as respondents gained more practice with the task. The difference in latencies was in fact the same for early and later trials. Furthermore, the hypothesis that the exclusion of feelings is more difficult than the exclusion of reasons, again, would not explain why the difference in latencies was more pronounced for moderate assessments than for more extreme responses. Our account does.

Our results seem to contradict those of Schmidt-Atzert (1988), who observed that subjects were slightly slower when reporting their "emotional responses" to pictures than when reporting their "evaluation responses." There are, however, important methodological differences between the two studies. First, Schmidt-Atzert's "evaluation" instructions allowed judgments that were primarily based on affect, provided that they were object-centered (e.g., "This picture is disgusting") as opposed to person-centered (e.g., "I feel disgusted"). In contrast, our reason-based assessment instructions explicitly excluded affective considerations. The possible inclusion of affect-based judgments among Schmidt-Atzert's evaluation responses may have contributed to the lower evaluation latencies that he observed. Second, unlike in this study, Schmidt-Atzert did not actually measure the time respondents took to report their emotional feelings or their "evaluations." Instead, he measured the time respondents took to report that a feeling or an evaluation had taken place. Moreover, when reporting that an evaluation had occurred, subjects in his study were allowed to use any bases for evaluation (e.g., funny, disgusting, nice, ugly). In contrast, subjects in our reason-based assessment condition were instructed to map these assessments onto a single like-dislike/good-bad continuum. Because, as mentioned above, it is the integration of multiple judgmental dimensions into a single evaluative response that slows down reason-based assessments, Schmidt-Atzert's findings (1988) may underestimate the actual latencies of truly integrative reason-based evaluations. Finally, in his study respondents provided both emotional feeling and evaluation responses (in a within-subject design) and verbalized the basis for their response after each trial. These design characteristics were likely to have introduced noise and made it more difficult to detect differences (and in fact the difference in latencies between

respondents (hence replicates) was not perfectly matched across conditions, additional analyses show that the result holds even after controlling for the number of respondents.

emotional feeling and evaluation responses was only marginally significant).

We also found that feeling-based assessments exhibited consistently greater interpersonal agreement than did reason-based assessments. The finding held even when the same physical instrument (the DTI) was used to report the two types of responses. A remarkably high degree of interpersonal agreement was observed both for global feelings of pleasantness and for more specific emotional feelings (e.g., anger, disgust, pride). Again we attribute this high level of interpersonal consensus to the triggering of affective sensory-motor programs and emotional schemata that are largely shared, biologically and socially. In that sense, affect monitoring builds on a shared social reality and may represent a much more reliable and diagnostic basis for judgments than many have suggested. Reason-based assessments, on the other hand, typically involve idiosyncratic knowledge, preference weightings, and standards. Such reason-based assessments are likely to exhibit considerably less interpersonal consistency in direct proportion to the dimensionality of the object or issue being judged.

STUDY 2

This study investigates how feeling and reason-based evaluative responses relate to the number of valenced thoughts spontaneously generated by the same stimuli. As discussed earlier, there is ample reason to predict that spontaneous thought generation should be more strongly related to reason-based assessments than to feeling responses. We predicted, however, that spontaneous thought generation would in fact be more strongly related to the more rapidly generated (and possibly more motivationally relevant) feeling-based assessments than to reason-based assessments.

Method

Procedure. The 35 pictures used in study 1 were shown to two independent samples of undergraduates: one from the same student population (replication 1, $n = 96$) and one from a different university (replication 2, $n = 91$). Subjects were each exposed to half of the stimuli (plus trial pictures) to prevent fatigue. Subjects were instructed to list all the thoughts, feelings, and reactions that occurred to them in response to each picture. Each picture was shown for 10 seconds, followed by a blue screen. As soon as the blue screen appeared, subjects had 60 seconds to list all their thoughts, using a separate box for each thought. The procedure was repeated for every picture. After responding to all the pictures, subjects were asked to review the thoughts they had listed for each picture. For each thought listed, subjects were asked to express whether the thought was positive, negative, or neutral, by placing either a "+," "−," or "n" next to each box. Having subjects self-code their own thoughts eliminates potential misunderstanding of what subjects privately meant by each thought. In addition to categorizing the valence of their thoughts, subjects in replication 2 were also asked to rate whether each thought represented primarily a feeling or primarily an evaluative assessment of the stimulus. They reported these ratings on a 1 (feeling) to 7 (evaluation) scale anchored by graphical icons illustrating the distinction. They also received the following instructions about the distinction between the two types of thoughts: "Evaluative thoughts express some positive or negative judgment, opinion or appraisal of something you saw in that picture. Feeling thoughts, on the other hand, express how something has made us feel. These thoughts can reflect positive feeling responses, such as when we are smiling because something seems pleasant or we feel happy at the time. Feeling thoughts can also reflect negative responses, such as when we are frowning, because something seems unpleasant or we feel sad at the time."

Analyses. The main analyses were again performed at the aggregate level, using the stimuli as the units of analysis and subjects as replicates. Ninety-two respondents in replication 1 and 89 respondents in replication 2 provided complete sets of responses. Because each subject saw only half of the pictures, each stimulus data point represents an average across roughly half of the respondents (46 subjects in replication 1 and 45 subjects in replication 2). The main variables of interest were the number of positive thoughts elicited by each picture, the number of negative thoughts, and the balance of thoughts (number of positive minus number of negative). In replication 2, these numbers were further tabulated by type of thought: feeling or evaluative. Thoughts given a rating of 3 and below were categorized as "feeling" thoughts, whereas thoughts given a rating of 5 and above were categorized as "evaluation" thoughts. Across respondents and stimuli, 41.7 percent of the thoughts were categorized as feeling thoughts, 44.2 percent were categorized as evaluation thoughts, and 14.1 percent were considered ambiguous (rating of 4). As in study 1, we estimated the homogeneity of responses across respondents, then averaged each measure across respondents. We then examined the aggregate relations between these measures and the average affective and reason-based assessment responses—both DTI and P&P—generated in study 1.

Results

Preliminary Analyses. The degree of interpersonal agreement was again estimated via coefficient α calculated using subjects as replicates. The agreement coefficients reported in Table 4 show that there was again substantial agreement in terms of number of positive and negative thoughts evoked and the balance between these thoughts (α's ranging between .95 and .98). This substantial homogeneity justifies an aggregate level analysis of the thought responses. There was only a modest correlation between the balance of feeling thoughts and the balance of evaluation thoughts ($r = .28$). This result further supports the empirical dissociation between feeling and nonaffective bases of judgments.

TABLE 4

STUDY 2: RELATIONS BETWEEN FEELING AND ASSESSMENT RESPONSES, AND THOUGHT RESPONSES

	Replication 1 (n = 92)			Replication 2 (n = 89)			Replication 2 (n = 89)					
	Number positive thoughts	Number negative thoughts	Balance (positive-negative)	Number positive thoughts	Number negative thoughts	Balance (positive-negative)	Number positive feeling thoughts	Number negative feeling thoughts	Balance (positive-negative) feeling thoughts	Number positive evaluation thoughts	Number negative evaluation thoughts	Balance (positive-negative) evaluation thoughts
Interrespondent agreement (α)	.98	.97	.98	.98	.97	.98	.96	.93	.96	.91	.93	.95
Mean	1.65	1.58	.06	1.72	1.60	.12	.85	.78	.07	.69	.68	.00
Correlation with DTI feelings	.91***	−.95***	.95***	.92***	−.96***	.97***	.87***	−.90***	.94***	.88***	−.90***	.94***
Correlation with P&P feelings	.92***	−.96***	.96***	.92***	−.96***	.97***	.88***	−.92***	.96***	.89***	−.90***	.94***
Correlation with DTI assessments	.32*	−.25	.29*	.30*	−.30	.31	.29*	−.19	.26	.28	−.38**	.35**
Correlation with P&P assessments	.21	−.08	.15	.21	−.16	.19	.18	−.06	.13	.24	−.26	.27
Regression coefficient feeling	1.140***	−1.145***	1.285***	1.228***	−1.183***	2.412***	.652***	−.609***	1.262***	.456***	−.484***	.940***
Regression coefficient assessment	.025	.132*	−.107	.003	−.047	−.045	.005	−.097^{+}	−.102	.010	−.047	.058

$^{+}p < .10$.
*$p < .05$.
**$p < .01$.
***$p < .001$.

Relationship with Affective and Reason-Based Responses. Table 4 shows the correlations, across stimuli, between the thought measures in each replication and the feeling and reason-based responses assessed in study 1. The first six columns exhibit these correlations irrespective of types of thoughts (feeling or evaluative). As predicted, both the DTI and P&P measures of affective responses were highly correlated with the number of positive thoughts, number of negative thoughts, and the balance between positive and negative thoughts (absolute r's $> .90$). These strong correlations held across the two samples of respondents (replications 1 and 2). As expected, the two measures of reason-based assessment were far more weakly related to the number of positive thoughts, number of negative thoughts, and the balance between positive and negative thoughts (absolute r's $< .35$). Again, these weaker correlations were quite comparable across the two independent samples of respondents.

To test the differential association between affective and reason-based assessments and the various measures of cognitive responses, we averaged the two affective response measures and the two reason-based assessments after standardizing each DTI measure (the P&P measures, being factor scores, were already standardized). We then ran multiple regressions with each thought measure (number of positive thoughts, number of negative thoughts, thought balance) as the dependent variable, and the average affective assessment and average reason-based assessment as twin predictors. The results, reported in the last two rows of Table 4, show that affective responses were consistently better predictors of the valence of respondents' thoughts than were reason-based assessment responses. This effect held across the two replications of respondents.

The last six columns of Table 4 display these relationships, after separating feeling and evaluation thoughts (within replication 2). The pattern of correlations with the affective and reason-based assessment measures was almost identical for the two types of thoughts. Both feeling and evaluation thoughts were more strongly related to the valence of affective responses than to the favorability of reason-based assessments. Again, multiple regression analyses show that, when effects of affective and reason-based assessment responses are jointly estimated, the valence of affective responses was a consistently more potent predictor of the number of valenced thoughts—regardless of their type—than was the favorability of reason-based assessments.

Discussion

The results support the prediction that feeling responses to moderately complex stimuli are better predictors of the spontaneous generation of valenced thoughts than are reason-based assessments. This effect seems quite robust: (1) it was replicated across two samples of respondents from two different universities; (2) it held across two measures of feelings and reason-based assessments; and (3) it held even when self-coded feeling and evaluation thoughts were considered separately. This finding supports, at a process level, Wilson and his colleagues' (e.g., Wilson et al. 1993) hypothesis that reason-based modes of evaluation often call to mind judgment dimensions that are different from those that would be otherwise used in more spontaneous modes of evaluations. This finding also extends their research by showing that the conscious monitoring of feelings is intimately related to people's spontaneous thoughts about attitudinal objects.

We believe that the high correlation observed between feeling responses and spontaneous thoughts (feeling and evaluation thoughts alike) occurred because initial feeling responses to a stimulus often trigger further thoughts that are either consistent with the initial feelings or are helpful in interpreting the experienced feelings. Because of their motivational significance feelings may direct our attention and thinking in a way that cold assessments do not. Of course, one could argue that the direction of causality is the reverse: feelings and thoughts might be strongly correlated because the thoughts that are spontaneously triggered by a stimulus are, in fact, the antecedents of the feeling responses. However, additional analyses show that the number of positive or negative thoughts triggered by a picture was negatively related to the time it took to report a feeling response ($\beta = -.356$, $p = .06$). In other words, subjects took less time to report their feelings for pictures that elicited a greater number of thoughts than for pictures that elicited fewer thoughts. This finding is inconsistent with the view that the thoughts preceded the feelings. It is, however, consistent with our stirring explanation that more rapidly experienced affect directs attention and effort, instantiating thought processes useful in reducing uncertainty about the nature and meaning of the feeling state.

STUDY 3

The next two studies investigate the judgmental properties of feelings in the context of responses to television commercials. Aside from their obvious relevance to consumer research, commercials offer unique opportunities for studying the theoretical issues discussed above. First, commercials are very rich semantically and perceptually, moving our analyses still further away from the relatively impoverished stimuli used in affect priming and mere exposure studies. Second, they are dynamic; their information content is continuously changing, as with many real world stimuli. Third, because their pacing cannot be adjusted by perceivers, they often pose challenges to on-line evaluation processes and allow the comparative study of how feelings and reason-based assessments are adjusted in real time. Television commercials are also of theoretical interest because they often elicit feeling and reason-based responses that are highly intertwined. As a result, consumers may have greater difficulty disentangling these alternative bases of evaluations, especially when required to provide real-time assessments.

The design was the same as in study 1. Eight groups of subjects were asked to report either their feeling responses or their reason-based assessments of a series of commercials.

Some of the respondents did so using the DTI, the others using P&P measures. Subjects were further divided into two replication sets of respondents. As in study 1, we used an aggregate-level MTMM analysis to examine the validity of the various measures and the possible dissociation between the two types of responses. The main objective, however, was to test two predictions. The first prediction was that feeling responses to the commercials would be more volatile than reason-based assessments of the same commercials, reflecting the advantage of this judgmental process in responding rapidly to changing circumstances. As study 1 showed, feelings—though potentially subject to further appraisal processes—can be registered faster than reason-based assessments. Feeling responses should, therefore, be adjusted more frequently as a function of changes in the stream of information. This prediction was tested through individual-level analyses of respondents' DTI responses. The second prediction was that respondents would again exhibit greater interpersonal agreement in their feeling-based responses than in their reason-based assessments, suggesting that the former possess greater reliability than had been imagined. Greater interpersonal consistency of feelings would be particularly meaningful considering the substantial richness and complexity of commercials. We tested this prediction by analyzing the correlation among respondents' feeling and reason-based assessments within all eight conditions.

Method

Stimuli. Nine commercials that were expected to elicit a wide range of feeling and reason-based assessments were selected. The commercials involved different product categories (e.g., insurance, airline, jeans) and were unfamiliar to most subjects. Because commercials evoke dynamic responses that are much more complex than those from pictures (e.g., a sad beginning and a happy ending), we expected considerable heterogeneity among both types of responses to each commercial. Since it would not be meaningful to summarize these responses with a single data point representing the entire commercial, each commercial was divided into cohesive segments, determined through discussions among the researchers. Each segment was expected to elicit homogeneous responses that could be adequately summarized by either a P&P measure or a DTI measure. The units of analysis were 36 segments from the nine commercials.

Subjects and Design. In total, 289 undergraduates were exposed to the commercial stimuli under one of four conditions. Half of the respondents were asked to report their feeling responses to the commercials; the other half were asked to report their reason-based assessments of the commercials. One hundred and one respondents reported these responses using P&P measures. One hundred and eighty-eight respondents reported these responses with the DTI. The DTI data were collected in two settings: 113 subjects responded to six commercials (set 1: 25 segments) and 75 subjects responded to the remaining three commercials (set 2: 11 segments). For replication purposes, respondents were randomly assigned to one of two replication sets within each of the four instrument/response-type conditions.

Response Types and Procedure. All subjects first received the same instructions about the distinction between the two types of responses as in study 1. Subjects in the DTI condition then received additional instructions on how to operate the dial to report either type of response. Unlike study 1, they were instructed to adjust the dial continuously to reflect changes in their responses. Although the commercials were presented in their full length, for analysis purpose the responses were summarized by segment. For each subject, we computed the average position of the dial during each predesignated segment.

Subjects in the P&P conditions reported their responses using the same scales as in study 1. These responses were collected segment by segment. Respondents in these conditions saw the commercials in their original form with pauses after each segment, which allowed them to respond to the clearly defined segment they had just seen. Although the order of the commercials was counterbalanced across sessions, the segments of each commercial were always presented in their original order to preserve the commercial's meaning.

Results

Response Validation and Covariation. The two P&P measures were summarized by factor-analyzing the average responses (across subjects) to each measure's multiple items. A principal component analysis of the feeling items revealed a single factor of pleasantness accounting for 83 percent of the variance in both replications. A similar analysis of the commercial assessment items showed that a single overall liking factor accounted for 84 and 81 percent of the variance in replication 1 and 2, respectively. A factor score for pleasantness of feelings and a factor score for liking were computed for each segment.

Table 5 shows the correlations (computed across segments) among the eight measures. The same-trait/same method correlations show that all measures had high reliability (r's $> .90$). The same-trait/different method correlations indicate that the two feeling measures had adequate convergent validity (r's around .80). This convergence is noteworthy given the use of independent groups of respondents and the dramatic differences in the way the two types of measures were collected and summarized (a truly continuous measure and a paper-and-pencil retrospective measure). The convergent validity coefficients for the commercial assessment measures were much lower (r's between .43 and .63). Although reliable, the P&P and DTI commercial assessment measures showed only a modest degree of trait overlap. The discriminant validity coefficients indicate that the two commercial assessment measures did not converge because the continuous DTI measure captured a substantial degree of feeling responses (different-trait/same-method r's

TABLE 5

STUDY 3 (COMMERCIAL STIMULI): CORRELATIONS AMONG MEASURES AND FACTOR LOADINGS

	Feeling, DTI (R1)	Feeling, DTI (R2)	Feeling, PP (R1)	Feeling, PP (R2)	Assessment, DTI (R1)	Assessment, DTI (R2)	Assessment, PP (R1)	Assessment, PP (R2)	Factor 1 (68%)	Factor 2 (19%)
Feeling, DTI (R1)									.93	.22
Feeling, DTI (R2)	**.99*****								.94	.20
Feeling, PP (R1)	.80***	.82***							.91	.10
Feeling, PP (R2)	.79***	.81***	**.98*****						.91	.11
Assessment, DTI (R1)	*.84****	*.84****	*.64****	*.64****					.78	.46
Assessment, DTI (R2)	*.74****	*.75****	*.58****	*.59****	**.94*****				.70	.53
Assessment, PP (R1)	*.43***	*.41**	*.38**	*.39**	.57***	.63***			.25	.94
Assessment, PP (R2)	*.31*	*.29*	*.25*	*.25*	.43***	.45***	**.91*****		.09	.95

NOTE.—Bold type indicates stability coefficients, underscoring indicates convergent validity coefficients, and italics indicates discriminant validity coefficients.
*$p < .05$.
**$p < .01$.
***$p < .001$.

between .74 and .84; different-trait/different-method *r*'s between .58 and .64). By contrast, the retrospective P&P commercial assessment measure was relatively immune to contamination from feelings (different-trait/same-method *r*'s between .25 and .39; different-trait/different-method *r*'s between .29 and .43). A principal component analysis of the eight summary measures uncovered two factors, accounting jointly for 81 percent of the variance. After a VARIMAX rotation, the four measures of feelings loaded on one factor, and the two P&P commercial assessment measures loaded on the other factor (see loadings in Table 5). However, the continuous DTI commercial assessment measure loaded on both factors, with stronger loadings on the first (feeling) factor than on the second like/dislike factor. An oblique (PROMAX) rotation revealed a more substantial correlation between the feeling and the commercial assessment factors ($r = .42$), confirming that the two types of responses were more intertwined than in the picture studies.

These results indicate that people reporting real-time commercial assessments had difficulty separating out feeling-based responses. In contrast, subjects using P&P measures were able to do so, presumably because they were given greater opportunity to reflect about each segment. The type of covariation reflected in the real-time DTI assessment measures is probably representative of consumers' natural responses to everyday stimuli—responses that combine reason-based assessments and feelings rather than focus on either type of information exclusively.

Response Volatility. The relative volatility of both types of responses was examined by comparing the continuous responses provided by the DTI-feeling group and the DTI-commercial assessment group. The comparison should be regarded as exploratory because the DTI-commercial assessment measures appeared to capture a substantial amount of feeling responses. This overlap should produce conservative estimates of true differences between these two types of responses, making any observed differences meaningful.

From each respondent's original continuous responses (recorded every 200 msec), we computed two measures of volatility. The first measure of volatility was the standard deviation of the each person's response over time. These standard deviations were calculated within each cohesive segment rather than across entire commercials. An average (segment-level) standard deviation was calculated for each subject. The second measure of volatility was the number of runs (directional changes) within each commercial. A run was defined as a change in either direction (i.e., from increasing to decreasing or vice versa) of the dial by a magnitude of two points or more (on the 1–100 scale). Because the commercials had varying lengths, we calculated the number of runs per 30-second segment for each subject.

The number of runs and average standard deviations were submitted to a MANOVA with type of response (feeling or reason-based assessment) and commercial set (1 or 2) as between-subjects factors. The analysis revealed a main effect of set (Wilk's $\lambda = .926$, $F(2, 183) = 7.28$, $p < .001$), suggesting that there were significant differences in response volatility across sets of commercials. More important, there was also a marginally significant main effect of type of response (Wilk's $\lambda = .972$, $F(2, 183) = 2.628$, two-tailed $p < .08$, one-tailed $p < .04$). This effect was not qualified by an interaction with set (Wilk's $\lambda = .997$, $F < 1$), suggesting that it was consistent across sets of commercials. Univariate analyses show that, as predicted, the average standard deviation by segment was significantly greater among subjects reporting their feeling responses ($M = 4.68$) than among subjects reporting their commercial assessment responses ($M = 4.06$; $F(1, 184) = 5.13$, $p < .03$). Similarly, the number of runs was slightly, but not significantly, greater among subjects reporting their feeling responses ($M = 1.69$) than among subjects reporting their commercial assessment responses ($M = 1.37$; $F(1, 184) = 1.79$, two-tailed $p = .18$). The results, which (as suggested above) are likely to be conservative, partially support the prediction that, in a dynamic environment, feeling responses are more volatile than reason-based assessments.

Interpersonal Agreement. The degree of response commonality or consensus, as defined by interpersonal

agreement in subjects' responses, was calculated as in the first two studies. The agreement coefficients are reported in Table 6. Overall, the degree of interpersonal agreement was high (average $\alpha = .92$), justifying the aggregate-level analyses reported above. More important, as in study 1, the agreement coefficients for the feeling responses (average $\alpha = .97$) were consistently higher than those for commercial assessment responses (average $\alpha = .85$). An ANOVA of the 36 coefficients (after Fisher's z transformation) shows that this difference was significant ($F(1, 34) = 78.75$, $p < .0001$). These results support the prediction that consciously monitored feelings exhibit greater response commonality (i.e., interpersonal consistency) than reason-based assessments, even in the context of richly textured stimuli such as television commercials. The only exception to this pattern was the DTI commercial assessment measure, which elicited unexpectedly high interpersonal agreement, presumably because it also captured a substantial degree of on-line affect monitoring, as discussed above.

TABLE 6

STUDY 3: INTERPERSONAL AGREEMENT COEFFICIENTS

	Replication 1		Replication 2	
Item	*n*	α	*n*	α
Feeling DTI:				
Pleasantness	28 and 17*	.98	27 and 17*	.98
Feeling P&P:				
Angry	28	.95	28	.96
Disgusted		.91		.93
Fearful		.96		.96
Feel bad		.97		.98
Feel good		.98		.98
Sad		.97		.98
Happy		.98		.98
Joyful		.98		.98
Pleasant		.98		.98
Unpleasant		.98		.98
Assessment DTI:				
Liking	29 and 20*	.96	29 and 21*	.97
Assessment P&P:				
Like/dislike	22	.86	23	.81
Good/bad		.81		.80
Interesting/ uninteresting		.83		.82
Satisfactory/ unsatisfactory		.80		.81
Favorable/ unfavorable		.86		.80
Valuable/worthless		.90		.89

*Sample sizes for commercials set 1 and set 2, respectively.

Discussion

The results support the prediction that people are more likely to display judgmental consensus in their consciously monitored feeling responses to commercials than in their reason-based assessments of the same commercials. The findings suggest that the reliability of feelings extends beyond the realm of moderately complex static stimuli such as magazine pictures to even richer dynamic stimuli such as television commercials.

The results also offer some support for the prediction that, in a dynamic response environment, feelings are more responsive/volatile than reason-based assessments. These results, however, need to be interpreted with some caution. First, the evidence was not strong statistically, perhaps because it was difficult and artificial for people to separate out their feeling-based responses when using the real-time DTI-commercial assessment measure. Second, when operationalized by the standard deviation of the response over time, volatility is not solely a function of the speed of response but is also a function of the range of the response. Although the range of responses was partially controlled by performing the analyses at the segment level, this control was not perfect. Therefore, compared to reason-based evaluations, feeling responses may be more volatile because of their greater relative speed and/or their greater average amplitude. These issues need further investigation.

Although feeling and reason-based assessments appear to be empirically distinguishable even in the context of television commercials (where the two types of responses are likely to be closely intertwined), subjects seemed to have considerable difficulty doing so when making on-line commercial assessments. Unlike in study 1, the DTI-commercial assessment measure seemed to capture a mixture of both feeling and reason-based assessments, with a stronger component of the former. As continuous responding using the DTI probably placed much heavier processing burdens on subjects, it may have been especially difficult for people to ignore their feeling-based responses, particularly the more rapid and automatic Type I and Type II feelings. To the extent that the processing demands of this task map on to those typically faced when the flow of information is not under the recipient's control (e.g., interpersonal influence), people appear to face the daunting task of interpreting and integrating these two streams of information. Subjects in the DTI-commercial assessment condition might have consciously resorted to monitoring their feelings as a way of simplifying their evaluation task. This would be consistent with the evidence that feelings provide a fast and efficient means of valuation. Subjects in the P&P-commercial assessment conditions were better able to report their feeling-free evaluations because these evaluations were collected segment by segment, giving them more ample time to reflect and record their reason-based assessments of the commercials. Consistent with this interpretation, Siemer and Reisenzein (1998) recently found that increased time pressure and competing task demands both make people more likely to rely on their mood in evaluative judgments.

STUDY 4

This study reexamines, in the context of television commercials, the relationship between people's feeling responses and reason-based assessments and their spontaneous thought generation. As in study 2, we predicted that the valence of people's thoughts would be more strongly related

to the valence of their feeling responses than to the favorability of their reason-based assessments.

Method

The design closely matched that of study 2. The 36 segments (nine commercials) used in study 3 were shown to two independent samples of undergraduates, hereafter called replication 1 ($n = 72$) and replication 2 ($n = 68$). Within each replication, one group of subjects saw half of the stimuli and another group saw the other half to prevent fatigue. The stimuli were presented segment by segment, as they were in the P&P conditions of study 3. After viewing each segment, subjects were given 45 seconds to list all the thoughts, feelings, and reactions that occurred to them during the segment. After completing the viewing/responding, they were then asked to reread the thoughts they had listed and code each as positive, negative, or neutral. Subjects in replication 2 were additionally asked to rate whether each thought represented primarily a feeling or an evaluative assessment of the stimulus. These subjects received the same instructions about this distinction as those used in study 2. They reported their ratings on the same 1 (feeling) to 7 (evaluation) scale. Again, thoughts rated 3 and below (40.4 percent of the thoughts) were categorized as feeling thoughts; those rated 5 and above (47.4 percent of the thoughts) were categorized as evaluative assessment thoughts; and those rated 4 (12.2 percent of the thoughts) were considered ambiguous.

Seventy-one respondents in replication 1 and 65 respondents in replication 2 provided complete sets of responses. The analyses were once more performed at the aggregate level, with each segment data point representing an average across roughly half of the respondents within each replication. As in study 2, the main variables of interest were the number of positive thoughts elicited by each segment, the number of negative thoughts, and the balance between these two numbers. These various thought measures were again related to the feeling and reason-based assessment measures assessed in study 3. Feeling and evaluation thoughts were also analyzed separately.

Results

As reported in Table 7, when the thoughts were tabulated irrespective of type, response homogeneity for the various thought measures was quite high (α ranging between .85 and .92), justifying an aggregation of the data across subjects. However, when the thoughts were separated into feeling and evaluative assessment thoughts, the latter exhibited a significantly lower degree of intersubject consistency (α between .60 and .64). Aggregate analyses of the evaluative assessment thoughts thus call for caution. There was a strong correlation between the balance of feeling thoughts and the balance of evaluation thoughts ($r = .78$, $p < .0001$), confirming that the two types of inputs were more intimately related than in the picture studies. Table 7 shows the raw correlations, across segments, between the thought measures in each replication and the feeling and commercial assessment responses measured in study 3.

When the type of thought is ignored (cols. 1–6), the various thought measures again correlated more strongly with the feeling measures than with the commercial assessment measures. Although the difference in correlations (between thoughts and feelings and between thoughts and assessments) appears to be smaller than that observed in study 2, the raw coefficients probably underestimate the true difference in correlation. This is because, as observed in study 3, the DTI commercial assessment measure also captured a substantial amount of feeling responses. To obtain cleaner estimates of the actual relations between thoughts and feelings and between thoughts and reason-based assessments, for each segment we computed factor scores of feelings and reason-based assessments based on a principal components analysis of the eight measures used in study 3. We then submitted each thought measure to a multiple regression with these two factor scores as predictors. The results are reported in the last two rows of Table 7. As in study 2, they again show that feeling responses were consistently better predictors of the valence of respondents' thoughts about the commercials than were reason-based assessments.

Similar regression analyses were performed within replication 2 for feeling and evaluation thoughts, separately. Feeling thoughts were again much better predicted by feeling responses than by reason-based assessment responses. The results for evaluation thoughts were more ambiguous. Positive evaluation thoughts were equally well predicted by reason-based assessment responses ($\beta = 0.153$, $p < .01$) and by feeling responses ($\beta = 0.101$, $p < .05$); the difference was not significant ($F < 1$). Negative evaluation thoughts were better predicted by feeling responses ($\beta = -0.136$, $p < .01$) than by reason-based assessment responses ($\beta = -0.027$, NS); the difference was significant ($F(1, 33) = 4.03$, $p < .06$). Finally, the balance of positive and negative evaluation thoughts was slightly, but not significantly ($F < 1$), better predicted by feeling responses ($\beta = 0.237$, $p < .005$) than by reason-based assessment responses ($\beta = .180$, $p < .02$).

Discussion

The results replicate study 2's finding that the number of people's valenced thoughts about stimuli is more strongly related to the valence of their feeling responses than to the valence of their reason-based assessments. This was certainly the case when all types of thoughts were considered together, and when feeling thoughts were considered separately. The pattern was not as clear when evaluation thoughts were considered separately. Perhaps the closer people get to an overall evaluation, the more likely they are to incorporate both types of information, especially in contexts—such as the viewing of commercials—that make it difficult to separate them. Alternatively, the number of evaluation thoughts, which was less consistent across respondents, may not have been reliable enough for aggregate-level analyses. Again, the lower interpersonal consistency

TABLE 7

STUDY 4: RELATIONS BETWEEN FEELING AND ASSESSMENT RESPONSES, AND COGNITIVE RESPONSES

	Replication 1 ($n = 71$)			Replication 2 ($n = 65$)			Replication 2 ($n = 65$)					
	Number positive thoughts	Number negative thoughts	Balance (positive-negative)	Number positive thoughts	Number negative thoughts	Balance (positive-negative)	Number positive feeling thoughts	Number negative feeling thoughts	Balance (positive-negative) feeling thoughts	Number positive evaluation thoughts	Number negative evaluation thoughts	Balance (positive-negative) evaluation thoughts
Interrespondent agreement (α)	.90	.86	.92	.85	.85	.88	.85	.82	.86	.60	.64	.63
Mean	1.46	1.14	.31	1.47	1.20	.38	.72	.52	.20	.63	.58	.04
Correlation with DTI feelings	.65***	−.63***	.69***	.61***	−.62***	.65***	.60***	−.55***	.67***	.36**	−.46**	.47**
Correlation with P&P feelings	.75***	−.83***	.85***	.72***	−.83***	.81***	.74***	−.81***	.88***	.41**	−.57***	.56***
Correlation with DTI assessments	.55***	−.51***	.58*	.57***	−.59***	.61***	.47**	−.52**	.56***	.44**	−.47**	.52***
Correlation with P&P assessments	.51**	−.31*	.36**	.51**	−.32*	.46**	.40**	−.33**	.42**	.52***	−.20	.44**
Regression coefficient feeling	.517***	−.431***	.948***	.457***	−.401***	.859***	.307***	−.221***	.528***	.101*	−.136**	.237**
Regression coefficient assessment	.029**	−.079	.359*	.276**	−.092	.368*	.101	−.056	.157+	.153**	−.027	.180*

+$p < .10$.
*$p < .05$.
**$p < .01$.
***$p < .001$.

in evaluation thoughts is consistent with the hypothesis that people exhibit greater commonality in affect-monitoring than in their reason-based assessments.

GENERAL DISCUSSION

The Forgotten Path

One cannot fully appreciate the psychology of evaluation without understanding the principles underlying feeling-based judgment. Some of these principles have begun to surface in recent studies on affect as information. However, previous research on affect as information has not examined the properties of integral feelings, leaving the operation of these feelings in judgment unclear. Examining responses to moderately complex and consciously accessible stimuli, this research has identified three important properties of consciously monitored feelings compared to cold evaluative assessments.

Integral Feelings Are Monitored Rapidly. The conscious monitoring of feelings can be significantly faster than the cold, reason-based assessment of the stimulus's qualities. Although this property has been widely assumed, it had not been demonstrated under conditions that were meaningful for consumer research and much of daily life. That feelings can be monitored relatively rapidly, even in the context of moderately rich stimuli, was clear from the lower response latencies of subjects' feeling responses to the pictures in study 1. It was also reflected—albeit not as clearly—in the greater volatility of subjects' continuous feeling responses to the commercials in study 3.

As discussed earlier, this finding cannot be explained by postulating that there was a greater affinity between feelings and the nonverbal DTI or that excluding feelings was more difficult than excluding reason-based assessments, though that may be the case under more demanding conditions. Nor does it seem likely that subjects were faster in reporting their feelings because they felt less accountable for their feelings than for their reason-based assessments. Had subjects reporting their feelings felt less accountable than those reporting their reason-based assessments, the former should have felt free to provide idiosyncratic responses. Instead, subjects reporting their feelings were more consistent in their responses than those reporting their reason-based assessments. The greater speed of affect monitoring comes, we believe, from the wide applicability of virtually hardwired affective sensory-motor programs and acquired emotional schemata whose associative processes can be triggered very rapidly, making feeling responses quickly available for conscious interpretation.

We do not believe, however, that the conscious monitoring of feelings is always faster than other processes of judgment. Much depends on the nature of feelings being monitored and the judgmental criteria to which this monitoring is compared. Although the conscious monitoring of Type I and Type II feelings can be very fast, the monitoring of Type III feelings, which involve significant cognitive appraisal activity, is probably relatively slow (Cohen and Areni 1991; LeDoux 1996). Similarly, although truly integrative reason-based assessments seem relatively slow (see study 1), reason-based assessments based on very few judgmental considerations (e.g., a single criterion) can be very rapid (see Schmidt-Atzert 1988), especially if these are guided by a top-down process or compared to a preexisting evaluative reference point such as a prior attitude (Lingle and Ostrom 1979).

It may be argued that even if feelings are monitored rapidly, evaluations based on these feelings need not be faster than evaluations based on reason-based assessments. It could be that feelings require more time-consuming interpretation before they can be translated into a summary evaluation. Recent findings suggest otherwise. Siemer and Reisenzein (1998) found that time pressure makes people more likely to rely on their (incidental) mood when making overall evaluations. Similarly, in study 3, subjects who were explicitly instructed to report their cold assessments of the commercials tended to rely on their (integral) feelings when they had to provide their responses in a real-time manner but not when they could provide their assessments in a retrospective manner. These results show that time pressure increases people's reliance on both incidental and integral feelings in judgments that need not be based on such feelings. Had the translation of feelings into overall evaluations been slow (despite their greater speed of monitoring), their use would not increase under time pressure. Thus, at least under some conditions, summary evaluations based on feelings may indeed be faster than those based on reasons.

Integral Feelings Elicit Agreement. A less intuitive property of feelings is that people may agree more about how a target makes them feel than they do about their reason-based assessments of the target. The greater interpersonal consistency of feelings was observed for both (*a*) global feeling (pleasant-unpleasant) versus overall assessment (like-dislike) responses and (*b*) specific feeling ("Made me feel angry") versus specific assessment ("Is interesting") responses. This finding is noteworthy considering the richness and complexity of the stimuli used in this research (e.g., pictures of news events, TV commercials). It contradicts the widespread assumptions that affective judgments are inherently subjective and contextually labile—hence unreliable—and that reason-based judgments have greater "objectivity," making them a more dependable source of evaluative information. Again, we attribute the higher interpersonal consistency of feelings to the broad applicability and inherent stability of universal affective sensory-motor programs and culturally shared emotional schemata. By comparison, reason-based evaluation processes, because of their greater flexibility, offer greater room for idiosyncrasy (e.g., Epstein and Pacini 1999).

Our results help explain some recent findings in the decision literature. Studies have shown that people disagree widely with respect to the amount of damages that they award in personal injury cases (Kahneman, Schkade, and Sunstein 1998) and their willingness to contribute financially to environmental causes (Kahneman et al. 1993). On the

other hand, they agree much more in their affective judgments about how outrageous a personal injury case is or how upsetting an environmental problem is. Kahneman and his colleagues attribute the inconsistency of the awards and the volunteered contributions to the nature of the dollar scales, which are unbounded, compared with affect scales, which are bounded. We believe that this is more than a scaling phenomenon. Interpersonal inconsistency may characterize most types of reason-based evaluation judgments, regardless of the scales used, because reasons are inherently more idiosyncratic. Consistent with our explanation, Kahneman et al. (1993, p. 312) found that judgments of "personal satisfaction expected from making a voluntary contribution"—an assessment-type of judgment—also showed low interpersonal agreement despite being collected on a bounded scale.

It would be premature, of course, to conclude that feeling responses are always more consistent across individuals than are reason-based judgments. We would predict, for instance, that Type III feeling responses would be idiosyncratic and contextually labile. In fact, much of the evidence viewed as supporting the contextual malleability of affective responses (e.g., Medvec et al. 1995) has focused on attribution- and counterfactual-dependent (i.e., Type III) affective responses, rather than Type I and Type II affective responses. Even Type I and Type II feeling responses need not elicit greater interpersonal agreement than reason-based assessments if the latter are based on widely shared attitudes and stereotypes (e.g., "French wines are better than California wines") and normative criteria (e.g., expected value). It is also not clear that, in the context of aesthetic judgments (e.g., music, art, food), feeling responses would be more interpersonally consistent than reason-based assessments. Finally, recall that our findings pertain only to integral feelings toward the target, not to incidental (mood-related) feelings, which are bound to be contextually labile.

Integral Feelings Are Potent Predictors of Thoughts. Perhaps the most surprising property of the conscious monitoring of target-induced feelings is that these feelings are remarkable predictors of the number and valence of people's spontaneous thoughts about the target—much better predictors than people's cold assessments of the target. This property was observed across two types of stimuli (pictures and commercials), replicated with independent groups of respondents for each type of stimuli, and appeared to hold—at least in study 2—even when feeling- and evaluation-types of thoughts were considered separately. Thus, initial feelings toward the target may have judgmental value not just because they are relatively fast and consistent but also because they direct thinking toward motivationally relevant properties of the stimuli.

Feelings are often instantiated immediately upon exposure to a target. Once instantiated, these feelings then frame subsequent thought generation through the spontaneous priming of feeling-consistent cognitions and the controlled retrieval of knowledge that helps explain the initial feeling response. A review of neurophysiological evidence has led Damasio (1994) to a similar conclusion: "Somatic states, negative or positive, caused by the appearance of a given representation, operate not only as a marker for the value of what is represented, but also as a booster for continued working memory and attention" (p. 198). The remarkable ability of feelings to predict spontaneous thoughts helps explain why immediate judgments based on very brief exposure to other individuals' nonverbal cues can be highly predictive of judgments based on much more extensive information about these individuals (Ambady and Rosenthal 1993). Such nonverbal cues are likely to produce an initial feeling response that can trigger or frame subsequent thought about the target.

Thinking Back, Looking Forward

Any piece of research makes a certain number of compromises. First, this research focused on the properties of feelings as they are being monitored. It did not explicitly examine how these feelings are eventually translated into overall judgments and decisions. Fortunately, the latter issue has been receiving a substantial amount of attention in the affect-as-information literature—albeit mostly in the context of incidental feelings (e.g., Martin et al. 1997; Ottati and Isbell 1996; Raghunathan and Pham 1999). Our research thus complements the existing literature on affect as information. Still, much more research is needed about how feelings signal and guide problem recognition, planning, evaluation, and choice and how feelings are used to assess self-efficacy and regulate progress (or lack of) toward goal regions (e.g., Carver and Scheier 1990).

Second, although our findings are based on two somewhat different sets of stimuli, their generality should not be assumed. As we have noted, feeling and reason-based assessments might behave differently in the context of aesthetic stimuli, stimuli that are amenable to stereotypical judgments, and when feeling-based assessments remain ambiguous in the absence of further interpretation.

Third, in study 1 and study 3, respondents were instructed to focus on one type of response—either feelings or reason-based assessments—to the exclusion of the other. The intent was to examine the properties of feelings and reason-based assessments in their pure forms. However, to the extent that these two types of responses co-occur in everyday evaluations, our judgmental task was somewhat artificial. It would be useful to study feelings and reason-based responses in a context where subjects are not explicitly instructed to exclude the other type of response. It would also be useful to compare the judgmental responses of subjects making natural, unconstrained summary evaluations to those of subjects explicitly instructed to monitor their feelings or to record their reason-based assessments. One could then infer which judgment process—the exclusive monitoring of feelings or a cold, reason-based assessment—is more similar to the evaluation process people naturally follow in similar circumstances. Based on Wilson's work (e.g., Wilson and Schooler 1991; Wilson et al. 1993), one might speculate that people's natural (default) process of evaluation resembles more the monitoring of feelings than the cold, reason-based

assessment of the target. Recent evidence suggests a more elaborate hypothesis. Using a paramorphic design that closely resembles the one described above, Pham (1998, experiment 2) found that the explicit monitoring of feelings mimics decisions made under consummatory motives (whether to see a movie for the experience itself) but not decisions made under instrumental motives (whether to see a movie in order to write a term paper about it). On the other hand, reason-based assessments mimic decisions made under instrumental motives, but not decisions made under consummatory motives. It is our position that both feeling-monitoring and reason-based assessments intervene in natural processes of evaluation, with one type of process or the other being more prominent depending on the judgment to be made.

That either feelings or reason-based assessments may become more prominent in natural processes of evaluation raises important questions about when and how the monitoring of feelings becomes more prominent. Wyer et al. (1999) recently proposed that in natural processes of evaluation feelings are monitored and reasons are assessed in parallel. Whichever basis of evaluation yields a suitable response first wins the race. This proposition is interesting in light of the finding that people exhibit great flexibility in including or excluding their feelings, depending on their relevance (Gorn et al. 2001; Pham 1998; Raghunathan and Pham 1999). To account for this flexibility using a race model such as Wyer et al.'s (1999), one would need to posit that some relevance check is performed late in the evaluation process. That is, when monitored feelings win the race (which, presumably, occurs frequently), they are then assessed for their relevance. If the feelings are relevant, they are finally incorporated into the overall evaluation. If they are not relevant, the overall evaluation is based on the runner-up input, a reason-based assessment. This might be termed a *late inclusion hypothesis*. Alternatively, feelings could be monitored and reasons assessed, not in parallel, but in a lexicographic fashion, with the order determined by the perceived relevance (or efficiency) of the two bases of evaluation. That is, the relevance check would be performed early in the evaluation process, before either evaluation subroutine (feeling monitoring or reason-based assessment) is launched. This would be an *early selection hypothesis*. These hypotheses merit empirical testing. Not only are they interesting per se, they also speak to a fundamental question: Are feeling and reason-based assessments just alternative inputs within a single evaluative system (e.g., Wyer et al. 1999), or do they embody separate evaluation pathways drawing on largely independent systems (e.g., Epstein and Pacini 1999; Zajonc 1980)?

Over the past 50 years, the cognitive revolution's notions of automaticity, heuristics and biases, and bounded rationality have repeatedly challenged traditional conceptions of human reason. Yet, emotions and feelings—historically considered the root of irrationality—were largely ignored in the study of valuation and decision making. By the mid-1970s, research on classically conditioned affect, mood effects, and affect priming had shown that incidental affect could alter judgments both by coloring perceptions of the information and by influencing the extent and nature of its processing. Still, feelings were assigned a heuristic or peripheral role, which was thought to diminish as judgments became more consequential. Today, research on affect as information and affect as part of a self-regulatory motivational system is leading to a fundamental shift in our conception of the role of feelings in judgment and decision making. Our research helps explain why feeling-based information is likely to play a prominent role in the assessment of meaningful perceptual inputs and why it deserves a central place in research on evaluation and choice. Although many scholars may still believe in the ancient dichotomy between emotion and reason, in light of our findings, the reliance on feelings in judgments and decisions seems rather sensible.

[*Received April 1999. Revised January 2001. David Glen Mick served as editor, and Hans Baumgartner served as associate editor for this article.*]

REFERENCES

Ambady, Nalini and Robert Rosenthal (1993), "Half a Minute: Predicting Teacher Evaluations from Thin Slices of Nonverbal Behavior and Physical Attractiveness," *Journal of Personality and Social Psychology*, 64 (March), 431–441.

Anderson, Norman H. (1981), *Foundations of Information Integration Theory*, New York: Academic Press.

Bargh, John A., Shelly C. Chaiken, Paula Raymond, and Charles Hymes (1996), "The Automatic Evaluation Effect: Unconditional Attitude Activation with a Pronunciation Task," *Journal of Experimental Social Psychology*, 32 (January), 104–128.

Batra, Rajeev and Olli T. Athola (1990), "Measuring the Hedonic and Utilitarian Sources of Consumer Attitudes," *Marketing Letters*, 2 (2), 159–170.

Bettman, James R., Mary Frances Luce, and John W. Payne (1998), "Constructive Consumer Choice Processes," *Journal of Consumer Research*, 25 (December), 187–217.

Biernat, Monica and Melvin Manis (1994), "Shifting Standards and Stereotype-Based Judgments," *Journal of Personality and Social Psychology*, 66 (January), 5–20.

Blaney, Paul H. (1986), "Affect and Memory: A Review," *Psychological Bulletin*, 99 (March), 229–246.

Bornstein, Robert F. (1989), "Exposure and Affect: Overview and Meta-Analysis of Research, 1968–1987," *Psychological Bulletin*, 106 (September), 265–289.

Breckler, Steven J. and Elizabeth C. Wiggins (1989), "Affect versus Evaluation in the Structure of Attitudes," *Journal of Experimental Social Psychology*, 25 (May), 253–271.

Buck, Ross (1985), "Prime Theory: An Integrated View of Motivation and Emotion," *Psychological Review*, 92 (July), 389–413.

Cacioppo, John T. and Wendi L. Gardner (1999), "Emotion," in *Annual Review of Psychology*, Vol. 50, ed. Susan T. Fiske, Stanford, CA: Annual Reviews, 191–214.

Carlston, Donal E. and Eliot R. Smith (1996), "Principles of Mental Representation," in *Social Psychology: Handbook of Basic Principles*, ed. E. Tory Higgins and Arie W. Kruglanski, New York: Guilford, 184–210.

Carver, Charles S. and Michael F. Scheier (1990), "On the Origins of Positive and Negative Affect: A Control-Process View," *Psychological Review*, 97 (January), 19–35.

Clore, Gerald L. (1992), "Cognitive Phenomenology: Feelings and the Construction of Judgment," in *The Construction of Social Judgments*, ed. Leonard L. Martin and Abraham Tesser, Hillsdale, NJ: Erlbaum, 133–164.

Cohen, Joel B. (1982), "The Role of Affect in Categorization: Toward a Reconsideration of the Concept of Attitude," in *Advances in Consumer Research*, Vol. 9, ed. Andrew Mitchell, 94–100.

——— (1990), "Attitude, Affect, and Consumer Behavior," in *Affect and Social Behavior*, ed. Bert S. Moore and Alice M. Isen, Cambridge: Cambridge University Press, 152–206.

——— and Charles S. Areni (1991), "Affect and Consumer Behavior," in *Handbook of Consumer Behavior*, ed. Thomas S. Robertson and Harold H. Kassarjian, Englewood Cliffs, NJ: Prentice-Hall, 188–240.

Damasio, Antonio R. (1994), *Descartes' Error: Emotion, Reason, and the Human Brain*, New York: Putnam.

Derbaix, Christian and Lennart Sjoberg (1994), "Movie Stars in Space: A Comparison of Preference and Similarity Judgments," *International Journal of Research in Marketing*, 11 (June), 261–274.

Eagly, Alice H. and Shelly Chaiken (1993), *The Psychology of Attitudes*, New York: Harcourt Brace Jovanovich.

Edwards, Kari (1990), "The Interplay of Affect and Cognition in Attitude Formation and Change," *Journal of Personality and Social Psychology*, 59 (August), 202–216.

Epstein, Seymour and Rosemary Pacini (1999), "Some Basic Issues Regarding Dual-Process Theories from the Perspective of Cognitive-Experiential Self-Theory," in *Dual-Process Theories in Social Psychology*, ed. Shelly Chaiken and Yaacov Trope, New York: Guilford, 462–482.

Etcoff, Nancy (1999), *Survival of the Prettiest*, New York: Doubleday.

Fishbein, Martin and Icek Ajzen (1975), *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*, Reading, MA: Addison-Wesley.

Gorn, Gerald J., Marvin E. Goldberg, and Kunal Basu (1993), "Mood, Awareness, and Product Evaluation," *Journal of Consumer Psychology*, 2 (3), 237–256.

———, Michel Tuan Pham, and Leo Yatming Sin (2001), "When Arousal Influences Ad Evaluation and Valence Does Not (and Vice Versa)," *Journal of Consumer Psychology*, 11 (July), 43–55.

Hastak, Manoj and Jerry C. Olson (1989), "Assessing the Role of Brand-Related Cognitive Responses as Mediators of Communication Effects on Cognitive Structure," *Journal of Consumer Research*, 15 (March), 444–456.

Hermans, Dirk, Jan de Houwer, and Paul Eelen (1994), "The Affective Priming Effect: Automatic Activation of Evaluative Information in Memory," *Cognition and Emotion*, 8 (November), 515–533.

Holbrook, Morris B. and Rajeev Batra (1987), "Assessing the Role of Emotions as Mediators of Consumer Responses to Advertising," *Journal of Consumer Research*, 14 (December) 404–420.

Isen, Alice M., Thomas E. Shalker, Margaret Clark, and Lynn Karp (1978), "Affect, Accessibility of Material in Memory, and Behavior: A Cognitive Loop?" *Journal of Personality and Social Psychology*, 36 (January), 1–12.

Izard, Caroll E. (1972), *Patterns of Emotions: A New Analysis of Anxiety and Depression*, New York: Academic Press.

Kahneman, Daniel, Ilana Ritov, Karen E. Jacowitz, and Paul Grant (1993), "Stated Willingness to Pay for Public Goods: A Psychological Perspective," *Psychological Science*, 4 (September), 310–315.

———, David A. Schkade, and Cass R. Sunstein (1998), "Assessing Punitive Damages (With Notes on Cognition and Valuation in Law)," *Yale Law Journal*, 107 (May), 2071–2153.

Keltner, Dacher, Phoebe C. Ellsworth, and Kari Edwards (1993), "Beyond Simple Pessimism: Effects of Sadness and Anger on Social Perception," *Journal of Personality and Social Psychology*, 64 (May), 740–752.

Kruglanski, Arie W. (1989), "The Psychology of Being 'Right': The Problem of Accuracy in Social Perception and Cognition," *Psychological Bulletin*, 106 (November), 395–409.

Lazarus, Richard S. (1991), *Emotion and Adaptation*, New York: Oxford University Press.

LeDoux, Joseph E. (1996), *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*, New York: Simon & Schuster.

Lerner, Jennifer S. and Dacher Keltner (2000), "Beyond Valence: Toward a Model of Emotion-Specific Influences on Judgment and Choice," *Cognition and Emotion*, 14 (July), 473–493.

Levine, Stephen R., Robert S. Wyer, and Nobert Schwarz (1994), "Are You What You Feel? The Affective and Cognitive Determinants of Self-Judgments," *European Journal of Social Psychology*, 24 (January–February), 63–77.

Lingle, John H. and Thomas M. Ostrom (1979), "Retrieval Selectivity in Memory-Based Impression Judgments," *Journal of Personality and Social Psychology*, 37 (February), 180–194.

Linville, Patricia W. (1982), "The Complexity-Extremity Effect and Age-Based Stereotyping," *Journal of Personality and Social Psychology*, 42 (February), 193–211.

Mandler, George, Yoshio Nakamura, and Billie J. Van Zandt (1987), "Nonspecific Effects of Exposure on Stimuli That Cannot Be Recognized," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13 (October), 646–648.

Martin, Leonard L., Teresa Abend, Constantine Sedikides, and Jeffrey D. Green (1997), "How Would I Feel If . . .? Mood as Input to a Role Fulfillment Evaluation Process," *Journal of Personality and Social Psychology*, 73 (August), 242–253.

Medvec, Victoria Husted, Scott F. Madey, and Thomas Gilovich (1995), "When Less Is More: Counterfactual Thinking and Satisfaction among Olympic Medalists," *Journal of Personality and Social Psychology*, 69 (October), 603–610.

Metcalfe, Janet and Walter Mischel (1999), "A Hot-Cool-System Analysis of Delay of Gratification: Dynamics of Willpower," *Psychological Review*, 106 (January), 3–19.

Murphy, Sheila T. and Robert B. Zajonc (1993), "Affect, Cognition, and Awareness: Affective Priming with Suboptimal and Optimal Stimulus," *Journal of Personality and Social Psychology*, 64 (May), 723–739.

Ottati, Victor C. and Linda M. Isbell (1996), "Effects of Mood during Exposure to Target Information on Subsequently Reported Judgments: An On-Line Model of Misattribution and Correction," *Journal of Personality and Social Psychology*, 71 (July), 39–53.

Pham, Michel Tuan (1998), "Representativeness, Relevance, and the Use of Feelings in Decision Making," *Journal of Consumer Research*, 25 (September), 144–160.

Plutchik, Robert (1966), "Multiple Rating Scales for the Mea-

surement of Affective States," *Journal of Clinical Psychology*, 22 (July), 423–425.

——— (1980), *Emotion: A Psychoevolutionary Synthesis*, New York: Harper & Row.

Raghunathan, Rajagopal and Michel Tuan Pham (1999), "All Negative Moods Are Not Equal: Motivational Influences of Anxiety and Sadness on Decision Making," *Organizational Behavior and Human Decision Processes*, 79 (July), 56–77.

Schmidt-Atzert, Lothar (1988), "Affect and Cognition: On the Chronological Order of Stimulus Evaluation and Emotion," in *Affect, Cognition, and Social Behavior*, ed. Klaus Fiedler and Joseph Forgas, Toronto: Hogrefe, 153–164.

Schwarz, Norbert (1990), "Feelings as Information: Informational and Motivational Functions of Affective States," in *Handbook of Motivation and Cognition*, Vol. 2, ed. E. Tory Higgins and Richard M. Sorrentino, New York: Guilford, 527–561.

——— (2000), "Social Judgment and Attitudes: Warmer, More Social, and Less Conscious," *European Journal of Social Psychology*, 30 (March–April), 149–176.

——— and Gerald L. Clore (1983), "Mood, Misattribution, and Judgments of Well-Being: Informative and Directive Functions of Affective States," *Journal of Personality and Social Psychology*, 45 (September), 513–523.

——— and Gerald L. Clore (1988), "How Do I Feel about It? Informative Functions of Affective States," in *Affect, Cognition, and Social Behavior*, ed. Klaus Fiedler and Joseph Forgas, Toronto: Hofgrefe International, 44–62.

——— and Gerald L. Clore (1996), "Feelings and Phenomenal Experiences," in *Social Psychology: Handbook of Basic Principles*, ed. E. Tory Higgins and Arie W. Kruglanski, New York: Guilford, 433–465.

Siemer, Matthias and Ranier Reisenzein (1998), "Effects of Mood on Evaluative Judgements: Influence of Reduced Processing Capacity and Mood Salience," *Cognition and Emotion*, 12 (November), 783–805.

Soldat, Alexander S., Robert C. Sinclair, and Melvin M. Mark (1997), "Color as an Environmental Processing Cue: External Affective Cues Can Directly Affect Processing Strategy without Affecting Mood," *Social Cognition*, 15 (Spring), 55–71.

Strack, Fritz (1992), "The Different Routes to Social Judgments: Experiential versus Informational Strategies," in *The Construction of Social Judgments*, ed. Leonard L. Martin and Abraham Tesser, Hillsdale, NJ: Erlbaum, 249–276.

Verplanken, Bas, Godelieve Hofstee, and Heidi J. W. Janssen (1998), "Accessibility of Affective versus Cognitive Components of Attitudes," *European Journal of Social Psychology*, 28 (January–February), 23–35.

Weiner, Bernard (1985), *An Attributional Theory of Motivation and Emotion*, New York: Springer.

Wilson, Timothy D., Douglas J. Lisle, Jonathan W. Schooler, Sara D. Hodges, et al. (1993), "Introspecting about Reasons Can Reduce Post-Choice Satisfaction," *Personality and Social Psychology Bulletin*, 19 (June), 331–339.

——— and Jonathan W. Schooler (1991), "Thinking Too Much: Introspection Can Reduce the Quality of Preferences and Decisions," *Journal of Personality and Social Psychology*, 60 (February), 181–192.

Winkielman, Piotr, Robert B. Zajonc, and Norbert Schwarz (1997), "Subliminal Affective Priming Resists Attributional Interventions," *Cognition and Emotion*, 11 (July), 433–465.

Wyer, Robert S., Gerald L. Clore, and Linda M. Isbell (1999), "Affect and Information Processing," in *Advances in Experimental Social Psychology*, ed. Mark P. Zanna, San Diego, CA: Academic Press.

Zajonc, Robert B. (1980), "Feeling and Thinking: Preferences Need No Inferences," *American Psychologist*, 35 (February), 151–175.

——— (1984), "On the Primacy of Affect," *American Psychologist*, 39 (February), 117–123.