2

The Distinction between Desire and Intention: A Folk-Conceptual Analysis

Bertram F. Malle and Joshua Knobe

Brian seems to be interested in asking Lisa to marry him, but Lisa is not sure how serious he is. He has been discussing the issue with friends, writing about it in his diary, and his face takes on a special glow whenever the topic of marriage comes up in conversation. As the days progress, Lisa may begin to ask herself whether he only has a *wish* to propose to her or whether he has he actually *decided* to propose.

Like Lisa in this example, people in general distinguish between desires (expressed in English with the verbs *wish, hope,* and *want*) and intentions (*decide, plan, intend*). People are often faced with the task of classifying a person's mental state into one of these two categories. In this chapter we propose a psychological theory about what people do under these circumstances—what criteria they use to classify mental states and how they use the resulting classification.

Philosophers have developed various analyses of the distinction between desire and intention (Audi 1988; Brand 1984; Bratman 1987; Davis 1984; Mele 1988; Velleman 1989). Typically, the aim of these philosophical analyses is to reach a more adequate understanding of what desires and intentions ultimately *are*—to better understand the nature of these mental states as they actually exist in human agents. We put these questions aside. Our aim is to understand how social perceivers use the concepts of desire and intention to interpret other people's behavior. Thus, our inquiry is one part of the broader attempt to better understand what has been called "naive psychology of action" (Heider 1958), "theory of mind" (Gopnik and Meltzoff 1997; Premack and Woodruff 1978; Wellman 1990), or "folk psychology" (Greenwood 1991)—the conceptual framework that helps people perceive, explain, predict, and change human behavior by reference to

mental states. We will be concerned with the specific question of how people come to treat certain mental states as desires and others as intentions. Developmental research has begun to explore how and when children distinguish these two mental states (Astington and Gopnik 1991; Astington, this volume; Moses, this volume; Perner 1991; Phillips, Baron-Cohen, and Rutter 1998; Schult 1996). However, a systematic account of the endpoint of this development—the adult distinction between intention and desire—has not been offered.

In providing such an account, we blend conceptual analysis with empirical study. We begin with a rough sketch of the desire-intention distinction, then construct specific hypotheses about the features that social perceivers use to distinguish desires from intentions. Each of these hypotheses is tested empirically, using naturally occurring data and controlled experiments.

A Rough Sketch

Desires and intentions are both representational states, and they both express a pro attitude toward the state of affairs they represent, frequently propelling the agent to act in such a way as to bring about that state. However, there is an important difference. When we say that an agent has a desire, we are not saying that she[1] has actually decided to do anything. She might have a desire for world peace even if she doesn't plan to take any steps to make this desire a reality. Similarly, an agent might have a desire to start screaming at her boss even though she has specifically decided not to do so. When, however, people adopt an intention, they are actually deciding to perform the action in question.

One might say, speaking roughly, that desires and intentions occupy different positions in the path that (typically) leads to action. Desires stand in the very beginning of the process. Before making a decision about how to act, the person needs to consider various desires, balancing them against each other and asking which of them can potentially be fulfilled. In the course of this reasoning process, the person arrives at an intention. This intention is an all-things-considered decision that takes into account the person's various desires. The intention, then, is just one step away from the action; all that remains is to put one's decision into motion.

   To clarify these initial intuitions, we constructed specific hypotheses about three criteria that social perceivers use to distinguish intentions from desires. These criteria are the *type of content*, the *function in reasoning*, and the *degree of commitment*.

Type of Content

A number of philosophers, including Baier (1970) and Castañeda (1972), have argued that, whereas desires can have many different types of content, intentions always have as their content an action performed by the person who holds that intention. As Aristotle remarks (1962, 1111b25), people may "wish for the victory of a particular actor or a particular athlete," but "no one chooses such things, for we choose only what we believe might be attained through our own agency." Thus, a person may want a lot of money, but she cannot intend a lot of money. What people intend is always their own action (to go jogging, to write a letter, to cook a nice meal) or an event directly controlled by their own action (to be polite, to be at the party tonight). People may intend to bring it about that an athlete be victorious (e.g., by paying for a world-class coach), but they cannot literally intend another person's action or an event outside of themselves.

   But what about such statements as "I intend to be the next president," "I intend you to marry my cousin's son," or "We intend Alex to go to kindergarten next fall"? Are these intention avowals not blatantly violating the proposed rule? And would not social perceivers who strictly adhere to the rule be utterly confused by these statements? As we will see shortly, such statements are extremely rare—in the vast majority of intention avowals, their content designates the agent's own action. The few exceptions may be addressed in various ways. One is to say that these expressions do not strictly follow the rules of proper English. A second option is to say that the English words in question (e.g., *decide*, *intend*) almost always refer to a single folk concept—here, the folk concept of intention—but that there are occasional cases in which they don't refer to this concept. These occasional cases may turn out to be extended or metaphorical uses of the standard term. For example, when a person says "I intend to be the next president," social perceivers may consider the claim to mean "I really want to . . . and

I will try everything I can to. . . ." The uses may also be shorthands for some longer sentence that involves the folk concept of intention in the standard way (e.g., "I intend you to marry my cousin's son" is a shorthand for "I intend to bring it about that you marry my cousin's son"). It may turn out that all of these options are valid, each applying to a different sort of counterexample. But the fundamental point is, we think, that a few cases in which people's uses of the words depart slightly from conceptual rules should not lead us to abandon the whole theory that identifies these rules. Surely, if a few years from now people invented a sport in which the word *intend* was used in a somewhat unusual way, we would not have to completely revise our theories of the folk concept of intention.

Our first hypothesis is, then, that people use the content of a pro attitude to identify it as a desire or an intention, with desires representing any content and with genuine intentions representing what we call *action content*. A pro attitude has action content when the content of the attitude is an action performed by the same person who holds that attitude. Linguistically, this content is most clearly displayed in English *that*-constructions, such as "Jones hopes that he will go jogging," but more commonly (Rosenbaum 1970) it is expressed in *to* + infinitive constructions, such as "Jones plans to go jogging."[2] What evidence do we have for the claim that desires can have any type of content whereas genuine intentions always have action content?

The content difference between intention and desire has grammaticalized in English such that there is a systematic difference in the syntactic complementation patterns for the verbs *want* and *intend*. According to a study by Aarts and Aarts (1991), the verb *want* is realized in three different ways: with *to* + *infinitive* (55 percent of the cases; e.g., "I just want to know"), with a direct object (28 percent; e.g., "Employees want a greater say at work"), or with a noun phrase (NP) and various complements (17 percent; e.g., "He wanted Browne dead"; "He wanted us to get going"—see Erdmann 1993). In contrast, the verb *intend* is primarily realized with *to* + *infinitive*. In a sample of 110 occurrences of *intend* or *intended* in American newspapers and magazines (extracted from the NEXIS database), we found that 97 percent of the active verb forms of *intend* were paired with *to* + *infinitive*.

In the realization of *intend* with *to* + *infinitive*, the grammatical subject of the *intend* verb is identical to the grammatical subject of the infinitive,

guaranteeing identity between the agent who intends and the agent of the intended action, as in "I intend to drive" (= I intend that I drive). Such identity is violated in occurrences of *want* when it is paired with NP complements (which are typically referring to another person or state of affairs). Thus, it is unproblematic to state "I want them out of this room right now" but problematic to state "I intend them out of this room right now" (except perhaps, metaphorically, to express an unusually strong desire or stern command).

To strengthen this point, we collected linguistic data on the distribution of contents for intention verbs (*intend*, *plan, decide*)[3] and desire verbs (*want*, *wish, hope*). Using the NEXIS database, we selected active-voice occurrences of each of the above verbs (30–100 per verb) and coded them for their content—referring to the agent or not. (For a subset of the verbs, two coders independently coded the agent/not-agent variable and reached an agreement of 93 percent.) The content of intention verbs referred to the agent in 98 percent of the cases,[4] whereas the content of desire verbs referred to the agent in 63 percent of the cases; $\chi^2(1, N = 396) = 82.4$, $p < 0.001$.

In addition, two coders also rated the controllability of the verbs' content (on a five-point scale from –2 to +2), for which they reached an agreement of $r = 0.87$. (The ratings were made while blind to the original verb; i.e., all verbs were replaced by the generic term *represent*.) The average controllability for all three desire verbs was uniformly low (from $M = -0.28$ for *wish* to $M = 0.24$ for *want*), whereas the average controllability for all three intention verbs was uniformly high (from $M = 0.94$ for *plan* to $M = 1.46$ for *decide*). The contrast of the three desire verbs against the three intention verbs was highly significant, $F(1, 391) = 104.9$, $p < 0.001$, $\eta^2 = 21$ percent. A strict test comparing only those verb forms that had action content still showed desire verbs to be paired with significantly less controllable actions ($M = 0.62$) than intention verbs ($M = 1.24$), with $p < 0.001$.

One implication of the rule that intentions refer to action content is that one cannot intend another person's actions, only one's own (Baier 1970; Brand 1984). To test this implication, we collected a small set of experimental data, using a multiple-choice verb completion design. Participants (109 college students) read several sentences that normally would contain a desire or intention verb, but the verb's position was left blank. Participants

were asked to choose the most appropriate verb from a list of five: want, hope, intend, plan, and decide. The results were as follows.[5]

(1) *Enja _____ her husband to pick up their daughter from the airport.*

wanted	85%	
hoped	0%	desire: 85%
intended	8%	
planned	6%	
decided	0%	intention: 15%

(2) *"I am serious: I _____ you to be back by midnight!"*

want	93%	
hope	2%	desire: 94%
intend	3%	
plan	3%	
decide	0%	intention: 6%

It should be noted that these items were presented as part of a 40-minute mass-testing questionnaire that certainly did not engage all participants. Moreover, the participants included about 5 percent (unidentified) non-native speakers. Thus, one should not overinterpret the small number of people who chose an intention verb to complete these sentences. What seems clear is that the vast majority of informants preferred the verb *want* to refer to a non-agent action content. This is particularly impressive for the second item, where the sentence was constructed to express a strong desire ("I am serious . . ."), conceived by some, including Rundle (1997, chapter 5), as closer to an intention. That same marker of intensity, when coupled with a content that describes *agent-controlled* action, led to 82 percent intention verb choices:

(3) *"I am serious: I _____ to be back by midnight!"*

want	15%	
hope	3%	desire: 18%
intend	45%	
plan	37%	
decide	0%	intention: 82%

We have argued that intentions are pro attitudes with action content[6] whereas desires may have any content (one's own action, another's action, an unlikely outcome, etc.). Thus, if the social perceiver sees that a pro attitude does not have action content, he can immediately assume that the attitude is a desire.

Now we would like to speculate briefly about why desires and intentions have different types of content. In folk psychology, intentional action is interpreted and explained in terms of mental states (Malle 1999). Some mental states are felt to be very close to action: They affect an action directly, without the intermediation of any other mental states. Intentions serve precisely this role. An intention to perform action A might lead directly to action A. The social perceiver need not posit an intermediary mental state that is affected by the intention and in turn affects the action. Other mental states are felt to be very distant from action: They can affect an action only through a long chain of causes and effects—e.g., by affecting another mental state, which then affects another mental state, and so on, until ultimately one reaches a mental state that affects the action directly. Thus, a desire for world peace does not directly cause the person to perform a particular action. It can lead to action only in conjunction with a number of other mental states (e.g., beliefs about which actions tend to promote world peace).

If we wanted to map out the whole process that leads to an action according to folk psychology, we might begin with fairly general beliefs, desires, and values. These mental states would serve as inputs to a reasoning process that ultimately led to an intention. The intention would then typically lead directly to an action. Now, the mental states that affect the action indirectly (the beliefs, desire, and values) may have any type of content. Even if they don't have action content, they can indirectly affect the agent's actions by affecting other mental states, as long as they eventually affect a mental state that specifies a particular action to be performed. However, the mental states that directly affect the action must always have action content. The content of these states is precisely the action they lead to. That is, in the chain of mental states that leads eventually to action A, the state just preceding A must specifically represent the action A. This final state is an intention.

Role in Reasoning

A person might have many pro attitudes with action content, but most of these will never be transformed into intentions. Maybe the person wants to

devote her life to charity. Maybe she wants to slap her annoying co-worker. Social perceivers need to know whether these pro attitudes with action content are just desires or whether they are genuine intentions (which are more predictive of action). To do so, they make use of additional criteria.

One of the additional criteria is the role that the attitude plays in the agent's reasoning. For example, a desire to eat chocolate is not based on any reasoning, but it may set into motion a reasoning process about how to acquire some chocolate. An intention to go to the corner store, in contrast, is likely to be based on reasoning about what one needs (e.g., chocolate) and whether the store carries it. These prototypical cases suggest that desires are typically inputs to reasoning, whereas intentions are typically outputs of reasoning. At least that is how people treat desires and intentions: According to the folk concept of intentionality, the reasoning process underlying intentional action takes desires as input and uses them to generate intentions (Malle and Knobe 1997a). By looking at a pro attitude's role within the agent's reasoning, then, social perceivers can distinguish between intentions and desires (even if both of them have action content).

Philosophers have argued that, in actual fact, intentions are sometimes inputs to reasoning (Bratman 1999) and desires are sometimes outputs of reasoning (Harman 1976; Schueler 1995). This, however, is no objection to the hypothesis we are advancing. We do not say that each attitude can occupy one and only one predetermined role in reasoning. Our point is simply that there is a tendency for social perceivers to infer that an attitude is a desire if it is the input of reasoning and to infer that it is an intention if it is the output of reasoning. The pro attitude's role in reasoning thus serves as a *clue* for perceivers as to the nature of the attitude.

To explore this hypothesis, we again used people's verb choices as measures of their desire or intention inferences. We held the action content of the stimulus sentences constant but varied the linguistic and conceptual context to indicate either a reasoning input phase (deliberation) or a reasoning output phase (decision). A sample of 206 undergraduate students completed these vignettes in a mass-testing questionnaire, choosing the verb they considered most appropriate from a list of six.

In the first vignette, the target sentence "Carl [verb] to make a large charity donation" was either paired with a reason explanation (4), rendering the target sentence a reasoning output, or with an indicator of early delib-

eration (5), rendering the target sentence a reasoning input. Clearly, people preferred intention verbs when the target sentence was a reasoning output, compared to the same sentence functioning as reasoning input.

(4) *Carl _____ to show his relatives that he is a generous person, . . .*

want	49%	
hope	8%	
need	7%	desire: 64%
intend	25%	
plan	4%	
decide	8%	intention: 36%

. . . so he _____ to make a large charity donation.

want	1%	
hope	1%	
need	1%	desire: 3%
intend	15%	
plan	33%	
decide	49%	intention: 97%

(5) *Carl _____ to make a large charity donation, but he didn't know much about charity organizations.*

want	52%	
hope	4%	
need	2%	desire: 57%
intend	13%	
plan	14%	
decide	16%	intention: 43%

A second vignette confirmed this difference with even stronger effect sizes:

(6) *In her dire financial situation Beth _____ to make a lot of money fast, . . .*

want	13%	
hope	17%	
need	61%	desire: 91%
plan	0%	
intend	4%	
decide	5%	intention: 9%

. . . so she _____ to buy high-risk stock.

need	0%	
hope	3%	
want	8%	desire: 11%
plan	16%	
intend	9%	
decide	64%	intention: 89%

(7) *Beth _____ to buy high-risk stock, but she was afraid of losing a lot of money.*

need	3%	
hope	4%	
want	64%	desire: 71%
plan	11%	
intend	13%	
decide	6%	intention: 30%

In a second test, we inspected naturally occurring instances of intention and desire verbs, hoping to find a similar trend that people would use desire verbs more often to describe what they consider inputs to reasoning and intention verbs more often to describe what they consider outputs of reasoning. Unfortunately, the NEXIS verb sample described above contained very few explicit indications of reasoning. Among 183 desire verbs, 15 showed explicit indications of their function in reasoning (e.g., *therefore*, *in order to*, *because*), and 9 of those were reasoning inputs (60 percent). Among 217 instances of intention verbs, 42 showed explicit indications of their function in reasoning, and 38 of those were reasoning outputs (90 percent), χ^2 (1, $N = 400$) = 16.0, $p < 0.001$. (For example, one desire-as-input

expression was the following: "Amy Regan-Axelson wants you to make time to come riding with her. That is why she founded Women's Mountain Biking." One intention-as-output expression was "The government intends to introduce annual council elections to make financial decision-making more responsive to the electorate.")

From these verb distributions and from the experimental data earlier we can conclude that there is a strong trend for intentions to be treated as outputs of reasoning and a somewhat weaker trend for desires to be treated as inputs to reasoning. Both of these trends are consistent with the folk concept of intention, which requires intentions to be based on (belief-desire) reasoning whereas no such requirement accompanies the concept of desire (Malle and Knobe 1997a).

We would now like to explore why intentions and desires might perform their different functions in the practical reasoning chain, at least as seen by folk psychology. Suppose that an agent has a desire for outcome O and a belief that action A will lead to outcome O. As a result of this belief-desire pair, she might acquire a desire to perform action A. Under those conditions, why can't she simply act on her desire to A? What would be the point of further deliberation in which the desire to A serves as input to a reasoning process that ultimately yields an intention to A as output?

The answer is that, before a pro attitude with action content becomes a full-fledged intention, the person needs to ask herself (a) whether she is capable of performing the action and (b) whether she has other desires that outweigh her desire to perform the action.

Suppose that the agent wants to climb a mountain on Saturday. She will not immediately decide to climb the mountain. She must first engage in a certain amount of reasoning. She has to make sure that she is capable of performing the action in one way or another. Moreover, she has to figure out whether she has any other desires that outweigh her desire to perform the action (e.g., whether there is something else that she would rather do on Saturday). If she concludes that she is capable of performing the action and has no outweighing desires, she may form a corresponding intention (Davidson 1980b). Otherwise, she may retain her desire to perform the action even though she forms no intention to act on that desire.

According to folk psychology, then, intentions serve to fulfill desires by identifying a course of action that is feasible to implement for the agent and

is compatible with the agent's other desires. Because of the closeness of intentions to actions in the world, intentions have to fit the world (in the sense that the agent would be capable of implementing them), and because of the consequences of actions for the agent's welfare, intentions have to fit the agent's desires. This fit is sometimes expressed as the "consistency criterion" for intentions (Moses, this volume). The agent's reasoning checks for such consistency, and social perceivers expect an agent to have engaged in this sort of reasoning when forming an intention.

Degree of Commitment

Even when there is evidence that an agent engaged in practical reasoning, there is no guarantee that the agent has formed an intention. Maybe she is still unsure. Or maybe she has concluded that, although there would be many advantages to performing the action, she is not going to form an intention to do so. The social perceiver needs a way to distinguish among these various possibilities. Thus, a third criterion is necessary: *commitment*.

When we say, in this technical sense, that an agent is "committed" to an action, we do not mean that the agent has made a promise to anyone else, nor even that she has made some sort of promise to herself. An agent is said to be committed when she has made up her mind, when she has settled on a particular path of action (Bratman 1987; Mele 1992a). Thus, suppose that the agent is wondering whether to order a hamburger or a slice of pizza. At first, she faces a number of competing desires, accompanied by various beliefs about the possible consequences of her action and the situation at hand. Then, after a period of reasoning, she settles on ordering the hamburger. At that point, she considers the question closed. When the waiter asks for her order, she does not begin deliberating all over again. She simply acts on the intention she had already formed. This settledness, the conclusion of the reasoning process, is what we mean by commitment.

Social perceivers consider the agent's degree of commitment to distinguish intentions from desires. However, they cannot directly perceive commitment, so they make inferences on the basis of at least three indicators. First, an agent indicates commitment by making *early investments* (e.g., buying a concert ticket in advance). Second, an agent indicates commitment by accepting *opportunity costs* of not pursuing alternative courses of action

(best indicated when such alternatives become in fact available). Third, an agent indicates commitment by accepting (sometimes inviting) *sanctions* from others in case she does not fulfill the intention (e.g., public announcements of intentions, which put the agent's credibility on the line).

Take, for example, the intention to stay with a romantic partner. A strong commitment is expressed through early investments (e.g., joint purchases), accepting opportunity costs (e.g., not dating others), and inviting sanctions (e.g., introducing the partner to friends and family), whereas fear of commitment manifests itself in the absence of these signals (few joint purchases, insistence on dating others, reluctance to meet friends and family).

Now we turn to our empirical data on commitment, which illustrate people's sensitivity to indicators of commitment when inferring intention.

To begin, when a speaker is committed to a course of action, others can rely on it. In (8), a speaker tries to communicate this reliability to an addressee, and consequently social perceivers ($N = 109$ college students) judged that the speaker would use an intention verb. Conversely, in (9) an indication of *not* being committed to a course of action led people to choose desire verbs (even though the content considered was clearly an action).

(8) *Sheila _____ to go to the Thai restaurant on 29th Street; you can find her there.*

wanted	3%	
hoped	1%	desire: 4%
intended	8%	
decided	63%	
planned	25%	intention: 96%

(9) *Sarah _____ to go to a Thai restaurant perhaps, or pretty much any Asian cuisine.*

wants	74%	
hopes	19%	desire: 93%
intends	3%	
plans	3%	
decides	1%	intention: 7%

Speakers themselves often indicate their commitment in avowals of intention, and in (10) commitment is indicated in a global way ("seriously"), which is sufficient to convince most social perceivers (N = 206 college students) that an intention verb would be most appropriate. In contrast, in (11) the speaker expresses doubt about the intention's fulfillment, and perceivers are significantly less likely to infer an intention:

(10) *The teacher said "Seriously, I _____ to return your papers tomorrow."*

need	5%	
hope	12%	
want	3%	desire: 20%
plan	39%	
intend	40%	
decide	2%	intention: 80%

(11) *The teacher said "I am not sure it's going to work, but I _____ to return your papers tomorrow."*

need	0%	
hope	43%	
want	10%	desire: 53%
plan	22%	
intend	23%	
decide	2%	intention: 47%

The next item illustrates the power of a specific commitment indicator, lack of) early investments to go to a concert, which allows an inference about the agent's (lacking) intention.

(12) *He likes U2, but I doubt he _____ to go to the concert—he hasn't bought tickets yet.*

need	2%	
hope	0%	
want	15%	desire: 17%

plan   41%
intend   34%
decide   9%   intention: 83%

The availability of multiple indicators of commitment (investments, opportunity costs, etc.) offers the possibility that, in social perception, commitment is inferred not as an on-off state but in degrees. Thus, the more commitment an agent shows toward a course of action, the more inclined people are to infer an intention rather than a desire. Initial evidence for a continuum of commitment comes from the next vignette, in which subjects were presented with a potential action of moving to Europe under three conditions that decreased in commitment: (a) with opportunity costs clearly indicated, (b) with a time index signaling reasoning but only vague commitment, and (c) with a strong pro attitude but a lack of commitment. The results show that the corresponding inference of intention (rather than desire) drops precipitously from 84 percent to 41 percent to 15 percent.

(13a) *I know she _____ to move to Europe; she's already given notice.*

need   2%
hope   6%
want   8%   desire: 16%
plan   35%
intend   29%
decide   20%   intention: 84%

(13b) *I know she _____ to move to Europe next year.*

need   0%
hope   17%
want   42%   desire: 59%
plan   23%
intend   14%
decide   4%   intention: 41%

(13c) *I know she _____ to move to Europe some day; it's been her child-hood dream.*

need	0%	
hope	55%	
want	30%	desire: 85%
plan	7%	
intend	8%	
decide	0%	intention: 15%

A Model of Desires and Intentions

We have now arrived at a three-part model of how social perceivers distinguish between desires and intentions. For one thing, they examine the content of the agent's pro attitude. If the content is not the agent's own action (but rather someone else's action, an object, or an outcome), social perceivers will assume that the attitude is a desire. If the pro attitude has action content, it may be classified as an intention. However, further information may be necessary: Perceivers also examine the role that the attitude played in the agent's reasoning. If the attitude is not based on any reasoning, it is classified as a desire; if it is based on reasoning, it may be classified as an intention (but further information might still be necessary). Finally, social perceivers assess the degree to which the agent is committed to the represented action. If the agent shows no indications of commitment, the attitude will be classified as a desire; if the agent does show commitment, the attitude will be classified as an intention.

Social perceivers might proceed through these three steps in order, always tracking content and beginning a search for other features only if the content is an action (otherwise, no search is necessary, because the pro attitude must be a desire). To warrant an intention classification, all three features must be identified, otherwise the search ends with a desire classification. At times, perceivers may look immediately at the feature of commitment and, if it is present, infer an intention (because commitment to action entails action content and is likely to be based on reasoning). This confirmatory strategy may be used when perceivers strongly expect an intention.

Alternative Distinguishing Features
The philosophical literature makes reference to additional features that potentially distinguish between desires and intentions.

A first possible criterion is the controllability of the intended behavior. Some philosophers, including Baier (1970), have argued that an agent cannot intend to perform a behavior unless it is controllable. One should distinguish, however, between the agent's belief that the behavior is controllable and the actual controllability of that behavior. An agent might decide to start her car even though, unbeknownst to her, the alternator is malfunctioning and the act of starting is therefore uncontrollable. A social perceiver who knows about the malfunction will ascribe a genuine intention to the agent, even though he is certain that the agent's intention will not be realized. If, however, the agent believes that the alternator is malfunctioning and that she therefore cannot start the car, the perceiver will not say that the agent *intended* to start it but only that she *wished* to. Here again, it is not clear precisely how confident the agent must be in the controllability of her behavior before the social perceiver will ascribe an intention. Still, we can be relatively certain that a perceiver will not ascribe an intention if the agent is confident that the behavior is uncontrollable.

Second, some scholars, including Davis (1984) and Velleman (1989), have argued that intending to A entails believing that one will A whereas wanting to A does not entail that. Others, including Anscombe (1957) and Thalberg (1972), claim that an agent can intend to A even when she is certain that she will not A. Some have taken an intermediate position: An agent need not believe that she will A (with any kind of certainty), but she cannot specifically believe that she will not A (Mele 1992a, chapter 8). Although we have no empirical data on this issue of belief, we did demonstrate that social perceivers require a certain amount of commitment before classifying a mental state as an intention, and a belief that one will not perform a particular action seems incompatible with having such a commitment. It would be perplexing for someone to say "I intend to go to the party tonight but I'm sure that I won't" (since, clearly, the speaker shows a lack of commitment). Thus, our criterion of commitment suggests that social perceivers will not ascribe an intention to an agent who simultaneously believes that she will not perform the action in question. However, further empirical research may clarify whether perceivers require a specific degree of confidence (e.g., being fairly sure that one will perform the action) before classifying a pro attitude as an intention.

Finally, some authors, including Harman (1986) and Mele (1992a), have claimed that intentions to A entail a plan for A-ing. The truth of this claim depends on the required detail of such a plan. On the one hand, during the reasoning process the agent checks for the action's controllability and its compatibility with other desires. Thus, the agent takes the first steps of planning by narrowing down the possible courses of action to a few that are feasible and compatible with other goals. On the other hand, reasoning toward an intention does not necessarily lead to a plan of action. Indeed, the very act of forming an intention may instigate the process of working out a plan of action (Bratman 1987). So we suspect that social perceivers sometimes ascribe intentions even in the absence of a plan. For example, people may say that a presidential candidate intends to reform the prison system even though they recognize that the candidate has not yet worked out a plan of how to implement these reforms. As long as the agent's intention is based on reasoning that supports the feasibility of the intended action (i.e., the agent has *some idea* of how to perform the action), social perceivers may not require plans as a necessary component of intention.

Other Meanings of Intention

The intention concept described in our model is confirmed by the clear and consistent everyday use of the word *intend* and the expression of avowed intentions ("I will A"). However, related words, such as *intended* (the passive participle or adjective) and *intention*, can have different meanings and functions, and our model suggests that they in fact refer to desires. For example, much philosophy has been written about the expression *doing A with the intention of doing B*. In some cases, the *intention of doing B* is truly an intention (consistent with the three posited features), but in many cases it is not. For example, "She bought a ticket with the intention of winning the jackpot" does not reference an intention proper, because one should not infer that she intended to win the jackpot (representing an uncontrollable action). As a result, several authors have concluded that the phrase *with the intention of* in fact refers to a goal or aim—more a desire than an intention (Harman 1986, p. 93; Mele 1992a, chapter 8, fn. 22; Velleman 1989, p. 112).

Another misleading relative of *intend* is the passive form *intended for* (synonymous with *designed for* or *designated for*), as in "These planes were

intended for export." These uses take on desire content rather than intention content, reliably violating the content and commitment conditions. Similar considerations apply to the term *(un)intended consequences*. Because such consequences are by definition consequences of action and thereby outcomes, they violate the feature of action content. Of course, an agent may have thought hard about certain outcomes and may be committed to bringing them about, but that means that the agent intends to perform particular actions in order to fulfill her desire for these outcomes. Despite the etymological similarity between the verb *to intend* and the adjective *intended*, they perform different functions in the folk theory of mind and behavior. Whereas the distinction between intention and desire helps social perceivers predict other people's future behaviors, the distinction between intended and unintended outcomes assists—at least in part—in the assignment of responsibility and blame for past behaviors, and it is therefore governed by different folk-conceptual rules than the notion of genuine intentions to act.

The Function of the Desire/Intention Distinction in Social Perception
It seems appropriate here to speculate briefly about the function of the intention/desire distinction in social perception. It lies, we think, in the different roles occupied by the concepts of desire and intention in the social perceiver's attempt to predict, explain, and influence others' actions. It is useful to have information about the agent's intentions because such information allows the greatest possible predictive accuracy. If the social perceiver is trying to make predictions about what the agent is going to do next, it might be helpful to know the agent's general preferences, habits, beliefs, etc., but it would be most helpful to know what she intends to do in this specific case. Even if the social perceiver knows that the agent has a desire to perform a particular action, he still won't be nearly so sure of his predictions as he would be if he knew that the agent had actually formed an intention.

Still, the social perceiver cannot base all his predictions on knowledge of the agent's intentions, because he cannot possibly know what the agent intends to do in every specific case. In the vast majority of cases, he will have to predict the agent's actions on the basis of more general information. In any given case, this information will not give the social perceiver

nearly as much predictive accuracy as he could have obtained from knowledge of the agent's specific intentions, but he will be able to use this information in a wider variety of cases. For example, Ben's knowledge that Anne wants to make a lot of money is of limited use in predicting the particular behaviors she is going to perform at this evening's party (for that, Ben would need to know Anne's specific intentions), but it does allow him to deal with her successfully in many different situations. Moreover, knowledge of desires provides a broad guideline for supporting or opposing others' actions. Rather than respond only to concrete intentions and actions, one can respond to the agent's overarching goal and, for example, suggest alternative paths to the goal or devise ways to block all paths to it.

People also use desires more often than intentions when they are trying to explain an action that has already been performed. If someone asks "Why did the agent perform action A?" it usually isn't very helpful to respond "Because she intended to perform action A." Ideally, one would want a response of greater explanatory power, i.e., a response that not only explained this one action but also provided a more general insight into the agent's goals or the demands of the situation. Precisely because desires are part of the reasoning input to intentions, they provide illuminating background information about an action's general purpose and meaning. If we hear someone say "She performed A because she wanted to obtain outcome O," we are gaining information that subsumes this action under a broader principle, which can then be used to predict, explain, and perhaps influence a variety of further actions.

Psychological Implications of the Desire/Intention Distinction

Assuming that social perceivers classify an agent's attitude as either a desire or an intention, one may ask how this classification affects social perceivers' own behaviors toward the agent. We wish to advance three hypotheses.

The Role of Intention Ascriptions in Persuasive Communication

Since information about intentions offers greater predictive accuracy, people may emphasize that the agent has formed an intention when they wish to convince others that the agent will actually try to perform the behavior: "She has made a firm decision; she definitely plans to do it." Conversely,

when people wish to indicate uncertainty about the agent's future trying, they may emphasize that the agent has a desire but not an intention: "She would really like to do that, but she hasn't reached a decision yet. . . ."

Intentions Are More Open to Charges of Incoherence

If a person believes her desires to be incompatible, people don't usually feel that she is making any kind of error. They just say "Well, that's what she happens to want." Consider a person who wants to go to a colloquium at 3:00 P.M. and also wants to go running at 3:00 P.M. Even though she knows that her two desires are incompatible, she doesn't seem to be guilty of any kind of irrationality. If, in contrast, she adopts a set of intentions that she knows to be incompatible (e.g., intending to go to the talk at 3:00 P.M. and also intending to go running at 3:00 P.M.), she definitely seems to be making some sort of error (Harman 1976; Moses, this volume).

This difference between desires and intentions stems from their differing roles in the reasoning process. The agent is presumed to have a vast array of conflicting desires. In the process of reasoning, she is supposed to sift through these desires and formulate a set of consistent intentions. Any inconsistencies in the agent's intentions are therefore regarded as evidence of an error in reasoning.

Intentions Are More Open to Debate

If social perceivers want to change an agent's intentions, they may present complex arguments (pointing out inconsistencies in the agent's intentions, drawing the agent's attention to disadvantages of the action intended, etc.); if they wish to change an agent's desires, they are more likely to forgo rational arguments in favor of other techniques. This is because intentions typically function as reasoning output and are therefore amenable to rational arguments, whereas desires typically function as reasoning inputs. If, for example, an agent announces her intention to date her secretary, the social perceiver may point to the disadvantages of doing so, or he may argue that the agent would be unlikely to succeed in implementing her intention. In short, he provides a series of arguments designed to dissuade the agent from her intention, and if she retains her intention despite powerful arguments against it he may feel that she is behaving irrationally. If, however, the agent merely announces a desire ("Sometimes I feel an urge to date Harry . . ."),

the social perceiver may feel that no arguments are possible. No matter how many arguments the social perceiver amasses, the agent can always say "I understand all that, but even so, I keep feeling this urge." Social perceivers who want to influence the agent's desires are therefore likely to look beyond rational argument to other influence techniques (e.g., portraying the secretary in an unflattering way).

Further studies are needed to test and refine these hypotheses, and other psychological implications of the desire/intention distinction will have to be explored. But we hope that our more general point has been successfully argued: that any satisfactory theory of social perception must account for the distinction people make between desire and intention.

Acknowledgments

Notes

1. We use feminine pronouns for agents and masculine pronouns for social perceivers.

2. Cases of intending "to be polite" or "to be at the house at 7" can also be classified as action content because they refer to events directly controlled by the agent's actions (to act politely, to come to the house at 7).

3. We excluded occurrences of *to plan* when the verb was used in the sense of planning as an activity, not as an intention state (e.g., planning around the events; planning for the next 15 years; planning a vacation). Similarly, we excluded instances of the verb *to decide* when it was used in the sense of judging or concluding (e.g., "In 1963 he decided that Macmillan should step down as premier.")

4. The exceptions were "intended his interview as an overture," "intended the bomb to be used for . . . ," and "did not intend any disrespect."

5. The results are displayed with verbs ordered by type (desire or intention) and endorsement frequency. In the original questionnaire, their order was randomized across items, with three verbs of one type never succeeding one another.

6. There is a complication that we have not discussed, and it concerns intentions of group agents. Some scholars have argued that it is natural to speak of "our intention" when, say, two people plan to sing a duet together or a football team intends

to execute a pass play together (Bratman 1997; Gilbert 1989; Searle 1995). But it is not clear whether an individual person can intend something that she knows depends significantly on others' actions (Velleman 1997), and in fact the rule of intention content we have proposed would deny such a possibility. Perhaps an individual can only *desire* that the entire group perform a certain action (Zaibert, in press) and *intend* to do his or her share. Or perhaps the entire group can say "We intend to *A*" (referring to the group, not any individual), just as social perceivers are comfortable ascribing such mental states to other groups (O'Laughlin and Malle 2000). In the case of "We intend to *A*," the agent who intends is identical to the agent who performs the intended action, in line with the proposed rule of intention content.