HONESTY AND INTELLIGENCE JUDGMENTS OF INDIVIDUALS AND GROUPS: THE EFFECTS OF ENTITY-RELATED BEHAVIOR DIAGNOSTICITY AND IMPLICIT THEORIES

John J. Skowronski
Northern Illinois University

Two experiments examined trait judgments made from behaviors. Results from an initial experiment suggest that the informativeness of the behaviors, and not peoples' affective responses or approach/avoidance tendencies, best account for the impact of inconsistent information on trait judgments. The results of a second experiment yielded additional support for this idea by showing that when a target's behaviors were inconsistent in their trait implications: (1) negativity effects emerged in judgments of the target's honesty; (2) positivity effects emerged in judgments of the target's intelligence; (3) these effects diminished when the targets were groups rather than individuals; and (4) judgments made by entity theorists were less affected by inconsistent behavior than judgments made by incremental theorists. These effects did not emerge when a target's behaviors were consistent in their implications. Collectively, these results suggest that the usefulness of a behavior for a trait judgment depends on whether the judgment situation prompts participants to evaluate the extent to which a target fits a single trait category or forces a calculation of the target's relative fit to alternative trait categories. The results also suggest that some aspects of trait judgments cannot be explained by evaluative responses to behaviors or to targets.

Evaluations of an actor are more influenced by the actor's negative attributes than by their positive attributes (Kanouse & Hanson, 1972; Klein, 1991,

1996; Vonk, 1993; Vonk & Van Knippenberg, 1994; Wojciszke, Bazinska, & Jaworski, 1998; Wojciszke, Brycz, & Borkenau, 1993; Ybarra, 2001). It has been suggested that this negativity bias in evaluations occurs because an actor's negative attributes have stronger implications for a perceiver than the actor's positive attributes (Peeters, 1971, 1983; Vonk, 1999), or because the perceiver's negative affect system responds more intensely to stimuli than their positive affect system (Cacioppo, Gardner, & Berntsen,1997).

These affect–based theories have been applied to personality trait judgments (see Ito, Larsen, N.K. Smith, & Cacioppo, 1998; Peeters & Czapinski, 1990), but conceptual problems plague these applications. Traits have informational properties that may be independent of the evaluative implications of behaviors. For example, traits such as honesty may be strongly related to evaluations of people; other traits, such as intelligence, may not be strongly related to such evaluations. The selection of honesty and intelligence to illustrate this point is not accidental: The distinction between social traits and competence traits has a long history. Results from multidimensional scaling studies yielded two dimensions that supposedly provide the conceptual foundation of traits: An affectively–laden social dimension and an intellectual dimension (Rosenberg, Nelson, & Vivekananthan, 1968). Even earlier, Heider (1958) distinguished between social traits and competence traits, arguing that judgments about ability–based traits (those that involve the idea of "can," such as intelligence) might have different properties than judgments about social traits (those that involve the idea of "try," such as honesty).

Echoes of these ideas are found in an article by Reeder and Brewer (1979; also see Reeder, 1993), who proposed that expectations about behavior–trait relations for the honesty trait dimension differ from those for the intelligence trait dimension. They argued that perceivers believe that dishonest people can behave in both honest and dishonest ways, but honest people are thought to rarely perform dishonest behaviors. In comparison, beliefs about the relation between trait valence and behavior variability were thought to differ on the intelligent trait dimension: People who are smart are thought to behave in both smart and not–so–smart ways; people who are not–so–smart are thought to rarely perform smart behaviors.

THE DIAGNOSTICITY–BASED INFORMATIVENESS VIEW

Skowronski (1985) and Skowronski and Carlston (1987, 1989) translated these ideas into the language of categorization theory. They suggested

that trait judgment might sometimes be likened to a category decision task: When perceivers are confronted with contradictory behaviors, they must determine which of the two available trait categories is a better "fit" to the behavior set (for recent connectionist views, see Kashima & Kerekes, 1994; E.R. Smith & DeCoster, 1998).

Research shows that such categorical decisions are affected by the *category diagnosticity* of the cues that are used to make the categorization. Category diagnosticity refers to the extent to which a cue simultaneously predicts membership in one category and non–membership in the other. Cues that are frequently and exclusively associated with members of a category are highly diagnostic of category membership; cues that are infrequently or inexclusively associated with members of a category are less diagnostic of category membership. Because cues that are highly diagnostic are better predictors of category membership than less–diagnostic cues, highly diagnostic cues have more impact on categorical decisions.

Skowronski and Carlston argued that the patterns of behavior–trait relations discussed by Reeder and Brewer affect a behavior's diagnosticity for making trait decisions. The greater exclusivity of dishonest behaviors relative to honest behaviors render those dishonest behaviors more diagnostic for decisions about whether a person is honest or dishonest. The greater exclusivity of intelligent behaviors relative to unintelligent behaviors render those intelligent behaviors more diagnostic for decisions about whether a person is intelligent or unintelligent. These diagnosticity patterns suggest that when contradictory information needs to be reconciled in making a trait judgment, there should be a negativity bias in honesty judgments and a positivity bias in intelligence judgments. These predictions have often been confirmed (e.g., Betz, Gannon, & Skowronski, 1993; Gannon, Skowronski, & Betz, 1994; Hess, Bolstad, Woodburn, & Auman, 1999; Reeder, 1997; Reeder & Coovert, 1986; Singh & Teoh, 2000).

The definition of diagnosticity used by Skowronski and Carlston refers to one particular informational property of behavior: its ability to predict membership in one of two alterative trait categories. One corollary to this idea is that when an actor's attributes do not prompt a judge to weigh an actor's membership in alternative trait categories, but instead allow a focus on only a single trait category, diagnosticity ought to be irrelevant to judgment. This idea is central to Kahneman and Tversky's (1973) *representativeness heuristic*. For example, if a person is described by four unintelligent behaviors, perceivers judge the person

by comparing this description to their ideas about unintelligent people: Ideas about the characteristics of intelligent people need not come to mind. Results reported by Lupfer, Weeks, and Dupuis (2000) support this notion. In their research, negativity and positivity biases did not appear in judgments made from consistent behavior sets (also see Birnbaum, 1972; 1973; 1974). Furthermore, this notion is similar to the *positive test* or *pseudodiagnostic testing* strategies described in the hypothesis testing literature. In fact, some theorists have explicitly linked these hypothesis testing strategies to the use of representativeness in social judgments (see Trope & Liberman, 1996, p. 242).

The essence of the Skowronski and Carlston argument, then, is that negativity and positivity biases occur because the informativeness of behaviors for trait judgments varies. When an actor is described by either single behaviors or by multiple behaviors that all imply the same trait, the judgment is made via representativeness (calculating fit between a behavior and a single activated trait category). In comparison, when an actor is described by multiple and contradictory behaviors, the judgment is made by calculating the fit between the target description and each of two trait categories that are activated and by selecting the better fit (which is influenced by the diagnosticity of the behaviors in the description).

LINKING DIAGNOSTICITY TO THE ENTITY-RELATED: INDIVIDUALS VS. GROUPS

Negativity and positivity biases in trait judgments are smaller when people judge groups than when they judge individuals. Diagnosticity explains this result well. Groups are expected to show more behavioral variability than individuals (Coovert & Reeder, 1990; Welbourne, 1999). For example, a rater's judged likelihood that a member of an honest group would perform a dishonest action is greater than their judged likelihood that an honest person would perform a dishonest action. Because exclusivity contributes to diagnosticity, these perceived likelihoods logically cause the category diagnosticity of dishonest behaviors to be *lower* for trait judgments of groups relative to judgments of individuals.

However, these entity–related diagnosticity changes have not (to my knowledge) been directly measured. Furthermore, the diminished negativity and positivity biases in impressions of groups relative to impressions of individuals can potentially be explained by other theories. For example, Welbourne (1999) postulated that the behaviors of incon-

sistently–behaving individuals should be perceived as more inconsistent than the behaviors of inconsistently–behaving group members. This difference causes inconsistent behaviors to be processed differently for individuals and for groups. In Welbourne's view, this processing difference is the cause of the diminished negativity and positivity biases in judgments made about groups.

Another alternative explanation for diminished biases in judgments of groups can be derived by extending theories that emphasize the role of affect in trait judgment. The *dual–affect model* (see Ito et al., 1998) argues that negativity biases in honesty judgments occur because the negative affect system responds more strongly to stimuli than the positive affect system. However, because of the low affective content of behaviors on the intelligence dimension, this heightened negative response does not occur to behaviors on that dimension. Instead, a tendency to evaluate stimuli positively in the absence of negativity (the *positivity offset*) leads to a positivity bias in intelligence judgments. A slightly different affect–based view is *the risk–reward model* (Peeters 1971; 1983; also see Peeters & Czapinski, 1990, Vonk 1999). This model suggests that negativity biases occur in trait judgments because negative behaviors are especially threatening to a perceiver. Even hypothetical interaction with someone who performs dishonest behaviors would be perceived as risky, prompting avoidance responses (including negative evaluations) that are strong enough to overcome potential rewards derived from approaching someone who exhibits honest behavior. This heightened risk is not present for less social traits (such as intelligence), so a positivity bias reflecting an inherent tendency to seek rewards by approaching stimuli emerges in intelligence judgments.

If these affect–based explanations for biases in trait judgments account for the lowered biases in judgments of groups relative to judgments of individuals, it must be the case that affective responses to behaviors must be diminished when those behaviors are attributed to groups (or group members) than to individuals. Given the relatively automatic nature of peoples' evaluative responses to stimuli (Fazio, 2000), the idea that evaluative responses to behaviors will differ when the behavior is attributed to groups rather than to individuals seems implausible. Nonetheless, to my knowledge there has been no research directly addressing whether such entity–driven evaluative shifts occur, nor whether they can account for the differing valence biases observed in judgments of groups and individuals.

EXPERIMENT 1

Experiment 1 attempts to provide initial evidence that might discriminate between the diagnosticity–based informativeness and affect–based explanations for the negativity bias in honesty judgments, the positivity bias in intelligence judgments, and the diminution of those biases when judgments are made about groups rather than about individuals. In Experiment 1, four types of behaviors (intelligent, unintelligent, honest, dishonest) were attributed to an individual or to one of two group members (family member, social club member). The *usefulness* of each behavior for a trait decision about the person, family, or social club was measured, as was the *affective response* provoked by each behavior and *perceptions of risk or reward* associated with each behavior.

The diagnosticity–based informativeness model predicts that dishonest behaviors should be seen as more useful to decisions about whether an entity was honest or dishonest than honest behaviors, and intelligent behaviors should be seen as more useful to decisions about whether an entity was intelligent or unintelligent than unintelligent behaviors. The informativeness model also predicts that behaviors should be seen as less useful for judgments about groups than for judgments about individuals. The dual affect and risk–reward models also predict these effects, but suggest that identical patterns should emerge on other measures. The dual affect model predicts that utility judgments and affect judgments should show parallel patterns of results. The risk–reward model predicts that utility judgments and risk–reward judgments should show parallel patterns of results. In comparison, the informativeness model predicts that the effects observed on the affect or risk–reward ratings will not parallel those appearing in the utility ratings.

METHODS

Experiment 1a
Experiment 1 was conducted as three sub–experiments. In Experiment 1a participants were given a stimulus booklet containing two lists of behaviors. These behaviors had been pretested for use in previous research (Skowronski & Carlston,1992; also see Betz et al. 1993; Gannon et al., 1994). One list contained 20 honest behaviors and 20 dishonest behaviors. The second list contained 20 intelligent behaviors and 20 unintelligent behaviors.

Diagnosticity has previously been measured by collecting an array of probability judgments used to calculate the *cue validity* of a behavior for a trait decision (see Skowronski & Carlston, 1987). In those studies participants were forced to think about the probability that a person with a trait would perform a behavior that was not consistent with the trait (e.g., how likely is it that an honest person would cheat at poker?). Experiment 1a used a method that similarly induced participants to think about the implications of behaviors for alternative traits. Participants were told that people sometimes behaved inconsistently, and that the experimenters were interested in participants' perceptions of how much each of the behaviors would help them to make a trait decision in the face of that inconsistency. Participants were told that they would rate how useful behaviors would be for the purpose of making choices about the personality traits possessed by people or groups. However, rather than consider behaviors in isolation, participants were to always assume that the target had performed other behaviors, and that some of these other behaviors were inconsistent with the one that they were reading. Participants were asked to think about how they would make a trait judgment using the target behavior in conjunction with these other, often–inconsistent behaviors. These instructions were intended to induce participants to avoid using the representativeness heuristic to think about the behavior, and instead to think about the implications of the behavior for possible membership in both relevant trait categories.

Participants were told that sometimes they would be making decisions between the traits of honesty and dishonesty, and that sometimes they would be making decisions between the traits of intelligence and unintelligence. Behaviors related to the honesty trait dimension were presented in one list and behaviors related to the intelligence dimension were presented in a second list. Whether these judgments were made for the behaviors on the intelligence list first or for the behaviors on the honesty list first was counterbalanced across participants.

Three entities were the targets of the judgments: individuals, families, and social clubs (for other research using such groups, see Hamilton & S.J. Sherman 1996; Lickel, Hamilton, & S.J. Sherman, 2000; McConnell, S.J. Sherman,& Hamilton, 1997). The use of two different groups ensures that any differences between individuals and groups were not caused by idiosyncratic characteristics of a particular group. After reading a behavior and imagining that the behavior was presented in the context of other behaviors, participants rated the usefulness of each behavior to
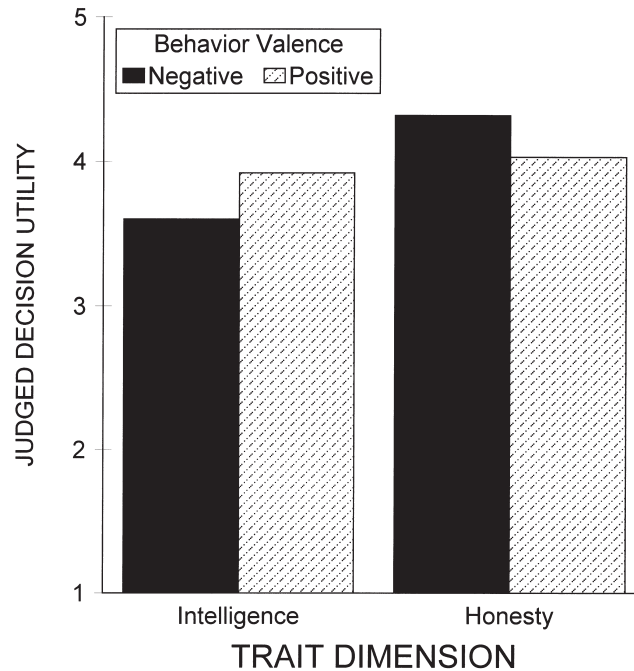
FIGURE 1. Experiment 1: The interaction between trait dimension and behavior valence on the judged utility of behaviors for trait category decisions in a behaviorally-inconsistent environment.

their decision about which of the two traits (honest vs. dishonest, intelligent vs. unintelligent) the individual, family, or social club was likely to have.[1] Participants' judgments were made on a 7–point scale anchored by *not at all useful* at the low end and *very useful* at the high end.

Experiment 1b
The design of this experiment was similar to that used in Experiment 1a, but the entity rated variable was between–participants rather than within–participants. There were three groups of participants and each of them received a different booklet. One booklet described the behaviors of individuals, the second described the behaviors of family members, and the third described the behaviors of social club members. After

---

1. Weaker, but similar results emerged from a study in which entity judged variable was a between–participants variable.

reading each behavior, participants were asked to report their evaluative reaction to the behavior. This rating was made on a 9–point bipolar scale anchored by *very bad* at the low end and *very good* at the high end. The composition of the booklets was otherwise identical to those used in Experiment 1a.

Experiment 1c

The design of, and the materials used in, this experiment were almost identical to those used in Experiment 1b. However, in Experiment 1c participants provided a rating of how potentially risky or potentially rewarding it would be to interact with the entity (individual, family, social club) described by the behavior. This rating was made on a 9–point bipolar scale anchored by *very risky* at the low end and *very rewarding* at the high end.

Participants

The participants were undergraduates in an introductory psychology course. They received credit toward fulfillment of a course research requirement as compensation for their participation. There were 92 participants in Experiment 1a, 142 participants in Experiment 1b, and 126 participants in Experiment 1c.

RESULTS

Experiment 1a

Twelve means were created from each participant's responses. These means represent all combinations of item type (honest, dishonest, intelligent or unintelligent) and target type (families social clubs, and individuals). These twelve means were entered into an Entity Judged (individual, family, social club) $\times$ Trait Dimension (honesty, intelligence) $\times$ Behavior Valence (positive, negative) repeated–measures ANOVA.

The means in Figure 1 show that dishonest behaviors were judged to be more useful than honest behaviors in the decision about whether an entity was honest or dishonest, but intelligent behaviors were judged to be more useful than unintelligent behaviors to the decision about whether an entity was intelligent or unintelligent (Trait Dimension $\times$ Behavior Valence interaction: $F(1, 1001) = 22.06, p < .0001$). This interaction is similar in form to the interaction for the cue validity measure reported by Skowronski and Carlston (1987), suggesting that both measures similarly assess the diagnosticity of behaviors for trait category decisions. In

addition, the judged decision utility of the behaviors was higher for decisions about individuals (M = 5.48) than for decisions about either families (M = 3.37) or social clubs (M = 3.03), F (2, 1001) = 600.60, p < .0001.[2] Preplanned comparisons revealed that the mean in the individual condition was significantly higher than the mean in either of the two group conditions.

Experiment 1b

Four means reflecting responses to each item type (honest, dishonest, intelligent, or unintelligent) were calculated from each participant's evaluative responses to the behaviors. These four means were entered into an Entity Judged (individual, family, social club) × Trait Dimension (honesty, intelligence) × Behavior Valence (positive, negative) mixed ANOVA with repeated–measures on the latter two variables.

Evaluative responses were more extreme to dishonest (M = –2.91) than to honest (M = 2.67) behaviors and to intelligent (M = 2.24) than to unintelligent (M = –1.74) behaviors (reflecting a Trait Dimension × Behavior Valence interaction: F (1, 139) = 121.59, p < .0001, as well as trait dimension and behavior valence effects: F (1, 139) = 53.45, p < .0001, F (1, 139) = 1650.58, p < .0001). As predicted by the dual–affect model, these data roughly parallel the utility judgments observed in Experiment 1a.

This parallelism was broken for the entity judged effect. Evaluative responses to behaviors on the honesty dimension were only slightly more extreme when those behaviors were enacted by individuals (M_{honest}= 2.77, $M_{dishonest}$= –2.93) than when the behaviors were enacted by family members (M_{honest}= 2.52, $M_{dishonest}$= –2.88) or social club members (M_{honest}=

---

2. Two other unanticipated effects emerged from the analysis. The means for the Behavior Valence × Entity Rated interaction, F (2, 1001) = 7.07, p < .001, show that negative behaviors (M = 3.52) were rated as more useful than positive behaviors (M = 3.23) for trait decisions about families, but that positive behaviors were rated as more useful than negative behaviors for trait decisions about social clubs ($M_{positive}$= 3.08, $M_{negative}$= 2.99) and individuals ($M_{positive}$= 5.61, $M_{negative}$= 5.35). The theoretical importance of this effect is unclear given that these means average across the trait dimensions. The means for the unanticipated trait dimension effect, F (1, 1001) = 42.82, p < .0001, show that honest and dishonest behaviors were generally seen as more useful for judgments about membership in the honest or dishonest trait categories (M = 4.16) than intelligent and unintelligent behaviors were for judgments about membership in the intelligent or unintelligent trait categories (M = 3.76). This outcome may indicate that people are more willing to make trait category assignments from honesty–implicative behaviors than from intelligence–implicative behaviors. Alternatively, the outcome could simply reflect minor differences in the extremity of the items used in the experiment.

2.72, $M_{\text{dishonest}}$= −2.91). However, on the intelligence dimension the strongest evaluative responses were provoked by behaviors that were enacted social club members (social club $M_{\text{intelligent}}$= 2.52, $M_{\text{unintelligent}}$= −2.05; individual $M_{\text{intelligent}}$= 2.15, $M_{\text{unintelligent}}$= −1.72; family $M_{\text{intelligent}}$= 2.05, $M_{\text{unintelligent}}$=−1.47). These means reflect a significant Entity Judged × Trait Dimension × Behavior Valence interaction, F (2, 139) = 3.13, p < .05.

Experiment 1c
Four means reflecting responses to each item type (honest, dishonest, intelligent, or unintelligent) were calculated from each participant's perceptions of the potential risk or reward that was perceived in interacting with the entity who performed each behavior. These four means were entered into an Entity Judged (individual, family, social club) × Trait Dimension (honesty, intelligence) × Behavior Valence (positive, negative) mixed ANOVA with repeated–measures on the latter two variables.

There were more perceived risks in approaching dishonest entities (M = 2.53) than there were rewards for approaching honest entities (M = 2.21), and there were more perceived rewards for approaching intelligent entities (M = 2.15) than there were risks for approaching unintelligent entities (M = −1.40). These means reflect a significant Trait Dimension × Behavior Valence interaction, F (1, 123) = 44.59, p < .0001, as well as behavior valence (F (1, 123) = 1158.71, p < .0001) and trait dimension (F (1, 123) = 65.09, p < .0001) effects. As predicted by the risk–reward model, these data roughly parallel the utility judgments observed in Experiment 1a. However, this parallelism was broken for the entity judged effect. Regardless of trait dimension, the perceived risks and rewards associated with the behaviors did not vary across individuals, families, and social clubs (entity rated effect: F (2, 123) = .29, p > .75; Entity Rated × Behavior Valence interaction: F (2, 123) = 1.60, p > .20).

DISCUSSION

The results of Experiment 1 showed that: (1) dishonest behaviors were perceived to be more useful than honest behaviors for honest vs. dishonest trait decisions; (2) intelligent behaviors were perceived to be more useful than unintelligent behaviors for intelligent vs. unintelligent trait decisions, and (3) the perceived usefulness of behaviors was higher for decisions about an individual's traits than about a group's traits. However, the reduction in the perceived utility of behaviors for

judgments about groups was not paralleled by a similar reduction in evaluative responses to those behaviors. Similarly, the entity judged variable was unrelated to the risk–reward ratings. Hence, the data are most consistent with the notion that, when behaviors contradict each other, it is the informativeness of behaviors, and not evaluative responses to those behaviors, that contribute to the impact of those behaviors on judgments.

However, some might claim that the data in Experiment 1 are misleading with respect to the dual–affect position. This claim is derived from the fact that participants were asked to report their evaluative reaction to each behavior, and not to evaluate the entity who performed the behavior. In this regard, it should be noted that Ito et al. (1998) claim that the negativity bias occurs early in processing, when a stimulus is first categorized as positive or negative, and not during later processing stages, such as during response production. Because the evaluation of a target might involve these later–stage processes, assessment of evaluative reactions to the behavior seemed to offer an appropriate test of the dual–affect model. Nonetheless, it is reasonable to suggest that evaluation of the target might better parallel the impact of a behavior on trait judgments. This possibility is examined in Experiment 2. Experiment 2 also offered the opportunity to examine Welbourne's (1998) claim that the reduction in positivity and negativity effects in trait judgments for groups is caused by the fact that people are less likely to try to reconcile incongruity in groups than in individuals.

EXPERIMENT 2

Goals and Overview

Experiment 2 was designed to show that the information value of behaviors was the primary determinant of the impact of those behaviors on trait judgment and to illustrate some of the situational and personal factors that affect that information value. The experiment explored the impact of trait–implicative behaviors on trait judgments of individuals and groups when behaviors were presented in sets rather than as single units. It was expected from diagnosticity theory that trait judgments of targets who behaved inconsistently would show evidence of a negativity bias in judgments of honesty, a positivity bias in judgments of intelligence, and a diminution of these biases in judgments made about groups relative to judgments made about individuals. Whether these effects could plausibly be explained by peoples' affective responses to the

targets or to their perceptions of the extent to which behaviors were inconsistent was examined. Trait judgments made about targets who behaved consistently were also examined: Diagnosticity theory suggests that the negativity bias in judgments of honesty, the positivity bias in judgments of intelligence, and the diminution of these biases in judgments made about groups relative to judgments made about individuals ought to be limited to cases in which the target behaves inconsistently. Finally, it might be expected that individual differences in the perceived information value of behaviors for trait judgments ought to affect the impact of those behaviors on trait judgments, and Experiment 2 explored this possibility.

Examining the Role of Evaluation in Trait Judgments

In addition to trait judgments, in Experiment 2 participants' evaluative judgments of the targets were obtained. If evaluative mechanisms are responsible for the effects observed in the trait judgments, then the patterns of data observed for the evaluative judgments ought to parallel the patterns observed in the trait judgments. Moreover, a mediational analysis should show that the evaluative judgments mediate the trait judgments. From a diagnosticity–based informativeness perspective, no such mediation should occur.

Examining the Role of Perceptions of
Behavior Inconsistency in Trait Judgments

Welbourne's (1999) argument that lowered perceptions of inconsistency among behaviors play a causal role in the smaller judgment biases evinced in judgments of groups was examined by assessing perceptions of inconsistency among the behaviors describing a target. If inconsistency resolution mechanisms are responsible for the effects observed in the trait judgments, then the patterns of data observed for the inconsistency judgments ought to parallel the patterns observed in the trait judgments. Moreover, a mediational analysis should show that the inconsistency judgments mediate the trait judgments. From a diagnosticity–based informativeness perspective, no such mediation should occur.

Examining Effect Generality: Trait Judgments
From Consistent Behavior Sets

The generality of negativity and positivity biases in trait judgments was also examined in Experiment 2. A number of theoretical mechanisms

(see Baumeister, Bratslavsky, Finkenauer,& Vohs, 2001; Rozin & Royzman, 2001) lead to the expectation that judgments made from negative behaviors should be more extreme than judgments made from positive behaviors, even when the behaviors in a set are all internally consistent. A different prediction is derived from Experiment 1's finding that more extreme evaluations are produced by dishonest than honest behaviors, but that stronger evaluations are produced by intelligent than unintelligent behaviors. If affect is the basis for trait judgments, there should be a negativity bias in honesty judgments and a positivity bias in intelligence judgments, even for these consistent behavior sets.

However, Skowronski and Carlston's (1989) reasoning about how trait judgments are made implies a third pattern. Recall that Skowronski and Carlston assume that the category diagnosticity of a behavior should affect judgments only when behaviors suggest possible membership in either of two trait categories. When actor descriptions contain behaviors that are all consistent with the same trait, diagnosticity should be irrelevant. Hence, in the consistent–behavior conditions, the extremity of judgments based on items of one valence, regardless of trait dimension, should be equal to the extremity of judgments based on the same number of items of the opposite valence.

An article by Lupfer et al. (2000) already reported such a result. However, the present article also explores whether there are differences in the extremity of judgments made about individuals and of groups for these consistent item sets. If such judgments are made using the representativeness heuristic, all participants should simply fit the behavior set to their cognitive representation of the matching trait category (e.g., how the set of honest behaviors map onto their idea of an "honest" person or an "honest" group). These trait representations should be quite similar for individual targets and for group targets. Hence, the fit of the behaviors to these cognitive representations should be much the same, regardless of whether the entity rated is a group or an individual, and should yield roughly equivalent judgments. Note that this differs from the prediction derived from the idea that behaviors are simply "less informative" for judgments of groups than of individuals. This general informativeness notion suggests that when described by sets of consistent behaviors, judgments of groups should be less extreme than judgments of individuals.

Implicit Theories and Trait Judgments
Experiment 2 also examines how behavior informativeness might be affected by implicit theories of personality stability. A series of recent arti-

cles (Dweck, Chiu, & Hong, 1995; for a review, see Levy, Plaks, Hong, Chiu, & Dweck, 2001) have explored judgments made by entity theorists (who believe that personalities are stable) and incremental theorists (who believe that personalities are changeable). Research shows that when a target's behaviors are inconsistent, the judgments of entity theorists will be more influenced by the general trend of a target's behaviors, while the judgments of incremental theorists will be more influenced by rare inconsistent actions (see Plaks, Stroessner; Dweck, & J.W. Sherman, 2001).

This result might be viewed in terms of the perceived informativeness of behaviors. Compared to incremental theorists, entity theorists may believe that an inconsistent behavior provides relatively little information about the entity's personality. Hence, the effect of theorist type on trait judgment might be traced to informativeness effects similar to those responsible for the differing judgments made about groups and individuals. In contrast, theories that might try to account for such effects in terms of affect must assume that incremental and entity theorists have different affective reactions to the events.

Furthermore, as with diagnosticity–driven informativeness effects, the theorist type effect may be limited to circumstances in which the behaviors in a set are inconsistent. When behaviors are consistent, all perceivers should rely on the representativeness heuristic. Hence, in such circumstances the judgments of entity theorists and incremental theorists should be similar. This prediction might seem to be inconsistent with the results of studies reporting that the judgments of entity theorists tend to be more extreme than the judgments of incremental theorists (Levy, Stroessner & Dweck, 1998), but this inconsistency may be illusory. The studies that obtain this extremity effect typically present judges with a mix of trait–implicative and trait–neutral items (but see Chiu, Hong, & Dweck, 1997). Because the inclusion of neutral items tends to weaken the implications of the trait–implicative items (a phenomenon called the dilution effect; Nisbett, Zukier, & Lemley, 1981), the differential extremity of the judgments rendered by incremental and entity theorists might be the result of the tendency to use (or, in the case of entity theorists, to not use) any behavior that does not fit with a trait—even a neutral one. It is an open question as to whether such effects emerge when the items in a set consistently imply a single trait category.

Finally, theorist type research has focused on how implicit theories affect judgments made about individuals. Such theories should also apply

to judgments of groups. To explore this possibility, the individual–focused scale used to explore implicit theories (see Levy & Dweck, 1997) was modified for some participants so that it measured implicit theories about a specific group. This modification allowed relations between theorist type and trait judgments to be examined for each entity (individual, family, social club) used in the experiment.

METHODS

Participants
Students enrolled in an introductory psychology class participated in the experiment. The 130 participants received credit toward completion of a course research option as compensation. All participants were randomly assigned to between–participant conditions.

Implicit Theory Measure
The existing implicit theory questionnaire (e.g., Dweck & Levy, 1997) was used unaltered in the individual target condition. This measure consists of eight statements, such as "People can change even their most basic qualities." Participants' responses reflect their agreement with each statement. The questionnaire was altered to reflect the group entity conditions. Participants in the family condition responded to items such as "Families can change even their most basic qualities," while those in the social club condition responded to items such as "Social clubs can change even their most basic qualities." After appropriate reverse scoring and averaging, separate median splits were conducted on scores within each condition. These were used to identify each participant as an entity theorist or as an incremental theorist. Though median splits were conducted separately for each group of participants, the median value was between 23 and 24 for all three groups.

Procedure
Participants were told that the experiment was concerned with making judgments about others' characteristics based on a small sample of their behaviors. Participants then received a stimulus booklet that described behaviors and made judgments from those behaviors. The target of the judgment differed (individual, family, social club) for different groups of participants. The behaviors were identical for all three entities, but in the family and social club conditions each behavior was described as coming from a different group member.

After rating all 16 targets on several different response scales, participants completed a set of individual difference measures. Among these measures was a version of the implicit theorist type scale. Participants who made judgments about individuals completed the individual–focused version of the measure, participants who made judgments about families completed the family–focused version of the measure, and participants who made judgments about social clubs completed the social club–focused version of the measure. After completing all the individual difference measures, the participants were debriefed and dismissed.

Stimulus Behavior Sets
Eight behavior sets had implications for a target's honesty or dishonesty. The other eight had implications for a target's intelligence or unintelligence. The trait implications of these behaviors were known from pretests conducted for other experiments. The behaviors were selected so that the average trait implications of the positive and negative behaviors were about equally extreme for the behaviors on each trait dimension. Examples of these behaviors are: *Stole a copy of old exams from the professor's office* (dishonest); *Found an expensive Swiss watch in a rental car and tracked down the owner to return it* (honest); *Got Fs on most final exams in high school* (unintelligent); *Always wins when playing Trivial Pursuit* (intelligent).

These honesty–implying and intelligence–implying behavior sets were intermingled in the stimulus booklet. Four of the eight sets that pertained to each trait dimension contained behaviors that all had implications for one trait (e.g., all 5 were honest, all 5 were dishonest, all 5 were intelligent, or all 5 were unintelligent). The other four sets that pertained to each trait dimension contained four behaviors that implied one trait and a single inconsistent behavior that implied the opposite trait (e.g., 4 honest and 1 dishonest behavior; 4 dishonest and 1 honest behavior; 4 intelligent and 1 unintelligent behavior; 4 unintelligent and 1 unintelligent behavior). Placement of the inconsistent behavior (first or last) was counterbalanced across trials.

Dependent Measures
Participants rated the likability, goodness, intelligence, honesty, and predictability of the entity described by each set of behaviors. Participants also rated the extent to which the behaviors in a set were consistent or inconsistent with each other. Participants' responses were made on 9–point scales with –4 at the low end and +4 at the high end. A negative descriptor anchored the low end of each scale (e.g., *very bad*, *very dishon-*
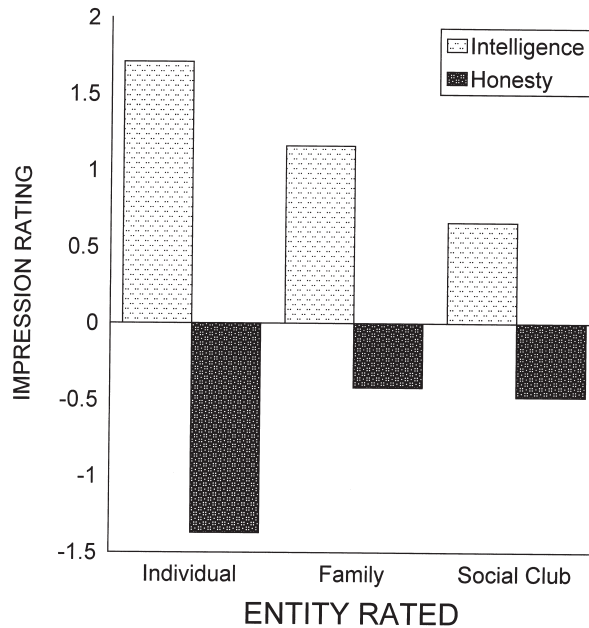
FIGURE 2. Experiment 2: Negativity effects in honesty judgments and positivity effects in intelligence judgments are reduced when groups are judged.

est) and a positive descriptor anchored the high end (e.g., *very good, very honest*). The midpoint of each scale was also labeled (e.g., *neither bad nor good*; *neither dishonest nor honest*).

RESULTS

Trait Ratings

   Inconsistent Behavior Sets. The honesty trait ratings for targets described by the honesty–implying inconsistent behavior sets and the intelligence ratings for targets described by the intelligence–implying inconsistent behavior sets were examined for evidence of negativity biases in honesty judgments, positivity biases in intelligence judgments, and a diminution of these biases when judgments were made about groups rather than about individuals. An additional interest was whether the judgments of incremental theorists would be more affected by the inconsistent behavior in the set than the judgments of entity theorists. These possibilities were evaluated by entering the trait ratings into
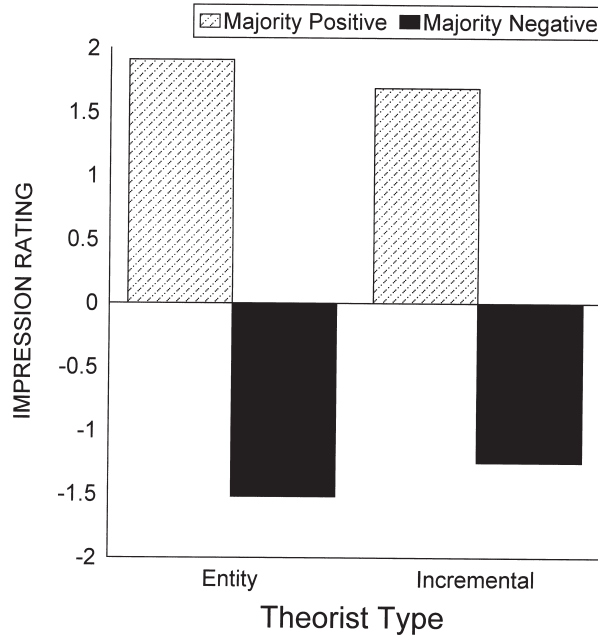
FIGURE 3. Experiment 2: The trait judgments of incremental theorists are more influenced by the implications of the inconsistent behavior than the judgments of entity theorists.

a mixed ANOVA. The variables in the ANOVA were trait dimension (honesty, dishonesty), valence of the behavior majority (positive, negative), placement of the inconsistent behavior (first, last), target judged (individual, family, social club), and theorist type (entity theorist, incremental theorist). The former three variables were within–participant variables; the latter two were between–participant variables.[3]

Although judgments made from descriptions containing a majority of positive behaviors were more positive ($M = 1.81$) than judgments derived from descriptions containing a majority of negative behaviors (M

---

3. Reviewers suggested that a number of results should be omitted from this article. Accordingly, I do not report effects involving the position of the deviant behavior variable, nor do I report analyses of the predictability ratings. Furthermore, because the same trait rating scales were used for all behavior sets, ratings of intelligence were obtained for sets that were intended to have implications for a target's honesty and ratings of honesty were obtained for sets that were intended to have implications for a target's intelligence. Results for these "crossed" trait ratings are not presented.
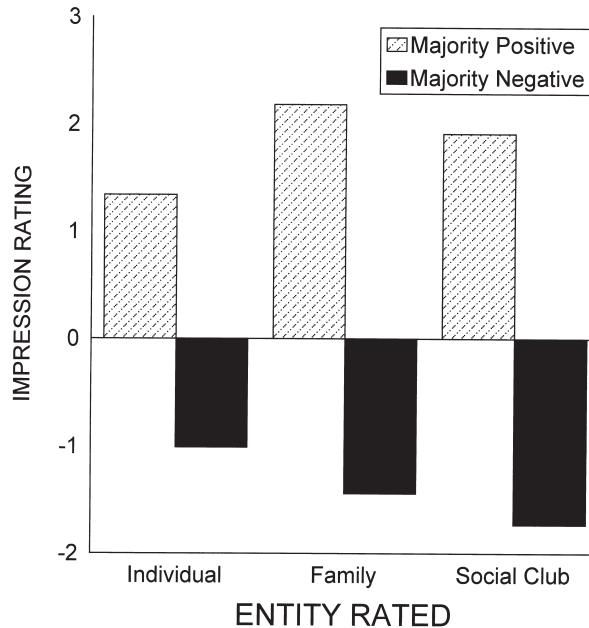
FIGURE 4. Experiment 2: The trait judgments of individuals are more affected by the implications of the inconsistent behavior than the trait judgments of groups.

$= -1.39$), F (1, 121) = 573.79, $p < .0001$, there was also a negativity bias in judgments of honesty ($M = -.76$) and a positivity bias in judgments of intelligence ($M = 1.18$), F (1, 121) = 200.87, $p < .0001$. These biases were smaller when judgments were made about groups than when judgments were made about individuals (see Figure 2), F (2, 121) = 18.18, $p < .0001$. In addition, the judgments of entity theorists were more likely to be influenced by the valence of the majority of the behaviors in the set than were the judgments of incremental theorists (see Figure 3), F (1, 121) = 4.13, $p < .05$.[4] Finally, (see Figure 4), judgments about individuals were

_____

4. Although easier to present, median split analyses may be less desirable than analyses that use the full range of a predictor (Bissonnette, Ickes, Bernstein, & Knowles, 1990). Analyses treating implicit theorist type as a continuous variable yielded results that did not alter the conclusions derived from the median split analyses. Another alternative analytic method discards data from participants with middling scores (between 3 and 4) on the theorist type measure. Analyses using this technique strengthened the theorist type effects that were significant in the median split analysis but did not otherwise affect the results for the theorist type variable.

more affected by the inconsistent item in the set than were judgments about groups (Valence of the Behavior Majority × Target Judged interaction: $F(2, 121) = 10.50$, $p < .0001$). These latter four outcomes can all be accounted for by the idea that the informativeness of a behavior affects the impact of that behavior on judgments.

Consistent Behavior Sets. It was expected that the effects observed in the ratings of targets described by inconsistent behavior sets would not emerge in the ratings of the targets described by consistent behavior sets. To examine this possibility, four averages were calculated reflecting the four behavior types (honest, dishonest, intelligent, unintelligent). These four averages were entered into a mixed ANOVA. The variables in the ANOVA were trait dimension (honesty, dishonesty), valence of the behavior set (positive, negative), target judged (individual, family, social club), and theorist type (entity theorist, incremental theorist). The former two variables in the list were within–participant variables; the latter two were between–participant variables.

The data convincingly show that the effects that were observed in the trait ratings made about inconsistently–behaving targets did not emerge in the trait ratings made about consistently–behaving targets. Neither a negativity bias in honesty judgments not a positivity bias in intelligence judgments emerged in ratings of these consistently–behaving targets. Moreover, examination of the data indicated that judgments of groups based on these consistent behavior sets were more extreme, not less extreme, than judgments of individuals made from the same sets, although the Group Judged × Valence of Behavior Set Interaction was not significant, $F(2, 123) = 1.82$, $p > .17$.[5] Thus, these data support the proposition that diagnosticity is irrelevant when all the behaviors in a set implicate the same trait.

The data also have implications for the simple theoretical position that "Bad is stronger than good" (e.g., Baumeister et al., 2001). This position suggests that judgments derived from negative behaviors should be more extreme than judgments derived from positive behaviors. A subsidiary analysis in which the data for the negative events were reverse–coded indicated that bad was not stronger than good in the trait ratings made about the consistently–behaving targets. There were no

---

5. The degrees of freedom that are reported vary slightly because of occasional missing data or uncodable responses. All the responses that could be used were used in each of the analyses. Hence, the number of cases that were entered into each analysis varied slightly.

significant differences in the extremity of the trait ratings given to consistently–behaving negative targets and consistently–behaving positive targets, $F(1, 123) = .79$, $p > .38$. This pattern did not significantly vary by trait dimension, $F(1, 123) = 1.39$, $p > .25$, showing that sets of negative and sets of positive behaviors produced equally extreme ratings on the honesty and intelligence dimensions.[6]

Target Evaluations

An evaluative index was constructed by averaging the likability and goodness ratings given to each target. Averaging was justified by high correlations between these ratings, regardless of whether they came from targets described by the inconsistent behavior sets or from targets described by the consistent behavior sets (both rs = .71, ps < .0001). Of particular interest was whether the patterns of data in these evaluative ratings paralleled those observed in the trait judgments. Such similarity would be consistent with the idea that evaluative judgments mediate trait judgments. To examine this possibility, two ANOVAs were conducted on the evaluative ratings. One ANOVA examined ratings given to targets who behaved inconsistently, and the second examined ratings given to targets who behaved consistently. The variables used as predictors in these analyses were identical to the variables used in the analyses of the trait ratings that were reported earlier in this article.

Inconsistent Behavior Sets. The analysis of the inconsistent behavior sets revealed a positivity bias in evaluations of targets who performed intelligence–relevant behaviors ($M = 1.02$) and a negativity bias in evaluations of targets who performed honesty–relevant behaviors ($M = -.62$), $F(1, 121) = 177.52$, $p < .0001$. Another result paralleling that obtained in the trait judgments shows that evaluations of groups tended to be less influenced by a deviant behavior than evaluations of individuals (majority positive: individual $M = .39$, family $M = 1.51$, social club $M = 1.18$; majority negative: individual $M = -.56$, family $M = -.56$, social club M = -.80;

_____

6. The only other significant effect in the analysis of the trait judgments made from the consistent behavior sets indicated that intelligence judgments were more positive ($M = .18$) than honesty judgments ($M = -.19$), $F(1, 123) = 25.47$, $p < .0001$. While this result might be consistent with the presence of a positivity bias in intelligence judgments and a negativity bias in honesty judgments, it may also result from minor differences in the extremity of the behaviors used to make up the behavior sets on the intelligence and honesty dimensions.

Valence of the Behavior Majority × Target Judged interaction: F (2, 121) = 15.23, $p < .0001$).[7]

However, other results demonstrate non–parallelism between the evaluations and the trait judgments. For example, one effect that was not significant in the trait judgments (F (1, 121) = .07, $p > .78$) was the interaction between trait dimension and the valence of the majority of behaviors in a set, which was significant for the evaluations (F (1, 121) = 60.94, $p < .0001$). The means for this effect show that the difference in liking for targets described by mostly positive behavior sets and for targets described by mostly negative behavior sets was much larger for targets described in terms of honesty–relevant behaviors (Ms = .54 and –1.79) than for targets described in terms of intelligence–relevant behaviors (Ms = 1.52 and .52). This outcome makes sense in terms of Rosenberg et al.'s (1968) contention that honesty is a trait that has strong evaluative implications, while intelligence is a trait with weaker evaluative implications. Furthermore, it should be noted that targets described by sets of mostly unintelligent behaviors were evaluated positively (M = .52) despite the fact that the intelligence rating of these targets was negative (M = –.41). This directional disparity in the two judgment types strongly argues against the idea that the intelligence judgments are based on evaluative reactions to the events.

Another interaction that emerged in the evaluations of the inconsistently–behaving targets, but not in the trait ratings, involved trait dimension, valence of the majority of behaviors in a set, and target judged, F (2, 121) = 3.75, p < .03. The means for this interaction reveal that the differences in evaluations made from mostly intelligent behavior sets and evaluations made from mostly unintelligent behavior sets were relatively consistent across the different targets (individual Ms = 1.52 and .52; family Ms = 1.95 and .7; and social clubs Ms = 1.41 and .25), but the differences in the evaluations made from mostly honest behavior sets and the evaluations made from mostly dishonest behavior sets were smaller for individuals than for groups (individual

---

7. The results also yielded a target judged main effect, F (2, 121) = 6.63, $p < .002$, indicating that judgments of families and social clubs are more positive than judgments of individuals. This result is obviously meaningless in light of the interaction. In comparison, the fact that evaluations were more positive when those evaluations were derived from behavior sets that were mostly positive than when the behavior sets contained mostly negative behaviors is interpretable, F (1, 121) = 305.46, $p < .0001$, but relatively trivial.

Ms = –.43 and –1.72, family Ms = 1.08 and –1.82, social club Ms = .96 and –1.85).

Consistent Behavior Sets. The argument against mediation of trait judgments by evaluative responses is also bolstered by the results of the analysis of the evaluations of consistently–behaving targets. Targets who consistently behaved honestly were liked (M = 2.58) and those who behaved dishonestly were disliked (M = –2.46). In comparison, targets who consistently behaved in an intelligent manner were liked (M = 2.10), but those who behaved in an unintelligent manner were evaluated neutrally (M = .17). This pattern of means reflects a significant Valence of Behavior Set × Trait Dimension interaction, F (1, 119) = 337.51, p < .0001. Two important points should be made about the means for this interaction. First, the evaluations of targets described by honest behaviors and targets described by dishonest behaviors were about equally extreme. Such equivalence is not consistent with those who might claim that negativity biases in evaluations or trait judgments are caused by stronger reactions to negative behaviors than to positive behaviors. Second, the neutral evaluation of targets who consistently behaved in an unintelligent manner (M = .17) makes it unlikely that such evaluations can account for the extremely unintelligent (M = –2.65) ratings given to such targets.

Behavior Consistency Ratings

The primary purpose in obtaining ratings of the consistency of the behaviors in a set was to assess the extent to which such ratings might account for the effects found in the trait judgments. To examine this possibility, two ANOVAs were conducted on the consistency ratings. One ANOVA examined ratings given to targets who behaved inconsistently, and the second examined ratings given to targets who behaved consistently. The variables used as predictors in these analyses were identical to the variables used in the analyses of the trait ratings that were reported earlier in this article.

Inconsistent Behavior Sets. From the perspective of Welbourne (1999), the behaviors should have been seen as more inconsistent when they came from the same individual than when they came from groups. Although the means were consistent with this idea (individual M = –.50, family M = –.15, social club M = –.11) the differences among these means were not significant, F (2, 120) = 1.35, p > .27. The target rated did enter into an interaction with trait dimension and the valence of behavior majority, F (2, 120) = 5.03, p < .009. However, inspection of the means

for this interaction again provide only modest support for Welbourne's assertion. When most of the behaviors in the set were intelligent, individuals ($M = 1.17$) were not seen as behaving less consistently than social clubs ($M = 1.03$), although both were seen as less consistent than families ($M = 1.68$). When most of the behaviors in a set were unintelligent, individuals ($M = -1.17$) were not seen as less consistent then families ($M = -.98$), although both were seen as somewhat less consistent than social clubs ($M = -.62$). Consistency judgments from targets who mostly behaved honestly provided the strongest support for Welbourne's assertion, with individuals ($M = -1.25$) clearly perceived to be less consistent than families ($M = -.36$) or social club members ($M = -.03$). However, this lowered consistency was not present when most of the behaviors in the set were dishonest (individual $M = -.74$, family $M = -.94$, social club $M = -.82$). On the whole, then, these data are not very consistent with Welbourne's claim that inconsistency perceptions mediate the diminution of positivity biases in intelligence judgments and in negativity biases in honesty judgments when those judgments are made about groups rather than about individuals.

Given the theoretical underpinnings of the trait theorist variable, one might have expected that ratings of behavior consistency should be lower for entity theorists than for incremental theorists. The means were consistent with this idea (entitativity theorist $M = -.44$, incremental theorist $M = -.07$), but the difference between these means merely approached significance, $F (1, 120) = 2.85$, $p < .09$.

Finally, an interesting and unanticipated Trait Dimension $\times$ Valence of the Behavior Majority interaction, $F (1, 120) = 57.10$, $p < .0001$, suggests that perceived inconsistency among behaviors is not defined solely by the objective mix of items in a stimulus set. Those who exhibited mostly honest behaviors ($M = -.56$) and those who exhibited mostly dishonest behaviors ($M = -.83$) were both seen as inconsistent, while the behaviors of targets who exhibit mostly intelligent behaviors were seen as consistent ($M = 1.29$) and the behaviors of those who exhibit mostly unintelligent behaviors were seen as inconsistent ($M = -.93$). Hence, the trait dimension that is implied by the behaviors in a set affects perceptions of inconsistency.[8]

---

8. As suggested by the means described in this paragraph, the main effect for the valence of the majority of behaviors in the stimulus set, $F (1, 120) = 102.43$, $p < .0001$, and the trait dimension main effect, $F (1, 120) = 46.64$, $p < .0001$, were also significant in this analysis.

Consistent Behavior Sets. A possible corollary to Welbourne's thesis is that the consistent behaviors of individuals should be seen as more consistent than the consistent behaviors of groups. However, although the target rated effect was significant in the analysis of the consistency judgments made about these targets, $F(2, 117) = 5.88$, $p < .004$, the means for this effect surprisingly show that the behaviors were seen as most consistent when they were performed by family members ($M = 2.03$), with individuals showing middling consistency ($M = 1.56$) and social club members the lowest consistency ($M = 1.26$). The only other significant effects in this analysis were a Trait Dimension × Valence of the Behavior Set interaction, $F(1, 117) = 4.59$, $p < .04$, and a main effect for the valence of the behavior set, $F(1, 117) = 143.10$, $p < .0001$. Positive behaviors were generally perceived to be more consistent than negative behaviors, but this difference was larger on the intelligence dimension (intelligent $M = 2.71$, unintelligent $M = .45$) than on the honesty dimension (honest $M = 2.54$, dishonest $M = .79$).

Mediational Analyses
While the analyses of the evaluative ratings and the behavior consistency ratings suggest that these variables do not mediate the effects observed in the trait judgments, this possibility was more formally evaluated in two mediational analyses. In one analysis, the trait judgments about the inconsistently–behaving targets were analyzed using the same factors described earlier in this article, but the evaluative ratings were also entered into the analysis as an additional predictor. In a second analysis, the trait judgments about the inconsistently–behaving targets were again analyzed using the same factors described earlier in this article, and the evaluative ratings were entered into the analysis as an additional predictor.

The evaluative ratings were strongly related to the trait ratings, $F(1, 120) = 2238.21$, $p < .0001$, as were the consistency ratings, $F(1, 121) = 781.43$, $p < .0001$. However, in both analyses there was still a negativity bias in honesty judgments, a positivity bias in intelligence judgments, and a diminution of these biases when judgments were made about groups (smallest $p < .0005$). Entry of the evaluative judgments into the model did cause judgments of incremental theorists to no longer be significantly different from the judgments of incremental theorists (Valence of the Behavior Majority × Theorist Type interaction: $F(1, 120) = 2.39$, $p > .12$), as did entry of the consistency ratings into the model (Valence of the Behavior Majority × Theorist Type interaction, $F(1, 120) =$

3.16, $p > .08$). However, because this interaction in the trait judgments was weak, claims that evaluations or consistency judgments mediated this effect are dubious.

DISCUSSION

The results from Experiment 2 show that when targets behave inconsistently, negativity biases emerge in judgments of a target's honesty and positivity biases emerge in judgments of a target's intelligence. The results also show that these biases diminish in judgments of groups relative to judgments of individuals. Both of these outcomes support predictions derived from category diagnosticity theory. The absence of target judged effects in the trait judgments made from sets of consistent behaviors has similar implications. According to Skowronski (1985), when the behaviors in a description are all internally consistent, representativeness should tend to govern trait judgments, regardless of whether judgments are made about individuals or groups. Assuming that the cognitive representations of trait–possessing groups do not substantially differ from the representations of trait–possessing individuals, then trait judgments about groups should not differ from trait judgments about individuals—and that is exactly what was observed.

From one point of view, many of the informativeness–supportive findings on the consistent behavior trials involve the absence of differences between conditions. Such null effects are often thought to be uninformative. A reasonable response to such a comment would be that null effects can be informative if they emerge in response to manipulations that "turn off" significant effects (as is the case in Experiment 2). Moreover, alternative explanations for these null effects encounter serious difficulties. For example, some might suggest a ceiling effect in judgments made about the consistently–behaving targets. This ceiling effect argument is confronted with several difficulties. The first of these difficulties comes from the Lupfer et al. (2000) data, which similarly found no valence effects. The smaller set size (three behaviors) used by Lupfer et al. works against a ceiling effect interpretation. The second difficulty is that a ceiling effect argument might apply only when the ratings given to targets were close to the endpoints of the response scale. This does not fit the data from Experiment 2: The average trait ratings in the consistent behavior conditions were often a full scale point or more below the scale endpoints.

Moreover, the data from Experiment 2 pose serious challenges to those theories that have attempted to account for trait judgment effects via affective mechanisms. The positivity bias in intelligence judgments for inconsistently–behaving targets and the absence of valence effects for consistently–behaving targets contradict those who offer a blanket claim that negative behaviors are more informative for trait judgments than are positive behaviors because of factors such as the non–normativeness of such behaviors (Ybarra & Stephan, 1999). If this were the case, then negativity effects should have emerged for both honesty and intelligence judgments regardless of whether the actor behaved consistently or inconsistently.

More generally, the results of Experiment 2 show that these trait judgment effects cannot be fully accounted for by affective responses to the targets (Baumeister, et al., 2001; Ito et al. 1998; Peeters & Czapinski, 1990). Mediational analyses do not support the notion that trait judgments are derived from affective responses to the targets, and the patterns of results observed in the evaluative judgments often differed from the patterns observed in the trait judgments. A particularly striking example of this non–parallelism was that participants provided a relatively neutral evaluation of the target who performed five unintelligent acts, despite rating the target as extremely unintelligent. From an affect–based judgment perspective this result is particularly surprising given that in Experiment 1 each of the unintelligent behaviors was perceived to be evaluatively negative. The usual result is for judgments to become more extreme as behaviors are added to a consistent set (see Levin & Kaplan, 1974). However, repeated encounters with a target's unintelligent behaviors may have brought to mind an individual who has a mental impairment, such as Down's syndrome, which might alter reactions to the behavior set. This possibility is supported by our post–experiment discussions with participants, several of whom mentioned the idea that some of the targets in the study may have had a mental impairment. This result emphasizes the importance of understanding the cognitive representations that are activated by behavioral stimuli and the way in which those behaviors fit the activated representations.

The point that affect–based theories in this area focus on the value of individual stimuli and tend to ignore variables affecting how those stimuli are used to make judgments can be made in another way. For example, consider the proposition from approach/avoidance theory that the avoidance gradient is steeper than the approach gradient. As in many

studies of prospect theory (Kahneman & Tversky, 1990), the implication is that a loss (e.g., losing a $100 bet) will be more aversive than an equally extreme gain (e.g., winning a $100 bet). Thus, a person might feel a bit negative if they place two bets, winning $100 in one and losing $100 in the other.

However, this is not the situation that exists in most trait judgment experiments: Stimuli are typically pretested and selected so that the initial reaction to a negative event is equal in extremity to the reaction to a positive event. Hence, assuming that a $115 win is equal in psychological extremity to a $100 loss, a hypothetical trait judgment equivalent might be to ask people how they feel after two bets, one in which they won $115 and the other in which they lost $100. Thus, a priori differences in the intensity or extremity of positive and negative events may certainly exist, but these are already controlled for by the pretesting that goes on in most trait judgment experiments. Effects such as the negativity bias in honesty judgments or the positivity bias intelligence judgments must be explained by factors that go beyond the initial response produced by each behavior. That's where the concept of diagnosticity plays its role: it suggests that behaviors that have a more exclusive relationship with a trait category will have more impact on trait judgments, in part, by increasing the likelihood that the person will be placed in the trait category that is implied by highly diagnostic behavior.

A similar point is made from the trait theorist data. Compared to incremental theorists, entity theorists may believe that an inconsistent item contains relatively little information about the personality of the entity. In support of this idea, trait judgments of entity theorists were more affected by the general trend in behaviors, while the trait judgments of incremental theorists were more affected by the rare, deviant behavior. This effect was not paralleled by a similar effect for evaluations, providing further evidence that trait judgments were not driven by evaluative responses to the target. Moreover, beliefs about personality stability and change only affected judgments about the inconsistently behaving targets. When targets behaved consistently, there was no relation between theorist type and trait judgment. This latter outcome again suggests that the informativeness of a behavior for trait judgments varies depending on whether one is making judgments about inconsistently–behaving targets or about consistently–behaving targets.

The results of Experiment 2 also suggest that the diminution in negativity and positivity biases in trait judgments about groups cannot

be accounted for by the different ways in which inconsistent information is processed for individuals and for groups (Welbourne, 1999). It is not claimed that such processing effects do not occur: It is clear that forming impressions about groups sometimes involves different mental processes than forming impressions about individuals. Rather, the claim here is that such differences do not account for the differences in group and individual trait judgments that emerged in Experiment 2. This claim is bolstered by the fact that information about families and social clubs appear to provoke processing that is similar to the way in which information about individuals is processed (McConnell, et al., 1997). The fact that negativity and positivity biases were diminished in judgments about such groups suggests that informativeness, and not processing, is responsible for these reductions.

CONCLUSIONS

People make trait judgments about others, and may often do so spontaneously from a single behavior (Carlston & Skowronski, 1994; Uleman, Hon, Roman, & Moskowitz, 1996). However, one of the tasks that people often confront in real life is to make trait judgments about others based on a collection of behavioral observations. Juries make decisions about defendants based on the evidence presented in a case. Faculty members make decisions about potential colleagues based on multiple facts presented in a curriculum vita or gleaned from an applicant's academic record. Voters make decisions about candidates for office based on a plethora of information obtained during the course of a campaign.

It is now known that there are a great number of factors that affect such judgments. Many of these factors are rooted in motivation or affect. Certainly, one's mood (Isbell & Wyer, 1999) as well as one's political beliefs (Uhlaner & Grofman, 1986) will color the interpretation and use of behavioral evidence that is gleaned about candidates. Moreover, feelings of personal responsibility, accountability (Tetlock, 1983), and the need to make a decision rapidly (Richter & Kruglanski, 1998) will also affect judgments.

However, neither affect nor processing mechanisms can provide a general, unified explanation for the data described in the present paper. Instead, these data highlight the role that informativeness plays in trait judgments. Moreover, the data suggest that informativeness is not fixed, but is a dynamic quality of behaviors. Four diagnosticity–driven out-

comes, the positivity bias in intelligence judgments, the negativity bias in honesty judgments, the diminution of these biases when the judgment target is a group rather than an individual, and individual differences in the extent to which people incorporate inconsistent information into their impressions, emerged only when targets behaved inconsistently. When a target's behaviors were all consistent with membership in a trait category, diagnosticity was irrelevant to the informativeness of a behavior for a trait judgment. Given that the concept of diagnosticity refers to the extent to which a behavior allows one to make a choice between membership in alternative categories, this is a sensible outcome: In the consistent–behavior case there is no such choice to be made. Instead, in consistent–behavior cases the perceived informativeness of a behavior is likely derived from the extent to which the behavior maps on to the cognitive representation of a single trait.

Thus, while behaviors have information value to people who make trait judgments, that information value is not fixed. It will vary across entities, across situations and will even vary with the other behaviors that are included in a target description. While researchers have made considerable progress toward understanding factors, such as diagnosticity, that affect the information value of behaviors for such trait judgments, much remains to be done. To understand the causes of informativeness, future research in the area must continue to explore the cognitive representations of trait concepts, the personal and social origins of such representations, and the cognitive mechanisms that are used to fit descriptions to trait categories and to make decisions about trait category membership.

REFERENCES

Baumeister, R.F., Bratslavsky, E., Finkenauer, C., & Vohs, K.D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*, 323–370.

Betz, A.L., Gannon, K.M. & Skowronski, J.J. (1993). The moment of tenure and the moment of truth: When it pays to be aware of recency effects in social judgments. *Social Cognition*, *10*, 397–413.

Birnbaum, M.H. (1972). Morality judgments: Tests of an averaging model. *Journal of Experimental Psychology*, *93*, 35–42.

Birnbaum, M.H. (1973). Morality judgments: Tests of an averaging model with differential weights. *Journal of Experimental Psychology*, 99, 395–399.

Birnbaum, M.H. (1974). The nonadditivity of person impressions. *Journal of Experimental Psychology*, *102*, 543–561.

Bissonnette, V., Ickes, W., Bernstein, I.H.,& Knowles, E. (1990). Personality moderating

variables: A warning about statistical artifact and a comparison of analytic techniques. *Journal of Personality*, *58*, 567–587.

Cacioppo, J.T., Gardner, W.L.,& Berntson, G.G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality & Social Psychology Review*, *1*, 3–25.

Carlston, D.E., & Skowronski, J.J. (1994) Savings in the relearning of trait information as evidence for spontaneous inference generation. *Journal of Personality & Social Psychology*, *66*, 840–856.

Chiu, C., Hong, Y., & Dweck, C.S. (1997). Lay dispositionism and implicit theories of personality. *Journal of Personality & Social Psychology*, *73*,19–30.

Coovert, M.D., & Reeder, G.D. (1990) Negativity effects in impression formation: The role of unit formation and schematic experiences. *Journal of Experimental Social Psychology*, *26*, 49–62.

Dweck, C.S., Chiu, C., & Hong, Y. (1995). Implicit theories and their role in judgments and reactions: A world from two perspectives. *Psychological Inquiry*, *6*, 267–285.

Fazio, R.H. (2000). On the automatic activation of associated evaluations: An overview. *Cognition & Emotion*, *15*, 115–141.

Gannon, K.M., Skowronski, J.J., & Betz, A.L. (1994). Depressive diligence in social information processing: Implications for order effects in impressions and for social memory. *Social Cognition*, *12*, 263–280.

Hamilton, D.L. & Sherman, S.J. (1996). Perceiving persons and groups. *Psychological Review*, *103*, 336–355.

Heider, F, (1958). *The psychology of interpersonal relations*. New York: Wiley.

Hess, T.M., Bolstad, C.A. Woodburn, S.M. & Auman, C. (1999). Trait diagnosticity versus behavioral consistency as determinants of impression change in adulthood. *Psychology & Aging*, *14*, 77–89.

Isbell, L. & Wyer, R.S., Jr. (1999). Correcting for mood–induced bias in the evaluation of political candidates: The roles of intrinsic and extrinsic motivation. *Personality & Social Psychology Bulletin*, *25*, 237–249.

Ito, T.A., Larsen, J.T., Smith, N.K., & Cacioppo, J.T. (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality & Social Psychology*, *75*, 887–900.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237–251.

Kahneman, D., & Tversky, A. (1990). Prospect theory: An analysis of decision under risk. In P.K. Moser, (Ed.), *Rationality in action: Contemporary approaches* (pp. 140–170). New York, NY,: Cambridge University Press.

Kanouse, D.E., & Hanson, L.R. (1972). *Negativity in evaluations*. New York: General Learning Press.

Kashima, Y., & Kerekes, A.R.Z (1994). A distributed memory model of averaging phenomena in impression formation. *Journal of Experimental Social Psychology*, *30*, 407–455.

Klein, J.G. (1991). Negativity effects in impression formation: A test in the political arena. *Personality & Social Psychology Bulletin*, *17*, 412–418.

Klein, J.G. (1996). Negativity in impressions of presidential candidates revisited: The 1992 election. *Personality & Social Psychology Bulletin*, *22*, 288–295.

Levin, I.P., & Kaplan, M.F. (1974). The set–size effect in personality impression formation is not an artifact. *Bulletin of the Psychonomic Society*, *3*, 187–188.

Levy, S.R.,& Dweck, C.S. (1997). Trait– versus process–focused social judgment. *Social Cognition*, *16*, 151–172.

Levy, S.R., Plaks, J.E.,Hong, Y.,Chiu, C,. & Dweck, C.S. (2001). Static versus dynamic theo-

ries and the perception of groups: Different routes to different destinations. *Personality & Social Psychology Review*, *5*, 156–168.

Levy, S.R., Stroessner, S.J, & Dweck, C.S. (1998). Stereotype formation and endorsement: The role of implicit theories. *Journal of Personality & Social Psychology*, *74*, 1421–1436.

Lickel, B., Hamilton, D.L., & Sherman, S.J. (2000). Elements of a lay theory of groups: Types of groups, relationship styles, and the perception of group entitativity. *Personality & Social Psychology Review*, *5*, 129–140.

Lupfer, M.B.,Weeks, M., & Dupuis, S. (2000). How pervasive is the negativity bias in judgments based on character appraisal? *Personality & Social Psychology Bulletin*, *26*, 1353–1366.

McConnell, A.R., Sherman, S.J., & Hamilton, D.L. (1997). Target entitativity: Implications for information processing about individual and group targets. *Journal of Personality & Social Psychology*, *72*, 750–762.

Nisbett, R.E., Zukier, H., & Lemley, R.E. (1981). The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology*, *13*, 248–277.

Peeters, G. (1971). The positive–negative asymmetry: On cognitive consistency and positivity bias. *European Journal of Social Psychology*, *1*, 455–474.

Peeters, G. (1983). Relational and informational patterns in social cognition. In W. Doise & S. Moscovici (Eds.), *Current issues in European Social Psychology* (Vol. 1, pp. 201–237). Cambridge: Cambridge University Press.

Peeters, G., & Czapinski, J. (1990). Positive–negative asymmetry in evaluations: The distinction between affective and informational negativity effects. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology*, (Vol. 1, pp. 33–60). Chichester: Wiley.

Plaks, J.E., Stroessner, S.J.,Dweck, C.S., & Sherman, J.W. (2001). Person theories and attention allocation: Preferences for stereotypic versus counterstereotypic information. *Journal of Personality & Social Psychology*, *80*, 876–893.

Reeder, G.D. (1993). Trait–behavior relations and dispositional inference. *Personality & Social Psychology Bulletin*, *19*, 586–593.

Reeder, G.D. (1997). Dispositional inferences of ability: Content and process. *Journal of Experimental Social Psychology*, *33*, 171–189.

Reeder, G.D., & Brewer, M.B. (1979). A schematic model of dispositional attribution in interpersonal perception. *Psychological Review*, *86*, 61–79.

Reeder, G.D., & Coovert, M.D. (1986). Revising an impression of morality. *Social Cognition*, *4*, 1–17.

Richter, L. & Kruglanski, A.W. (1998). Seizing on the latest: Motivationally driven recency effects in impression formation. *Journal of Experimental Social Psychology*, *34*, 313–329.

Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality & Social Psychology*, *9*, 283–294.

Rozin, P. & Royzman, E. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, *5*, 296–320.

Singh, R.& Teoh, J.B. P. (2000). Impression formation from intellectual and social traits: Evidence for behavioural adaptation and cognitive processing. *British Journal of Social Psychology*, *39*, 537–554.

Skowronski, J.J. (1985). Explaining negativity and extremity biases in impression formation: A categorization–cue validity approach. *Dissertation Abstracts International*, *46*, 2–B.

Skowronski, J.J., & Carlston, D. E. (1987). Social judgment and social memory: The role of

cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology*, *52*, 689–699.

Skowronski, J.J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, *105*, 131–142.

Skowronski, J.J., & Carlston, D.E. (1992). Caught in the act: When impressions based on highly diagnostic behaviors are resistant to contradiction. *European Journal of Social Psychology*, *22*, 435–452.

Smith, E.R., & DeCoster, J. (1998). Person perception and stereotyping: Simulation using distributed representations in a recurrent connectionist network. In S.J. Read & L.C. Miller (Eds.), *Connectionist models of social reasoning and social behavior* (pp. 111–140). Mahwah, NJ: Lawrence Erlbaum Associates, Inc., Publishers.

Tetlock, P.E. (1983). Accountability and the perseverance of first impressions. *Social Psychology Quarterly*, *46*, 285–292.

Uhlaner, C.J., & Grofman, B. (1986). The race may be close but my horse is going to win: Wish fulfillment in the 1980 Presidential election. *Political Behavior*, *8*, 101–129.

Uleman, J.S., Hon, A., Roman, R.J., & Moskowitz, G.B. (1996). On–line evidence for spontaneous trait inferences at encoding. *Personality & Social Psychology Bulletin*, *22*, 377–394.

Vonk, R. (1999). Effects of other–profitability and self–profitability on evaluative judgments of behaviours. *European Journal of Social Psychology*, *29*, 833–842.

Vonk, R. (1993). The negativity effect in trait ratings and in open–ended descriptions of persons. *Personality & Social Psychology Bulletin*, *19*, 269–278.

Vonk, R.& Van Knippenberg, A. (1994). The sovereignty of negative inferences: Suspicion of ulterior motives does not reduce the negativity effect. Social Cognition, *12*, 169–186.

Welbourne, J.L. (1999). The impact of perceived entitivity on inconsistency resolution for groups and individuals. *Journal of Experimental Social Psychology*, *35*, 481–508.

Wojciszke, B., Bazinska, R., & Jaworski, M. (1998). On the dominance of moral categories in impression formation. *Personality and Social Psychology Bulletin*, *24*, 1251–1263.

Wojciszke, B., Brycz, H., & Borkenau, P. (1993). Effects of information content and evaluative extremity on positivity and negativity biases. *Journal of Personality & Social Psychology*, *64*, 327–336.

Ybarra, O. (2001). When first impressions don't last: The role of isolation and adaptation processes in person impressions. *Social Cognition*, *19*, 547–586.

Ybarra, O., & Stephan, W.G. (1999). Attributional orientations and the prediction of behavior: The attribution–prediction bias. *Journal of Personality & Social Psychology*, *76*, 718–727.