Available online at www.sciencedirect.com
ScienceDirect
Cognition 106 (2008) 594–632
COGNITION
www.elsevier.com/locate/COGNIT

What do you want? How perceivers use cues to make goal inferences about others

Joseph P. Magliano [a,*], John J. Skowronski [a], M. Anne Britt [a], C. Dominik Güss [b], Chris Forsythe [c]

[a] *Department of Psychology, Northern Illinois University, DeKalb, IL 60115, USA*
[b] *North Florida University, USA*
[c] *Sandia National Laboratories, USA*

Received 28 August 2006; revised 1 February 2007; accepted 17 March 2007

Abstract

Variables influencing inferences about a stranger's goal during an unsolicited social interaction were explored. Experiment 1 developed a procedure for identifying cues. Experiments 2 and 3 assessed the relative importance of various cues (space, time, characteristics of oneself, characteristics of the stranger, and the stranger's behavior) for goal judgments. Results indicated that situational context cues informed goal judgments in ways that were consistent with diagnosticity ratings and typicality ratings of those cues. Stranger characteristics and stranger behaviors affected goal judgments more than would be expected from these quantitative measures of their informativeness. Nonetheless, the results are consistent with a mental model view that assumes perceivers monitor situational cues present during interactions and that goal inferences are guided by the informativeness of these cues.

Keywords: Goals; Situation models; Inference

* Corresponding author.
E-mail address: jmagliano@niu.edu (J.P. Magliano).

0010-0277/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2007.03.010

1. Introduction

The comprehension, prediction, and explanation of social interactions, and subsequent emotional and behavioral responses, are often contingent on inferences that are made about others' goals (see Hassin, Aarts, & Ferguson, 2005; Read & Miller, 1993). One approach that can help to understand where such inferences come from is derived from studies of narrative processing. Results from such studies suggest that comprehension of fictive social interactions rests upon inferring the goals of characters (e.g., Copeland, Magliano, & Radvansky, 2006; Long & Golding, 1993; Magliano, Radvansky, & Copeland, in press; Magliano, Taylor, & Kim, 2005; Suh & Trabasso, 1993; Trabasso, van den Broek, & Suh, 1989). This research shows that causal inferences linking a character's goals to their actions are an important basis for establishing coherence in understanding. Moreover, goal inferences may have a special status in event understanding: Trabasso et al. (1989) found that causal relations derived from goals are perceived to be stronger than relationships derived from other psychological causes (also see Magliano & Radvansky, 2001). The present research explores the extent to which approaches used in the study of narrative understanding can apply more broadly to perceivers' understanding of the goals of others.

2. The narrative perspective: The elements of mental models

Discourse processing researchers generally agree that deep comprehension of narratives arises via the construction of a mental model for the situation described in the text, which has been called the *situation-model* (e.g., van Dijk & Kintsch, 1983; for reviews see Graesser, Millis, & Zwaan, 1997; Zwaan & Radvansky, 1998). In order to construct a situation model for a narrative text, readers use their general knowledge of the real world to construct an understanding and representation of how events unfold within the fictive world. A central assumption is that readers build situation models by monitoring and indexing how events are related along multiple dimensions of continuity. These dimensions include time and place in which the events unfold (*the spatio-temporal framework*); the *entities* in the scenario (people, objects, ideas); the *properties of those entities* (e.g., colors, sizes, emotions, goals); and *relational information* (spatial, temporal, causal, ownership, kinship, social, etc.) specifying the nature of the linkages among the spatio-temporal framework, the entities, and their properties (Magliano, Miller, & Zwaan, 2001; Magliano, Zwaan, & Graesser, 1999; Radvansky, 1998; Zwaan, Magliano, & Graesser, 1995; Zwaan & Radvansky, 1998).

Research generally indicates that some cues carry more import in a situation model and are monitored more closely by the reader than others (see Zwaan & Radvansky, 1998). For example, some research has shown that characters play a central role in situation models, (Ozyurek & Trabasso, 1997; Scott-Rich & Taylor, 2000), which makes sense to the extent that situation models are about people. Research also indicates that the implied causal relationships between non-intentional events, and between goals and behaviors, carry considerable predictive power and are mon-

itored closely (Trabasso et al., 1989; Zwaan, Magliano, et al., 1995; Zwaan, Langston, & Graesser, 1995). In contrast, readers do not closely monitor other aspects of narrative space, such as spatial relationships between objects, unless: (1) it is highly causally relevant (Sundermeier, van den Broek, & Zwaan, 2005); (2) they have a specific goal to do so (Zwaan & van Oostendorp, 1993), (3) they have extensive prior knowledge of the fictive space (e.g., Morrow, Greenspan, & Bower, 1987), or (4) they have an opportunity to reread a narrative (Zwaan, Magliano, et al., 1995). In particular, aspects of the spatial dimension appear to be subordinate to other situational dimensions (e.g., time) even for modalities that are inherently spatial, such as film (Magliano et al., 2001).

This narrative perspective can be extended to the study of real-world events. We assume that when understanding real-world situations people construct situation models that capture .information along the dimensions specified by narrative researchers (*the spatio-temporal framework*, the *entities* in the scenario, the *properties of those entities*, and *relational information* specifying the nature of the linkages among the dimensions; see Copeland et al., 2006; Magliano et al., in press; also see Pennington & Hastie, 1993). However, just as it is the case with narrative texts, it is critical to assess the relative importance of the different dimensions in understanding and representing real world events.

Thus, the first goal of the research described in the present article is to explore the extent to which an actor's goal might be implied by other cues specified by situation models (time, location, stranger's actions, stranger's appearance, self-state). Another way to phrase this is to say that the present research asks the question of the extent to which cues, either singly or in combination, are perceived to be informative with respect to understanding an actor's goals. For example, when a person approaches at midnight in a dark alley holding a gun in his hand, you might make the inference that the person wants to rob you. The fact that one will infer the actor's goal from this cue array implies that one sees something in the actor's behavior and characteristics in that time and situation as *informative* with respect to the goal. However, where do such perceptions of informativeness come from?

3. Cue-level sources of informativeness in inference-making

We assume that people possess generic knowledge structures that reflect different types of social interactions (e.g., Schank & Abelson, 1977). These mental representations capture the probabilistic relationships between elements of a situation, such as between cues and inferences. As such, goal inferences should be influenced by the extent to which perceivers monitor various cues that are present in an unfolding event and the probabilistic relationships between those cues and the goal inferences associated with those cues.

Theoretical perspectives exploring probabilistically based category judgments and hypothesis testing (Garcia-Marques, Sherman, & Palma-Oliveira, 2001; for a review Trope & Liberman, 1996), as well as some approaches to similarity-based judgment processes (e.g., Kahneman & Tversky, 1972) suggest that at least two sources of

information can contribute to the perception that a cue is informative with respect to its ability to predict membership in a social category, such as an actor's goal. The first of these is the *diagnosticity* of the cue. Diagnosticity refers to the extent to which a cue is associated with membership in one category, but not with membership in other categories (see Rosch & Mervis, 1975). For example, because only birds, and not mammals or reptiles, have feathers, one should be able to infer with a high degree of confidence that a creature that possesses feathers is a bird. Such logic has previously been applied to social inferences (e.g., see Skowronski & Carlston, 1999; Trope & Mackie, 1987). If people think that brandishing a gun is *uniquely* associated with the goal of robbery, but not with other goals (such as helping, romance, or starting a friendship), then gun-brandishing ought to be perceived as an especially powerful cue leading to an inference of a robbery goal.

However, it is also the case that people often make judgments based on typicality: If a target's feature matches one of the typical features contained in the mental representation of a comparison category, people will see that feature as informative with respect to category membership. As illustrated by the phenomenon of confirmation biases in social hypothesis testing, (Davies, 2003; Devine, Hirt, & Gehrke, 1990; Strohmer & Shivy, 1994), this will occur, even if the feature is relatively non-diagnostic of category membership. Hence, even if a target's feature is not diagnostic of category membership (in the sense that it does not distinguish membership in one category from membership in other categories), the feature will tend to be perceived as informative for category judgments if the feature frequently co-occurs with category membership. For example, in the course of describing a creature one might describe the creature as having two legs. Even though there are innumerable creatures that possess that feature (i.e., it is low in diagnosticity), if one is entertaining the hypothesis that the creature is a bird, that feature should influence judgment because it helps to enhance the perceived "fit" between the target and the mental representation of the category. In the social domain, while a stranger who approaches you may have many goals in mind, one might see that approach behavior as supporting the notion that the stranger's goal is robbery because stranger approach behavior is thought to be frequently associated with robbery.

Kahneman and Tversky's approach to similarity-based judgments (1972) suggests that the degree to which a target and a representation are perceived to overlap is derived, in part, from the number of features that overlap and the extent to which those features are "essential" to the category. This suggests that both typicality and diagnosticity will independently contribute to the judgment process: typicality because it is related to the number of features that overlap, and diagnosticity because it is related to the "essentialness" of the features in the representation. However, while both typicality and diagnosticity effects have been separately obtained in prior research, to our knowledge few researchers have documented the separate and simultaneous effects of these cue characteristics on judgment. Accordingly, a second goal of the present research is to show that a cue's diagnosticity and its typicality can be measured, and that these separately and independently contribute to the impact of that cue on goal inferences. Cues that are more typical, and that are more diagnostic, should have a heightened impact on goal judgments.

4. How important are actor behaviors to goal inferences?

One can clearly use ideas about diagnosticity and typicality to argue that some cues should be more informative to goal inferences than other cues. However, one might wonder whether such ideas will apply to *classes* of cues, as well as to individual cues. That is, one might wonder whether it is the case that one cue class (e.g., an individual's behavior) is more important to goal inferences than other cue classes. As mentioned above, research on narrative comprehension strongly suggests a hierarchy of cue importance (Zwaan & Radvansky, 1998). One might also wonder whether such effects can be entirely accounted for by the perceived informativeness of the cues (as measured by the cue's typicality and diagnosticity), or whether some cue categories have a special status with respect to goal judgment that extends beyond these statistical measures.

Two lines of scholarship lead one to expect that an actor's behavior should be particularly important to goal inferences. The first of these is derived from research on situation models in narratives that shows that implied causal relationships between goals and behaviors carry considerable predictive power and are monitored closely (Trabasso et al., 1989; Zwaan, Magliano, et al., 1995; Zwaan, Langston, et al., 1995). The second line of scholarship is derived from the work of those social psychologists who follow Heider's (1958) famous dictum that 'behavior engulfs the field.'' From this perspective, inferences about others ought to be primarily determined by their actions. Situational factors should play a secondary role in judgment, causing inferences about others to be augmented (as when a person overcomes numerous obstacles to succeed) or discounted (as when a person's failure can be explained away by an illness). Many studies of trait attribution attest to the operation of such processes (for examples, see Kruglanski, Schwartz, Maides, & Hamel, 1978; McClure, 1998; Wells & Ronis, 1982).

However, there are reasons to be cautious in blindly accepting the pre-eminence of behavior with respect to inference-making. First, a closer look at Heider's writings suggests that his thinking about the relations between behaviors and situations is more complex than is typically portrayed. For example, Heider (1958) noted (p. 38) that ''the ambiguity of behavior as a local stimulus is reduced when it is seen in a situational context.'' This may occur because (p. 116) ''what the other person actually did is not only perceived on the basis of local cues given by movements of the person, but also on the basis of what we think we know about the situation.'' Put bluntly, sometimes one needs to know the context in which a behavior occurs in order to understand what the behavior means for the internal characteristics (e.g., goals) of the actor. Hence, this perspective implies that our knowledge of situations can sometimes have a substantial impact on the extent to which a given behavior is thought to imply a goal (see Trope, 1998 & Trope & Gaunt, 1999, for elaboration of these ideas in a more contemporary context).

This view makes considerable sense. For example, consider how the same behavior (someone approaches you) has different goal implications in different contexts. In one context, imagine that you have just entered a department store at mid-day. A male stranger approaches. He is dressed in a suit and tie and carrying a clipboard.

In this particular set of circumstances, given the setting, the time, the person's appearance, and the object being carried, one reasonable inference is that the approaching stranger might want to ask you to apply for a credit card. The inference made about the stranger's goals might be somewhat different if the same behavior (the male stranger approaches) was exhibited by an actor who wore all-black casual clothes, who was carrying a gun, and who approached at midnight in an alley.

This situationally grounded view of goal inferences fits well with the results of research conducted by Barker, a founder of Ecological Psychology (Barker, 1968; Barker & Wright, 1955). Barker's 20-year research program at his "Midwest Psychological Field Station" demonstrated that behaviors are highly constrained by the context (i.e. time and place) in which those behaviors occur. For example, the behaviors and goals of someone in a supermarket (e.g., selecting and purchasing food) are likely to be different than they would be in a classroom (e.g., listening and taking notes), in part because those places are designed to promote different kinds of behaviors. If perceivers are sensitive to these situational constraints, then Barker's research implies that settings, such as time and place, may play more than an incidental role in the process of making goal inferences from others' behaviors.

Given these conflicting views, the third major goal of the research described in the present paper was to explore the relative power of various cue classes to influence goal inferences. In particular, we were interested how impactful actor behaviors were to goal inferences relative to the other cue categories. A secondary interest was in exploring whether any differential strength effects that emerged were solely determined by the diagnosticity and typicality of the cues within each of the categories, or whether the impactfulness of cue categories went beyond these objective measures of a cue's informativeness. Situation model theories suggest that cue categories should have an impact over and above statistical measures of informativeness.

5. Experiment 1

Experiment 1 was designed to assess two of the properties of cues that contribute to the informativeness of those cues to goal judgments. The first of these is the extent to which the cue is perceived to be diagnostic of a goal. In Experiment 1, diagnosticity was measured by assessing the extent to which a cue was perceived to *uniquely* covary with a goal. The second attribute that was measured was a cue's *typicality*. This was measured in Experiment 1 by assessing the *relative frequency* with which a cue is perceived to covary with a goal.

In order identify the cues and measure their informativeness in predicting goals, we needed to develop a semantic space that represented this information. This semantic space was generated in Experiment 1 by asking participants to consider real or hypothetical situations in which they interacted with a stranger who had a specific goal (e.g., initiate a romantic relationship, ask for a jump start, rob you, have his/her picture taken). They then answered questions (see Appendix A) that were designed to elicit specific kinds of information that were associated with situations in which that goal was pursued (e.g., the time, the location). The frequency with which a cue was

produced across the goals under consideration was used to calculate the uniqueness of a cue with respect to the goals and the frequency with which a cue was produced for a given goal relative to other goals.

The cues that were solicited in Experiment 1 were largely determined by the types of cues that people are thought to use when constructing situation models (e.g., see the event-indexing model of Zwaan & Radvansky, 1998). These were: (1) the time at which an event occurred, (2) the location at which an event occurred, (3) the behavior of the stranger, (4) the stranger's appearance, and (5) self-characteristics, such as one's appearance and one's own emotional state. We note that this last dimension is not strictly specified by the event-indexing model. The reason is straightforward: The reader of a fictional narrative is generally not a part of the narrative's events. However, in a real-life situation, the self (appearance, behavior) is a part of the context for an actor's behavior. Accordingly, we reasoned that it was important to assess self-characteristics to see if such characteristics were incorporated into peoples' models of an actor's behavior in the real world.

The measures that emerge from Experiment 1 will be used to provide typicality and diagnosticity values for each cue. These measures were later used in Experiments 2 and 3 to explore the impact of a cue on goal judgments. However, these values can also provide insight into the question of why some cue categories may be more important to goal judgments than others. One possible outcome from Experiment 1 is that behavior cues may have higher diagnosticity and typicality values than other cue categories. If the results of later studies show that an actor's behavior is more important to goal judgments than are other cue categories, that heightened importance may be at least partially traced to the greater perceived informativeness (higher typicality and diagnosticity) of behavior cues relative to other cues.

5.1. Methods

5.1.1. Participants

One hundred and sixty participants enrolled in introduction to psychology or research methods courses at Northern Illinois University participated in the experiment. They received course credit as compensation for their participation.

5.1.2. Materials

Twenty-four goals were originally selected for the experiment. These goals were selected based on the results of a pilot study in which approximately 75 students enrolled in a Critical Thinking course were asked to list reasons why strangers may interact with them. We selected the original 24 goals based on whether: (1) they were produced by more than one participant and (2) they represented a variety of different types of goals. Specifically, we chose goals related to establishment of social relationships (e.g., romantic relationships), solicitations (e.g., asking for money), recruitment efforts (e.g., filling out a survey), assistance provision (e.g., providing medical assistance), and attempts at harm (e.g., robbing).

Using these goals, a questionnaire was constructed to solicit situational cues for the spatial–temporal frameworks, the self characteristics, the stranger's characteristics,

and the stranger's actions that are associated with each of the goals. A sample question set is shown in Appendix A. Participants were asked to either imagine an actual or hypothetical situation in which they were interacting with a stranger who had a particular goal (e.g. Imagine someone you do not know approaches you with the goal to get money from you). Participants then answered a series of open-ended questions regarding the details of that interaction.

There were four parts to the questionnaire. Each part was designed to elicit information about the one of the four different situational dimensions. Part 1 asked questions about the spatial–temporal framework (e.g., Describe the location; What time is it?). Part 2 asked questions about self characteristics (e.g., Describe how you are dressed; Describe how you feel.). Part 3 asked questions about characteristics of the stranger (e.g., Describe how he is dressed; What specific facial expression, if any, does the stranger have?). Part 4 asked questions about the actions of a stranger (e.g., List the sequence of actions taken by the stranger and by you over the course of the interaction.).

5.1.3. Procedure

The 24 question sets, each assessing a different goal, were used to create eight questionnaires. Each questionnaire contained three question sets, each of which pertained to a different goal. Any given participant considered only three goals. The three-goal sets were randomly determined. The goal sets in a given questionnaire were fixed, but the order of presentation was counterbalanced across participants.

Participants completed these questionnaires in a large classroom in groups of ten to thirty. They were told that the researchers were interested in learning about social interactions that involve strangers. Participants were asked to consider real or hypothetical interactions with strangers who have specific goals for initiating an interaction and were told to assume that the goal mentioned was the only goal that the stranger had. They were instructed that they would be asked to answer questions that were designed to reveal the details of the interaction. As such, the participants were instructed to imagine the interaction in as much detail as possible. They were given as much time as needed to complete the three questionnaires.

5.1.4. Situational cues and measures of informativeness

Although data were collected for 24 goals, subsequent analyses were based on only 10 goals. This was done in order to make the analyses more manageable. The 10 goals were chosen after preliminary examination of the data suggested that they provided a relatively equal representation of the five different goal types mentioned above. The social relationship goals were establishing a platonic friendship and establishing a romantic relationship; the solicitation goals were asking for money, asking for a photograph to be taken, and asking for a jump-start of a car; the recruitment goals were religious conversion and filling out a survey; the assistance goal was offering medical assistance; and the harming goals were robbing and killing.

We identified situational cues that were produced as an answer to a given question across all 10 goals. A cue was included in this list of cues only if it was produced by two or more participants. A matrix of 141 cues $\times$ 10 goals was assembled from this

list. Each cell of the matrix contained a frequency count reflecting the number of participants who produced the cue as an answer to a given goal-related questionnaire item. Each frequency could range in value from 0 to the number of participants who completed questionnaires that included that goal.

This matrix was used to derive two measures of the extent to which a cue informs a perceiver about the extent to which a stranger possesses a given goal. These were the cue's *relative frequency of association* with the goal (a measure that we think reflects the cue's *typicality*) and the extent to which the cue was *uniquely associated* with a given goal (a measure that we think reflects the cue's *diagnosticity*). The following equations were used to derive these measures:

$$\text{Relative frequency of association(FA)} = (\text{Freq of cue}_i \text{ for goal}_n)/(\text{total number of participants who considered goal}_n) \quad (1)$$

$$\text{Uniqueness of associations(UA)} = (\text{Freq of cue}_i \text{ for goal}_n)/(\text{Sum of the frequencies for cue}_i \text{ for goals}_{1-10}) \quad (2)$$

For example, 17 participants each responded to a questionnaire assessing the platonic friendship goal. Of those, 12 produced classroom as a location for the interaction. This location was produced a total of 16 times in participants' responses across all ten goals. Thus, the FA score for this goal-cue pairing would be $12/17 = .71$ (indicating that the cue was typical of the goal); the UA score would be $12/16 = .77$ (indicating that the cue was diagnostic of the goal). Note that it is possible for the FA and UA scores to be quite different. For example, 13 participants who saw the friendship questionnaire produced the cue "I'm dressed casually." This cue was produced 102 times across all 10 goals. As such, the FA score was .76 (highly typical), but the UA score was .13 (not very diagnostic).

5.2. Results

Each of the 141 cues that were identified in preliminary data reduction was placed into one of four categories: spatial–temporal cues, entity–self cues, entity–stranger cues, and strangers' actions. Spatio-temporal cues were responses that referred to the time and/or place in which the interaction occurred. Entity–self cues referred to characteristics of the self prior to the interaction (e.g., appearance, emotional state). Entity–stranger cues referred to the characteristics of the stranger prior to the interaction (e.g., he was tall). Stranger–action cues referred to the behaviors of the stranger in the interaction (he ran toward me; he asked if he could use my phone). The FA and UA scores were calculated for each goal-cue pair. Overall, FA and UA scores were moderately correlated ($r = .49$, $p < .05$). However, there was considerable variation in the strength of the correlations across cue categories. The spatial–temporal ($r = .51$, $p < .05$) entity–self ($r = .29$, $p < .05$), entity–strategy ($r = .48$, $p < .05$) all evinced moderate correlations between FA and UA scores. However, there was a high correlation between FA and UA scores for the stranger's action cues ($r = .76$, $p < .05$).

Table 1
Mean FA and UA scores as a function of situational dimension

Situational dimension	Measure of informativeness	
	FA	UA
Spatial–temporal	.30 (.11)	.24 (.06)
Entity–self	.18 (.07)	.23 (.11)
Entity–stranger	.37 (.08)	.35 (.05)
Stranger's actions	.34 (.13)	.77 (.16)

Note: standard deviations are shown in the parentheses. FA refers to a cue's relative frequency of association (or typicality) and UA refers to a cue's unique association with a goal (or diagnosticity).

Mean FA and UA scores were calculated for each of the four cue types. This was done separately for each of the ten goals. The means are shown in Table 1. These mean scores were entered into a Score (FA, UA) × Cue Type (spatio-temporal, entity–self, entity–stranger, stranger's actions) repeated measures ANOVA. The results of this analysis indicated that UA scores ($M = .40$) were higher than FA scores ($M = .30$), $F(1,9) = 26.87$, $MSE = 0.007639$, $p < .05$. There was also a main effect of cue category, $F(3,27) = 41.34$, $MSE = 0.001118$, $p < .05$. Post hoc analyses (LSD) revealed that mean informativeness scores were highest for the stranger's behaviors ($M = .55$), followed by the entity–stranger cues ($M = .36$), the spatial–temporal cues ($M = .27$), and the entity–self cues ($M = .21$), $F(3,27) = 41.34$, $MSE = 0.011175$, $p < .05$.

However, as indicated by the means in Table 1, these main effects were qualified by a significant Score × Cue Type interaction, $F(3,27) = 31.36$, $MSE = 0.008471$, $p < .05$. In decomposing this interaction, we were primarily interested in assessing whether FA and UA scores differed within each situational cue. Subsequent analyses (LSD tests) revealed that FA scores were significantly higher than UA scores for spatio-temporal cues. Conversely, UA scores were significantly higher than FA scores for the stranger's action cues. For both the entity–self cues and entity–stranger cues, the FA and UA scores did not significantly differ.

5.3. Discussion

Participants in this study were asked to imagine that a stranger had a given goal (e.g., making friends with the participant) and provided open-ended reports as to where and when that request would be made, what the characteristics of the self would be prior to the stranger's behavior, what the characteristics of the stranger would be, and what the behaviors of the stranger would be. These responses were used to calculate two different indices of the extent to which a cue category implied a goal category. One index was based on how often participants' generated that cue in response to a given goal (FA, a measure of typicality); the second index was how often that cue was produced for a given goal relative to the frequency with which it was produced for other goals (UA, a measure of diagnosticity).

These results suggest that some cues are more important to goal inferences than others. In line with the thinking of many social psychologists, the data suggest that a stranger's actions should be more informative to goal inferences than other cue types. However, the data also suggest that other cues should substantially affect such inferences. The stranger's appearance, self-characteristics, and the interaction's setting (time and place) were all associated with robustly positive typicality and diagnosticity scores, implying that they all would contribute to goal inferences. However, these results also suggest that the extent to which a cue's typicality and diagnosticity imply a given goal might vary substantially across cue types. That is, the data suggest that typicality of spatio-temporal cues might be more important to the impact of those cues on goal judgments than the diagnosticity of those cues. Conversely, the data suggest that the diagnosticity of a stranger's actions is likely to be more important to the impact of those cues on goal judgments than the typicality of those actions.

When considering these results, it would be useful to keep in mind that the semantic space derived by the procedure used in Experiment 1 may partially reflect naïve theories about the situations that co-occur with the goals chosen for the study. Furthermore, one should be cognizant of the fact that the space contains information that was explicitly available in working memory and codeable in language. Accordingly, it is an open question as to whether those same cues would be derived from, and would have similar diagnosticity and typicality values, in situations that people actually experienced.

However, in regards to this issue it should be noted that our approach for deriving situational constraints for goals has considerable precedent elsewhere. A number of studies (e.g., McRae, de Sa, & Siedenberg, 1997; McRae, Ferretti, & Amyote, 1997) have used similar methods to derive the situational, thematic, and semantic constraints associated with different lexical (e.g., verbs) or conceptual (abstract versus concrete) classes of words. For example, Wiemer-Hastings and Xu (2005) asked participants to exhaustively produce features and situations associated with concrete nouns and abstract nouns. Using this approach, they found that abstract and concrete nouns differed in terms of the situational constraints in which they occur (e.g., time, place, agency, etc.).

Given such precedents, we would argue that the methods that we used in Experiment 1 provided a reasonable starting point for extracting cues that are relevant to the goals in which were interested and for developing measures of the typicality and diagnosticity of those cues. Moreover, the semantic space extracted from Experiment 1 provides a basis for assessing the extent to which the measured diagnosticity and typicality of the cues can predict variance in judgments regarding the likelihood that stranger will have a goal. That is the issue to which we turn in Experiment 2.

6. Experiment 2

Experiments 2 and 3 were designed to assess the extent that the measures of informativeness derived from the cue–goal semantic space could predict the likelihood that a stranger would have a goal during a situated social interaction. In Experiment

2, participants were given some of the cues generated in Experiment 1, one at a time. In response to each cue, participants were asked to make a judgment about the extent to which a stranger possessed a given goal. These goal judgments were examined to see if some cues (e.g., the behaviors of a stranger) were more powerful predictors of goal judgments than other cues (e.g., the spatio-temporal context). In addition, goal judgments were examined in regards to the extent to which cue typicality and cue diagnosticity predicted those judgments. These two principles have both been invoked to explain why cues are informative for judgments, but to our knowledge, have rarely been examined in the context of the same experiment. Hence, it is of interest to see if typicality predicts judgments controlling for diagnosticity, and vice versa. Moreover, one can examine the data to see if, as implied by Experiment 1, typicality will be a more powerful predictor of the impact of spatio-temporal context on goal judgments, while diagnosticity will be a more powerful predictor of the impact of the stranger's behaviors on goal judgments.

6.1. Methods

6.1.1. Participants

One hundred and thirty Northern Illinois University undergraduates enrolled in an introduction to psychology course participated in the experiment. They received course credit as compensation for their participation.

6.1.2. Procedure

Introductory instructions to the experiment indicated to participants that the experimenters were interested in learning about how people infer the goals of strangers. They were informed that there are many cues surrounding interactions that could help someone figure out a stranger's goal. They were then given a specific example of what the experimenters meant by a cue. They were told to imagine a situation in which a stranger approached with a leaky pen. They were then asked to consider the likelihood that the stranger wanted to borrow a pen.

Participants were then told that they would see a series of cues, presented one at a time, that could potentially occur in a situation in which they are interacting with a stranger. They were told that such cues could vary in how informative they are of a goal. To illustrate this point, participants were told to consider the cue of being in a café when the stranger approaches. It was suggested that this cue may not be very informative of the stranger's intent to ask to borrow a pen.

After the instructions had been read, participants were ushered to personal computers and used them to complete the study. Presentation of stimuli and response recording were accomplished by the *E-Prime* computer program. Each participant had their own personal computer and completed the study in a cubicle that was isolated from other participants.

Each participant was given 141 cues, one at a time. The cues were presented in the context of the generic statement "Imagine that a stranger approaches you and <CUE>. How likely is that the stranger would want <GOAL STATEMENT>? The order of cue presentation was randomly determined for each subject.

There were 10 groups of participants; each made responses to a different goal. The computer program determined which goal the participant would be using and identified that goal to participants shortly before they began the experiment. Participants made their judgments on a five-point scale, which is shown below:

0	1	2	3	4
Not at all Likely or unrelated	Slightly Likely	Moderately Likely	Very Likely	Virtually Certain

After all cues were read and rated, participants were debriefed and dismissed.

6.2. Results

A series of multiple regression analyses assessed the relationship between each cue's FA and UA scores (determined from Experiment 1's data) and goal likelihood judgments. The main goal of the analyses was to assess how well the FA and UA scores for each cue type predicted goal judgments. Accordingly, the unit of analyses for these multiple regression analyses was the average judgment for each goal-cue pair. That is, for each pair, an average likelihood judgment was calculated from all participants' responses. This average judgment for each pair was the criterion variable for the analyses. There were 141 cues and 10 goals, which yielded 1410 items: 469 items came from spatio-temporal cues, 161 came from entity–self cues, 361 came from entity–stranger cues, and 419 came from stranger action cues.

Four regression analyses were conducted. In one analyses, the FA score for each cue–goal pair was entered into a regression model predicting goal judgments. A second analysis used a two-step procedure in which the UA score was entered on the first step, and the FA score on the second. Two additional analyses similarly explored the predictive power of the UA score. In one analysis it was the sole predictor of goal judgments; in the second analysts the predictive power of UA was examined controlling for FA. Hence, we obtained results exploring how each cue's UA and FA predicted goal judgments, both when considered in isolation, and when considered in the context of the other score.

Table 2 contains the results of these analyses. The first part of the table shows the simple effects for each measure for each cue type. The second part of the table depicts the unique variance accounted for by each measure as estimated by the simultaneous regression analyses.

6.2.1. Diagnosticity effects

The data in Table 2 show that UA scores for three out of the four cue types were significant predictors of goal judgments. Goal likelihood judgments increased significantly as a function of a cue's diagnosticity for spatial–temporal, entity–stranger, and stranger action cues. Although the predictive power of the UA of each of these cue types was reduced when the FA of each cue was in the model, the UA of the cues

Table 2
β-Weights and R^2 for FA and UA scores for each of the four situational cues for the first and second steps of the hierarchical regression analyses

Cue type	Typicality (FA)			Diagnosticity (UA)		
	β	R^2	F	β	R^2	F
Simple regression (*first step*)						
Spatial–temporal	0.34	0.11	60.58*	0.36	0.13	70.41*
Entity–self	0.23	0.06	9.48*	0.13	0.02	2.56
Entity–stranger	0.44	0.19	86.73*	0.55	0.30	156.97*
Stranger's actions	0.56	0.31	192.96*	0.58	0.34	222.93*
Unique variance explained (*second step*)						
Spatial–temporal	0.21	0.03	17.75*	0.26	0.05	26.76*
Entity–self	0.22	0.04	7.37*	0.06	0.00	0.62
Entity–stranger	0.23	0.04	22.63*	0.44	0.15	82.64*
Stranger's actions	0.28	0.03	22.47*	0.36	0.06	36.83*

* $p < .05$.

continued to predict goal judgments, even when the FA of the cues was accounted for. The UA for the entity–self cues did not predict goal judgments.

Examination of the variance accounted for shows that the cue types have different predictive power, but the nature of the effect varies with the analysis conducted. In the simple regression analyses, the most powerful UA predictors of a stranger's goals (as revealed by the variance accounted for) are the stranger's actions, followed by the stranger's characteristics and the spatio-temporal framework. However, when FA is entered into the model, the UA for the stranger's characteristic category is the most powerful predictor of goal inferences. Finally, the data show that the predictive power of the UA of each cue was reduced when the FA of each cue was in the model.

6.2.2. Typicality effects

The data in Table 2 show that the FA scores for each cue type were significant predictors of goal judgments, regardless of cue type and regardless of the other variables entered into the regression models. As with the UA judgments, however, the exact pattern of results varies with the analysis conducted. In the simple regression analyses, the most powerful FA predictor of a stranger's goals (as revealed by the variance accounted for) is the stranger's actions category, followed by the stranger's characteristic category, the spatial–temporal framework category, and the entity–self category. However, when UA is entered into the model, the FA score accounts for a comparable amount of variance in all cue categories.

6.2.3. Combined effects

The variance explained by the full regression analyses provided estimations of the combined impact of diagnosticity and typicality in predicting goal judgments. Both UA and FA accounted for a significant 16% [$F(2,467) = 45.34$, $p < .05$], 6% [$F(2,159) = 5.04$, $p < .05$], 35% [$F(2,359) = 94.52$, $p < .05$] and 37% [$F(2,417) = 123.16$, $p < .05$] of the variance for spatial–temporal, entity–self, entity–stranger, and

stranger's actions, respectively. For all measures, the combined variance explained by both measures was greater than the unique variance accounted for by the individual measures, suggesting some overlap in the variance accounted for by the measures.

One can also get a sense of the predictive power of a given cue type by comparing results obtained from regression models that include both FA and UA for each of the four cue types. The results of these analyses show that while goal judgments are most powerfully predicted by a stranger's behaviors, such judgments were only slightly better predicted than models that included cues from the entity–stranger category. In comparison, both spatial–temporal cues and entity–self cues were less powerful predictors of goal judgments. However, the informativeness measures associated with the cues included in these less powerfully predictive categories, in particular the spatial–temporal cues, did still predict a respectable amount of variance.

6.3. Discussion

The results of Experiment 2 suggest that the kinds of cues that are derived from situation models designed to explain how people understand narratives can apply to an understanding of a perceiver's goal judgments. The spatial–temporal elements of a situation, the characteristics of the self, the characteristics of the stranger, and the behavior of the stranger all influenced the goal judgments that people made.

However, the results of Experiment 2 also suggest that different kinds of cues have differential abilities to predict goal judgments. In some ways, that such an outcome should occur is not a surprise. Following Heider (1958), many social psychologists believe that behavior is the key to understanding perceiver inferences. This idea was also implied by the results of Experiment 1. However, the data suggest that an over-emphasis on behavior may be a mistake. The data from Experiment 2 suggest that a stranger's characteristics can play a role that is almost as powerful as that played by behavior in a perceiver's inferences about the goals of a stranger.

The data from Experiment 2 also confirmed the notion that the informativeness of a cue for a goal judgment can come from two separate sources: A cue's typicality and a cue's diagnosticity. Theories of judgment need to incorporate both a cue's typicality and diagnosticity into their conceptions of how goal judgments are made from cues. Connectionist-based theories are one of the theory types that have this ability (see Van Overwalle & Labiouse, 2004). Practically speaking, measures that assess only a cue's typicality (Rosch & Mervis, 1975) or diagnosticity (Skowronski & Carlston, 1987) are likely to miss at least some of the variance that is predictable in goal judgments.

The results of Experiment 2 are consistent with Experiment 1's results in that they suggest that diagnosticity and typically may be differentially informative across cue categories. For example, the FAs of entity–self cues were significant predictors of judgments, whereas the UAs of such cues were not. Conversely, for entity–stranger cues the UA scores accounted for substantially more variance that FA scores. The same trend was found for spatial–temporal and stranger–action cues, but the differences in the variance explained between the two measures of informativeness were marginal.

7. Experiment 3

The goal judgments in Experiment 2 were derived from an experiment that was relatively impoverished with respect to the presences of cues: participants encountered only one cue at a time. More important are cases in which participants have access to multiple cue categories: It is these settings that come closest to reality. After all, in the real world, prior to encountering the behaviors of a stranger, one is in a time and place, one is in a self-state, and one may see the physical characteristics of the stranger. Accordingly, in Experiment 3 participants were sometimes asked to make goal judgments when two or more of these cues were present.

One goal in Experiment 3 was, again, to examine whether, and the extent to which, cues that were exemplars of the cue categories derived from situation models predicted goal judgments. The results of Experiment 2 suggested that these cues did predict goal judgments, although some cues were more powerful predictors than others. The results of Experiment 3 will allow us to see if these results replicate in a context in which cues are presented in combination rather than singly.

There is reason to believe that they may not. For example, if it is the case that "behavior engulfs the field," (Heider, 1958), in the context of behavioral information non-behavior cues may play only a minor role in goal judgments made about a target. Moreover, it may be the case that the addition of any new cue, behavioral or otherwise, to a cue combination that already is indicative of a given goal may have only a limited impact on judgments, regardless of cue type. That is, a highly informative cue might produce a 90% match to the mental representation of a goal when the cue is presented by itself, but may add only a limited amount to that match when four other cues have themselves already produced a 90% match to the goal (i.e., there is only 10% left to account for) have already been encountered. This notion is often implemented in connectionist and PDP approaches that model the extent to which a cue array matches a category (see Van Overwalle & Labiouse, 2004).

A second purpose of Experiment 3 was to again verify the idea that diagnosticity and typicality both independently predict goal judgments. Such independence suggests that goal judgments can be better predicted from regression models that incorporate both measures of diagniosticity and typicality than from models that include only one of those two variables.

A third purpose of Experiment 3 was to see whether any differences in the power of cues from different cue categories to influence goal judgments can be entirely accounted for by the typicality and diagnosticity differences among cues, or whether cues differ in their ability to influence goal judgments even when differences in the typicality and diagnosticity of different cue types are accounted for. This latter possibility is suggested by studies suggesting that variables such as the attention paid to cues (e.g., Fiske, 1980) or the extent to which people can construct an intuitive theory about the situations and goals (e.g., Wright & Murphy, 1984) are related to the extent to which cues affect judgments. This possibility would also lend credence to the argument that situation model theories developed in the context of text comprehension can also account for event understanding (Copeland et al., 2006; Magliano et al., in press, 2005).

7.1. Methods

7.1.1. Participants

One hundred and thirty Northern Illinois University undergraduates enrolled in an introductory psychology course participated in the experiment. They earned course credit for their participation.

7.1.2. Cue sets

Each scenario presented to participants in Experiment 3 contained from one to five cues. Each cue reflected a different category of information (time, location, state of the self, stranger's behavior, stranger's appearance). The cues always appeared in a set order, with location and time cues appearing first, then the self cue, then stranger's appearance cue, then the stranger's behavior cue. This order was chosen because it approximates the order in which people typically encounter different cue types in a given situation. The ordinal relations among the cues were preserved on all trials, even those on which some cues were not presented (e.g., time cues always came before self, appearance, and behavior cues, self cues always came before appearance and behavior cues, etc.)

In response to each cue set, people judged, on the same five-point scale used in Experiment 2, the likelihood that the stranger depicted in the scenario had a given goal. There were 10 goals used in the experiment. Every participant made judgments about the goal of 123 targets, each of which was judged with respect to the same goal.

Each target was described by a different cue set. The cue sets are schematically depicted in Appendix B. To construct these sets, we first classified cues within each cue type as *high*, *low*, or *null* in informativeness. To obtain items for the *high* and *low* categories, we first calculated an informativeness score by computing an average of the FA and UA scores. For each of the 10 goals, we selected cues that had an informativeness score greater than zero (i.e., produced as a cue associated with a goal in Experiment 1). We then conducted a median split based on the informativeness scores and assigned cues to the *high* or *low* category based on the results of the split. This procedure was done separately for each cue category. Cues in the null category for a given goal were not produced as cues for that goal in Experiment 1 and therefore, had an informativeness score of 0. Only those cues for which it was possible for them to co-occur with a given goal were considered for inclusion in the materials for Experiment 3.

We emphasize that because the splits were done separately within each category type, the *high* and *low* classifications are not equivalent across cue types. That is, an average score that would cause a temporal cue to be classified as *high* might cause a stranger's behavior cue to be classified as *low*. To illustrate this point, the average combined FA and UA scores for each cue category for the selected cues are shown in Table 3. Because of this non-equivalence, these classifications were used only to construct the cue sets, and were not used in the data analyses. Those analyses used the diagnosticity and typicality scores calculated for the cues, and not the average of the two or the category to which the cue was assigned, to predict goal judgments.

Table 3
Mean average informativeness score [(relative frequency + uniqueness)/2] as a function of situational dimension of the cues used in Experiment 3

Situational dimension	Degree of informativeness (FA + UA)/2		
	High	Moderate	Null
Location	.60	.40	0.00
Time	.34	.15	0.00
Entity–self	.44	.21	0.00
Entity–stranger	.64	.37	0.00
Stranger's actions	.69	.41	0.00

Nonetheless, using these informativeness scores allowed us some measure of control over the types of cues that were combined to make the cue sets. This control was even more important given there were simply too many possible Cue Number (1–5) × Cue Type (time, location, state of the self, stranger's behavior, stranger's appearance) × Cue Level (high, low, or absent) combinations for a participant to make judgments about all possible cue sets. Because we could not present all possible cue combinations, we constructed two subgroups of cue sets. We reasoned that the two separate subgroups of cue sets would allow us to examine hypotheses about the extent to which goal judgments were affected by the number of cues that were presented, the type of cue presented, and the cue's diagnosticity and typicality. These subgroups are presented in Appendix B.

The cue sets in this first subgroup (item sets 1–93 in Appendix B) were intended to allow examination of the extent to which a given cue type affected goal judgments in the context of a differing number of other cue types. For example, it might be the case that action cues are very important to goal judgments when they are the only cue presented, but are much less important in the context of the other four cue types. Accordingly, cue sets in this subgroup were constructed so that each cue contained from one to five cues on each trial. For each goal, we selected five cues, one for each of the five categories of cues (location, time, entity- self, entity–stranger, stranger's actions). This was done for each level informativeness (high, low, null). Within each of the three informativeness levels, all possible combinations of cue categories were presented. As can be seen in Appendix B, the application of these rules yielded 93 possible Cue Number × Cue Type combinations.

The second subgroup of target descriptions (item sets 94–123 in Appendix B) was designed to provide more precise information about the relative impact of a cue's diagnosticity and typicality on goal judgments when all five of the elements that are typically present in an interaction scenario are present (which some would argue is the most ecologically valid case). Accordingly, as illustrated by comparing items sets 94 and 95 in Appendix B, some cue sets held constant the average informativeness level of the location, time, self, and stranger's appearance cues, but varied the informativeness of the stranger's action cues. Note that sets 94 and 95 reflect only two of the possible 3 levels of the stranger's action variable: the third level was already generated in the first subset of cues (see item set 31). The systematic

manipulation of cue level across all cue types yielded 33 cue type combinations that had not already been generated in the first subset.

While it would have been best to randomly select from our pretested pool of cues to fill out these 123 cue sets for each of the 10 goals, the nature of some of the cues did not allow the use of such random selection. For example, some of the cues were directly contradictory to other cues or to the goal under consideration (e.g. the location "bad neighborhood" and "you are happy" may seem contradictory); other cue combinations did not make any sense to present together (e.g., the location "classroom" and "stranger is wearing a swim suit" are not likely to coincide). Hence, potential cue combinations were rejected if they were contradictory or did not make sense in combination.

7.1.3. Procedure

The same basic procedures and measures that were used for Experiment 2 were adopted for Experiment 3. The instructions were quite similar, as well. The only major change to the instructions used in Experiment 3 was that participants were told that they would be given descriptions requiring that they consider multiple cues at once, and that the number of cues present would vary from trial to trial. They were also told that many of the sets of cues would seem similar and that it was important to carefully consider all of the cues presented to them on any given trial.

7.2. Results

7.2.1. Preliminary data reduction

To simplify data analyses, the responses for each participant were averaged with the responses of those other participants who judged the same targets with respect to the same goals. Thus, the criterion variable in all the regressions was an average judgment for a given target. Readers should note the implications of this type of dependent measure. In the usual regression analyses, variance is obtained by examining responses across participants. That is not the case in the present analyses, in which variance is obtained by examining responses across targets. We also note that goal type was treated as a between-participants variable (participants received only one goal) and is controlled for in all analyses.

Our preliminary data set was composed of 1230 observations: average judgments made about 123 targets for each of 10 goals. However, these were analyzed in two separate sets of analyses, each of which analyzed only a subset of the data. The first subset allowed an examination of how cue number affected goal judgments; the second set focused only on how judgments were affected when all five cue types were used to describe a target and a situation.

7.2.2. Analyses involving the cue number variable

7.2.2.1. The data sets and analyses overview. Regression analyses were conducted on each of five data sets, each designed to separately explore effects within each cue type. Each analysis examined judgments for trials in which a given cue type was a part of

the stimulus set. Each data set contained 780 average judgments that were influenced by a given cue type.

Two of the predictors of each judgment were the typicality (relative frequency) and diagnosticity (uniqueness) scores for the cue that contributed to the judgment. A third predictor in the model was the number of events that contributed to each judgment (1–5). A final predictor in each model was the goal that was judged. Because we had no hypotheses about between-goal differences, we do not report such effects, other to note that such effects were always significant [smallest $F(9, 764) = 13.71$, $p < .0001$]. More important to our purposes, however, is that the other significant effects that we report control for between-goal effects.

We analyzed each data set in two ways. The first of these involved relatively simple regression models that contained only two predictors. One of these was always the goal that was judged and the other was one of the three main predictors of interest (cue number, diagnosticity, typicality). The objective of these analyses was to explore whether each of the variables predicted goal judgments without controlling for other effects (except for goal). The second set of analyses were run as simultaneous regressions in which the relation between each predictor and goal judgments were examined in the context of all the other predictors of goal judgments.[1]

Results from both sets of models are necessary to appropriately interpret effects. For example, an effect in the simultaneous regressions might not be significant for either of two reasons: (1) the effect simply does not predict judgments, or (2) the variance accounted for by the effect overlaps with that accounted for by another effect. In the former case, the effect would be significant in neither set of analyses; in the latter case the effect would be significant in the simple analyses but not in the simultaneous analyses.

7.2.2.2. Diagnosticity effects. One of the issues explored in the experiment was whether the relation between a given cue and goal judgments depended on that cue's measured uniqueness (diagnosticity). The data presented in Table 4 (top half) show that when the relation between diagnosticity and goal judgments was evaluated without taking the impact of typicality and number of cues into account, diagnosticity always robustly, and positively, predicted goal judgments. However, the relation between the diagnosticity of a cue and that cue's impact on judgments was not consistent across all cues. The diagnosticity of the stranger's behavior and appearance was more strongly related to goal judgments (e.g., accounted for more variance) than the diagnosticity of the time, place, and self cues.

As reflected in the bottom half of Table 4, the ability of a cue's diagnosticity to predict goal judgments was dramatically less robust when the predictive impact of diagnosticity was evaluated in the context of the predictors of typicality and cue number. In fact, when these other variables were taken into account, diagnosticity

[1] We also conducted various hierarchical regression analyses (e.g., Cohen & Cohen, 1983) to explore interactions among the three predictors of interest. While some interactions did emerge, they tended to do so inconsistently across analyses and accounted for relatively little variance. Hence, we omit discussion of the results of these analyses.

Table 4
Experiment 3: Analyses with varying cue numbers showing that a cue's diagnosticity predicts its impact on goal judgments

	Cue type				
	Time	Place	Self	Stranger's appearance	Stranger's behavior
Simple regressions					
β	.51	.46	.50	.60	.67
R^2 change	.24	.17	.23	.37	.43
$F(1,769)$	273.19*	179.03*	266.06*	536.59*	736.87*
Simultaneous regressions					
β	.31	.06	.37	.50	.47
R^2 change	.0003	.0006	.04	.11	.13
$F(1,764)$	1.01	.73	67.19*	204.21*	280.09*

* $p < .05$.

did not predict goal judgments for time and place cues. This pattern of results is indicative of the fact that there was considerable overlap among the predictors in the goal judgment variance accounted for by each.

7.2.2.3. Typicality effects. One of the issues explored in the experiment was whether the relation between a given cue and goal judgments depended on that cue's measured relative frequency (typicality). The data presented in Table 5 (top half) show that when the relation between typicality and goal judgments was evaluated without taking the impact of diagnosticity and number of cues into account, typicality always robustly, and positively, predicted goal judgments. However, the relation between the typicality of a cue and that cue's impact on judgments was not consistent across all cues. The typicality of the stranger's behavior and appearance was more strongly related to goal judgments (e.g., accounted for more variance) than the time, place, and self cues.

Table 5
Experiment 3: Analyses with varying cue numbers showing that a cue's typicality predicts its impact on goal judgments

	Cue type				
	Time	Place	Self	Stranger's appearance	Stranger's behavior
Simple regressions					
β	.51	.52	.48	.55	.64
t R^2 change	.24	.20	.20	.27	.36
$F(1,769)$	272.61*	218.67*	216.95*	320.83*	525.99*
Simultaneous regressions					
β	.20	.47	.17	.15	.34
R^2 change	.0003	.03	.01	.01	.06
$F(1,764)$	.41	43.24*	13.37*	15.32*	127.51*

* $p < .05$.

However, as reflected in the bottom half of Table 5, the ability of a cue's typicality to predict goal judgments was dramatically less robust when the predictive impact of typicality was evaluated in the context of the predictors of diagnosicity and cue number. In fact, when these other variables were taken into account, typicality did not predict goal judgments for time cues. This pattern of results is indicative of the fact that there was considerable overlap among the predictors in the goal judgment variance accounted for by each.

7.2.2.4. Combined effects. As was the case with Experiment 1, the variance explained by the combination of UA and FA was significantly greater than the unique variance explained by both measures. The combination of UA and FA accounted for a significant 24% [$F(2,777) = 127.87$, $p < .05$], 20% [$F(2,777) = 61.83$, $p < .05$], 24% [$F(2,777) = 120.22$, $p < .05$], 38% [F (2,777) = 210.15, $p < .05$] and 49% [F (2,777) = 331.50, $p < .05$] of variance in goal judgments for time, space, entity–self, entity–stranger, and stranger's actions, respectively. These analyses also provide evidence with respect to the relative informativeness of the different cue categories. Cues relating to the stranger's behaviors accounted for more variance than cues in the entity–stranger category. Accordingly, cues from the entity–stranger category accounted for more variance than cues in either the time, space, or entity–self categories, which were comparable in their judgmental impact.

7.2.2.5. Cue number. One of the issues explored in the experiment was whether the relation between a given cue and goal judgment depended on the number of other cues that were presented along with that cue. Unsurprisingly, the data presented in Table 6 show that cue number always predicted goal judgments, regardless of whether the effect was examined in simple or simultaneous regression models.

However, the pattern of results depicted in Table 6 shows that the cue number variable accounted for much more variance for some cues (time, place, self) than for others. This pattern suggests that some cues are more important to goal judgments than other cues. The data in Fig. 1 make this quite clear. With all five cues

Table 6
Experiment 3: Goal judgments' increase with increases in cue number; mean judgment (adjusted for other variables entered in the simultaneous regression model) for each cue type for cue numbers ranging from 1 to 5 are presented in Fig. 1

	Cue type				
	Time	Place	Self	Stranger's appearance	Stranger's behavior
Simple regressions					
R^2 change	.19	.18	.19	.10	.02
$F(1,766)$	50.99*	45.77*	50.43*	23.17*	5.13*
Simultaneous regressions					
R^2 change	.19	.18	.19	.10	.02
$F(1,764)$	76.17*	62.97*	75.79*	43.71*	12.04*

* $p < .05$.

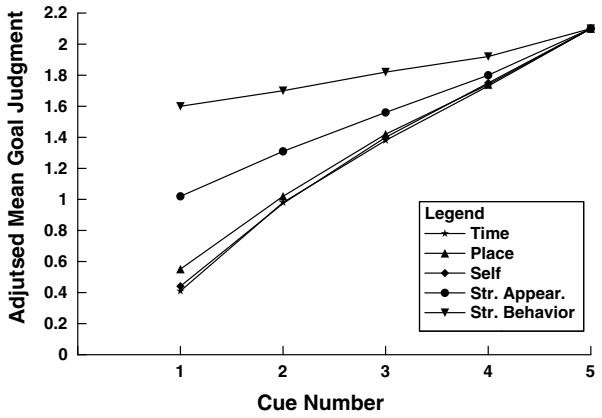

Fig. 1. Least-squares means (e.g. means adjusted for other variables entered into the regression) depicting how judgments were affected when a given cue type was included in the set of cues presented (adjusting for the cue's typicality and diagnosticity).

in the model, the average goal judgment (adjusted for other terms in the regression model) is 2.10. The goal judgment mean decreases as one deletes a cue, so that only four cues are in the array. However, it decreases *more* when one of the cues in the array is *sure to be* a time cue than when one of the cues in the array is sure to be a stranger's behavior cue. This trend continues as cues are removed, so that when only one cue is presented there are relatively large differences in the extent to which the different cue types predict goal judgments. Moreover these effects occur in simultaneous regression models in which the diagnosticity and typicality of the key fixed cue type are accounted for. Hence, these differences in the relation between cue type and goal judgment *cannot* occur because some cues have greater diagnosticity and/or typicality values than other cues; they occur *even when these differences are taken into consideration*.

7.2.3. Analyses comparing predictive power of diagnosticity and typicality scores across cue types

7.2.3.1. The data set. Regression analyses were conducted on the data set containing only trials that contained goal judgments made when all five cue types were present. There were 330 such trials. The predictors of interest on these trials were the diagnosticity and typicality values for each of the five cue types.

We analyzed each data set in two ways. The first of these involved relatively simple regression models that contained only two predictors. One of these was always the goal to be judged (which was always a significant predictor, but which is not of theoretical interest in this paper), and the second was a diagnosticity or typicality value for one of the five cue types. The objective of these analyses was to explore whether each of the variables predicted goal judgments without controlling for other effects (except for goal). The second set of analyses were run as simultaneous

regressions in which the relation between each predictor and goal judgments were examined in the context of all the other predictors of goal judgments.

7.2.3.2. Diagnosticity effects. One of the issues explored in the experiment was whether the relation between a given cue and goal judgments depended on that cue's measured uniqueness (diagnosticity). The data presented in Table 7 (top half) show that when the relation between the diagnosticity of a single cue and goal judgments was evaluated without accounting for the diagnosticity and typicality of the other cues, diagnosticity always robustly, and positively, predicted goal judgments. This effect is especially impressive given that these effects reflect the impact of the diagnosticity of a cue on goal judgments in the context of the four other cues that were present. The fact that this effect emerged for all five cues suggests that all five are taken into consideration when making goal judgments. However, the relation between the diagnosticity of a cue and that cue's impact on judgments was not consistent across all cues. The diagnosticity of the stranger's behavior and appearance was more strongly related to goal judgments (e.g., accounted for more variance) than the time, place, and self cues.

The ability of a cue's diagnosticity to predict judgments was weakened when the typicality and diagnosticity of the other cues were added to the regression model. As reflected in the bottom half of Table 7, only the diagnosticity of the stranger's appearance and the diagnosticity of the stranger's behavior predicted goal judgments independently of the other cues. The effect for the space cue was also statistically significant, but that effect is discounted. As reflected in Table 7, the sign of the diagnosticity beta for the space variable changed from the simple to the simultaneous model, suggesting the action of a suppressor variable. Overall, this pattern of results is indicative of the fact that there was considerable overlap among the predictors in the goal judgment variance accounted for by each.

7.2.3.3. Typicality effects. A similar story emerges when one examines the extent to which the typicality of each cue predicts goal judgments. The data presented in

Table 7
Experiment 3: Analyses of judgments based on all 5 cues showing that a cue's diagnosticity predicts its impact on goal judgments

	Cue type				
	Time	Place	Self	Stranger's appearance	Stranger's behavior
Simple regressions					
β	.44	.43	.46	.58	.63
R^2 change	.19	.15	.20	.32	.38
$F(1,319)$	87.69*	64.93*	97.23*	191.33*	247.41*
Simultaneous regressions					
β	.13	−.25	−.03	.21	.28
R^2 change	.00	.009	.00	.002	.04*
$F(1,310)$	0.11	9.29*	.24	16.41*	42.27*

* $p < .05$.

Table 8
Experiment 3: Analyses of judgments based on all 5 cues showing that a cue's typicality predicts its impact on goal judgments

	Cue type				
	Time	Place	Self	Stranger's appearance	Stranger's behavior
Simple regressions					
β	.45	.48	.44	.52	.59
R^2 change	.19	.17	.17	.23	.31
$F(1,319)$	88.40*	79.58*	79.57*	115.47*	168.00*
Simultaneous regressions					
β	−.01	.29	.16	.11	.25
R^2 change	.000	.01	.01	.004	.03
$F(1,310)$	0.00	12.00*	8.09*	4.54*	33.09*

* $p < .05$.

Table 8 (top half) show that when the relation between the typicality of a single cue and goal judgments was evaluated without accounting for the diagnosticity and typicality of the other cues, typicality always robustly, and positively, predicted goal judgments. This effect is especially impressive given that these effects reflect the impact of the typicality of a cue on goal judgments in the context of the four other cues that were present. The fact that this effect emerged for all five cues suggests that all five are taken into consideration when making goal judgments. However, the relation between the typicality of a cue and that cue's impact on judgments was not consistent across all cues. The typicality of the stranger's behavior and appearance was more strongly related to goal judgments (e.g., accounted for more variance) than the time, place, and self cues.

The ability of a cue's typicality to predict judgments was weakened when the typicality and diagnosticity of the other cues were added to the regression model. As reflected in the bottom half of Table 8, while only the typicality of the time cues failed to predict goal judgments in the simultaneous regression, the amount of variance accounted for was dramatically reduced compared to the amount of variance accounted for by each of the cue categories in the simple regression models. This pattern of results is again indicative of the fact that there was considerable overlap among the predictors in the goal judgment variance accounted for by each.

It is important to note that the unique variance accounted for by UA and FA for each cue category was considerably less than the variance accounted for by all five cue categories. That is, the UA and FA scores for all five categories accounted for 64% of the variance [$F(10,1219) = 218.06$, $p < .05$].

7.3. Discussion

One goal in Experiment 3 was to examine whether, and the extent to which, cues that were exemplars of the cue categories derived from situation models predicted

goal judgments. The results of Experiment 3 suggested that these cues did predict goal judgments, although some cues were more powerful predictors than others. Behavior cues were particularly powerful predictors of goal judgments.

Some of this predictive power comes from the fact that behavior cues tend to have greater diagnosticity and typicality values than other cues. To the extent that a cue's information value tends to come from these sources, it makes sense for cues that rate high on these attributes to have particularly strong influences on judgments.

However, that does not tell the whole story. For example, in Experiment 3 it is the case that variation in the diagnosticity and typicality of behavior cues was more important to judgments than variation on these attributes for the other cue types. One way to understand this is that it is as if people paid more attention to the information value of behavior cues than to the information value of other cues. While attention may not be the mechanism underlying this effect, that mechanism is suggested by studies of narrative processing indicating that agents and their behaviors are more strongly attended to than other story elements (Magliano et al., in press, 2005; Scott-Rich & Taylor, 2000).

Attention may also be implicated in the fact that behavior cues had more impact on goal judgments than other cues, even controlling for the diagnosticity and typicality of the cues. The attentional mechanism may again help to explain this finding. Research in social cognition suggests that information that captures attention is related to heightened impact of that information on social judgments (e.g., Fiske, 1980).

While it is the case that behaviors did have greater impact on judgments than other cue types, it would be going too far to suggest that our study shows that "behavior engulfs the field." When behaviors are presented alone, they can be very powerful predictors of goal judgments. However, they add only a limited amount of additional variance when numerous other cues are already available and predict the same goal implied by the behavior. This result would seem to dovetail well with the kinds of judgment patterns that are predicted by connectionist models. In fact, such models have already done relatively well in predicting patterns that emerge for other kinds of social judgments, such as trait attributions (see Van Overwalle & Labiouse, 2004).

These connectionist approaches also make allowances for the fact that cues can have information value for either of two reasons: their typicality and their diagnosticity. A final purpose of Experiment 3 was to verify the idea that diagnosticity and typicality both independently predict goal judgments. Our results were moderately supportive of this idea. The analyses that were performed on judgments made from the five-cue sets showed that a cue's typicality and diagnosticity sometimes predicted judgments, even when the diagnosticity and typicality of all the other cues were accounted for. This was especially true for the diagnosticity and typicality of the stranger's appearance and behavior cues.

However, some might note that this was a highly conservative test of the independence hypothesis, for it evaluates independence of typicality and diagnositicity with regard to the same cue as well as the independence of that cue's typicality or diagnosticity with respect to the typicalities and diagnosticities of other cues. Perhaps

a more appropriate test was provided in the analyses conducted on judgments made from varying number of cue sets. That analysis examined the impact of a cue's diagnosticity controlling only for that cue's typicality (and vice versa). In these analyses, only time cues showed strong evidence of a lack of independence; for all other cues, either typicality or diagnosticity (or both) predicted goal judgments in the simultaneous regressions.

Why might time cues have evinced this lack of independence? It may have been an artifact of the time cues that we chose. Correlational analyses suggested that the typicality and diagnosticity values were quite highly correlated for these cues ($r = .99$). If subsequent studies can do a better job of breaking this correlation, even time cues may yield evidence of the independent effect of a cue's typicality and diagnosticity on goal judgments.

8. General discussion

In this research we explored factors that facilitate goal inferences in the context of social interactions with a stranger. Although goals have a great deal of explanatory and predictive power with respect to behavior (e.g., Trabasso et al., 1989), there has been relatively little research to explore the extent that the situational contexts constrain the process of inferring the goals of others. We adopted a mental model perspective of social inferences that was largely influenced by situation model theory from the discourse processing field (e.g., Zwaan & Radvansky, 1998).

One characteristic of social interactions is that they occur in larger situational contexts that involve, time, place, actors, and their behaviors. We assumed that this context constrains the inference processes, as is the case when experiencing fictive social interactions (Graesser et al., 1997). The results of the research described in this article support this situated view of social inferences. That is, the spatial–temporal frameworks, the state of the self, the stranger's appearance, and the stranger's actions are all informative of the stranger's goal. However, the present study illustrates that various aspects of the context have differing degrees of informativeness with respect to the goals of a stranger.

The appearance of a stranger and behaviors of a stranger were used in a manner suggesting that people viewed these cues as highly informative of the stranger's goal. However, the spatial–temporal framework and the state of the self prior to the interaction also had a significant impact on goal judgments. In fact, behaviors contributed relatively little to goal judgments when they occurred in a fully situated context. Thus, this study suggests that it is somewhat inaccurate to say that behavior engulfs the field (Heider, 1958). While behaviors do tend to have considerable power to predict goals, it is also the case that behaviors occur in the context of a field or situation, and are interpreted as such (see also Barker, 1968; Barker & Wright, 1955).

In this context, it is interesting to note that some streams of social psychology research continue to focus on behaviors relatively independently of the context in

which those behaviors occur. For example, there is a large body of recent work examining the spontaneity of trait inferences that heavily relies on behavior descriptions that are presented relatively free of context (for an example see Carlston & Skowronski, 2005). Similarly, many studies of explicit social judgment have used behavior descriptions that are similarly unconstrained by context (e.g., see Skowronski, 2002).

Other recent work has broken from this context-free behavior focus. For example, recent approaches to the attribution process suggest that situational information has an impact on attributions early in information processing by altering behavior interpretation (Trope & Gaunt, 1999) An additional example comes from research on jury decision-making in real world contexts suggesting that people use a diverse array of multiple cues to try to "make sense" of a crime. This sense-making activity is often viewed in the context of story-telling activity (e.g., Pennington & Hastie, 1993), which provides a natural link to the studies described in the present paper. A similar emphasis on the use of situational cues comes from research using parallel processing concepts that focus on issues of explanatory coherence (e.g., Kunda & Thagard, 1996; Read & Miller, 1993). Clearly, while still presenting relatively impoverished stimuli, the results reported in the present research are more philosophically aligned with such studies, as well as with the ecological view of decision-making pioneered by Brunswik (1956) and more recently championed in social psychology by Zebrowitz (e.g., Zebrowitz & Monteparre, 2006).

This perspective is also quite consistent with theories of narrative comprehension (Graesser, Singer, & Trabasso, 1994; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). For example, the event-indexing model (Zwaan & Radvansky, 1998) assumes that narrative situations are comprised of a spatial–temporal framework, entities and agents, and linking relations, such as causality. The event indexing model assumes that a reader's inferences are constrained by his or her understanding of the dynamically unfolding situation. The results of the current study strongly suggest that social inferences are similarly constrained by the situational cues present at the time of the interaction. The results of this study also suggest that theories of situation model construction are general theories of event understanding and are not specific to text comprehension (Copeland et al., 2006; Magliano et al., 2001; Magliano et al., in press, 2005). However, an even stronger test of this idea would involve having participants actually engage in social interactions that examine goal inferences made from a stranger's behavior in the presence of various sets of situational cues.

We adopted a theoretical rationale suggesting that the utility of cues for judgments was, in part, a function of both a cue's typicality and diagnosticity. We assume that people possess mental representations of events, and that these representations contain situational cues that are probabilistically associated with other elements of these representations (e.g., goal inferences). The informativeness of a cue for a goal judgment comes from these probabilities, which are derived from computations of diagnosticity and typicality. We consequently adopted a methodology that allowed the simultaneous computation of measures reflecting a cue's diagnosticity and typi-

cality for every goal-relevant cue under consideration in this study, and that could also simultaneously evaluate the independent impact of these variables on goal judgment. This is one of the novel aspects of this study. That is, few, if any, studies exist that have *simultaneously* assessed the role of diagnosticity and typicality in categorical judgments or social inferences, examining the relative impact of each while controlling for the other.

Indeed, the results of all three experiments suggest that the informativeness of a cue is a combination of both its diagnosticity (i.e., UA score) and typicality (i.e., FA score). However, while they have independent relations to goal judgments, the combined impact of both measures was almost always greater than their individual contribution for all cue categories. This was particularly striking in Experiment 3, in which multiple cues were presented. Thus, these findings suggest that although diagnosticity and typicality carry unique variance, they also account for considerable shared variance in goal judgments. Such results are consistent with judgment models, such as connectionist models (see Van Overwalle & Labiouse, 2004), that are able to simultaneously account for the influence both dimensions of cue informativeness.

The results of this study also suggest that diagnosticity and typically may have differential relationships to goal judgments based on cue category. Diagnosticity carried more predictive weight than typicality for stranger appearance and actions, but not for spatial–temporal and entity–self cues. This could imply that there are factors other than diagnosticity and typicality that could constrain goal judgments. Indeed, the event indexing model (Magliano et al., 1999; Zwaan & Radvansky, 1998) assumes that readers attend to and represent changes in some dimensions more closely than others. For example, they routinely monitory shifts in causality, but only monitor spatial relationships when they have the explicit goal to do so (see Zwaan & Radvansky, 1998, for an extensive review). It may be the case that individuals intuitively attend to some dimensions more closely than others when making social inferences. Additionally, a cue's salience in the context could have implications on how much weight is placed upon its diagnosticity and typicality with respect to a goal inference. As such, the stranger and his/her actions may be more salient in these hypothetical interactions than the spatial–temporal framework or the state of the self. This again suggests the need to examine these issues in the context of real interactions instead of in hypothetical form.

It is important to note that the relative importance of the cues (and their salience) may be influenced by how they are presented. This is particularly important to consider with respect to Experiment 3, in which the cues were presented in combination. It was not possible to present all cue combinations, nor was it tenable to present all possible orderings with the limited set that we chose for this experiment. The cues were presented in an order that we believed reflected the order in which people would become aware of them as an interaction unfolds. That is, we assumed that people would first be aware of the spatial–temporal framework and how they are feeling. As the interaction unfolds, they would become aware of the appearance of the stranger, and his/her actions. In this regard, it is important to note that the results of

Experiment 3 illustrate that the stranger's appearance and action carry more weight than the other cues, which suggests that the results of Experiment 3 are not driven by the primacy effects that so often characterize studies using similar methodologies (Hogarth & Einhorn, 1992). Nonetheless, the possibility of an experiment's results being caused by order effects is another argument favoring a move to real-life interactions. Similarly, the data also suggests that recency effects were minimal – if they were present, then adding behavior cues to the four other cues that were already present in the cue array should have had a larger effect than in evidenced in Fig. 1.

However, even in real life, cue ordering might be important and not artifactual. In real life, a prior context may "set the stage" for behavior, facilitating interpretations and inhibiting others. This is what Heider (1958) implied when he wrote that "the ambiguity of behavior as a local stimulus is reduced when it is seen in a situational context." The results of Experiment 3 support this claim. As context provides increasing support for a given goal inference, later-occurring cues have less of an impact on judgments than they otherwise would have if they occurred alone. The data in Table 6 are quite striking in this regard: Stranger actions carried relatively little predictive weight when they occurred in a fully situated context, whereas they were the strongest predictor of a stranger's goal when they occurred alone. Again, this supports the conclusion that behavior does not engulf the field, but is rather part of it.

9. Implications for future research

The approach we have taken affords several directions for future research. One important direction would be to assess whether goal inferences are generated "online" while engaged in social interactions. There is a substantial body of evidence to suggest that they are when reading (e.g., Hastie & Park, 1986; Suh & Trabasso, 1993) or viewing (Magliano et al., 2005) a narrative. Of particular interest would be the extent to which different cues facilitate, or are sufficient for producing, goal inferences. The present data suggests that spatial–temporal frameworks may narrow the field for potential goals, but they do not provide enough sufficient context to support a specific inference, which contradicts Barker and Wright's (1955) claim that location highly constrains behavior. On the other hand, our data suggests that actions are highly informative of a goal and should be sufficient to elicit the inference. Such research would be necessary to develop a processing model of goal inferences.

Another direction of interest would be to explore circumstances in which cues conflict. As noted in the Method section, we constructed cue sets so that the cues were not inconsistent with each other. However, in the real world, cues are not always so consistently aligned. What kinds of effects do such misalignments have on the inference process? Useful insight into such questions might come from incorporating such notions as the story model of judgment (e.g., Pennington & Hastie, 1993) or constraint satisfaction models (Kunda & Thagard, 1996) into

the research program. These models have differing implications for how incongruent cues might affect the judgment process. The story model suggests that people might counterargue or discount cues that do not fit a mental representation developed to accommodate ideas about an actor's cues. Alternatively, constraint satisfaction models suggest that the presence of such incongruous cues might sometimes cause a perceiver to substantially alter the goal inferences that would be considered to be plausible.

Addressing both of these research questions is contingent on having a well-developed semantic space. Encouragingly, the semantic space that we developed for this study arguably captures only a subset of the knowledge that supports goal inferences and yet, using this potentially constrained space, we were able to account for a large amount of variance in goal likelihood judgments.

Nonetheless, it is both empirically and theoretically important to examine the extent to which measures derived from the explicit knowledge represented in this space can account for less constrained inferential processes. The literature elsewhere provides reasons to be optimistic in this regard. For example, McRae and colleagues used an approach similar to ours in order to create a semantic space for word meaning that involved having participants list features associated with a large set of concepts. (McRae, de Sa et al., 1997; McRae, Ferretti et al., 1997). This space enabled them to derive measures capturing the relationships of the features within and across concepts. Importantly, these measures were not only predictive of conscious judgments about the features and concepts in the space (e.g., typicality judgments; McRae, Cree, Westmacott, & de Sa, 1999; McRae, de Sa et al., 1997; McRae, Ferretti et al., 1997), but also predicted semantic priming of a concept given the presence of a feature (McRae, Ferretti et al., 1997; see also Randall, Moss, Rodd, Greer, & Tyler, 2004).

As noted in the discussion of Experiment 1, it is also useful to be cognizant of two additional limitations. One of these is that the semantic space derived by the procedure used in Experiment 1 may partially reflect naïve theories about the situations that co-occur with the goals chosen for the study. A second is that the semantic space developed in Experiment 1 reflected information that was explicitly available in working memory and codeable in language. Accordingly, it is an open question as to whether those same cues would be generated, and would have similar diagnosticity and typicality values (and similar goal-predictive abilities), in situations that people actually experienced. While such issues limit the generalizability of the current studies, they also suggest research opportunities. Such studies could systematically manipulate the nature of situational cues present in various settings (computer simulations, videotaped presentations, or even real-life interactions) to explore whether those situation cues were related to goal inferences in the same way that we observed in our experiments. Such experiments would fit in well with a Brunswik's (1956) notion of an ecological approach to social judgment.

One additional direction for future research could be cross-cultural assessments of cues during goal inferences. Misunderstandings in multi-cultural interactions might be based on different interpretations of cues. It might not be surprising that

people from different cultures learn different concepts of social interactions and appropriate behaviors in the process of enculturation, but a detailed study of different cues in ambiguous multicultural settings could shed light on the question why exactly misunderstandings occur between members of specific cultural groups.

While such theoretically driven studies are of obvious interest, it is also important to focus on the practical issues involved in making judgments about a stranger's goals. Such judgments must be made frequently, often with little time to think about or reflect on the information that is available. Police officers are often placed into exactly such situations, and need to make rapid decisions about a stranger's intents prior to either helping the stranger or shooting him (see Correll, Park, Judd, & Wittenbrink, 2002). Hence, it is not an exaggeration to say that understanding the goals of another person can be a matter of life or death. It is hoped that the research described in the present article can in some way contribute to an understanding of how such judgments are made and how critical errors based on faulty judgments can be prevented.

Acknowledgements

We would like to thank three anonymous reviewers for their helpful comments on an earlier version of this manuscript. This work was funded by Sandia National Laboratories. Sandia National Laboratories is a multi-program laboratory operated by the Sandia Corporation, a Lockheed-Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

Appendix A

Imagine someone you do not know approaches you to <**GOAL**>. Think of one possible situation. We would like you to image the encounter in great detail and answer the following questions about it. Please consider each question very carefully and list all the possibilities. For some of these items you may not have anything to fill in. In those cases, please write "not relevant" (NR).

Part 1: Please think about the location of this situation and answer the following questions:

1. Where are you when the stranger approaches? Describe the location.
2. What time is it?
3. What day is it?
4. What season of the year is it?
5. Are you alone? If not, who may be with you?
6. Is the stranger alone? If not, who may be with the stranger?

7. What other activities or events are going on around you?
8. Please list anything that might have happened immediately prior to the event?
9. Any other descriptions that you think are important?

Part 2: Describe yourself in this situation. We need to know a basic description of you before the event takes place.

1. Describe how are you dressed (e.g. jacket, pants, skirt, hat, shoes).
2. Describe your appearance (e.g. clean, disheveled, tidy).
3. Describe how you feel (e.g. nervous, excited, angry, happy, fearful, relieved, disgusted).
4. What specific facial expression, if any, do you have?
5. What gestures or movements, if any, are you making?
6. List any of your characteristics that you believe are relevant to the situation (e.g., gender, race, age, hair).
7. What, if anything, are you holding?
8. What may be the stranger's impression of you?
9. Mention anything else you think is relevant?

Part 2: Describe the stranger in this situation. We need to know a basic description of the stranger before the event takes place.

1. Describe how the stranger is dressed (e.g. jacket, pants, skirt, hat, shoes).
2. Describe the stranger's appearance (e.g. clean, disheveled, tidy, stocky)?
3. Describe how the stranger may feel (e.g. nervous, excited, angry, happy, fearful, relieved, disgusted)?
4. What specific facial expression, if any, does the stranger have?
5. What gestures or movements, if any, is the stranger making?
6. List any of the stranger's characteristics that you believe are relevant to the situation (e.g., gender, race, age, mustache).
7. What, if anything, is the stranger holding (e.g. weapon, purse, map, food)?
8. What is your impression of the stranger?
9. Is the stranger a member of an organization or profession related to the encounter? If so, what is that?
10. Mention anything else you think is relevant (e.g. speech. mannerisms).

Part 2: Listing the stranger's goals, plans, and actions.

1. List all of the reasons you can think of for why the stranger wants to <**GOAL**>?
2. List all of the ways you can think of that the stranger plans to <**GOAL**>?
3. List out the sequence of actions taken by the stranger and by you over the course of the interaction.

Appendix B

Note: H, high informativeness; M, moderate informativeness, and L, low informativeness.

Item	Location	Time	Entity–self	Entity–stranger	Stranger's action
Set 1					
1	H				
2		H			
3			H		
4				H	
5					H
6	H	H			
7	H		H		
8	H			H	
9	H				H
10		H	H		
11		H		H	
12		H			H
13			H	H	
14			H		H
15				H	H
16	H	H	H		
17	H	H		H	
18	H	H			H
19	H		H	H	
20	H		H		H
21	H			H	H
22		H	H	H	
23		H	H		H
24		H		H	H
25			H	H	H
26	H	H	H	H	
27	H	H	H		H
28	H	H		H	H
29	H		H	H	H
30		H	H	H	H
31	H	H	H	H	H
32	M				
33		M			
34			M		
35				M	
36					M
37	M	M			

(*continued on next page*)

Appendix B (*continued*)

Item	Location	Time	Entity–self	Entity–stranger	Stranger's action
38	M		M		
39	M			M	
40	M				M
41		M	M		
42		M		M	
43		M			M
44			M	M	
45			M		M
46				M	M
47	M	M	M		
48	M	M		M	
49	M	M			M
50	M		M	M	
51	M		M		M
52	M			M	M
53		M	M	M	
54		M	M		M
55		M		M	M
56			M	M	M
57	M	M	M	M	
58	M	M	M		M
59	M	M		M	M
60	M		M	M	M
61		M	M	M	M
62	M	M	M	M	M
63	L				
64		L			
65			L		
66				L	
67					L
68	L	L			
69	L		L		
70	L			L	
71	L				L
72		L	L		
73		L		L	
74		L			L
75			L	L	
76			L		L
77				L	L
78	L	L	L		

Appendix B (*continued*)

Item	Location	Time	Entity–self	Entity–stranger	Stranger's action
79	L	L		L	
80	L	L			L
81	L		L	L	
82	L		L		L
83	L			L	L
84		L	L	L	
85		L	L		L
86		L		L	L
87			L	L	L
88	L	L	L	L	
89	L	L	L		L
90	L	L		L	L
91	L		L	L	L
92		L	L	L	L
93	L	L	L	L	L
Set 2					
94	H	H	H	H	M
95	H	H	H	H	L
96	H	H	H	M	H
97	H	H	H	L	H
98	H	H	M	H	H
99	H	H	L	H	H
100	H	M	H	H	H
101	H	L	H	H	H
102	M	H	H	H	H
103	L	H	H	H	H
104	M	M	M	M	H
105	M	M	M	M	L
106	M	M	M	H	M
107	M	M	M	L	M
108	M	M	H	M	M
109	M	M	L	M	M
110	M	H	M	M	M
111	M	L	M	M	M
112	H	M	M	M	M
113	L	M	M	M	M
114	L	L	L	L	H
115	L	L	L	L	M
116	L	L	L	H	L
117	L	L	L	M	L

(*continued on next page*)

Appendix B (*continued*)

Item	Location	Time	Entity–self	Entity–stranger	Stranger's action
118	L	L	H	L	L
119	L	L	M	L	L
120	L	H	L	L	L
121	L	M	L	L	L
122	H	L	L	L	L
123	M	L	L	L	L

References

Barker, R. G. (1968). *Ecological psychology*. Stanford, CA: Stanford University Press.

Barker, R. G., & Wright, H. F. (1955). *Midwest and its children: The psychological ecology of an American town*. Evanston, IL: Row, Peterson.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments* (2nd ed.). Berkeley, CA: University of California Press.

Carlston, D. E., & Skowronski, J. J. (2005). Linking versus thinking: Evidence for the different associative and attributional bases of spontaneous trait transference and spontaneous trait inference. *Journal of Personality & Social Psychology, 89*, 884–898.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Copeland, D. E., Magliano, J. P., & Radvansky, G. A. (2006). Situation models in comprehension, memory, and augmented cognition. In M. Bernard, J. C. Forsythe, & T. Goldsmith (Eds.), *Human Cognitive Models in System Design* (pp. 37–66). Malwah, NJ: Erlbaum.

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality & Social Psychology, 83*, 1314–1329.

Davies, M. F. (2003). Confirmatory bias in the evaluation of personality descriptions: Positive test strategies and output interference. *Journal of Personality & Social Psychology, 85*, 736–744.

Devine, P. G., Hirt, E. R., & Gehrke, E. M. (1990). Diagnostic and confirmation strategies in trait hypothesis testing. *Journal of Personality & Social Psychology, 58*, 952–963.

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality & Social Psychology, 38*, 889–906.

Garcia-Marques, L., Sherman, S. J., & Palma-Oliveira, J. M. (2001). Hypothesis testing and the perception of diagnosticity. *Journal of Experimental Social Psychology, 37*, 183–200.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology, 48*, 163–189.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*, 371–395.

Hassin, R. R., Aarts, H., & Ferguson, M. J. (2005). Automatic goal inferences. *Journal of Experimental Social Psychology, 41*, 129–140.

Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review, 93*, 258–268.

Heider, F. (1958). *The psychology of interpersonal relations*. Oxford, UK: Wiley.

Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology, 24*, 1–55.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*, 430–454.

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review, 103*, 284–308.

Kruglanski, A. W., Schwartz, J. M., Maides, S., & Hamel, I. Z. (1978). Covariation, discounting, and augmentation: Towards a clarification of attributional principles. *Journal of Personality, 46*, 176–189.

Long, D. L., & Golding, J. M. (1993). Superordinate goal inferences: Are they automatically generated during comprehension. *Discourse Processes, 16*, 55–73.

Magliano, J. P., Miller, J., & Zwaan, R. A. (2001). Indexing space and time in film understanding. *Applied Cognitive Psychology, 15*, 533–545.

Magliano, J. P., & Radvansky, G. A. (2001). Goal coordination in narrative comprehension. *Psychonomic Bulletin and Review, 8*, 372–376.

Magliano, J. P., Radvansky, G. A., & Copeland, D. E. (in press). Beyond language comprehension: Situation models as a form or autobiographical memory. In: F. Schmalhofer & C. Perfetti (Eds.) *Higher level language processes in the brain: Inference and comprehension processes*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Magliano, J. P., Taylor, H. A., & Kim, H. J. (2005). When goals collide: Monitoring the goals of multiple characters. *Memory & Cognition., 33*, 1357–1367.

Magliano, J. P., Zwaan, R. A., & Graesser, A. C. (1999). The role of situational continuity in narrative understanding. In S. R. Goldman & H. van Oostendorp (Eds.), *The construction of mental representation during reading* (pp. 219–245). Mahwah, NJ: Erlbaum.

McClure, J. (1998). Discounting causes of behavior: Are two reasons better than one?. *Journal of Personality & Social Psychology 74*, 7–20.

McRae, K., Cree, G. S., Westmacott, R., & de Sa, V. R. (1999). Further evidence for feature correlations in semantic memory. *Canadian Journal of Experimental Psychology, 53*, 360–373.

McRae, K., de Sa, V., & Siedenberg, M. S. (1997). On the nature and scope of feature representation of word meaning. *Journal of Experimental Psychology: General, 126*, 99–130.

McRae, K., Ferretti, T. R., & Amyote, L. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes, 12*, 137–176.

Morrow, D. G., Greenspan, S. L., & Bower, G. H. (1987). Accessibility and situation models during comprehension. *Journal of Memory and Language, 26*, 165–187.

Ozyurek, A., & Trabasso, T. (1997). Evaluation during the understanding of narratives. *Discourse Processes., 23*, 305–335.

Pennington, N., & Hastie, R. (1993). Reasoning in explanation-based decision making. *Cognition, 49*, 123–163.

Radvansky, G. A. (1998). The organization of information retrieved from situation models. *Psychonomic Bulletin & Review., 5*, 283–289.

Randall, B., Moss, H. E., Rodd, J. M., Greer, M., & Tyler, L. K. (2004). Distinctiveness and Correlation in Conceptual Structure: Behavioral and Computational Studies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 393–406.

Read, S. J., & Miller, L. C. (1993). Rapist or "regular guy": Explanatory coherence in the construction of mental models of others. *Personality & Social Psychology Bulletin, 19*, 526–540.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: studies in the internal structure of categories. *Cognitive Psychology, 7*, 573–605.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Erlbaum.

Scott-Rich, S., & Taylor, H. A. (2000). Not all narrative shifts function equally. *Memory & Cognition, 28*, 1257–1266.

Skowronski, J. J. (2002). Honesty and intelligence judgments of individuals and groups: The effects of entity-related behavior diagnosticity and implicit theories. *Social Cognition, 20*, 136–169.

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 & Social Psychology, 52*, 689–699.

Skowronski, J. J., & Carlston, D. E. (1999). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin, 105*, 131–142.

Strohmer, D. C., & Shivy, V. A. (1994). Bias in counselor hypothesis testing: Testing the robustness of counselor confirmatory bias. *Journal of Counseling & Development, 73*, 191–197.

Suh, S. Y., & Trabasso, T. (1993). Inferences during reading: Converging evidence from discourse analysis, talk-aloud protocols, and recognition priming. *Journal of Memory and Language, 32*, 279–300.

Sundermeier, B. A., van den Broek, P., & Zwaan, R. A. (2005). Causal coherence and the availability of locations and objects during narrative comprehension. *Memory & Cognition, 33*, 462–470.

Trabasso, T., van den Broek, P., & Suh, S. (1989). Logical necessity and transitivity of causal relations in the representation of stories. *Discourse Processes, 12*, 1–25.

Trope, Y. (1998). Dispositional bias in person perception: A hypothesis-testing perception. In J. M. Darley & J. Cooper (Eds.), *Attribution and social interaction: The legacy of Edward E. Jones* (pp. 67–126). Washington, DC, US: American Psychological Association.

Trope, Y., & Gaunt, R. (1999). A dual-process model of overconfident attributional inferences. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 161–178). New York: Guilford Press.

Trope, Y., & Liberman, A. (1996). Social hypothesis testing: Cognitive and motivational mechanisms. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 239–270). New York, NY, US: Guilford Press.

Trope, Y., & Mackie, D. M. (1987). Sensitivity to alternatives in social hypothesis-testing. *Journal of Experimental Social Psychology, 23*, 445–459.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Van Overwalle, F., & Labiouse, C. (2004). A recurrent connectionist model of person impression formation. *Personality & Social Psychology Review, 8*, 28–61.

Wells, G. L., & Ronis, D. L. (1982). Discounting and augmentation: Is there something special about number of causes?. *Personality & Social Psychology Bulletin 8*, 566–572.

Wiemer-Hastings, K., & Xu, X. (2005). Content differences for abstract and concrete concepts. *Cognitive Science, 29*, 719–736.

Wright, J. C., & Murphy, G. L. (1984). The utility of theories in intuitive statistics: the robustness of theory-based judgments. *Journal of Experimental Psychology: General, 113*, 301–322.

Zebrowitz, L. A., & Monteparre, J. (2006). First impressions from facial appearance cues. In J. J. Skowronski & N. Ambady (Eds.), *First impressions* (pp. 81–113). New York: Guilford Press.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science, 6*, 292–297.

Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 386–397.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*, 162–185.

Zwaan, R. A., & van Oostendorp, H. (1993). Do readers construct spatial representations in naturalistic story comprehension?. *Discourse Processes 16*, 125–143.