

Moral judgment is a highly complex mental process, and the features that drive it have been the center of research. Classical models assumed that intentionality of an action guides moral judgments on that action. However, recent findings suggest that moral judgment could inversely affect judgment on intentionality. Specifically, most people attribute intentionality to morally bad side-effects, but not to morally good side-effects. In the present study, we focused on this phenomenon, known as 'the Knobe effect', and examined the neural mechanisms underlying this effect using functional magnetic resonance imaging (fMRI) by comparing adults with autism spectrum disorder (ASD) and neurotypical adults (NT). We prepared various types of scenarios, including not only negative but also positive scenarios, such as intended, attempted, and accidental harms/helps. Furthermore, we prepared two distinct sets of 20 negative and 20 positive side-effects in contrast that previous studies examined the Knobe effect with only few typical scenarios, so that we could quantify the effect and eliminate scenario-specific factors. Twenty ASD and 20 NT judged the degree of morality and intentionality of the protagonist's action for each scenario in the scanner. We found that both ASD and NT showed significant Knobe effects. Interestingly, attribution of intentionality for negative side-effects was significantly smaller in ASD than NT, with no significant difference for positive side-effects, indicating that the Knobe effect is attenuated in ASD. In addition, moral judgments for attempted harms were less severe in ASD than NT. We hypothesized that this decreased dependence of moral judgment on intention is related to the diminished Knobe effect in ASD. We quantified how much moral judgment of each individual depends on intention (the degree of intentionism: DOI) by comparing moral judgments for attempted helps to attempted harms. We could derive DOI from these scenarios, since protagonists brought the same morally neutral outcomes with either good or bad intentions there. Interestingly, as with the Knobe effect, DOI was significantly lower in ASD than NT. Moreover, DOI showed a significant positive correlation with the strength of the Knobe effect, that is, those who depends more on intention in moral judgment showed stronger Knobe effects. These results suggest that the Knobe effect is subserved by a bidirectional interaction between intentionality and morality. We will also discuss the neural basis of the Knobe effect and its change in ASD.