

*Word count: 10920 (including footnotes and references)

**Intentional Action and Moral Judgment
in Asperger Syndrome and High-Functioning Autism***

Tiziana Zalla, Edouard Machery, Marion Leboyer

Abstract. While evidence shows that people with Asperger Syndrome or High Functioning Autism (AS/HFA), a mild form of autism, have an impaired capacity for mindreading, little is known as to whether their understanding of the intentional status of actions is also impaired. In this work, we report the results from two experiments that address this question. In the first experiment, by using Machery's free cup and extra-dollar cases, we show that individuals with AS/HFA seem to have difficulty understanding the intentional status of means that are negatively valued, in contrast to what is the case in the general population (Machery, 2008). In the second one, by using Knobe's harm and help cases and his murder and bull's-eye cases, we show that individuals with AS/HFA are sensitive to moral considerations and that these inform their judgment about the intentional status of actions, as is also the case in the general population (Knobe, 2003). In light of these findings, we hypothesize that because of their mindreading impairments, individuals with AS/HFA have difficulty judging that some actions are intentional—in particular, those actions that are not intrinsically valuable, but that are desired as means for something else—and that to compensate for their impairment, they use a moral heuristic to decide whether an action is intentional: If an action violates a moral norm, then it is likely seen as intentional.

* The authors wish to thank Pierre Jacob, Paul Egré, Joshua Knobe, Shaun Nichols, Maria Grazia Rossi, and Liane Young for their insightful comments on a previous version of this article. The present findings have been presented at the "Winter Workshop 2008 on Games, Experiments and Philosophy" at the Max Planck Institute of Economics, in January 2008 in Jena, Germany. This work was made possible by the Fondation France Telecom and the Fondation FondaMental.

Address for correspondence: Tiziana Zalla, Institut Jean Nicod- CNRS, Ecole Normale Supérieure, 29 rue d'Ulm 75005 Paris, France. **E-mail:** tiziana.zalla@ens.fr

1. Introduction

The concept of intentional action plays a central role in social cognition. We distinguish between intentional and non-intentional actions when we explain people's behavior: We appeal to causes to explain non-intentional actions, but to reasons to explain intentional ones (Bratman, 1987; Malle & Knobe, 1997b). We also rely on this distinction when we evaluate people's behavior in light of our moral or conventional norms: Wrong intentional actions are judged to be worse than similar non-intentional actions (Lagnado & Shannon, 2008).

While much is known about the concept of intentional action in typical populations (for review and discussion of the literature in the 1980s and 1990s, see Malle & Knobe, 1997a; for more recent work, see Knobe, 2003, 2006, 2007; Nadelhoffer, 2004, 2006a, b; Mele & Cushman, 2007; Nichols & Ulatowski, 2007; Phelan & Sarkissian, 2008, 2009; Machery, 2008; Mallon, 2008; Wright & Bengson, 2009), our understanding of this concept among people with atypical social cognition, particularly people with autism spectrum disorders, remains incomplete (Phillips, Baron-Cohen & Rutter, 1998; Russell & Hill, 2001; Zelazo et al., 2002; Leslie, Mallon, & DiCorcia, 2006; Zalla et al., 2009).

In this article, we focus on people with Asperger Syndrome (AS) or High Functioning Autism (HFA), and we examine how they conceptualize the intentional status of actions. We provide evidence that people with AS/HFA's judgments about intentional actions differ from typical individuals' when deciding whether an action is intentional requires complex mindreading processing—particularly, when one has to judge that a side effect or a means is intentionally brought about in spite of not being intrinsically valued. Furthermore, we provide evidence that to compensate for their difficulties, people with AS/HFA rely on a simple moral heuristic: If an action violates a moral norm, then it is likely seen as intentional.

Here is how we will proceed. In Section 2, we will develop our hypotheses and review the existing literature. In Section 3, we will describe the participants in our two studies. In Section 4, we will present the first study, which shows that people with AS/HFA conceptualize the intentional status of some means (*viz.* means that are negatively valued) differently from typical individuals. In

Section 5, we present some evidence that just like typical individuals, people with AS/HFA use the moral valence of side effects to evaluate their intentional status. We also show that people with AS/HFA's judgments of blame and praise appear to be disconnected from their judgments about the intentional status of actions. We conclude that people with AS/HFA have difficulty evaluating the intentional status of actions when doing so require some complex mindreading processes (as is the case for those actions that are only desired as means for something else) and that to compensate for these difficulties, they have developed a simple heuristic that relies on their heightened sensitivity to norms and norm violations.

2. Hypotheses

2.1 The Compositional Structure of Actions

Many events are concomitant with any intended action. For example, when someone enters in her own apartment (the intended action), she inserts the key in the keyhole, turns the key, pushes the door, and steps in (the concomitant events). Concomitant events include means as well as side effects (e.g., making someone else startle because of the noise produced by turning the key in the lock). When people perceive *a token action* or are told about it, they do not merely register the sequence of events; rather, they structure this sequence into a representation of the token action: One of the events is represented as the intended outcome, while other events are represented as means for this outcome and yet other events as side effects. Concomitant events stand in temporal relation with one another and with the intended action: The action consisting in inserting the key in the keyhole precedes the action consisting of turning the key. They can also stand in causal relation to one another and to the intended action. Suppose someone buys a soft drink in a vending machine, an action that includes the actions consisting in inserting some coins in the machine and in getting the soft drink. The first event is causally related to the second one, since it is a means to an end. Similarly, our conceptual knowledge of *action types* involves representations of the organization of the sequence of concomitant events. According to Schank and Abelson (1977), actions or events are represented in memory as scripts—viz. as organized sequences of goal-directed events. A script

specifically refers to the temporal and the causal links between the concomitant events and to the hierarchical order of goals and sub-goals—that is, how the sub-goals contribute to bringing about the ultimate goal.

People typically judge that bringing about one's ultimate goal (e.g., entering one's apartment) is intentional. They also judge that *some* subordinate events that are causally and hierarchically related to this outcome (e.g., putting the key in the lock)—*but not all* (e.g., making someone else startle while turning the key)—are intentional. What we are interested in here is how people decide whether the concomitant events are intentional. Consider for instance the following situation. Someone wants to buy a smoothie and is told that the price of the drink she desires has increased by one dollar. She pays this extra dollar. Ask yourself: Did she pay it intentionally? Consider now the following situation. Someone wants to buy a smoothie and is told that her smoothie will be served in a commemorative cup (with no additional cost). She receives the commemorative cup. Ask yourself: Did she receive a commemorative cup intentionally? When participants were presented with vignettes describing these two situations (see Section 4), 95% of them judged that paying an extra dollar was intentional, while only 45% judged that obtaining the commemorative cup was intentional (Machery, 2008). How did they come to these judgments?

There is little doubt that there are several distinct strategies available to judge whether a concomitant event is intentional. However, we propose that judging that a concomitant event is intentional typically involves the following representations and inferences. For the sake of simplicity, we illustrate this with the extra-dollar and free-cup situations. To determine whether getting a free cup and paying an extra dollar are intentional, participants (the ascribers) typically construct at least three consecutive representations of the events described in the vignettes: (1) a causal representation, (2) a valence-based representation, and (3) a mentalistic representation. A causal representation represents the causal relations between events (represented by arrows in the figures below). For instance, in the extra-dollar situation, the event *paying an extra dollar* causes

the occurrence of the event *getting a smoothie* (Figure 1).¹ A valence-based representation represents the valence of the different events for the agent described in the vignette (the ascriber). Valences can be positive (represented by a “+” in the figures below), negative (represented by a “-”), neutral (represented by a “N”), or underdetermined (represented by a “?”). Valences are evaluated by reference to what the agent values intrinsically (in contrast both to what he or she might desire for instrumental reasons and to what the ascriber—i.e. the participant reading the vignette—values). If the agent in the extra-dollar case desires his thirst to be quenched, the event consisting in the agent’s thirst being quenched is positively valued (Figure 2). Note that even though the agent’s values are supposed to be the ones that are relevant to determining the valence of the events under consideration, people may accidentally project their own values onto the values of the agents described in the vignettes. A mentalistic representation represents the events as being the objects of the agent’s mental states, in particular as being the objects of the agent’s instrumental or ultimate desires. For instance, in the extra-dollar case, a mentalistic representation represents the event *paying an extra dollar* as being the object of a purely instrumental desire (Figure 3). It is *purely* instrumental because the agent negatively values this event. It would be simply instrumental if the agent valued the event as something good and wanted it to occur for instrumental reasons.²

Let us highlight an important feature of purely instrumental desires. When people attribute a purely instrumental desire to an agent, their representation of how the agent values an event (viz. the object of the purely instrumental desire) diverges from their representation of her desires. The agent might value negatively paying an extra dollar, but she might nonetheless instrumentally desire paying this extra dollar as a means for reaching her ultimate goal (quenched her thirst). Thus, people need to clearly distinguish what an agent might value from what she might desire. Clearly,

¹ We use italics to name events.

² As we will explain in more detail below, we take that someone wants something for instrumental reason either when it is a means for fulfilling her ultimate desire or when it is a negatively valued event that is necessary for her ultimate desire to be fulfilled.

typically individuals are able to draw this distinction. This is however not an easy distinction to draw, and people with an impaired capacity for mindreading might have difficulty drawing it.

The three representations we have described are successively constructed, each on the basis of the previous one. The causal representation of the events under consideration makes use only of our capacity for causal reasoning. The valence-based representation makes use of some of our mindreading capacities: In order to produce this representation, people must be able to represent the agent as having preferences, likes and dislikes. Note that only some mindreading capacities—the ones involved in representing an organism as having likes and dislikes—are engaged by the production of the valence-based representation. By contrast, the third representation seems to appeal to the whole gamut of our mindreading capacities. First, ultimate and instrumental desires are ascribed to the agent (Figures 3 and 6 below). Furthermore, the ascription of a purely instrumental desire to an agent supposes the ascription of a belief. In the extra-dollar situation, the agent does not want to pay a dollar (Figure 2 below), but if we understand that the agent believes that it is necessary for him to pay this extra dollar in order to obtain what he wants (a smoothie), then we can ascribe to him the instrumental desire to pay an extra dollar.³ In other words, the ascriber views this instrumental desire as a cognitively derived desire generated by beliefs about what will lead to the fulfillment of the ultimate desire.

Thus, we propose that when someone is told about the extra-dollar situation, she successively produces the following three representations.

Figures 1-3 about here

By comparison, we propose that when someone is told about the free-cup case, she successively produces the following three representations.

Figures 4-6 about here

³ Of course, means are not always necessary for the ultimate desire to be fulfilled, since there are often several possible ways to fulfil an ultimate desire.

The intentional representations of the events described in the vignettes are used to determine whether the events under consideration—viz. paying an extra dollar in the extra-dollar situation and getting a free cup in the free-cup situation—are intentional. An event is judged to be intentional because it is the object of a desire. This analysis suggests that judging the intentional status of events that are concomitant with an intended action (e.g., the event of paying an extra dollar that is concomitant with the action consisting of buying an extra-large smoothie) involves understanding the structure of the whole action. This can be a complex process that involves understanding what the agent values and ascribing instrumental and ultimate desires to her.

2.2 Moral Cognition and Judgments about Intentionality

Malle and Knobe (1997a) have proposed a well-known account of the folk concept of intentional action. They propose that for an action to be judged intentional, it has to have the following five properties: The agent has to *desire* the outcome of this action; she also needs to *believe* that her action will bring about this outcome; in addition, she has to *intend* to perform the action; she needs to be *aware* of performing it while performing it; finally, she must have some *skill* to perform it.⁴ Intention is a necessary component of intentional action because an agent can desire an outcome x and know that an action y will bring about x , but do y (and thereby bring about x) by accident (Bratman, 1987). Suppose that John desired to kill his wife, that he knew that pouring some poison in her coffee would kill her, and that he poured some poison in her coffee without intending to do so (because the poison looked like sugar); then, we would judge that he did not intentionally kill his wife. Awareness is a necessary property of intentional actions because an agent might intend to do x , but nonetheless do x by accident (Searle, 1983). Suppose John intended to call successively his brother and his sister; suppose also that he dialed his sister's phone number while believing he was calling his brother; then, we would judge that he did not intentionally call his sister. Finally, people often judge that an agent does x unintentionally, even if she has the intention to do x , when she is

⁴ It also seems that taken together, these five properties are meant to be sufficient for being intentional.

merely lucky to do x —viz. when she has no skill to bring about x (O’Shaughnessy, 1980). Figure 7 summarizes Malle and Knobe’s model.

Figure 7 about here

Our account of the representations and inferences involved in judging whether a concomitant event is intentional differs from, but is consistent with, Malle and Knobe’s original account. In contrast to them, we have proposed a dynamic model of how people typically judge whether a concomitant event is intentional, while their account is static. Like Malle and Knobe, we also view (either instrumental or ultimate) desire as necessary for an action to be judged intentional.

Recent research has cast some doubts on whether desire and skill are really necessary for an action to be judged intentional (Knobe, 2003a, 2003b, 2006; Nadelhoffer, 2004; Mele & Cushman, 2007). Let’s focus first on desire. Knobe (2003b, 2006) designed a pair of cases (the harm and help cases) in which an agent brings about a foreseen side effect, while saying explicitly that he does not care about this side effect and thus expressing no desire to bring it about (but see below). In the harm case, the foreseen side effect (harming the environment) is plausibly morally bad, while in the help case it is plausibly morally good. Participants were significantly more likely to judge that the chairman intentionally harmed the environment (82%) than they were to judge that he intentionally helped the environment (23%). This effect is now known as “the side-effect effect.”⁵ The side-effect effect suggests that in addition to people’s beliefs about the psychological states of the agent, the moral value of her action influences people’s judgments about the intentional status of this action. Furthermore, it suggests that when an action is morally wrong, desire is not necessary for it to be judged intentional.

Let’s turn now to skill. Knobe has also argued that moral considerations play a role in cases where the agent lacks *skill* in accomplishing an intentional action (Knobe, 2003a; Nadelhoffer, 2004; Mele & Cushman, 2007). Knobe designed a pair of cases (the murder and the bull’s eye cases) in which an agent is trying to perform a behavior although he does not really have the skill to perform that

⁵ The side-effect effect has been replicated with numerous other pairs.

behavior in any reliable fashion. Ultimately, the agent only succeeds through mere luck. In the murder case, the agent's behavior is plausibly immoral, whereas in the bull's-eye case the agent's behavior is plausibly morally neutral. Knobe (2003a) found that the moral significance of the behavior influences people's intuitions about the intentional status of the action: Only 23% of the participants judged that the agent intentionally hit the target in the bull's-eye case, whereas 91% judged that the agent acted intentionally in the murder case.

It is currently unclear how to revise our understanding of the folk concept of intentional action in light of these findings (Nadelhoffer, 2006; Knobe, 2007; Mele & Cushman, 2007; Nichols & Ulatowski, 2007; Machery, 2008; Mallon, 2008; Phelan & Sarkissian, 2008, 2009; Wright & Bengson, 2009; Guglielmo, Monroe, & Malle, forthcoming; Guglielmo & Malle, ms). Particularly, Knobe's findings suggest that people are willing to judge some actions to be intentional even when the agent does not desire performing these actions. However, it is unclear whether in the harm case the CEO has really no desire whatsoever toward harming the environment. Machery (2008) has proposed that people judge negative side effects such as harming the environment to be intentional because they conceptualize harming the environment as a cost that the agent is willing to incur in order to get a benefit ("the trade-off hypothesis"). In the terminology developed in this article, this amounts to proposing that although it is not a means, the side effect in the harm case (but not in the help case) is viewed as the object of a purely instrumental desire (Figure 8). Thus, according to the trade-off hypothesis, the side effect in the harm case (harming the environment) and the means in the extra-dollar case (paying an extra dollar) are conceptualized similarly: Both are viewed as costs that an agent knows she must incur in order to get a desired benefit; both are viewed as objects of purely instrumental desires (compare Figures 3 and 8).

Figure 8 about here

It is worth noting here that we are broadening the concept of instrumental desire. It is typically thought that only means are the objects of such desires. By contrast, we think that in addition to means, any negatively valued event that is necessary for the fulfillment of an ultimate desire (including a necessary, foreseen side effect) can be the object of an instrumental desire. We are

aware that this conception of instrumental desire is controversial, and we hasten to add that in the experiment we report below (Section 4) the agent's desire would count as an instrumental desire under any plausible definition.

In any case, according to the trade-off hypothesis, the mentalistic relation of the agent to harming the environment is distinct from his mentalistic relation to helping the environment. Because he is willing to incur the cost of harming the environment, he has a pro-attitude toward harming the environment; more exactly, he has a purely instrumental desire to harm the environment. By contrast, he has no desire whatsoever to help the environment.

Evidence for the trade-off hypothesis is mixed. Consistent with this account, Guglielmo and Malle (ms) have recently found that people are willing to ascribe to the CEO some desire to harm the environment. Against the trade-off hypothesis (see also Phelan & Sarkissian, 2009; Wright & Bengson, 2009), Mallon (2008) found that participants judged that some negative side effects that cannot be interpreted as costs are intentionally brought about (in response, see Nanay, forthcoming).

This is not the appropriate occasion to resolve this controversy. It might be that contrary to what our description of the process by which we arrive at a judgment about the intentionality of concomitant events suggests and contrary to the trade-off hypothesis, some actions that are *morally wrong* are indeed judged intentional, although they are not the object of any desire whatsoever. But what matters for this article is the following point: When one has to judge whether *morally neutral* actions are intentional, it seems necessary that they be the object of some kind of desire. People must then be able to ascribe such desire when they judge that a morally neutral action is intentional.

2.3 Asperger Syndrome, Intentional Action, and Moral Judgments

Autism spectrum disorders (ASD) are pervasive developmental disorders characterized by abnormal social interaction, verbal and non-verbal communication problems, restricted interests, and disruptive stereotypic movements. Within the domain of ASD, "High Functioning Autism" (HFA) commonly refers to individuals meeting criteria for autism with normal intellectual functioning and a history of speech and language delay. Those at the higher-functioning end of the

HFA group, sometimes diagnosed with Asperger Syndrome (AS) show no evidence of impaired language function, and their intellectual abilities fall within the normal range⁶. As with other individuals within the ASD, the clinical features of HFA and AS include trouble forming friendships, difficulties with social cognition, inappropriate social interactions, restricted interests, diminished capacity for empathy, and poor communication (Frith, 1996, 2004; Bauminger & Kasari, 2000). Importantly, and contrary to the more severe forms of autism, individuals with AS may pass tests of Theory of Mind (ToM) and thus have some ability to attribute intentions, beliefs, and desires to others. Specifically, they can often solve first-order (e.g., “Sally thinks it’s *x*, when really it’s *y*”) and second-order false beliefs tests (e.g., “Sally thinks Mary thinks *x*, but both Sally and Mary are wrong”) (Bowler, 1992; Happé, 1995; Dahlgren & Trillingsgaard, 1996), although they may fail more advanced ToM tasks, based on detection of sarcasm, irony, or bluff (Happé, 1994) or on recognition of Faux Pas (Baron-Cohen et al., 1999; Zalla et al., 2009).

The understanding of volitional states, such as desires and intentions, which emerge earlier than beliefs in typical ontogenesis, appeared to be relatively preserved in autism, as many individuals with the disorder have been found to pass simple desire tests (Baron-Cohen, 1991). However, it has also been shown that more complex aspects of desire are not understood by pre-school children and many people with ASD (Phillips, Baron-Cohen, & Rutter, 1995) and that difficulties might arise when a desire is caused by a belief and not by an external event (Baron-Cohen, 1991). In addition, adults with HFA or AS may exhibit difficulties in the application of some mentalistic concepts, such as the concept of intentional action (Zalla et al., 2009).

As we saw earlier, judgments about the intentional status of side effects (such as getting a free cup in the free-cup case or harming the environment in the harm case) or means (such as paying an extra

⁶ Asperger syndrome is a pervasive developmental disorder included in the ICD-10 (World Health Organization 1992) and the DSM-IV-TR (American Psychiatric Association 2000). The validity of AS, however, as a distinct diagnostic entity from other pervasive developmental disorders, separate from high functioning autism, has been neither conclusively established nor refuted (Leekam et al., 2000; Wing, 2005).

dollar in the extra-dollar case) typically result from people constructing a mentalistic representation of these events. One judges that paying an extra dollar is intentional because one has represented this action as the object of a purely instrumental desire (Figure 3), while one judges that getting a free cup is not intentional because this action is not instrumental and one has not represented it as the object of any desire (Figure 6). Constructing the relevant mentalistic representations supposes that one is able to ascribe instrumental and ultimate desires to the agent, a process that might require the whole gamut of our mindreading capacities. As we argued above, to ascribe a purely instrumental desire about an action, one needs to ascribe a belief that this action can causally bring about the fulfillment of the ultimate goal of the agent. Ascribing such a belief allows the ascriber to understand why the agent desires something that she either has no ultimate desire for (if the value of the action for the agent is neutral) or even something she desires to avoid (if its value is negative).

The first hypothesis we put forward in this article is that ascribing purely instrumental desires is particularly challenging for individuals with impaired mindreading capacities, and that this might be the case for people with AS/HFA. The reason is that to ascribe a purely instrumental desire is a mindreading feat, as we just saw. If people with AS/HFA really have difficulty ascribing purely instrumental desires, they should then have difficulty understanding the intentional status of some actions—namely those actions that are not intrinsically desired, either because they are neutral or because they are negatively valued, but that are desired as means for a desired goal. That is, we expect people with AS/HFA to have difficulty with the intentional status of some actions when viewing these actions as intentional involves understanding that an agent can instrumentally desire something that she does not value.

Although research on how adults with AS/HFA conceptualize intentional actions is limited, some studies with people with ASD provide relevant evidence for our first hypothesis.⁷ According to

⁷ Carpenter, Pennington and Rogers (2001) have suggested that the comprehension of intentions is not severely impaired in people with ASD. They reported that young children with autism were able to understand children's understanding of others' unfulfilled intentions and the goal state of an intended action,

Philips, Baron-Cohen, and Rutter (1998), young participants with autism showed difficulties in distinguishing the unintended from the intended outcomes of their own actions as compared to a control group, when the unintended actions were positively valued. The experimental task involved firing a ray gun to hit one of several colored targets, and participants either hit the chosen target or accidentally hit a target with a different color. The value (and thus desirability) of the outcomes was manipulated in order to make some outcomes intended, but not desirable (the participant hit the chosen target, but its value was low), and others desirable, but unintended (the participant did not hit the chosen target, but the value of the target she accidentally hit was high). Participants were asked to judge whether they had “meant to” hit the target by comparing their goal and the outcome of their action. As expected, children’s understanding of the intended/unintended distinction improved with age: In the comparison group, older (5/6-year-old) children were better than younger (4-year-old) children at recognizing that hitting a desired target can be performed unintentionally. Moreover, compared with children with similar intellectual abilities, there was a tendency in young children with autism to find it difficult to understand that outcomes that are desirable might nevertheless be unintended. These findings suggest that, consistent with our first hypothesis, people with ASD might have difficulty judging that an outcome can be non-intentional when it is positively valued or conceptualizing an action as intentional when the agent negatively values it.⁸

As we have seen above, people typically use the moral value of actions as a cue to their intentional status. But what about people with AS/HFA? While relatively few studies have explored suggesting that although they have a slightly less complex understanding of others’ intentions, disturbances in this domain are not as marked as deficits in ToM and joint attention. However, the authors raised the possibility that their finding of no autism-specific deficits in this understanding may be a result of relatively poor performance by their control group. Moreover, for criticisms of the unfulfilled intention paradigm, see D’Entremont and Yazbek (2007) and Huang, Heyes, and Charman (2002).

⁸ Other findings suggest that children with ASD have a difficulty distinguishing intentional from accidental actions in others, but they are not directly relevant for hypothesis 1 (D’Entremont & Yazbek, 2007).

moral cognition in individuals with AS, existing evidence agreed about preserved abilities in reasoning about moral facts in people with ASD (Blair, 1996; Grant et al., 2005; Leslie et al., 2006). Blair (1996) showed that in spite of their impairments in theory of mind, children with autism are able to draw the distinction between moral and conventional transgressions and exhibit automatic physiological responses to the distress of others, when this distress results from moral violations (Blair, 1996, 1999). More recently, Leslie et al. (2006) have showed that, in individuals with autism, moral judgments are not the simple result of affective responses to others' distress, but that they also reflect some spared moral reasoning. Moreover, Grant and collaborators (2005) showed that children with high functioning autism can use motives in judgments of culpability when the outcome (others' distress) is negative. They could also distinguish between damage to property and damage to persons and between social-conventional rules and moral rules, and they judged damage to persons more severely than damage to properties, although their justifications were of poor quality, with fewer references to the protagonist's intention.

In a previous study using the Faux Pas task, Zalla and collaborators (2009) have showed that the application of the concept of intentional action in adults with AS/HFA is predominantly informed by an increased sensitivity to rule violation. The faux pas is a particular case of a non-intentional action: As defined by Baron-Cohen and collaborators (1999), a faux pas occurs when a speaker says something that is unpleasant to the listener, although the speaker never intended to do so. Detecting faux pas requires the ability to understand the speaker's state of mind and to appreciate the emotional impact of the statement on the listener. Participants were presented with twenty stories describing interpersonal interactions in everyday life situations in which a faux pas occurred and twenty control stories containing a minor conflict or accident.⁹ This study showed that individuals with AS/HFA have difficulty conceptualizing faux pas as non-intentional. Although they

⁹ For each story, the following six questions were asked: 1) the detection question: "In the story did someone say something awkward or something they should not have said?"; 2) the person identification question: "Who said something they should not have said?"; 3) the content question: "What did they say that they should not have said?"; 4) the explanation question: "Why shouldn't they have said it?"; 5) the belief question: "Did they know/remember that?"; 6) the empathy question: "How did the listener feel?"

considered faux pas a socially inappropriate behavior, they were generally unable to provide correct motives and appropriate reason explanations when required to explain their judgments. They tended to view faux pas as having been done intentionally. Typically, individuals with AS/HFA provided explanations in terms of malicious intentions, and they judged that the character committing the faux pas intended to humiliate and to offend the listener. In some cases, negative judgments about the character committing the faux pas were presented in terms of personality traits: She was said to be arrogant, selfish, or pretentious. By contrast, for control participants, a faux pas is a non-intentional by-product of an intentional act based on some false belief. Thus, as Zalla et al. (2009) hypothesized, when asked to judge actions that are socially inappropriate, people with AS/HFA might rely on the following heuristic: If an action is morally wrong, then the agent intentionally performed the action. Based on the findings just reviewed, one would expect people with AS/HFA to be responsive to the moral value of actions. Indeed, the second hypothesis investigated in this article posits that people with AS/HFA use the moral value of actions as a cue to their intentional status. As a result, although their judgments about intentional actions may differ from typical individuals' (our first hypothesis), both groups are bound to make similar judgments when the action is morally wrong, since individuals with typical development also rely on the moral value of an action to judge its intentional status, as Knobe's and others' research shows.

2.4 Predictions

The aim of the present study is to investigate whether judgments about intentional action by people with AS/HFA, who are thought to be responsive to moral considerations along with being somewhat impaired in ToM abilities, differs from those of people with typical development.

Study 1 examines this question by presenting participants with a scenario that involves neutral actions and that requires complex mindreading processing. We used the free-cup and the extra-dollar cases to assess whether people with AS/HFA have difficulty conceptualizing the intentional status of morally neutral instrumental actions, i.e., those actions that are means for another action, but that are not the object of an ultimate desire. Because understanding the intentional status of

actions that are the objects of purely instrumental desires requires extensive and complex mentalization and because the moral value of the action cannot be used as a cue when the action is morally neutral, we predicted that people with AS/HFA's judgments would differ from typical individuals'.

In the second experiment, we used two pairs of vignettes—Knobe's harm and help cases (Knobe, 2003b, 2006) and his murder and bull's-eye cases (Knobe, 2003a)—to assess whether moral considerations in individuals with AS inform their judgment about intentionality in two different circumstances: when the agent apparently does not desire a side effect of her action and when the agent lacks the required skill to bring about her goal. Based on previous findings, we predicted that people with AS would use the moral value of the outcomes as a cue to the intentional status of actions. As a result, their answers should be similar to the judgments made by typical individuals (Knobe, 2003a, b, 2006).

Importantly, the similarity between typical individuals and individuals with AS/HFA does not mean that the processes that underwrite judgments about intentionality are the same in people with AS/HFA and in typical individuals. People with AS/HFA might have difficulty judging the intentional status of actions when these judgments require complex processing, while relying on a compensatory heuristic based on normative reasoning when actions violate norms.

3. Participants

Demographic and clinical description of participants is detailed in table 1.

--- Table 1 about here ---

A group of twenty adults with a clinical diagnosis of AS or HFA according to DSM-IV R (American Psychiatric Association, 1994) and to ASDI (Asperger Syndrome Diagnostic Interview, Gillberg et al., 2001) were recruited from Albert Chenevier Hospital in Créteil, France. All diagnoses, made by clinicians experienced in the field of autism, were based on observations of the participants. Interviews with parents using the ADI-R (Autism Diagnostic Interview, Lord et al., 1994) confirmed the diagnoses; elevated scores indicated problematic behavior in the three

following areas: reciprocal social interaction, communication, and stereotyped behaviors. The cut-off points for the three classes of behavior are reciprocal social interaction 10, communication 8, and stereotyped behaviors 3, respectively. All participants scored above the cut-offs points. As part of the checking process, the French translation of A-TAC (Autism, tics, AD-HD and other comorbidities; Hansson et al., 2005) was completed by the parents. This screening questionnaire is focused on a number of abilities, conducts, and behaviors in a child's functioning as compared to his or her peers. Parents are asked to report any problem or specific characteristic during any period of life, even when this is no longer present.

All patients were medication-free at the time of testing and had no history of psychiatric disorder (other than AS/HFA), neurological disorder, or head injury. All participants were native French speakers, had normal/corrected to normal vision, and signed informed consent before volunteering for this study, in accordance with the local ethics committee and the Declaration of Helsinki.

Twenty-eight comparison participants were taken from the general population and were chosen to match the clinical group with respect to the characteristics of age, IQ, gender, and education. Participants with AS/HFA and CPs received Verbal and Performance IQs (WAIS-III, Wechsler, 1999) and the Faux Pas Detection Task (max score=60) (Baron-Cohen et al., 1999) to assess Theory-of-Mind abilities. Overall, unpaired t-test revealed that individuals with AS/HFA did not differ from the CPs on chronological age ($t(46) = -0.139, p = 0.89$), on education ($t(46) = 0.293, p = 0.77$), and on IQ level (Full-scale, Verbal and Performance: t-test: $t(39) = 0.22, p = 0.82$; $t(39) = 0.87, p = 0.38$; $t(39) = -0.82, p = 0.41$). On the Faux Pas test, the group with AS/HFA scored significantly lower than the comparison group, consistently with what is expected from the clinical presentation of the syndrome ($t(34) = 4.936, p < 0.0001$). Data on Verbal and Performance IQs could not be obtained for seven CPs and only a subgroup of sixteen CPs received the Faux Pas test. Here, AS/HFA individuals scored significantly lower than the CPs, consistently with what expected from the clinical presentation of the syndrome ($t(40) = -4.272, p < 0.0001$).

4. Study 1

4.1 Participants

A group of sixteen adults with AS/HFA (mean age = 28.1 ± 6.9) and twenty-eight adult comparison participants participated to this experiment. The AS/HFA group was taken from the larger sample described above in Section 3 (table 1) and did not differ on chronological age ($t(42) = 0.1$; $p = 0.92$), on education ($t(42) = 1.25$; $p = 0.21$), and on IQ level (Full-scale, Verbal, Performance: t-test: $t(42) = 0.22$, $p = 0.82$; $t(42) = 0.87$, $p = 0.38$; $t(42) = -0.82$, $p = 0.41$) from the CP group. They also scored significantly lower than the CPs on the Faux Pas test ($t(40) = -4.272$, $p < 0.0001$).

4.2 Procedure

Participants were presented with the French translations of the extra-dollar and free-cup vignettes written on a page placed in front of them throughout the questioning so that they did not have to remember the scenarios (see appendix for the French version of these vignettes).

The Extra-Euro Case

Joe was feeling quite dehydrated, so he stopped by the local smoothie shop to buy the largest sized drink available. Before ordering, the cashier told him that the Mega-Sized Smoothies were now one euro more than they used to be. Joe replied, ‘I don’t care if I have to pay one euro more, I just want the biggest smoothie you have.’ Sure enough, Joe received the Mega-Sized Smoothie and paid one euro more for it. Did Joe intentionally pay one euro more?

The Free-Cup Case

Joe was feeling quite dehydrated, so he stopped by the local smoothie shop to buy the largest sized drink available. Before ordering, the cashier told him that if he bought a Mega-Sized Smoothie he would get it in a special commemorative cup. Joe replied, ‘I don’t care about a commemorative cup, I just want the biggest smoothie you have.’ Sure enough, Joe received the Mega-Sized Smoothie in a commemorative cup. Did Joe intentionally obtain the commemorative cup?

The order of vignette presentation was counterbalanced between participants. After reading the extra-euro vignette, participants were asked to answer the following two questions:

- 1) Did Joe intentionally pay one euro more?
- 2) Was Joe's paying an extra euro blameworthy, praiseworthy, or neutral?

For the free-cup case, participants were asked to answer the following two questions:

- 1) Did Joe intentionally obtain the commemorative cup?
- 2) Was Joe's obtaining the commemorative cup blameworthy, praiseworthy, or neutral?

If hypothesis 1 is correct, then we would expect people with AS/HFA to differ from typical individuals when they are presented with the extra-euro and free-cup cases. To understand the sequence of events described in the extra-euro case correctly—particularly, to understand that paying an extra euro *is* intentional—one needs to ascribe a purely instrumental desire to the agent. Thus, we should expect people with AS/HFA to be less likely than people without AS/HFA to judge that paying an extra euro is intentional. By contrast, to understand the sequence of events described in the free cup case correctly—particularly, to understand that getting a free cup is *not* intentional—one does not need to ascribe a purely instrumental desire to the agent. Thus, we should not expect people with and without AS/HFA to disagree on the intentional status of getting a free cup.

4.3 Results: Intentionality

As expected, when we examined French comparison participants' answers, we found the same pattern of responses as the one previously reported by Machery (2008). 64.3 % of comparison participants reported that Joe intentionally paid an extra euro, while only 14.3% said that Joe intentionally obtained the commemorative cup. A chi-square test yielded a highly significant difference between the two cases for the intentionality question ($\chi^2(1) = 52.5, p < 0.0001$). By contrast, only 25% of participants with Asperger Syndrome said that Joe intentionally paid an extra euro to get a mega-sized smoothie, and 25% said that Joe intentionally obtained the commemorative

cup. The two groups significantly differed in their judgment of intentionality on the extra-euro case ($\chi^2(1) = 29.2, p < 0.0001$), whereas there was no difference on the free-cup case ($\chi^2(1) = 3.18, p = \text{ns}$).

4.4 Results: Praise and Blame

When asked to judge the praiseworthiness of Joe's action, most CPs judged in both cases Joe's action to be neutral: 96.4% and 100% of them considered his action as neutral in the extra-euro case and the free-cup case, respectively. Moral judgment in participants with AS/HFA slightly differed in that only 75% considered Joe's action in both cases to be morally neutral. 12.5% judged Joe's action in the extra-euro case as being praiseworthy and 12.5% as being blameworthy. This difference did not reach significance, although it was trending toward significance ($\chi^2(2) = 5.12, p = 0.07$). Conversely, in the free-cup case, 25% considered Joe's action as praiseworthy. This difference was significant ($\chi^2(1) = 4.97, p = 0.02$).

4.5 Discussion

Comparison participants evaluated differently the intentional status of paying an extra euro and of getting a free cup (see also Machery, 2008). By contrast, participants with AS/HFA treated them similarly, judging both to be non-intentional. These findings show first that adults with AS/HFA can distinguish intentional from non-intentional actions. Second, they support the hypothesis that people with AS/HFA have difficulty understanding the intentional status of those actions that people desire for purely instrumental reasons. Let us elaborate on this point in the remainder of Section 4.

Participants with AS/HFA and CPs judged that the target action in both vignettes was neutral. As a result, participants with AS/HFA were not in a position to rely on their sensitivity to normative violations to judge the intentionality of an action, as they might do when the action is normatively inappropriate. Thus, successful evaluation of the intentional status of the action in the extra-euro and free-cup cases requires participants with AS/HFA to correctly represent the structure of the action described in the vignettes.

Now, according to the proposed description of the representations and inferences needed to decide whether the events that are concomitant with an intended action (such as paying an extra euro or getting a free cup) are intentional (Section 1), one decides that an action is intentional if it is represented as being the object of an ultimate or instrumental desire. Comparison participants are supposed to represent paying an extra euro as the object of a purely instrumental desire and, as a result, as intentional (Figure 3). If participants with AS/HFA have difficulty representing neutral or negative actions as being desired for instrumental reasons, they should have difficulty conceptualizing paying an extra euro as the object of a purely instrumental desire; as a result, they might judge that the agent did not intentionally pay an extra euro to get an extra-large smoothie—exactly what we found. Our hypothesis is further supported by the fact that when asked why they believed that paying an extra euro was not intentional, many of them answered that Joe could not have acted differently or that it was not a tip but the price decided by the seller. They thus failed to conceptualize the action *paying an extra euro* as a means that the agent is willing to bring about in order to get something he desires.

By contrast, to characterize the intentional status of the action in the free-cup case correctly, one needs not represent *getting a free cup* as the object of an (instrumental or ultimate) desire, since it is neither a means for nor a necessary foreseen side effect of the agent's goal (Figure 6). The hypothesized difficulties of participants with AS/HFA to represent events as objects of purely instrumental actions would thus not prevent them from correctly judging that the agent did not intentionally get a free cup.

5. Study 2

5.1 Participants

The group of twenty adults with a clinical diagnosis of AS or HFA and twenty-eight adults with typical development that were described in Section 3 participated to this study.

5.2. Procedure

Participants were presented with two pairs of vignettes—the harm and help cases and the murder and bull’s-eye cases—written on a page placed in front of them throughout the questioning so that they did not have to remember the scenarios (see appendix for the French version of these vignettes). The order of vignette presentation was counterbalanced between subjects.

The Harm Case

The vice-president of a company went to the chairman of the board and said, ‘We are thinking of starting a new program. It will help us increase profits, but it will also harm the environment.’ The chairman of the board answered, ‘I don’t care at all about harming the environment. I just want to make as much profit as I can. Let’s start the new program.’ They started the new program. Sure enough, the environment was harmed.

The Help Case

The vice-president of a company went to the chairman of the board and said, ‘We are thinking of starting a new program. It will help us increase profits, and it will also help the environment.’ The chairman of the board answered, ‘I don’t care at all about helping the environment. I just want to make as much profit as I can. Let’s start the new program.’ They started the new program. Sure enough, the environment was helped.

After reading the harm and help cases, participants were asked the following two questions:

- 1) Did the chairman harm (or help) the environment intentionally? (Yes, No)
- 2) Was the chairman’s harming (or helping) the environment blameworthy, praiseworthy, or neutral?

The murder case and the bull’s eye case read as follows.

The Murder Case

Jake desperately wants to have more money. He knows that he will inherit a lot of money when his aunt dies. One day, he sees his aunt walking by the window. He raises his rifle, gets her in the sights, and presses the trigger. But Jake isn’t very good at using his rifle. His

hand slips on the barrel of the gun, and the shot goes wild... Nonetheless, the bullet hits her directly in the heart. She dies instantly.

The Bull's-Eye Case

Jake desperately wants to win the rifle contest. He knows that he will only win the contest if he hits the bulls-eye. He raises the rifle, gets the bull's-eye in the sights, and presses the trigger. But Jake isn't very good at using his rifle. His hand slips on the barrel of the gun, and the shot goes wild. Nonetheless, the bullet lands directly on the bull's-eye. Jake wins the contest.

For the murder and the bull's-eye cases, participants were asked the following two questions:

- 1) Did Jake shoot his aunt / hit the bull's-eye intentionally? (Yes, No)
- 2) Was Jake's behavior of shooting his aunt / hitting the bull's-eye blameworthy, praiseworthy, or neutral?

5.3 Results: Intentionality

When asked whether the chairman harmed or helped the environment intentionally, a pattern of response similar to that reported by Knobe (2003, 2004) was observed: Both CPs (82.1 %) and participants with AS/HFA (100%) said that the chairman harmed the environment intentionally, whereas only 17.9 % of CPs and 36.8% of people with AS/HFA reported that the chairman helped the environment intentionally. The difference between the two groups for both the harm case ($\chi^2(1) = 2.3, p = 0.12$) and the help case ($\chi^2(1) = 1.78, p = 0.18$) was not significant.

In the murder case, 96.4% of CPs and 85% of participants with AS/HFA judged that Jake shot his aunt intentionally, whereas in the bull's-eye case only 50% of CPs and 40% of the participants with AS/HFA thought that Jake hit the bull's-eye intentionally, which mirrors previous findings (Knobe, 2003a). The difference between the two groups for both the murder case ($\chi^2(1) = 1.99, p = 0.937$) and the bull's-eye case ($\chi^2(1) = 0.47, p = 0.49$) was not significantly different.

5.4 Results: Praise and Blame

When asked whether the agent's harming or helping the environment was blameworthy, praiseworthy, or neutral, almost all CPs (95.8%) judged that harming the environment was blameworthy, while 4.2% judged his/her action neutral. 58.4% of CPs judged that helping the environment was neutral, and 41.6 % as praiseworthy. All participants with AS/HFA answered that the agent's harming the environment was blameworthy. The difference between participants with AS/HFA and CPs was not significant for the harm case. However, in the help case, people with AS/HFA showed a pattern of response that differed from CPs': Helping the environment was judged to be praiseworthy by 83.3% and neutral by 16.7% of participants with AS. This comparison is statistically significant ($\chi^2(1) = 6.83, p = 0.007$).

When asked why the agent's action was morally neutral, almost all CPs answered that because helping the environment was not the chairman's main intent, but just a non-intended effect, this person cannot be praised for it. The few CPs who said that the chairman's action was praiseworthy justified their responses by appealing to the goodness of the chairman's involuntary outcome, even though they overtly acknowledged that it was somewhat inconsistent to praise someone for something that she did not accomplish intentionally.

In the murder case, when asked whether Jake's action of shooting his aunt was blameworthy, praiseworthy, or neutral, all participants in both groups judged that the agent's behavior was blameworthy. Conversely, in the bull's-eye case, 28.6% of CPs judged the agent's behavior as praiseworthy and 71.4% as neutral, while 60% of participants with AS/HFA judged the agent's action as praiseworthy and 40% as neutral. This contrast is statistically significant ($\chi^2(1) = 6.2, p = .01$).

5.5 Discussion

The aim of this second experiment was to investigate whether individuals with AS/HFA use the moral value of the agent's action as a cue to the intentional status of this action. Some interesting findings emerge from this study. First, as we found in our first study, adults with AS/HFA distinguished intentional from non-intentional actions. Second, as predicted, adults with AS/HFA's

judgments about the intentional status of side effects are responsive to moral considerations to the same extent as people with typical development (Knobe, 2003b, 2006). Almost all French matched comparison participants and all individuals with AS/HFA considered the negative side effect as intentional. Furthermore, judgments about the intentional status of lucky outcomes were also influenced by the moral value of the outcome. In the murder case and the bull's-eye case, the judgments of intentionality by both typical individuals and people with AS/HFA were affected by the moral appreciation of the agent's action.

Previous studies have shown that individuals with ASD with no language retardation can employ language strategies to pass mentalizing tasks (Happé, 1995; Tager Flusberg & Joseph, 2003). In accordance with these studies, one could flesh out a plausible interpretation of the combined results of Studies 1 and 2 by suggesting that because of difficulties in mindreading in some contexts (as shown by Study 1), individuals with AS/HFA have developed some compensation strategies based on their spared cognitive functions (as shown by Study 2). Previous studies have demonstrated that both moral reasoning (Blair, 1996; Grant et al., 2005; Leslie et al., 2006) and rule-learning process (Goldstein, Siegel, & Minshew, 1995) are relatively preserved in this population. Goldstein et al. (1995) suggested that individuals with ASD have no difficulty performing reasoning tasks when rule application is explicit, and several studies reported that an overt and effortful learned rule-based mechanism is often used to compensate for a diminished intuitive social understanding (Happé, 1995). It is thus possible that people with AS/HFA use a relatively intact function to learn to conform to social normative rules to compensate for their impairments in mindreading. Individuals with AS/HFA might employ it in everyday life situations to draw judgments about intentionality whenever mentalizing reasoning based on other cues is defective. They would thus rely on the following simple heuristic: If an action is wrong, then it is intentional. From this perspective, the side-effect effect reported in Study 2 reveals a moral bias in the assignment of an important psychological concept (viz. the concept of an intentional action), due to an increased responsiveness to normative violations.

A third interesting finding emerged from Study 2. The evaluation of the praiseworthiness of the side effect in the help case by people with AS/HFA differed from that of CPs. As expected, in the help case, the majority of CPs was particularly attentive to the fact that although helping the environment was a good action, the chairman did not help the environment intentionally; rather, helping the environment was an accidental effect of his main intention (*viz.* making profit). In accordance with their intentionality judgment, the majority of CPs considered the outcome in the help case to be neutral since the chairman did not really intend to improve the environment, and, consequently, his action could not be praised. A few CPs who said that the chairman's action was praiseworthy justified their response by appealing explicitly to the goodness of the side effect, but they were aware that it was inappropriate to praise someone for an action that she did not perform intentionally.

To rule out the possibility that participants understood the question as bearing on the action *qua type* (helping the environment in general), rather than on the *token* action performed by the agent, we made it clear that participants' judgments of intentionality and judgments of praise and blame should be about the particular action described in each vignette. The results show that the majority of people with AS/HFA, who like the CPs judged that the chairman did not help the environment intentionally, also judged that helping the environment was praiseworthy. This finding suggests that, in contrast to CPs, their judgment about the praiseworthiness of actions is not influenced by their judgment about the intentional status of these actions: That is, they seem to take an action to be praiseworthy if it is good. If this is correct, mentalistic judgments (*viz.* judgments about intentionality) are not integrated with judgments about the intrinsic value of the outcome to make praise judgments—an important type of moral judgments. A similar pattern was found in the bull's eye case. People with AS/HFA's judgments about the praiseworthiness of hitting the bull's eye differed from CPs': They judged that hitting the bull's eye by luck was praiseworthy. Thus, we found again that people with AS/HFA did not take into consideration whether the action was performed intentionally when they made praise judgments.

6. Conclusion

The results from the two studies presented in this article show that adults with AS/HFA's judgments about intentionality differ from the judgments made by typical adults in predictable ways. As established in Study 1, in contrast to typical individuals, people with AS/HFA seem to have difficulty judging that morally neutral actions that are negatively valued can be intentionally brought about when they are means for desired ends. Plausibly, this is due to the complexity of the mindreading processing required to make such judgments. As shown in Study 2, adults with AS/HFA exhibit the same asymmetry as typical individuals when they are asked to decide whether a morally bad and a morally good outcome are brought about intentionally. This suggests that, in individuals with AS/HFA, increased responsiveness to normative rule violations underlies a cognitive strategy that compensates for their diminished mindreading abilities. Furthermore, while typical people's moral judgments are sensitive to the agent's intentions, individuals with AS/HFA's judgments of praise do not seem to be influenced by the intentional status of actions. This suggests that, consistent with previous findings (Zalla et al., 2009), the moral and the mindreading reasoning systems are somewhat disconnected in individuals with AS/HFA.

References

- American Psychiatry Association 2000: *Diagnostic and Statistical Manual of Mental Disorders*, 4th Edition DSM-IV-TR (Text Revision), Washington, DC: American Psychiatry Association.
- Baron-Cohen, S. 1991: Do people with autism understand what causes emotion? *Child Development*, 62, 385-395.
- Baron-Cohen S. 1995: *Mindblindness. An Essay on Autism and Theory of Mind*. Cambridge: MIT Press.
- Baron-Cohen, S., O’Riordan, M., Stone, V., Jones, R. and Plaisted, K. 1999: Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders*, 29(5), 407-418.
- Blair, R. J. R. 1996: Brief report: Morality in the autistic child. *Journal of Autism and Developmental Disorders*, 26, 571-579.
- Blair, R. J. R. 1999: Psychophysiological responsiveness to the distress of others in children with autism. *Personality and Individual Differences*, 26, 477–85.
- Bratman, M. E. 1987: *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press.
- Carpenter, M., Pennington, B. F. and Rogers, S. J. 2001: Understanding of others’ intentions in children with autism. *Journal of Autism and Developmental Disorders*, 31, 589–99.
- D’Entremont, B. and Yazbek, A. 2007: Imitation of Intentional and Accidental Actions by Children with Autism. *Journal of Autism Developmental Disorders*, 37, 1665-1678.
- Frith, U. 1989: *Autism: Explaining the Enigma*. Oxford: Blackwell.
- Gillberg, C., Gillberg, C., Råstam, M. and Wentz, E. 2001: The Asperger Syndrome (and High-Functioning Autism) Diagnostic Interview (ASDI): A Preliminary Study of a New Structured Clinical Interview. *Autism*, 5 (1), 57 - 66.
- Guglielmo, S. and Malle, B. F. ms: Can unintended side effects be intentional? Solving a puzzle in people's judgments of intentionality and morality.

- Guglielmo, S., Monroe A.E. and Malle, B. F. Forthcoming: At the Heart of Morality Lies Folk Psychology, *Inquiry*, 52.
- Goldstein G., Siegel D.J., and Minshew N.J. 1995: Abstraction and Problem Solving in Autism: Further Categorization of the Fundamental Deficit. *Archives of Clinical Neuropsychology*, Volume 10, 4, pp. 335-335
- Grant C.M., Boucher J., Riggs K. J. and Grayson A. 2005: Moral understanding in children with autism. *Autism*, 9, 317-331.
- Happé, F. 1994: An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped and normal children and adults. *Journal of Autism and Developmental Disorders*, 24, 129–154.
- Happé, F. 1995: The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, 66, 843-855.
- ICD-10 1992: *Classification of Mental and Behavioural Disorders Diagnostic criteria for research*. World Health Organization.
- Knobe, J. 2003a: Intentional action in folk psychology: An experimental investigation. *Philosophical Psychology*, 16, 309-324.
- Knobe, J. 2003b: Intentional action and side effects in ordinary language. *Analysis*, 63, 190 - 193.
- Knobe, J. 2004: Intention, intentional action and moral considerations. *Analysis*, 64, 181 – 187.
- Knobe, J. 2005: Theory of Mind and Moral Cognition: Exploring the Connections. *Trends in Cognitive Sciences*, 9, 357-359.
- Knobe, J. 2006: The concept of intentional action: a case study in the uses of folk psychology. *Philosophical Studies*, 130, 2 , 203-231.
- Knobe, J. 2007: Reason explanation in folk psychology. *Midwest Studies in Philosophy*, 31, 90-106.
- Knobe, J. and Mendlow , G. 2004 : The good, the bad and the blameworthy: understanding the role of evaluative reasoning in folk psychology . *Journal of Theoretical and Philosophical Psychology* , 24 , 252 – 258.

- Knobe, J. and Burra, A. 2006: Intention and intentional action: a cross-cultural study. *Journal of Culture and Cognition*, 1(2), 113 – 132.
- Lagnado, D. A. and Channon, S. 2008: Judgments of cause and blame: The effects of intentionality and foreseeability. *Cognition*, 108, 754-770.
- Leekam, S., Libby, S., Wing, L., Gould, J., and Gillberg, C. 2000: Comparison of ICD-10 and Gillberg's criteria for Asperger syndrome. *Autism: Special Issue: Asperger syndrome*, 4, 11–28.
- Leslie, A., Knobe, J. and Cohen, A. 2006: Acting intentionally and the side-effect effect: 'Theory of mind' and moral judgment. *Psychological Science*, 17, 421 – 427.
- Leslie, A., Mallon, R. and DiCorcia, J. A. 2006: Transgressors, victims, and cry babies: Is basic moral judgment spared in autism? *Social Neuroscience*, 1, 270-283.
- Machery, E. 2008: Understanding the folk concept of intentional action: Philosophical and experimental issues. *Mind & Language*, 23, 165-189.
- Malle, B. F. and Knobe, J. 1997a: The folk concept of intentionality. *Journal of Experimental Social Psychology*, 33, 101-121.
- Malle, B. F. and Knobe, J. 1997b: Which behaviors do people explain? A basic actor-observer asymmetry. *Journal of Personality and Social Psychology*, 72, 288-304.
- Mallon, R. 2008: Knobe vs. Machery: Testing the trade-off hypothesis. *Mind & Language*, 23, 247-255.
- Mele, A. R. and Cushman, F. 2007: Intentional action, folk judgments, and stories: Sorting things out. *Midwest Studies in Philosophy*, 31, 184-201.
- Nadelhoffer, T. 2004: Blame, Badness, and Intentional Action: A Reply to Knobe and Mendlow. *Journal of Theoretical and Philosophical Psychology*, 24, 259-269.
- Nadelhoffer, T. 2006: Bad acts, blameworthy agents, and intentional actions: some problems for juror impartiality. *Philosophical Explorations*, 9, 203-219.
- Nanay, B. forthcoming: Morality or modality: What does the attribution of intentionality depend on? *The Canadian Journal of Philosophy*.

- Nichols, S. and Ulatowski, J. 2007: Intuitions and individual differences: the Knobe effect revisited. *Mind & Language*, 22, 346-365.
- O'Shaughnessy, B. 1980: *The will*. Cambridge: Cambridge University Press.
- Phelan, M. T. and Sarkissian, H. 2008: The folk strike back; or, why you didn't do it intentionally, though it was bad and you knew it. *Philosophical Studies*, 138, 291-298.
- Phelan, M. T. and Sarkissian, H. 2009: Is the trade-off hypothesis worth trading for? *Mind & Language*, 24: 164-180.
- Phillips, W., Baron-Cohen, S., and Rutter, M. 1995: To what extent do children with autism understand desires? *Development and Psychopathology*, 7, 151-170.
- Phillips, W., Baron-Cohen, S. and Rutter, M. 1998: Understanding intention in normal development and in autism. *British Journal of Developmental Psychology*, 16, 337-348.
- Russell, J. and Hill, E.L. 2001: 'Action-monitoring and Intention Reporting in Children with Autism', *Journal of Child Psychology and Psychiatry*, 42, 317-28.
- Searle, J. R. 1983: *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
- Schank R, and Abelson P. 1977: *Scripts, plans, goals and understanding*. Hillsdale: Erlbaum.
- Tager-Flusberg, H. and Joseph, R. M. 2003: Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society B*, 358: 303-314.
- Wing, L. 2005: Reflections on opening the Pandora's box. *Journal of Autism and Developmental Disorders*, 35, 197-203.
- Wright, J. and Bengson, J. 2009: Asymmetries in folk judgments of responsibility and intentional action. *Mind & Language*, 24: 24-50.
- Zalla T, Sav A-M., Stopin A, Ahade S, and Leboyer M. 2009: 'Faux pas detection and intentional action in asperger syndrome. A replication on a French sample, *Journal of Autism and Developmental Disorders*, 39:373-382.

Zelazo P. D., Jacques S., Burack J, and Frye D. 2002: The Relation between Theory of Mind and Rule Use: Evidence from Persons with Autism-Spectrum Disorders, *Infant and Child Development*, 11, 171–195.

Appendix

Le verre décoré (free-cup case)

Parce que Stéphane a très soif, il entre dans un café et commande un perrier. Le serveur lui répond que s'il commande un Perrier, son Perrier lui sera servi dans un verre spécialement décoré.

Stéphane réfléchit 5 secondes et répond: "Ce verre spécialement décoré n'a pas d'importance, je veux juste un Perrier." Naturellement, le Perrier lui est servi dans un verre spécialement décoré.

Un euro de plus (extra-euro case)

Parce que Stéphane a très soif, il entre dans un café et commande un Perrier. Le serveur lui répond que le Perrier a augmenté d'un euro. Stéphane répond: "Payer un euro en plus n'a aucune importance, je veux juste un Perrier." Naturellement, le Perrier lui est servi et Pierre paie un euro de plus.

Nuire à l'environnement (harm case)

Le vice-président d'une compagnie va voir le PDG et lui dit: "nous pensons démarrer un nouveau programme. Cela augmentera nos profits, mais cela nuira à l'environnement". Le PDG répond: "Cela m'est complètement égal de nuire à l'environnement. Je veux seulement réaliser autant de profits que possible. Démarrons le nouveau programme". Ils démarrent le nouveau programme. Comme prévu, cela nuit à l'environnement.

Améliorer l'environnement (help case)

Le vice-président d'une compagnie va voir le PDG et lui dit: "nous pensons démarrer un nouveau programme. Cela augmentera nos profits, mais cela améliorera la qualité de l'environnement". Le PDG répond: "Cela m'est complètement égal d'améliorer la qualité de l'environnement. Je veux seulement réaliser autant de profits que possible. Démarrons le nouveau programme". Ils démarrent le nouveau programme. Comme prévu, la qualité de l'environnement s'améliore.

L'héritage (murder case)

Jacques désire à tout prix avoir plus d'argent. Il sait qu'il héritera d'une grosse somme d'argent si sa tante meurt. Un jour, il voit sa tante passer par la fenêtre. Il prend sa carabine, ajuste sa tante dans le

viseur, et appuie sur la détente. Mais Jacques ne sait pas très bien utiliser sa carabine. Sa main glisse sur la crosse de la carabine, et le tir part dans le décor... Pourtant, la balle atteint sa tante en plein cœur. Elle meurt sur le champ.

Le concours de tir (bull's eye case)

Jacques désire à tout prix remporter le concours de tir à la carabine. Il sait qu'il ne remportera le concours que s'il atteint la cible en plein coeur. Il prend sa carabine, ajuste la cible dans le viseur, et appuie sur la détente. Mais Jacques ne sait pas très bien utiliser sa carabine. Sa main glisse sur la crosse de la carabine, et le tir part dans le décor... Pourtant, la balle atteint la cible en plein coeur. Jacques remporte le concours de tir.

Table 1. Demographic and neuropsychological data of both groups.

	Asperger Syndrome	Comparison
	(N=20)	(N=28)
Age (yrs/mths)	27.6 (\pm 6.7)	27.9 (\pm 7.6)
Gender F/M	4/16	7/21
Education (yrs)	13.7 (4)	13.4 (3.2)
Faux Pas test (total score)	40.1 (\pm 5.7)	56.5 (\pm 4.1)
Full scale IQ (WAIS-III)	93.5 (22.4)	97.7 (12.9)
Verbal IQ	99.4 (25)	98.3 (11.4)
Performance IQ	87.2 (16.2)	97.5 (13.9)

Figure 1. Causal Representation of the Events in the Extra-Dollar Case.

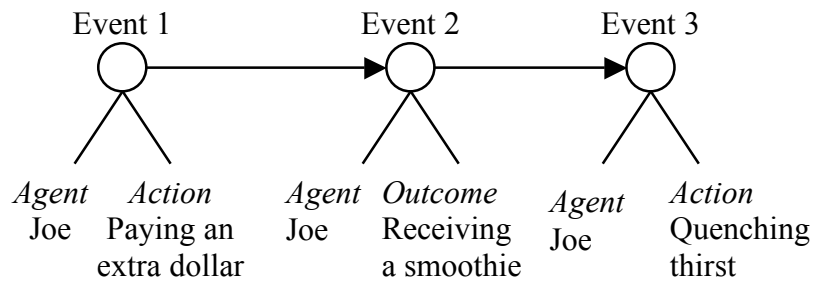


Figure 2. Valence-Based Representation of the Events in the Extra-Dollar Case.

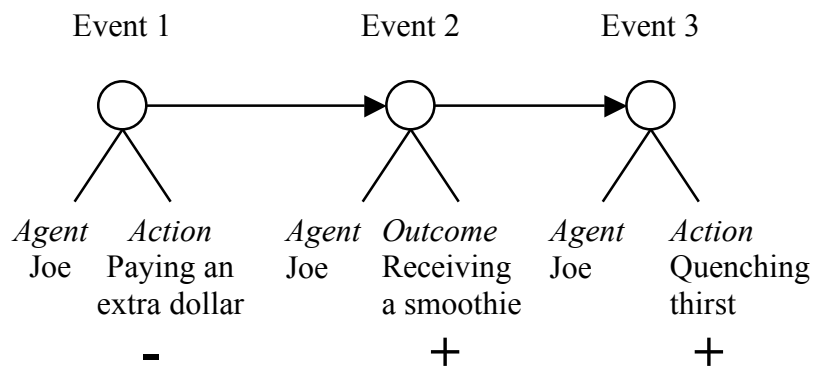


Figure 3. Mentalistic Representation of the Events in the Extra-Dollar Case.

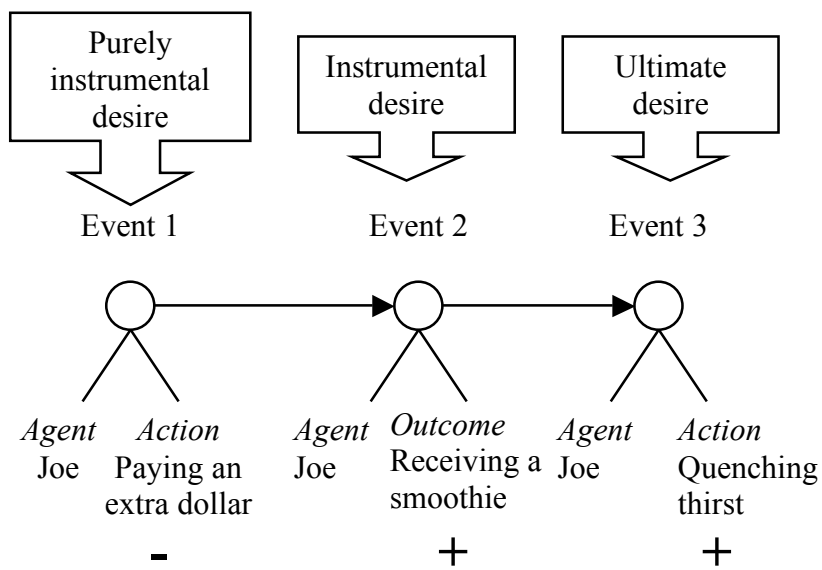


Figure 4. Causal Representation of the Events in the Free-Cup Case.

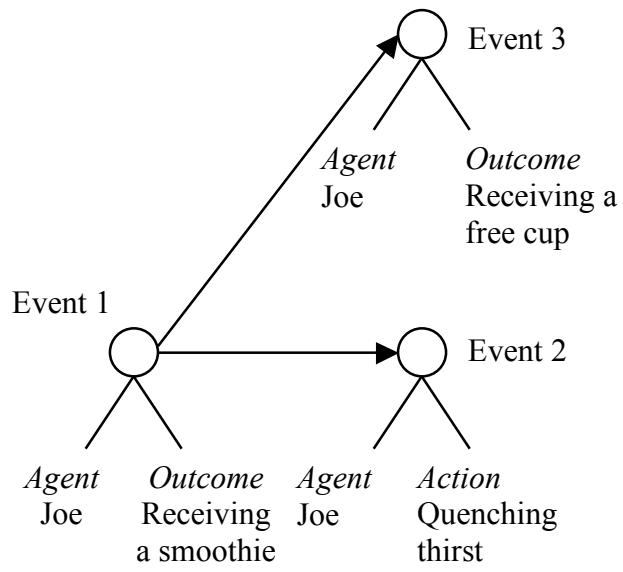


Figure 5. Valence-Based Representation of the Events in the Free-Cup Case.

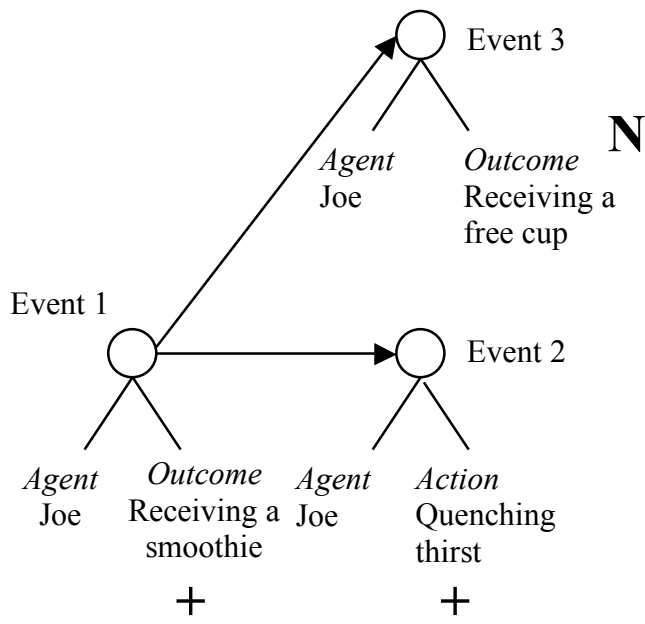


Figure 6. Mentalistic Representation of the Events in the Free-Cup Case.

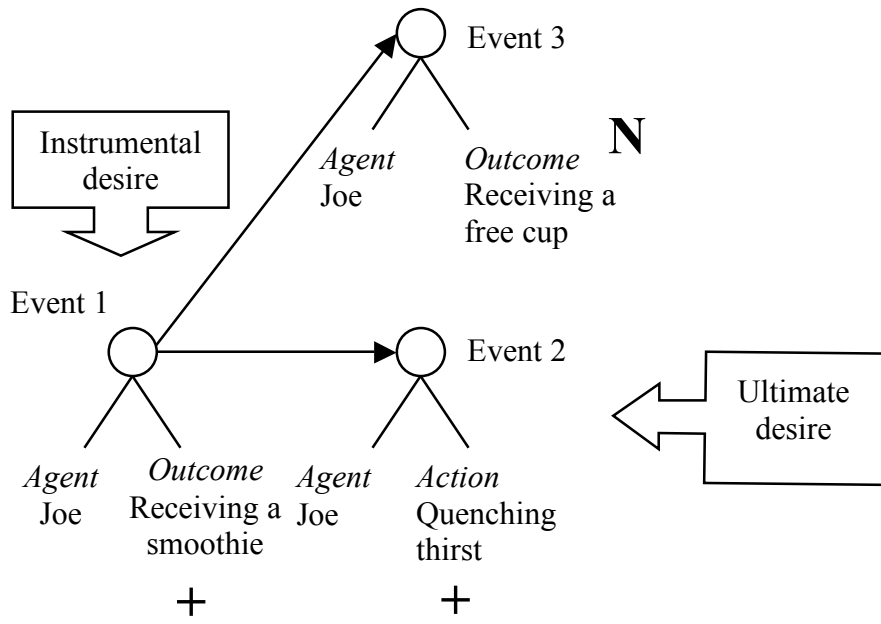


Figure 7. Malle and Knobe's (1997a) Model

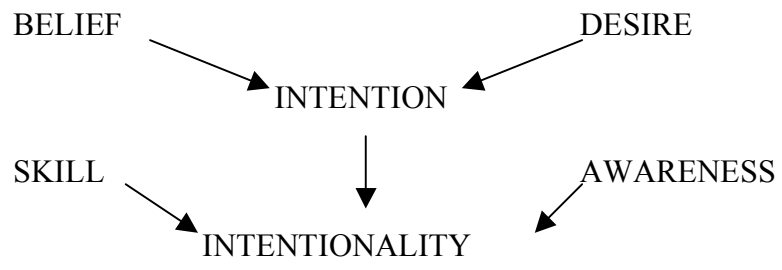


Figure 8. Intentional Representation of the Events in the Harm Case.

