



# Relations between self-understanding and other-understanding: similarities and interactions

Adrianna Smurzyńska 

Jagiellonian University

*adrianna.smurzynska@gmail.com*

Received 4 January 2019; accepted 18 November 2019; published 2 February 2020

## Abstract

The aim of this article is to examine the relation between self-understanding and other-understanding. In this context, three theories of mentalization will be considered: theory theory, simulation theory and the person model theory. Two types of issues will be analyzed: (I) whether the abilities involved in self- and other-understanding are identical or diverse and (II) whether the abilities of self- and other-understanding are distinct or intertwined. These questions will be applied to theory theory, simulation theory and the person model theory; they will also be placed in the context of selected research concerning this issue.

**Keywords:** mentalization; self-understanding; other-understanding; theory theory; simulation theory; person model theory

## 1. Introduction

The classical other-minds problem is summed up by the question “how can we know that there are other minds apart from our own and how can we learn anything about minds other than our own?” One of the solutions to that problem stated that knowledge about one’s own mind is the basis for inferring knowledge about others (Hyslop, 2016). In contrast to classical solutions of the other minds problem, there is no place in most contemporary theories for privileged and direct access to one’s own mind. Consequently, self-knowledge cannot be regarded as the basis for inferences about other minds.

Nowadays, philosophers and cognitive scientists are instead poised to ask, “how do we ascribe mental states to others?” Terms used to describe this capability are ‘theory of mind’, ‘mindreading’, and ‘mentalization’ (Baron-Cohen, 1995; Frith & Frith, 2006; Premack & Woodruff, 1978; Vogeley & Newen, 2009). In this article I will use the term ‘mentalization’ because it is, in my opinion, the most neutral. The term ‘theory of mind’ suggests a way of attributing

mental states by appealing to a theory of mental states, like in theory theory. The term ‘mindreading’ is associated with ‘reading’, thus suggesting that there are symbols which must be deciphered. ‘Reading’ rules out non-inferential processes and presupposes that attribution of mental states is inferential. In contrast, the term ‘mentalization’ does not presuppose the exact mechanism that forms the basis of this process.

The term ‘mentalization’ may be defined as various abilities: to possess a theory of the mental realm (Guerini & Marraffa, 2015); to form a judgement, belief, or representation about another’s mental state (Goldman, 2009); to understand what people are thinking, perceiving, and feeling (Robbins, 2004); to ascribe beliefs, desires, intentions, and other propositional attitudes to oneself and others (Robbins, 2004); to form a judgement about a current ‘state’ (Gordon, 2008); to possess the ‘cognitive capacity to attribute mental states to self and others’ (Goldman, 2012). All these descriptions include references to mental states and the process of forming statements about them. Mental states will be treated in this article as unquestioned entities, while the process of forming statements about them will be debated.

The process of attributing mental states to others can be modelled in different ways. In the philosophical literature it is sometimes explained as the process of building folk theories (theory theory, TT; Goldman, 2012), making simulations (simulation theory, ST; Goldman & Mason, 2007), establishing person models (person model theory, PMT; Newen, 2015), experiencing mental states directly (interaction theory, IT; Gallagher, 2008), or understanding them through narrative abilities (narrative practice hypothesis, NPH; Gallagher & Hutto, 2008).

In this article, I will consider the topic of the relationship between self-understanding and other-understanding in each of three theories: TT, ST and PMT. IT and NPH will be omitted as they are based on a non-mechanistic paradigm that is dissimilar to TT, ST and PMT. I differentiate two types of topic: the first is the question of whether abilities involved in self- and other-understanding are identical or diverse. In other words, are the mechanisms of these abilities (as the set of components, operations, and processes understood as the active implementation of these operations) identical or diverse? If they are diverse, I will trace the similarities and differences between them. The second type of topic is the question of interrelations between processes involved in self- and other-understanding. Here the question emerges: are the abilities of self- and other-understanding distinct or intertwined?

## **2. Relations between self- and other-understanding**

Current investigations do not deliver a straightforward solution and struggle to determine whether the abilities involved in self- and other-understanding are identical or diverse. In fact, they are quite divergent in their answers as to whether there are one or two different types of these abilities (Nichols & Stich, 2003). However, both these research traditions are worth considering. The first emphasizes the similarities between the mechanisms that govern self-understanding and other-understanding (e.g., Damon & Hart, 1982); the second stresses the differences between them (e.g., Harnik, Fearon, & Fonagy, 2009). In the analyses of the three theories of mentalization (TT, ST and PMT), I will focus on the way they describe the mechanisms of mental states’ attribution.

As outlined above, it is worth assessing the interrelations between the processes involved in self- and other-understanding. It can be stated that they are distinct or intertwined. Self- and other-understanding are distinct when it is possible to adequately ascribe mental states to oneself without engaging the ability to ascribe mental states to others. Simultaneously, it is possible to ascribe mental states adequately to others without engaging one's own knowledge about the self. In other words, these abilities are distinct when they may be applied separately to oneself and to another. Self- and other-understanding are intertwined when they interact during self-understanding and other-understanding. In this case, it is impossible, or at least difficult, to separate these processes when understanding oneself or the other. In the literature, arguments can be found for both the first (e.g., Nichols & Stich, 2003) and the second option (e.g., Brandt, Buttelmann, Lieven, & Tomasello, 2016). However, both traditions wonder if one of these processes may occur without the other, i.e., if there are cases of adequate self-understanding without adequate other-understanding, and *vice versa*. In this article, I will focus on the way in which these processes interact with each other in the context of three theories of ascribing mental states: TT, ST and PMT.

### **3. Theory theory**

According to TT, the basis for the attribution of mental states is a folk-psychological theory of mental states which contains terms such as 'beliefs', 'desires', and 'intentions' (Goldman, 2012; Newen, 2015). Some researchers regard this theory as innate, while others regard it as learned, but most of them have claimed that it develops throughout our life (at least, up to a specified moment) (Baron-Cohen, 1995; Gopnik, 1996; Krueger & Overgaard, 2012).

Understanding that other people have different representations of reality – that they may not know something we know – is the basic indicator of mentalization. This capability can be measured by the popular false-belief task (e.g., Gopnik & Astington, 1988; Wimmer & Perner, 1983). The standard test is as follows: person A puts thing X in location M. When person A leaves the room, person B takes thing X from location M to location N. Then, a child who has observed the whole process is asked "where will person A look for thing X when she/he goes back into the room?" (Wimmer & Perner, 1983). Research has shown that most children under 4 years of age have no ability to understand that person A does not know the object's new location. This result has been replicated in different experimental settings (Goldman, 2006). Another common experiment in this area of research considers the understanding of representational changes (Gopnik & Astington, 1988). Children are asked about their own and another's present and past thoughts about the contents of a box. The results of such experiments have shown that the development trajectories of self- and other-understanding are similar, although the ability to understand others starts earlier.

TT is the oldest and the most diverse approach within the debate on mental states' ascription. There are classical (e.g., Perner, 1991; Wellman, 1990) and modern versions (e.g., Carruthers, 2009; Goldman, 2006; Nichols, & Stich, 2003). The modern theories are hybrid, combining different theoretical approaches, e.g., elements of TT and ST, which are treated as cooperating with each other (Goldman, 2006).

In TT, the attribution of mental states to the self and to others is based on the same inferential mechanism (Bogdan, 2007; Carruthers, 2009; Gopnik, 1993; Newen & Schicht, 2009; Robbins, 2004). Mental states are inferred from the observation and recollection of behavior and folk-psychological generalizations of human functioning. Detecting one's own propositional attitudes, such as beliefs, desires, or intention, is supposed to be a result of the same mechanism behind detecting other's mental states. Supporters of this conception admit that we have a wider understanding of ourselves thanks to a wider scope of available data about ourselves, not because of any privileged access to our own minds (Carruthers, 2009; Robbins, 2004). They also claim that direct access to one's own mental states is an illusion and anything we assume is a result of interpretation (Gopnik & Meltzoff, 1994; Robbins, 2004). In non-orthodox versions of TT, it is claimed that understanding our own mind occurs prior to understanding others' minds, although this is also based on inferences (see Goldman, 2006).

If one questions direct and privileged access to one's own mental states, then the distinction between the first- and third-person perspective becomes blurred. In this case, there is no need to propose two different kinds of social understanding. Thus, mentalization may be treated as a monolithic ability that may then be separately applied to different kinds of data, i.e., information on oneself or on another.

In the conceptual frame presented in the previous section, self-understanding and other-understanding may be regarded as being based on the same mechanism and as distinct from one another. The mechanism, which is understood as a set of components and operations upon them, is conceived in TT as an inferential process based on theoretical assumptions. This mechanism may be applied separately to information on oneself or on another. It is not necessary to engage self-understanding when understanding others and *vice versa*.

#### 4. Simulation theory

In ST, the process of understanding others is described as running internal simulations of their mental states (Goldman & Mason, 2007). In classical versions, a simulation is regarded as automatic and unconscious (e.g., Gordon, 1995). In modern versions, there is a place for both low-level unconscious processes such as mirroring, and high-level conscious processes such as reflective and imaginative thinking (Goldman, 2006; Waytz & Mitchell, 2008). In ST, one's own mind is used as a tool to enable oneself to understand others (Goldman, 2006; Waytz & Mitchell, 2008).

In general, simulation is a process of mental pretense – the process of recreating or re-enacting the mental states of another person (Goldman, 2009). Understanding others results from integrating perceptions of behavior and an imagined simulation of mental states (Waytz & Mitchell, 2008). It leads to imagining the mental states of others and predicting their actions (Bogdan, 2007; Goldman, 2006; Krueger & Overgaard, 2012; Newen & Schlicht, 2009).

In mirroring, the perceptible cues of another person's mental states automatically trigger a similar experience in the perceiver's mind (Waytz & Mitchell, 2008). Recognizing the emotions of others is thought to be the result of mimicry of others' emotional expressions (Decety & Grezes, 2006). This process is enabled by activation of the anterior cingulate cortex as well as

the premotor, somatosensory, posterior parietal, insular and temporal cortices (Keysers & Gazzola, 2009). The same mechanism allows intentions to be mirrored and the actions of others to be predicted (Decety & Grezes, 2006). Such mirrored states are akin to one's own emotions and sensations (Keysers & Gazzola, 2009)

The interpretative aspects of simulation involve imagining a hypothetical scenario unrelated to one's current experience (Waytz & Mitchell, 2008); it can relate both to imagining oneself as another and to imagining oneself in a different setting and situation. This ability is examined using tasks that involve considering one's own mental states, imagining fictitious or future events, and recalling past experiences (Waytz & Mitchell, 2008). The neural network which underlies such deliberation comprises the medial prefrontal cortex, the precuneus, the posterior cingulate, and the lateral parietal cortex (Waytz & Mitchell, 2008). The results of such studies show activation of the same regions that are engaged when people consider either their own or another's mental states (Waytz & Mitchell, 2008).

In this theory, the interrelation between understanding ourselves and others can be investigated from different angles. The simulation of others' mental states is based on ascribing a simulated state to oneself and then attributing it to another person (Bogdan, 2007; Goldman, 2006; Krueger & Overgaard, 2012; Newen & Schlicht, 2009). Access to one's own mind, at least to some extent, is required. Some versions of this theory include the necessity of introspection (Goldman, 2006) or using self-understanding when understanding others (Mitchell, Macrae, & Banaji, 2006). Other proponents of ST indicate that ascribing mental states to oneself by imagination is based on the same mechanism used to ascribe mental states to others. Imagining another's experience resembles reflecting one's own past experiences or possible future events (Waytz & Mitchell, 2008). To enable simulation, access to one's own experience is necessary. In the case of a reflexive simulation of a hypothetical scenario, self-understanding is based on the same type of imagining process as other-understanding; however, simulating the mind of another requires at least modest access to one's own mind, therefore self-understanding and other-understanding are closely interconnected.

When it comes to the mechanism that underlies the ascription of mental states, it seems that self-understanding and other-understanding are similar to each other. In both situations, simulation may be based on low-level and high-level mechanisms; however, low-level mechanisms proceed differently when understanding oneself and understanding others. The first-person experience is different, especially when it comes to emotions. At the same time, it is hard to assume that prior self-experience does not affect one's understanding of others. Therefore, the processes of self-understanding and other-understanding seem to be interrelated and to possess similar mechanisms.

## **5. Person model theory**

PMT argues that we build person models when understanding oneself and others (Newen, 2015; Newen & Schlicht, 2009). Person models are understood as mental representations; they are unities of traits and properties which are used to register and evaluate the mental and physical properties of oneself and others. At first glance, TT and PMT seem quite similar, but PMT

puts a strong emphasis on the fact that person models can be based on parsimonious information and are not necessarily systematically interconnected. Thus, they cannot be called ‘theories’, although a person model can be developed into a theory after learning more about the other person (Newen, 2015). Moreover, person models are often different for different people, whereas the folk-psychological theory in TT appears to be universal and applicable to all other people.

The authors of PMT distinguish two types of person models: person images, which are explicit and easily accessible in the consciousness, and person schema, which are implicit, practical knowledge (Newen & Schlicht, 2009; Newen, 2015). Both types of person models may be built for individuals as well as for groups. Person schemas are implicit versions of person models and are based on intuitive processes. They enable social perception and advanced interactions by integrating the other’s appearance, personality traits, characteristic features and probable behaviors in one representation. Person images consist of similar elements, but they are consciously accessed and are developed through inferential processes. In this theory, it is claimed that previously built person schemata may develop and become a person image. This theory also states that people use different epistemic strategies to understand others, such as direct perception, interaction, simulation, and theory-based inferences.

In PMT, understanding oneself and others relies on the same mechanisms: epistemic strategies (Newen & Schlicht, 2009). More precisely, the mechanisms of mental states’ ascription to oneself and others are the same (direct perception, interaction, simulation and theory-based inferences) and are chosen according to the situation.

Self- and other-understanding may, but do not necessarily interact (see Newen, 2015). This means that person X may use another person’s model to improve her own understanding of herself and *vice versa*. In other words, the way of ascribing mental states to others may be enhanced in one’s self-model and *vice versa*, but it does not apply in every case.

## **6. Self- and other-understanding: identical or diverse?**

Most contemporary theories of self-understanding question the assumption of privileged and direct access to one’s own mind. They instead place a strong emphasis on similar, interpretative aspects of self- and other-understanding; however, it is worth asking whether these abilities are identical or diverse.

In the formulations of each theory analyzed in this work, it is stated that both self- and other-understanding have an interpretative aspect (Bogdan, 2005; Carruthers, 2009; Goldman, 2006; Gopnik, 1993; Robbins, 2004; Krueger & Overgaard, 2012; Newen & Schlicht, 2009; Waytz & Mitchell, 2008). Attribution of propositional attitudes engages interpretation of behavior and sensations in situational contexts. This characteristic captures both the self-attribution and other-attribution of mental states. The access to propositional attitudes is interpretative in both situations and is enhanced by imaginative and reflexive thinking on mental states. It is also worth mentioning that the same mental states are attributed to oneself and to the other. Both types of understanding encapsulate ascription of beliefs, desires, intentions, emotions and imaginings. Both self- and other-understanding have similar neural correlates (Decety & Grezes,

2006; Keyzers & Gazzola, 2009; Waytz & Mitchell, 2008). What is more, meta-analyses of research on mental states' attributions (e.g. Nichols & Stich, 2003) show that the developmental trajectories of both types of understanding are parallel.

At the same time, the differences between these types of understanding should not be ignored; they are embedded in the way in which agency is experienced. In cases of self-understanding a person perceives herself/himself as the agent, but in cases of other-understanding a person conceives another as a distinct and autonomous individual (see Gallagher, 2008). This difference induces another type of access to mental states. The accuracy of mental state attribution is better for self-understanding than other-understanding in cases of current judgements, images or feelings, but it is not necessarily better in cases of beliefs or intentions (Robbins, 2004). This may also happen because beliefs and intentions are more based on interpretation than are current judgements, images, or feelings. Self-awareness and agency over one's own mental states also enable consideration of the differences between one's own and another's mind, as well as reflection on such differences (Decety & Grezes, 2006). This is especially important as sharing emotions and experiences must be coordinated with self-awareness and agency (Decety & Grezes 2006). Such cooperation enables not only disentanglement between the self and others, but also helps avoid confusion between one's own mind and the imagining of another's mind.

Self-understanding and other-understanding differ from each other, but this does not mean they are two separate abilities. They are not identical because the aspect of agency affects the way of understanding. The differences may come from the fact that they represent one type of mechanism with two types of access (Carruthers, 2009). This type of agency-related access may then influence the way of perceiving mental states. Such considerations cannot solve the question of whether self-understanding and other-understanding are two abilities or two modes of one ability. It can be stated, though, that they are diverse due to the aspect of agency, and that this issue is not sufficiently considered in TT and PMT. The arguments presented here regarding the similarities and differences do not exhaust the topic but are proposed as an illustration of the issue under consideration.

### **7. Self- and other-understanding: distinct or intertwined?**

The interdependence between self- and other-understanding can be represented in three ways: it may be asked (I) whether they can exist separately, i.e., whether it is possible to ascribe mental states to others without being able to ascribe them to oneself and *vice versa*; it may be also asked (II) if one of them may be a basis for the other; ultimately, and most importantly for this article, it may be asked (III) whether the abilities of self- and other-understanding are distinct or intertwined.

(I) Self- and other-understanding may be treated as distinct abilities if there exist cases of double dissociation, i.e., when self-understanding appears without other-understanding and when other-understanding appears without self-understanding. Such an argument is based on research on disordered mentalization, e.g., in autism spectrum disorder and schizophrenia. Studies analyzed by Nichols and Stich (2003) have indicated that people with passive symptoms of schizophrenia have an intact ability to attribute mental states to others but an impaired

ability to attribute mental states to themselves. They deny control over their actions and thoughts and are unable to attribute their beliefs and intentions to themselves. At the same time, they pass tests for mentalizing others, such as false belief tasks. The opposite situation occurs with autistic people (see Nichols & Stich, 2003). They have access to their own mental states but they have particular problems with predicting the behavior and recognizing the emotions of others. However, it is worth asking whether the conclusions drawn from such observations should be extrapolated to the general population and should influence the way we understand the ability to mentalize. It may be that the situation is different in non-disordered mentalization.

(II) Questions about the interdependence of self- and other-understanding are thus worth asking. There are three options for such a situation: self-understanding may be treated as a basis for understanding others, other-understanding may be a basis for understanding oneself, or they may interact in both directions.

Some researchers claim that understanding ourselves is more basic than understanding others (e.g., Goldman, 2006). One's own thoughts and experiences are very often treated as the basis for understanding others (Waytz & Mitchell, 2008; Keysers & Gazzola, 2009). When recognizing intentions, emotions, and pain, the mental states that one has previously experienced are more recognizable in the reactions of others (Decety & Grezes, 2006). At the same time, a common egocentric bias which causes people to attribute their own feelings and beliefs to others may reduce the ability to recognize other's mental states (Boyer, Robbins, & Jack, 2005). With complex shared experiences, this kind of impact is more complex. Studies conducted by Hodges (2005) investigated differences in recognizing the mental states of people with similar experiences, such as new motherhood, alcoholism, parental divorce, and of people without common experiences in these respects. Similarities in personal history did not enhance the understanding of another person, although they did influence empathic concern, i.e., caring and sympathy, as well as the feeling of being understood.

It is also possible that the capacity for understanding others is grounded in more basic processes (see Carruthers, 2009; Mills & Danovitch, 2009). From this point of view, interpretations and inferences about one's own mental states originate from the way we interpret others' behavior. Mentalization about others is then internalized and used to understand oneself. From a developmental perspective, making inferences about mental states is easier when thinking about the beliefs and desires of others (Mills & Danovitch, 2009). Research by Brandt, Buttelmann, Lieven and Tomasello (2016) has also shown that understanding third-person complement-clause constructions is parallel to understanding one's own false beliefs. Thus, one can infer that inferential processes are originally used for others, and only afterwards do self-interpretations become possible.

(III) Such interactions may also go in both directions and be mediated by the ability to disentangle one's own and others' mental states. This may happen, for example, when two people are believed to be similar to one another (see Mitchell, Macrae, & Banaji, 2006; Newen, 2015). Both reflecting about oneself and about people believed to be similar to oneself coexist with activation in overlapping regions of the medial prefrontal cortex (Mitchell, Macrae, & Banaji, 2006). This situation is presented in PMT. A believed similarity in a relevant respect influences the process of forming a representation of another's personality. Such a belief of



similarity may rely on more general features, such as being an agent and a sentient human, or on more concrete features like age or socio-cultural background (Newen, 2015). It is worth asking if such a similarity may have a negative effect on the process of understanding the other. There is a chance that the other is similar in some respects (B, C, D) but not others (E, F, G). The similarity in respects B, C, and D may lead to the generalization and supposition that such a person is similar in other respects.

Conclusions about mutual interactions between self- and other-understanding can be also drawn from observations of social interactions. From such observations it is sometimes inferred that understanding of oneself and others develops through social interaction (see Damon & Hart, 1982; Hernik, Fearon, & Fonagy, 2009). The process of interaction improves children's understanding that their own mental states may be different from those of others, and that people may have different perspectives (Gopnik & Astington, 1988; Brandt, Buttelmann, Lieven, & Tomasello, 2016). Perspective-taking enables adequate attribution of mental states to both oneself and others (Fonagy & Luyten, 2009). The ability to understand how one's beliefs change is supposed to be a result of recognizing the difference between one's own mental states and those of others (Gopnik & Astington, 1988). What is more, problems with understanding oneself often coexist with impairments in understanding others, especially when it comes to problems with differentiating one's experience from those of others (Fonagy & Luyten, 2009).

All in all, in everyday situations the processes of understanding oneself and understanding others are closely intertwined, and their separation is demanding and complex. Taking the perspective of the other without relating it to our own is a difficult and deliberate operation (Boyer, Robbins, & Jack, 2005). In other words, the separation of these processes is possible but does not necessarily occur in most social situations. A large body of research indicates that the relation between understanding oneself and others is parallel and reciprocal (see Damon & Hart, 1982; Gopnik & Astington, 1988, Nichols & Stich, 2003). It is also doubted whether statements of double dissociation in autism and schizophrenia are correct (Carruthers, 2009). It may be that in these cases the abilities to understand both oneself and others are impaired. Undoubtedly, more empirical research on this subject is necessary.

The question about mutual influences between oneself and others are undeveloped in TT and PMT. TT is more concerned about separation between these processes, conceiving them as distinct and autonomous. PMT adds the possibility of their interaction by introducing the optional opportunity of using one's own person model when developing another's person model, or *vice versa*. The ST perspective is the most developed of the three on this issue and it indicates that understanding others involves a type of access to one's own mind. However, the question of to what extent these processes are intertwined is still open within ST.

The interrelation between self- and other-understanding in TT, ST, and PMT, as well as related empirical research has been more correlated with the question of whether one of them may occur without the other. Their interactions and mutual influences raise a question established in this article: are self- and other-understanding distinct or intertwined? The arguments presented here indicate that the second option is more plausible, but empirical research on this issue is necessary to establish specific conclusions. The ways in which these two types of understanding cooperate and affect one another is worthy of exploration.

## 8. Conclusion

In this article, the topic of the relation between self- and other-understanding was introduced by two questions: (I) are abilities involved in self- and other-understanding identical or diverse? (II) are these abilities distinct or intertwined? These questions were applied to theory theory, simulation theory, and the person model theory.

From the analyses presented in this work, it may be concluded that self- and other-understanding are similar but not identical. In both situations, the same kinds of mental states may be attributed. In both cases, the ascription of mental states engages interpretation of behavior and sensations and is enhanced by imaginative and reflexive thinking. What is more, the development and neural correlates of both types of understanding are similar to one another. The differences arise from the way of experiencing mental states related to agency. Compared to other-understanding, in self-understanding there is a different type of access to mental states and better accuracy of the attribution of some mental states, i.e., current judgements, images or feelings. At the same time, it is worth mentioning that such analysis cannot determine whether self-understanding and other-understanding are two abilities or two modes of one ability. It can only be stated that self-understanding and other-understanding are similar but not identical abilities. For the second question, it should be stated that the separation of these abilities is very difficult, if not impossible. However, when making such statements it is worthwhile establishing empirical research concerning this topic. In TT, ST, PMT, and related research, there has been more of a focus on the possibility of self-understanding occurring without other-understanding or *vice versa* than on their mutual interactions. However, addressing this issue seems to be essential for understanding social cognition and thus it is worthy of further theoretical and empirical exploration.

## Acknowledgement

This research was financially supported by the Jagiellonian University (grant DSC). I would like to thank anonymous reviewers for valuable comments on the earlier version of this paper.

## References

- Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism on Theory of Mind*. Cambridge: MIT Press.
- Bogdan, R. J. (2005). Why self-ascriptions are difficult and develop late. In B. F. Malle & S. D. Hodges (Eds.), *Other minds: How humans bridge the divide between self and others* (pp. 190-206). New York: Guilford Press.
- Boyer, P., Robbins, P., Jack, A. I. (2005). Varieties of self-systems worth having. *Consciousness and Cognition, 14*, 647–660.
- Brandt, S., Buttelmann, D., Lieven, E., Tomasello, M. (2016). Children's understanding of first- and third-person perspectives in complement clauses and false-belief tasks. *Journal of experimental child psychology, 151*, 131-143.

- Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and brain sciences*, 32, 121-138.
- Damon, W., Hart, D. (1982). The development of self-understanding from infancy through adolescence. *Child development*, 53, 841-864.
- Decety, J., Grèzes, J. (2006). The power of simulation: imagining one's own and other's behavior. *Brain research*, 1079, 4-14.
- Fonagy, P., Luyten, P. (2009). A developmental, mentalization-based approach to the understanding and treatment of borderline personality disorder. *Development and psychopathology*, 21, 1355-1381.
- Frith, C. D., Frith, U. (2006). The neural basis of mentalizing. *Neuron*, 50, 531 - 534.
- Gallagher, S. (2008). Direct perception in the intersubjective context. *Consciousness and Cognition*, 17, 535-543.
- Gallagher, S., Hutto, D. (2008). Understanding others through primary interaction and narrative practice. *The shared mind: Perspectives on intersubjectivity*, 12, 17-38.
- Goldman, A. (2006). *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading (Philosophy of Mind)*. Oxford: Oxford University Press.
- Goldman, A. I. (2009). Mirroring, mindreading, and simulation. In J. Pineda (Ed.) *Mirror Neuron Systems: The Role of Mirroring Processes In Social Cognition* (pp. 311-330). Humana Press.
- Goldman, A. I. (2012). Theory of Mind. In E. Margolis, R. Samuels & S. P. Stich (Eds.), *The Oxford Handbook of Philosophy of Cognitive Science* (pp. 402-424). Oxford University Press.
- Goldman, A., Mason, K. (2007). Simulation. In P. Thagard (Ed.), *Philosophy of Psychology and Cognitive Science* (pp. 267-293). Elsevier.
- Gopnik, A. (1993). How we know our minds: The illusion of first-person knowledge of intentionality. *Behavioral & Brain Sciences*, 16, 29-113.
- Gopnik, A. (1996). Theories and Modules; Creation Myths, Developmental Realities, and Neurath's Boat. In P. Carruthers & P. K. Smith (Eds.), *Theories of Theories of Mind* (pp. 169-183). Cambridge: Cambridge University Press.
- Gopnik, A., Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child development*, 59, 26-37.
- Gordon, R. M. (1995). Simulation without introspection or inference from me to you. In M. Davies & T. Stone (Eds.), *Mental Simulation* (pp. 53-67). Blackwell.
- Gordon, R. M. (2008). Beyond mindreading. *Philosophical explorations*, 11, 219-222.
- Guerini, R., Marraffa, M. (2015). Mentalization and Self-Identity: Philosophical and Psychological Perspectives. *Rivista Internazionale di Filosofia e Psicologia*, 6, 196-197.
- Hernik, M., Fearon, P., Fonagy, P. (2009). There must be more to development of mindreading and metacognition than passing false belief tasks. *Behavioral and Brain Sciences*, 32, 147-148.
- Hodges, S. D. (2005). Is how much you understand me in your head or mine. In B. F. Malle & S. D. Hodges (Eds.), *Other minds: How humans bridge the divide between self and others* (pp. 298-309). New York: Guilford Press.

- Hyslop, A. (2016). Other Minds. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. URL = <<https://plato.stanford.edu/archives/spr2016/entries/other-minds/>>.
- Keyzers, C., Gazzola, V. (2008). Unifying social cognition. In J. A. Pineda (Ed.) *Mirror neuron systems* (pp. 1-35). Humana Press.
- Krueger, J., Overgaard, S. (2012). Seeing subjectivity: Defending a perceptual account of other minds. *ProtoSociology: Consciousness and subjectivity*, 47, 239-262.
- Mills, C. M., Danovitch, J. H. (2009). Getting to know yourself... and others. *Behavioral and Brain Sciences*, 32, 154-155.
- Mitchell, J. P., Macrae, C. N., Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron*, 50, 655-663.
- Newen, A. (2015). Understanding Others: The Person Model Theory. In T. Metzinger & J. M. Windt (Eds.), *Open MIND: 26(T)*. Frankfurt am Main: MIND Group.
- Newen, A., Schlicht, T. (2009). Understanding other minds: A criticism of Goldman's Simulation Theory and an outline of the Person Model Theory. *Grazer Philosophische Studien*, 79, 209-242.
- Nichols, S., Stich, S. (2003). *Mindreading. An Integrated Account of Presence, Self-Awareness, and Understanding Other Minds*. Oxford: Clarendon Press.
- Perner, J. (1991). *Understanding the representational mind*. The MIT Press.
- Premack, D., Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. *Behavioral and brain sciences*, 1, 515-526.
- Robbins, P. (2004). Knowing Me, Knowing You. Theory of Mind and the Machinery of Introspection. *Journal of Consciousness Studies*, 11, 129-43.
- Vogele, K., Newen, A. (2009). Consciousness of oneself and others in relation to mental disorders. In S. J. Wood, N. B. Allen, C. Pantelis (Eds.) *The neuropsychology of mental illness* (pp. 408-413). Cambridge University Press.
- Waytz, A., Mitchell, J. P. (2008). Two mechanisms for simulating other minds: dissociations between mirroring and self-projection. *Current Directions in Psychological Science*, 20, 197-200.
- Wellman, H. M. (1990). *The child's theory of mind*. MIT Press.
- Wimmer, H., Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103-12.

**Adrianna Smurzyńska** is a PhD candidate at the Institute of Philosophy, Jagiellonian University (Krakow, Poland). Her main research interests involve philosophy of psychology and philosophy of psychiatry.

The editorial and publishing process of this publication has been financed by the Ministry of Science and Higher Education from the funds for the dissemination of research (DUN) within the framework of publishing activity, contract no. 711/P-DUN/2019, period of implementation: the years 2019–2020.