Measurement invariance and change of affective and cognitive Theory of Mind in mental health patients

Inaugural-Dissertation

zur Erlangung des Doktorgrades der Philosophie

der Ludwig-Maximilians-Universität

München

vorgelegt von

Magdalena Knopp

Wien

2024

Erstgutachter: PD. Dr. Manuel Sprung

Zweitgutachterin: Prof. Dr. Beate Sodian

Tag der mündlichen Prüfung: 25.4.2024

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisors, Manuel Sprung and Beate Sodian, for their unwavering support, invaluable guidance, and constructive feedback throughout the journey of this dissertation. Your expertise and insights into the realm of psychology have been indispensable for the realization of this work.

The research presented here was supported by the Society for Research Promotion Lower Austria (GFF). I would like to extend my sincere gratitude to the opportunities provided by the GFF.

I would also like to extend my sincere thanks to the entire faculty and staff of the Psychology/Ludwig-Maximilians-Universität München for providing a stimulating academic environment, essential resources, and continuous encouragement.

Moreover, my appreciation goes to the participants who were involved in this study. Without their time, commitment, and willingness to share personal experiences, this research would not have been possible.

I am profoundly grateful to my peers and colleagues who shared this academic journey with me, especially [Juliane Burghardt, Claudia Oppenauer, Jasmin Willinger, Silvia Gradl, Selina Persen, Anna Koestler], for the countless brainstorming sessions, discussions, and for always being there to offer a fresh perspective when challenges emerged. Special thanks go to Juliane Burghardt and Claudia Oppenauer for supporting me with their knowledge, expertise, and skills. I always loved and appreciated our chats, status updates and lunchtimes. Thank you for being such kind persons and great office colleagues.

Lastly, I want to express my grateful thanks to my family, [Jan Hronik, Ingrid Knopp], for their unconditional love, endless patience, and unwavering belief in my abilities. Their constant encouragement and faith pushed me forward, even in the most challenging times. This dissertation is not just a proof to my hard work, but also to the collective efforts of everyone who believed in and contributed to my goals.

Elaborate Summary

Theory of Mind (ToM) is one of the most investigated aspects of social cognition and refers to a range of mental processes that allow humans to interpret, perceive, and respond to social cues, while accounting for a specific situation. ToM is the cognitive ability to ascribe mental states such as desires, beliefs, intentions and emotions to oneself and others (Apperly, 2012). It allows humans to predict and explain behavior. ToM can also be construed as a part of empathy, in this case ToM is equated to one aspect of empathy termed cognitive empathy. The second aspect of empathy is affective empathy, which refers to affective contagion (i.e., sharing the emotional experiences of others). The two systems of empathy rely on separate but interacting brain networks (Shamay-Tsoory, 2011).

ToM is not a monolithic ability; it includes two different facets, affective and cognitive ToM (Maleki et al., 2020; Shamay-Tsoory et al., 2007; Zabihzadeh et al., 2017). Cognitive ToM involves representing thoughts, intentions, or beliefs by making inferences about mental states through interpretation or prediction of others' behaviors (i.e., understanding the intentions of others). Affective ToM involves representing emotions and feelings by decoding and discriminating the mental states of others based on available environmental information.

The importance of ToM is exemplified by the finding that ToM deficits have been reliably associated to mental disorders. For instance, individuals with alcohol use disorder (AUD) show medium to strong ToM impairments (Bora & Zorlu, 2017; Hanegraaf et al., 2021; Onuoha et al., 2016; Sanvicente-Vieira et al., 2017). ToM impairments are also common in patients with borderline personality disorder (BPD; Bora, 2021; Hanegraaf et al., 2021; Németh et al., 2018; Richman & Unoka, 2015). The ToM deficits shown by these groups are likely contributing to the interpersonal difficulties frequently seen among BPD and AUD patients (Hanegraaf et al., 2021). Further, both disorders are frequently comorbid. At the same time the two groups are marked by uniquely different socio-demographic characteristics. Even though both are marked by ToM deficits, the two disorders have been linked to different types of ToM deficits. People can show exceeding ToM, less ToM, or no ToM. Exceeding ToM implies over-interpretating others' behaviors. Reduced ToM and no ToM imply a limited tendency to ascribe mental states to others, which is often marked by literal understanding (Vegni et al., 2021). By examining these two distinct clinical samples,

which show characteristic symptoms, socio-demographic characteristics, and ToM deficits, a more comprehensive understanding of ToM abilities can be achieved.

Studies on socio-demographic characteristics such as sex or age have shown inconsistent results. There is some evidence that men outperform women (Russell et al., 2007), while in other studies sex differences are largely absent (Barrett et al., 1998; Derntl et al., 2010). Given the inconsistent results, further research is crucial to clarify sex influences on ToM. Studies about age indicate a decline in ToM in older adults (Bailey et al., 2008; Henry et al., 2013; Krych-Appelbaum et al., 2007). Hence, ToM development across the entire lifespan is relevant.

ToM research is plagued by the presence of a multitude of measures (Olderbak & Wilhelm, 2020; Quesque & Rossetti, 2020). According to Quesque and Rossetti (2020), ToM measures need to require test takers to represent others' mental states and to distinguish these from one's own. These two criteria are fulfilled by the Movie for the Assessment of Social Cognition (MASC). The MACS is consistently employed across all three studies presented in this dissertation. While many ToM measures have been criticized for lacking validity (Pabst et al.,2022), the MASC (Dziobek et al., 2006) is considered both valid and reliable (Benito-Ruiz et al., 2022; Fossati et al., 2018). Nevertheless, it has created inconsistent findings. To enhance the understanding of the MASC's psychometric properties, individual and situational factors that influence affective and cognitive ToM should be investigated to improve the understanding of ToM measurements in the clinical context. This dissertation presents findings from three articles. Article 1 tested invariance of the MASC over the duration of the test and depending on sex and age. Article 2 explored the impact of the interaction partner's gender per item within the ToM measure. Article 3 presents findings on changes of ToM during inpatient treatment.

Article 1 explored ToM in a large clinical sample, including patients with AUD and Personality Disorders (PD). The study assessed the changes in cognitive and affective ToM within the MASC over the course of a test session. Results indicated a decrease in cognitive ToM and an improvement in affective ToM performance as the test progressed. This effect was moderated by age, older participants showed a more pronounced trend in affective ToM than younger. Sex differences were also observed, with women displaying higher affective ToM skills. Furthermore, individuals with PD generally showed better ToM abilities than those with AUD. These findings underscore

the importance of considering individual (e.g., age, sex), and situational (test duration) variables when measuring ToM abilities, offering insight into potential reasons for inconsistent results in prior studies.

Article 2 examined ToM assessments by exploring how the gender of the interaction partners and the congruence of their perspective affect ToM measurement. Based on previous research, it was hypothesized that ToM scores vary with the social groups of the interaction partners. ToM performance is enhanced when the target and the perceiver share similar social groups and perspectives. Conversely, ToM performance is diminished when interaction partners have identical social groups but possess differing perspectives (Simpson & Todd, 2017). Article 2 assessed ToM in a large clinical sample, including individuals with AUD, and PD, and healthy controls (HC). A consistent pattern emerged, items with the same gender of interaction partners resulted in lower ToM abilities compared to items with different gender of interaction partners, within both clinical samples and HC. Within clinical samples, items with male targets resulted in better ToM performance compared to items with female targets, the smaller HC did not replicate this effect. The findings indicate that the gender of interaction partners serves as a significant moderator of ToM performance. However, the findings of this article need to be further investigated with studies in which the gender within items is systematically manipulated.

Article 3 focused on the effect inpatient psychotherapeutic treatment has on affective and cognitive ToM and related outcomes (like alcohol use and psychological symptoms) in individuals with AUD. This longitudinal study used the MASC to track changes in affective and cognitive ToM between admission and discharge. This longitudinal study was carried out within a naturalistic environment to increase the degree of clinical representativeness and the external validity of clinical care routine. All variables were assessed at admission and discharge of the inpatient stay. Article 3 investigated affective and cognitive ToM in a longitudinal study within individuals with AUD for the first time. It contributes to the expanding body of literature by demonstrating that only cognitive ToM improved following an eclectic abstinence-oriented inpatient treatment. It might be that more specific interventions are required to yield improvements in affective ToM. Moreover, the results of Article 3 reveal that initial ToM abilities positively correlated with a decrease in symptoms of depression and somatization, suggesting ToM's potential as a treatment target to improve

psychological health not only in patients with BPD (Kvarstein et al., 2020), but also in those with AUD. These findings highlight the need for further research on ToM's relationship with psychological symptoms in AUD, emphasizing the value of incorporating ToM training into treatment programs. This fits with the current literature highlighting ToM's vital role in the recovery process (Rupp et al., 2017) and the alleviation of comorbid symptoms in other mental disorders (Sondermann et al., 2020).

In conclusion, this dissertation explored the interactions of multiple individual and situational factors that influence ToM measurement. The studies showed that ToM measurements would be more consistent if age is accounted for. Future, ToM measures could balance the number of male and female items. The studies help to understand the heterogeneity in the previous findings. By differentiating between these two ToM facets within sizable clinical samples, while always relying on the MASC, this research enriches our understanding of the subject.

Previous findings highlight the significance of ToM across a spectrum of mental disorders, McLaren et al. (2022) provided a comprehensive examination of ToM exceeding in many mental disorders. Moreover, a systematic review by Cotter et al. (2018) emphasized the role of social cognitive processes as transdiagnostic clinical indicators across various clinical presentations, underscoring their importance in discerning disease progression, and treatment efficacy. Previous findings have led to the conclusion that ToM should be considered as a transdiagnostic factor essential for conceptualizing mental health. This was for instance recognized within the Research Domain Criteria (RDoC) framework (National Institute of Mental Health, 2020). This framework assesses dysfunctions across broad psychological and biological matrices and tries to overcome the limitations of categorical diagnostic models.

Especially the last article highlighted the potential relevance of ToM for therapeutic processes. To effectively integrate social cognition training within psychotherapeutic treatment programs, a more comprehensive understanding of ToM in clinical adult samples is necessary. Evidence supports that such integrative therapies can improve outcomes across various psychiatric disorders (Peyroux & Franck, 2014). This offers valuable insights for identifying and prioritizing therapeutic interventions in patients within clinical samples, specifically among patients with AUD and PD.

Deutsche Zusammenfassung

Die Theory of Mind (ToM) ist einer der am meisten untersuchten Aspekte der sozialen Kognition und bezieht sich auf eine Reihe von mentalen Prozessen, die es Menschen ermöglichen, soziale Hinweise zu interpretieren, wahrzunehmen und darauf zu reagieren, während sie eine bestimmte Situation berücksichtigen. ToM ist die kognitive Fähigkeit, sich selbst und anderen mentale Zustände wie Wünsche, Überzeugungen, Absichten und Gefühle zuzuschreiben (Apperly, 2012). Sie ermöglicht es dem Menschen, Verhalten vorherzusagen und zu erklären. ToM kann auch als Teil der Empathie verstanden werden; in diesem Fall wird die ToM mit einem Aspekt der Empathie gleichgesetzt, der als die kognitive Empathie bezeichnet wird. Der zweite Aspekt der Empathie ist die affektive Empathie, die sich auf die emotionale Ansteckung (d. h. das Teilen der emotionalen Erfahrungen anderer) bezieht. Die beiden Systeme der Empathie beruhen auf getrennten, aber interagierenden Arealen im Gehirn (Shamay-Tsoory, 2011).

ToM ist keine monolithische Fähigkeit; sie umfasst zwei verschiedene Facetten, die affektive und kognitive ToM (Maleki et al., 2020; Shamay-Tsoory et al., 2007; Zabihzadeh et al., 2017). Die kognitive ToM beinhaltet die Darstellung von Gedanken, Absichten oder Überzeugungen durch Rückschlüsse auf mentale Zustände durch Interpretation oder Vorhersage des Verhaltens anderer (d. h. das Verstehen der Absichten anderer). Die affektive ToM umfasst die Darstellung von Emotionen und Gefühlen durch Dekodierung und Differenzierung der mentalen Zustände anderer auf der Grundlage der verfügbaren Umgebungsinformationen.

Die Bedeutung der ToM wird durch Evidenzen veranschaulicht, dass ToM-Defizite zuverlässig mit psychischen Erkrankungen in Verbindung gebracht wurden. So zeigten beispielsweise Personen mit einer Alkoholerkrankung (AUD) mittlere bis starke ToM-Beeinträchtigungen (Bora & Zorlu, 2017; Hanegraaf et al., 2021; Onuoha et al., 2016; Sanvicente-Vieira et al., 2017). ToM-Beeinträchtigungen zeigten sich häufig auch bei PatientInnen mit einer Borderline-Persönlichkeitsstörung (BPD; Bora, 2021; Hanegraaf et al., 2021; Németh et al., 2018; Richman & Unoka, 2015). Die von diesen Gruppen berichteten ToM-Defizite tragen unter anderem zu den zwischenmenschlichen Schwierigkeiten bei, die häufig bei BPD- und AUD-PatientInnen zu beobachten sind (Hanegraaf et al., 2021). Zudem gehen beide Erkrankungen mit vielen Komorbiditäten einher. Gleichzeitig sind beide Gruppen

durch eindeutig unterschiedliche soziodemografische Merkmale gekennzeichnet. Auch wenn beide durch ToM-Defizite gekennzeichnet sind, wurden beide Erkrankungen mit unterschiedlichen Arten der ToM-Defizite in Verbindung gebracht. Menschen können eine exceeding ToM, eine geringere ToM oder gar keine ToM aufweisen. Die exceeding ToM bedeutet, dass das Verhalten anderer überinterpretiert wird. Die reduzierte ToM und keine ToM implizieren eine eingeschränkte Tendenz, anderen Menschen mentale Zustände zuzuschreiben. Dies ist oft durch wörtliches Verstehen gekennzeichnet (Vegni et al., 2021). Durch die Untersuchung dieser beiden unterschiedlichen klinischen Stichproben mit charakteristischen Symptomen, soziodemografischen Merkmalen und ToM-Defiziten, kann ein umfassenderes Verständnis der ToM-Fähigkeiten erreicht werden.

Studien zu soziodemografischen Merkmalen wie Geschlecht oder Alter führten zu widersprüchlichen Ergebnissen. Es gibt einige Hinweise darauf, dass Männer besser abschneiden als Frauen (Russell et al., 2007), während in anderen Studien Geschlechtsunterschiede weitgehend fehlen (Barrett et al., 1998; Derntl et al., 2010). Angesichts der widersprüchlichen Ergebnisse sind weitere Forschungsprojekte hinsichtlich der geschlechtsspezifischen Einflüsse auf die ToM unerlässlich. Studien über das Alter deuten auf einen Rückgang der ToM bei älteren Erwachsenen hin (Bailey et al., 2008; Henry et al., 2013; Krych-Appelbaum et al., 2007). Daher ist die Entwicklung der ToM über die gesamte Lebensspanne hinweg relevant.

Die ToM-Forschung steht, aufgrund einer Vielzahl von Messverfahren vor großen Herausforderungen (Olderbak & Wilhelm, 2020; Quesque & Rossetti, 2020). Nach Quesque und Rossetti (2020) beinhalten die grundlegenden Komponenten der ToM, die Darstellung der mentalen Zustände anderer und die Fähigkeit diese von den eigenen zu unterscheiden. Diese beiden Kriterien werden mit dem Movie for the Assessment of Social Cognition (MASC) erfasst, einem Instrument, das in allen drei Studien dieser Dissertation eingesetzt wurde. Während viele ToM-Messungen wegen mangelnder Validität kritisiert werden (Pabst et al., 2022), wurde der MASC (Dziobek et al., 2006) als valide und zuverlässig beschrieben (Benito-Ruiz et al., 2022; Fossati et al., 2018). Dennoch hat er zu inkonsistenten Ergebnissen geführt. Um das Verständnis der psychometrischen Eigenschaften des MASC zu verbessern, sollten individuelle und situative Faktoren, die die affektive und kognitive ToM beeinflussen, untersucht werden. Zudem sollte hierdurch das Verständnis der ToM-Messungen im

klinischen Kontext verbessert werden. In dieser Dissertation werden Ergebnisse aus drei Artikeln vorgestellt. Artikel 1 untersucht die Invarianz des MASC über die Dauer der Testung in Abhängigkeit von Geschlecht und Alter. Artikel 2 analysiert den Einfluss des Geschlechts der Interaktionspartner pro Item innerhalb der ToM-Messung. In Artikel 3 werden Ergebnisse zu Veränderungen der ToM während einer stationären Behandlung vorgestellt.

Artikel 1 untersuchte die ToM in einer großen klinischen Stichprobe, die PatientInnen mit AUD und Persönlichkeitsstörungen (PD) umfasste. Die Studie untersuchte die Veränderungen der kognitiven und affektiven ToM anhand des MASCs im Verlauf einer Testsitzung. Die Ergebnisse zeigten eine Abnahme der kognitiven ToM und eine Verbesserung der affektiven ToM-Leistung im Verlauf der Testung. Dieser Effekt wurde durch das Alter moderiert und war bei älteren TeilnehmerInnen besonders ausgeprägt. Es wurden auch geschlechtsspezifische Unterschiede beobachtet, wobei Frauen höhere affektive ToM-Fähigkeiten aufwiesen. Darüber hinaus zeigten Personen mit PD generell bessere ToM-Fähigkeiten als solche mit AUD. Diese Ergebnisse unterstreichen, wie wichtig es ist, bei der Messung von ToM-Fähigkeiten individuelle (z. B. Alter, Geschlecht) und situative (Testdauer) Variablen zu berücksichtigen und geben Aufschluss über mögliche Gründe für widersprüchliche Ergebnisse früherer Studien.

In Artikel 2 wurde untersucht, wie sich das Geschlecht der Interaktionspartner und die Kongruenz ihrer Perspektiven auf die ToM-Messung auswirken. Basierend auf der Literatur wurde die Hypothese aufgestellt, dass die ToM von sozialen Gruppen der Interaktionspartner beeinflusst wird. Die ToM-Leistung ist höher, wenn die Zielperson und der Wahrnehmende ähnliche soziale Gruppen und Perspektiven haben. Umgekehrt sinkt die ToM-Leistung, wenn die Interaktionspartner identische soziale Gruppen, aber unterschiedliche Perspektiven haben (Simpson & Todd, 2017). Artikel 2 untersuchte die ToM in einer großen klinischen Stichprobe, die sich auf Personen mit AUD und PD sowie gesunde Kontrollpersonen (HC), bezieht. Es zeigte sich ein konsistentes Muster: Items mit Interaktionspartnern desselben Geschlechts führten zu niedrigeren ToM-Fähigkeiten im Vergleich zu Items mit Interaktionspartnern unterschiedlichen Geschlechts. Dies zeigte sich sowohl in den klinischen Stichproben als auch in der Stichprobe der HC. In den klinischen Stichproben führten Items mit männlichen Personen zu einer besseren ToM-Leistung im Vergleich zu Items mit

weiblichen Personen. In der kleineren Stichprobe der HC konnte dieser Effekt nicht repliziert werden. Die Ergebnisse deuten darauf hin, dass das Geschlecht der Interaktionspartner ein signifikanter Moderator der ToM-Leistung ist. Die Ergebnisse dieses Artikels müssen jedoch durch Studien, die sich auf die systematische Manipulation des Geschlechts der Items konzentrieren, weiter untersucht werden.

Artikel 3 befasste sich mit den Auswirkungen einer stationären psychotherapeutischen Behandlung auf die affektive und kognitive ToM und die damit verbundenen psychologischen Symptome bei Personen mit AUD. In dieser Längsschnittstudie wurde der MASC verwendet, um Veränderungen in der affektiven und kognitiven ToM zwischen Aufnahme und Entlassung zu erfassen. Diese Längsschnittstudie wurde in einem naturalistischen Umfeld durchgeführt, um den Grad der klinischen Repräsentativität zu erhöhen und die externe Validität zu verstärken. Zudem wurden diese Variablen sowohl bei der Aufnahme als auch bei der Entlassung der Behandlung gemessen. Artikel 3 untersuchte erstmalig die affektive und kognitive ToM in einer Längsschnittstudie bei Personen mit AUD. Der Artikel erweitert das Verständnis der bestehenden Literatur und zeigt auf, dass die kognitive ToM durch eine eklektische abstinenzorientierte stationäre Behandlung verbessert werden kann. Es könnte sein, dass spezifischere Interventionen erforderlich sind, um Verbesserungen in der affektiven ToM zu erzielen. Darüber hinaus korrelierten die anfänglichen ToM-Fähigkeiten positiv mit einer Abnahme der Depressions- und Somatisierungssymptome.

Dies deutete darauf hin, dass ToM ein potenzielles Behandlungsziel zur Verbesserung der psychischen Gesundheit darstellt (Kvarstein et al., 2020). Diese Ergebnisse sprechen gemäß der aktuellen Literatur für den Bedarf an weiteren Forschungsprojekten im Zusammenhang zwischen der ToM und psychologischen Symptomen bei Personen mit AUD und unterstreichen den potenziellen Einfluss der ToM bei psychologischen und psychotherapeutischen Interventionen im Hinblick auf den Genesungsprozess (Rupp et al., 2017) und die Linderung von komorbiden Symptomen (Sondermann et al., 2020).

Zusammenfassend lässt sich sagen, dass in dieser Dissertation die Interaktionen mehrerer individueller und situativer Faktoren untersucht wurden, die die ToM-Messung beeinflussen. Die Studien zeigten, dass ToM-Messungen konsistenter sein sollten, wenn das Alter berücksichtigt wird. Zukünftige Studien sollten im Rahmen

der ToM-Messungen die Anzahl der männlichen und weiblichen Items ausgleichen. Die Studien liefern Hinweise, um die Heterogenität der bisherigen Ergebnisse besser zu verstehen. Die Differenzierung der beiden ToM-Facetten innerhalb großer klinischer Stichproben, die sich stets auf den MASC stützen, bereichert unser Verständnis der ToM -Forschung.

McLaren et al. (2022) haben eine umfassende Untersuchung der ToM-Defizite bei vielen psychischen Erkrankungen durchgeführt. Darüber hinaus betonte eine systematische Übersichtsarbeit von Cotter et al. (2018) die Rolle sozial kognitiver Prozesse als transdiagnostische Indikatoren über verschiedene klinische Erkrankungen hinweg und bekräftigte ihre Bedeutung bei der Erkennung des Krankheitsverlaufs und der Wirksamkeit der Behandlung. Frühere Befunde zeigten, dass die ToM als transdiagnostischer Faktor betrachtet werden sollte, der für die Konzeptualisierung psychischer Gesundheit wesentlich ist. Dies steht im Einklang mit der Research Domain Criteria (RDoC; National Institute of Mental Health, 2020). Dieses Model bewertet Erkrankungen über psychologische und biologische Ebenen hinweg und versucht, die Grenzen kategorischer Diagnosemodelle zu überwinden.

Insbesondere der letzte Artikel hat die potenzielle Bedeutung der ToM für therapeutische Prozesse hervorgehoben. Um ein soziales Kognitionstraining effektiv in psychotherapeutische Behandlungsprogramme zu integrieren, ist ein umfassenderes Verständnis der ToM in klinischen Erwachsenenstichproben notwendig. Es gibt Belege dafür, dass solche integrativen Therapien die Ergebnisse bei verschiedenen psychiatrischen Erkrankungen verbessern können (Peyroux & Franck, 2014). Dies bietet wertvolle Erkenntnisse für die Identifizierung und Priorisierung von therapeutischen Maßnahmen bei PatientInnen in klinischen Stichproben, insbesondere bei PatientInnen mit AUD und PD.

Table of Contents

1. General Introduction

1.1 What is a Theory of Mind?

1.1.1 Definition of Theory of Mind

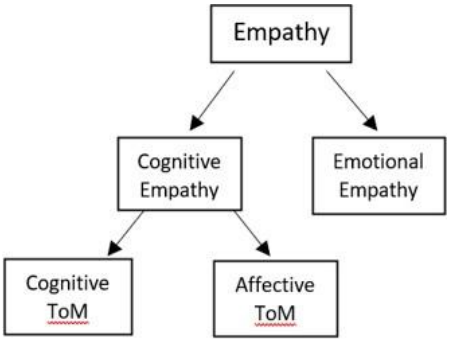
The term Theory of Mind (ToM) was first introduced in 1978 by Premack and Woodruff (Premack & Woodruff, 1978), while testing chimpanzees' abilities to understand false beliefs, which represent incorrect ideas. They examined the chimpanzees' ability to discern an actor's intention by presenting them with videos of humans facing various problems. Premack and Woodruff (1978) argued that attributing mental states requires theoretical knowledge, which is a *theory about the mind of another person (i.e., ToM)*. Noting that the chimpanzees chose solutions based on their understanding of interpersonal relationships within their environment (Dennett, 1978). These ideas were subsequently applied to children. Extensive research in developmental psychology, especially by Freidson (1958), revealed that children are innately egocentric, which leads to systematic ToM errors. By the age of four, they begin to adopt the perspectives of others (Marvin et al., 1976) and start to differentiate between their own and other's mental states. This is a core element of ToM, which is described as a cognitive ability to recognize others' mental states, and is the foundation for social functioning (Rosello et al., 2020; Sabbagh et al., 2009; Singer, 2006). Children learn to infer behaviors based on another's incorrect beliefs (Schaafsma et al., 2015; Wimmer & Perner, 1983). In children, ToM is linked to peer popularity and prosocial behavior, including acts of helping, cooperating and comforting (Imuta et al., 2016; Slaughter et al., 2015). A better ToM was also related to better friendships (Fink et al., 2015). These findings also expand to adults, ToM is a necessary prerequisite for understanding social interactions, which is the basis for socially adequate responses (Luke & Banerjee, 2013; Thirion-Marissiaux & Nader-Grosbois, 2008). In contrast, ToM deficits are associated with poor social adjustment (Adegboye et al., 2017). Thus, ToM development is essential for the development of social competencies like conflict management and communicative competencies.

Subsequent studies have investigated ToM and its deficits in human development. Most research has centered on children, however recent investigations have expanded their research to adults (Apperly, 2012) and those with mental disorders (Cotter et al., 2018), such as autism spectrum disorder (ASD; Andreou & Skrimpa, 2020; Baron-Cohen et al., 1985).

Following the original mentioning of ToM, research about this ability flourished. At the same time a plethora of overlapping concepts and connected constructs were introduced (Olderbak and Wilhelm, 2020; Quesque & Rossetti, 2020). Since their definitions are often vague, differentiating them is not entirely possible. Nevertheless, the following paragraph will present central aspects of ToM's nomological network, afterwords theories of ToM will be outlined. ToM ability is a social cognition ability, which refers to a range of mental processes that allow humans to interpret, perceive, and respond to social cues while accounting for a specific situation (Andreou et al., 2015; Chung et al., 2011; Dekker et al., 2021). Moreover, ToM is the cognitive ability to ascribe mental states such as desires, intentions and emotions to others and oneself (Apperly, 2012). It allows humans to explain and predict behavior. ToM implies the ability to distinguish between others' and one's own mental state (Quesque & Rossetti, 2020). ToM can also be understood as a part of empathy, in this case ToM is equated to one facet of empathy termed cognitive empathy. The second facet of empathy is affective empathy, which refers to affective contagion (i.e. sharing the emotional experiences of others). The two systems of empathy are depend on separate but interacting brain networks (Shamay-Tsoory, 2011). An illustration of the relationship between empathy and ToM is shown in Figure 1. Some researchers have used mentalizing (Frith et al., 1991) or mindreading as synonyms for ToM (Vogeley et al., 2001).

Figure 1

An illustration of empathy and the two ToM subsystems. Modelled after Dvash and Shamay-Tsoory (2014)



ToM is a multifaceted construct, which can be construed in three different ways: as a conceptual domain, a cognitive process and a social competence (Apperly, 2012). The conceptual domain approach describes concepts that need to development in

order to have a ToM like desire, belief, and intention. The cognitive process approach explores which processes are part of ToM. The social competence approach interprets ToM performance as a skill which varies between individuals.

1.1.2 Theories about Theory of Mind

Researcher have formulated theories to explain how ToM emerges during children's cognitive development. Reasoning about beliefs and desires begins very early, there are three prominent theories formed in the 1990s: Theory-Theory, Modularity Theory, and Simulation Theory, which will be summarized below.

1.1.2.1 Theory-Theory. The theory-theory as proposed by Wellman et al. (Gopnik, 2003; Gopnik & Wellman, 1992, 2012), suggests that individuals possess an implicit understanding of the human mind's function and structure. This coherent system of concepts aids in predicting and explaining behavior (Premack & Woodruff, 1978). Theory-theorists agreed on the emergence of an understanding of mental representation (Perner, 1991). In general, knowledge about the mind consists of domain-specific theories and conceptual changes and concerns the development of children's mental state understanding. In this view, children gather information akin to how scientists collect data to build a scientific theory. Conceptual change in children's ToM occurs when new evidence contradicts their current understanding. This is akin to the shift in scientific theories when the accumulation of contradictory evidence prompts a transition to a new theory, involving several gradual stages (Gopnik & Wellman, 1992). This approach posits that theories about the social world form a network of mental concepts, such as beliefs, desires, and plans. Children's experiences are crucial in shaping these concepts (Jenkins & Astington, 1996). An early developing concept is the ability to create models of reality or fictional situations in a nonrepresentational manner, evident in two-year-olds. By three years, children begin to show an understanding of beliefs in a similar nonrepresentational way, including those beliefs that accurately reflect the world. By the ages of four to five, children develop a representational understanding of the mind, focusing on representations of reality rather than reality itself (Gopnik & Wellman, 1992). Concepts like beliefs, which emerge later in development (Wellman & Liu, 2004) and understanding the representational relationship between reality and a model (Perner, 1991) are key milestones. The theory-theory is supported by changes in children's predictions about others' behavior. Thus, if evidence about others' actions cannot be

interpreted, the child's theory is extended. However, if neither the initial theory nor the extension explains the phenomenon, the initial theory is rejected and a new theory arises. This new theory can explain both the previously explained phenomena and the still unclear phenomena. This approach also posits a developmental link between understanding one's own mental states and those of others, predicting that both are based on the same conceptual system and should be comprehended around the same age (Sodian, 2005).

1.1.2.2 Modularity Theory. Modularity theories suggest that children's ToM development is driven by an innate neural ToM Mechanism (ToMM) for forming meta-representations (Leslie, 1994; Scholl & Leslie, 1999). The modularity of mind is characterized by features such as automatic activation, rapid processing, output generation, domain specificity, and a unique neural architecture (Fodor, 1983). Domain specificity, in particular, involves specialized types of representation (Scholl & Leslie, 1999). Although this mechanism is trained by experience, children have the ability from birth on (Baillargeon et al., 2010). ToMM is specifically designed for learning about mental states, indicating that ToM has a unique, innate foundation distinct from other cognitive domains (Scholl & Leslie, 2001). Children only a few months old can distinguish between agents and non-agents by using a rudimentary understanding of bodily mechanisms. During the first two years of life, the representation of intentional agents and the development of meta-representation operate. In the second year of life, ToMM arises, and later in childhood, advances in ToM are regulated by inhibitory control to manage the executive functions (EF) of ToM tasks. EFs are cognitive processes that help regulate behavior to achieve goals, particularly in novel or complex situations (Alvarez & Emory, 2006). This inhibitory control becomes more refined during the preschool years. The onset of childhood ASD has been linked to deficits in this mechanism (Leslie et al., 2004). The theory posits that the ToMM automatically ascribes beliefs to content it perceives as true. The ToMM is responsible for forming meta-representations of beliefs, and its selection process suppresses the dominant response. However, around the age of three, this process may encounter challenges in accurately assessing false beliefs. This could account for why children at this age have difficulties comprehending the content of another person's false belief, leading to difficulties in false belief tasks. Nonetheless, modularity theories face challenges in explaining how social experiences influence

these findings (de Villiers, 2005).

1.1.2.3 Simulation Theory. The core concept of simulation theory (Gallese & Goldman, 1998; Harris, 1992) is that adults' and children's interpretations of behavior are based on their experience of their own thoughts and feelings. This theory asserts that attributing mental states to others is contingent upon a simulation process. According to simulation theories, the method of reasoning about the minds of others involves adopting their perspective in the relevant context through a simulation process, utilizing one's own mind as a referential model. This simulation process occurs by projecting oneself into another person's perspective and simulating their mental activity. Children's experience is crucial as a way of practicing perspective-taking in order to improve simulation skills and ToM (Harris, 1992). The process of mental simulation is operationalized through mirroring mechanisms, enabling the mimicry of another person's mental state (Gallese & Goldman, 1998). Simulation theories argue that the ability for simulation is innate, present from birth (Harbers et al., 2012). Over time, children learn to modify their mental states during the simulation process. At around one year of age, children can discern others' intentions by observing their actions, which subsequently influences their own behavior. According to the theory, a younger child projects their own present mental state onto the other person, this is called egocentric simulation (Gopnik & Wellman, 1992). As they develop, children learn to simulate the perspectives of others without their behavior being directly influenced by these simulations. This is achieved by attributing the simulated perspective to the interaction partner, thus facilitating interaction. Children gradually acquire the ability to understand the resultant behavior of others, even when their own mental state differs from that of others, effectively setting aside their personal mental state. Later in development, children adopt an 'intentional stance', employing their imaginative capabilities even in the presence of counterfactual elements. This advancement enables them to navigate hypothetical scenarios that diverge from actual reality. Consequently, a failure in false belief tasks is indicative of a deficiency in the simulation process (Harris, 1992). Contemporary philosophical discourse within psychology suggests that the field requires a synthesis of theory-based knowledge (to anticipate actions) and simulation (for belief; Sodian, 2005).

1.1.3 Dissecting ToM

ToM is not a monolithic construct. Different parts of ToM have been proposed,

which were termed affective and cognitive ToM (Maleki et al., 2020; Shamay-Tsoory et al., 2007; Zabihzadeh et al., 2017), decoding and reasoning, or hot and cold aspects of ToM.

Affective ToM involves thinking about emotions and feelings by decoding and discriminating the mental states of others based on available environmental information. Cognitive ToM involves reasoning about thoughts, intentions, or beliefs by making inferences about mental states through interpretation or prediction of others' behaviors (i.e., understanding the intentions of others). Decoding involves spontaneous appraisal of one's environment. It is a rudimentary skill that does not require complex reasoning skills. Social information about the observable environment (e.g., facial expressions) is used to perceive the thoughts of others. Reasoning requires additional information about the context (Harkness et al., 2005). Affective ToM is equivalent to the social-perceptual aspect of ToM, whereas cognitive ToM corresponds to the social-cognitive aspect. This distinction is further clarified by the conceptualization of 'cold' and 'hot' facets of ToM (Brothers & Ring, 1992). 'Cold' ToM is oriented towards more cognitive processing, in contrast to 'hot' ToM, which is associated with emotional processing. Additionally, identifying two aspects of ToM, Njomboro et al. (2008) compared decoding aspects (attribution of emotional states) and reasoning aspects (understanding of intention).

Previous studies have shown that these subcomponents depend on overlapping and distinct neuronal structures. Both ToM facets are associated with the temporoparietal junction. Structural imaging studies in patients with dementia and functional imaging studies in healthy controls (HC) showed that affective ToM engages the amygdala, basal ganglia, inferior frontal gyrus and ventromedial prefrontal cortex, whereas cognitive ToM engages the dorsomedial prefrontal cortex and dorsolateral prefrontal cortex (Healey & Grossman, 2018; Schurz et al., 2021). A meta-analysis by Schurz et al. (2021) demonstrated that some ToM tasks measure cognitive (e.g., strange stories requiring participants to explain the speaker's intention where their utterance was not literally true; Happé, 1994) and affective processes (e.g., facial emotion recognition) concurrently. Tasks engaging both aspects (cognitive and affective ToM) show high clinical discrimination, but interindividual differences could reflect different sources. This has implications for task selection.

1.1.3.1 Error Types. Besides differentiating between affective and cognitive ToM,

ToM deficits can further be divided into different error types: exceeding ToM, less ToM, and no ToM. Exceeding ToM, which is also termed overmentalization (Vegni et al., 2021) or hypermentalization (Sharp & Vanwoerden, 2015), involves over-interpretating others' behaviors by interpreting incidental actions intentionally. A recent meta-analysis showed that hypermentalization is typical for many mental disorders (McLaren et al., 2022). Originally, ToM was often associated to borderline-personality disorder (BPD; Sharp & Vanwoerden, 2015). In contrast, alcohol use disorder (AUD) is typically associated to reduced or no ToM compared to HC (Onuoha et al., 2016).

1.2 Measurements

Despite the significance of ToM for social functioning, development and mental health, ToM and social cognition research are plagued by the existence of different concepts with similar meaning and different measurements for identical concepts (Olderbak and Wilhelm (2020). This phenomenon is called jingle and jangle fallacy by Olderbak and Wilhelm (2020). Jingle fallacies occur when measurements are assumed to assess the same construct but, in fact measure different constructs. Conversely, Jangle fallacies describe measurements which are supposed to assess different constructs but in fact measure the same construct. A plethora of similar constructs and a lack of specificity (e.g., Theory of Mind, perspective taking) lead to substantial differences in how authors measure the concepts (Olderbak & Wilhelm, 2020; Quesque & Rossetti, 2020). This goes hand in hand with ambiguous definitions and assessment methods with low validity. Nonetheless, there is consensus that the representation of another's mental state and its differentiation from one's own are core components of ToM (Quesque & Rossetti, 2020).

Currently many ToM tasks lack ecological validity (Pabst et al., 2022). This, in turn, has forced researchers to question the conclusions drawn from the available evidence. For instance, the 'Reading the Mind in the Eyes Test' (RMET) is commonly used to measure ToM, however, it was unclear whether it actually measures affective ToM or rather lower-level processes such as emotion recognition (Kittel et al., 2022; Quesque & Rossetti, 2020). Therefore, validating and exploring measures is crucial. Since the number of ToM measures would go beyond the scope of this dissertation, the following paragraphs will only describe the RMET, false belief tasks, Faux Pas and Movie for the Assessment of Social Cognition (MASC).

1.2.1 RMET

A very common and popular measure is the RMET (Baron-Cohen et al., 2001). It is often used to measures individual differences in ToM among adults and is frequently used in clinical samples because it is easy to implement. The RMET encompasses 36 pictures followed by one question in a multiple-choice (MC) format. Pictures present the eye region, including the eyebrows and the upper half of the nose, (Baron-Cohen et al., 1997). Participants choose one out of four adjectives (e.g., jealous, playful) to describe what the person in the picture is thinking or feeling. The pictures are balanced between women and men. The aim of the task seeks assesses the ability to perceive and interpret emotional cues.

The RMET has been used to assess mental state decoding by measuring what a person is perceiving when looking at the target's eye gaze. The RMET test does not include all ToM parts, such as inferring the content of the mental state (Baron-Cohen et al., 2001). The RMET has been criticized with respect to psychometrics (Vellante et al., 2013). Some studies argue that the RMET captures lower-level cognitive processes (Kittel et al., 2022; Oakley, 2016). With respect to the two criteria by Quesque and Rossetti (2020), to successfully pass as a ToM test, the RMET does not qualify as such. It should not be feasible by merely inferring another person's mental state based on one's current experience or lower-level processes, such as attention orientation. A meta-analysis described that the internal consistency of the RMET was acceptable, but the construct validity was questionable (Kittel et al., 2022). Performance on the test was highly correlated with emotion perception, and weakly to moderately related with cognitive and affective empathy (Kittel et al., 2022). While many previous studies have used the RMET to assess the affective ToM (Kopera et al., 2020; Rupp et al., 2021), other studies have used it to measure emotion recognition (Frileux et al., 2020; Maurage, Grynberg, Noël, Joassin, Hanak, et al., 2011) or empathy (Chapman et al., 2006).

1.2.2 False Belief tasks

False Belief tasks were the first ToM measures, at their center is the requirement to represent incorrect ideas that another person can hold about reality. In the false belief task modelled after Wimmer and Perner (1983), children are told stories, that require them to infer the false beliefs held by a character in the narrative without being distracted by knowing the correct information. One popular story is about

Maxi, a character who puts a chocolate bar in one location (A). In his absence, the chocolate bar is unexpectedly transferred by another character (Mother) from one location (A) to another (B). The second character leaves the room and Maxi returns. Subsequently, the child is asked where the character Maxi will look for the chocolate bar. Children of 3-years or younger often incorrectly answer that Maxi would look at location B (where the chocolate bar is). These young children predict Maxi's behavior based on their own knowledge. Most 4-years-olds answer correctly by suggesting that Maxi would look at location A (where he thinks the chocolate bar is). Children around the age of 3.5 to 4 years predict actions based on Maxi's false belief. This age-related improvement was supported in a meta-analysis (Wellman et al., 2001). A similar false belief task, developed for young children, involves a story about two characters, Sally and Andrew (Wimmer & Perner, 1983). Sally places her toy in the cupboard and then leaves. In her absence, Andrew moves the toy from the cupboard to the chest of drawers. After Andrew exits, Sally returns. The child is then asked where Sally will look for her toy. This variant of the false belief task is commonly referred to as a change-of-location task.

Representing false beliefs requires comprehending two distinct models concurrently and discerning the inconsistencies between propositions in one model versus those in the other (Wimmer & Perner, 1983). Another type of false belief paradigm is the unexpected-identity-false belief task. This task is less based on children's comprehension of the essential events of the story. The children are asked about the content of a familiar box, such as a tube filled with a particular brand of candy. Afterwards, the experimenter opens the box and exposes atypical content, like a pencil. The child is asked what someone who has not seen inside the box would think was in there. Most 3-year-old children explained that the other child would think there was a pencil in the box. Many 4-year-olds reported 'candy', reflecting that another child would have a false belief about the contents of the box (Perner et al., 1987). A significant improvement in children's abilities arise from 3-4 years to 4-5 years (Flavell et al., 1983). This means that younger children have difficulties differentiating between an object's appearance and its actual reality. Older children develop the capacity to understand the relation between appearance and reality. Another related concept is representational change. It requires the ability to represent past representations of the world and contrast them with representations, thus

understanding the process of change. Both understandings are necessary to understand false belief tasks, the appearance-reality distinction, and representational change (Gopnik & Astington, 1988).

Moreover, the third year of life constitutes a critical stage in children's socio-cognitive development, characterized by a shift in language ability and developments from an implicit understanding to a more explicit reasoning about mental states (Kaltefleiter et al., 2021). Language plays an essential role in the development of false belief representations and thus in ToM. Different false belief tasks impose different linguistic demands. The change-of-location task and the unexpected-identity task, involving a narrative, are both verbally demanding (Milligan et al., 2007). The language development of a 3-year-old child predicted their ToM in early adolescence (Ebert, 2020). Taking the perspective of another person requires first-order ToM. For example, a first-order question is 'Does Mom think Mary saw the puppy' (Longobardi et al., 2019). In contrast, taking the perspective of a character who is taking the perspective of another character requires application of second-order ToM by attributing a mental state to someone who is attributing another mental state. For example, a second-order question is 'What does Mom think Mary will tell her friends she's getting for her birthday?' (Longobardi et al., 2019). This is illustrated in Figure 2. Children begin to pass first-order ToM around the age of four. Different age-related performances have been reported regarding first- or second-order ToM. While 4-year-olds can pass first-order tasks. 5- to 7-year-old children can pass second-order ToM tasks (Perner & Wimmer, 1985).

Figure 2

Modelled after Felisberti and King (2017)

First-order ToM Second-order ToM

1.2.3 Faux Pas

The Faux Pas test presents brief stories about a character who commits a social faux pas. Participants indicate whether the story contains a faux pas and if so, explain what was inappropriate and why (Stone et al., 1998). One example is

mistaking a woman's sister for her daughter. The speaker did not say this out of any bad intention, but due to a misbelief. Originally, the Faux Pas task contained 10 vignettes; later, 10 vignettes without a faux pas were added to control for general response bias. This was followed by six questions about each story and two control items (Baron-Cohen et al., 1999b). The task requires substantial reading comprehension. This task requires knowledge about representations of mental states, perspective taking, and social norms, particularly recognition of transgressions of social norms (Osterhaus et al., 2016). It assesses ToM abilities at a higher level than first-or second-order false belief tasks (Brüne, 2005). The Faux Pas is one of the most frequently used ToM tests (Osterhaus & Bosacki, 2022). In this recent review items from many ToM measures were analyzed to identify factors underlying ToM: social reasoning, reasoning about ambiguity and transgression of social norms. Findings of this recent review found that the Faux Pas constitutes a measure of ToM that focuses on transgressions of social norms.

Performance on the false belief task requires a cognitive understanding of the difference between the speaker's knowledge and that of the listener. Faux Pas tasks require in addition an appreciation of the listener's emotional state (Baron-Cohen et al., 1999b), which is now understood as affective ToM. Firstly, comprehending that the speaker is unaware of the inappropriateness of their comment, and that the listener will likely feel insulted, involves both cognitive and affective components (Stone et al., 1998). As mentioned above, a ToM task should fulfill two criteria: being able to represent another mental state's 'mentalizing criterion' and distinguishing between one's own and the other's mental state 'nonmerging criterion' (Quesque & Rossetti, 2020). Both criteria are fulfilled by the Faux Pas (Baron-Cohen et al., 1999b).

Between the ages of 9 and 11 years, children develop the ability to understand and recognize Faux Pas (Stone et al., 1998). There is evidence that girls can pass the Faux Pas task by age 9, boys by age 11. Moreover, children with Asperger syndrome or high-functioning ASD are impaired in the Faux Pas test (Baron-Cohen et al., 1999b). Performance gains in the Faux Pas seem to be due to children's greater ease of applying their AToM knowledge. AToM performance correlates with the age at which children understand higher-order false beliefs and is influenced by information-processing skills such as language and intelligence (Osterhaus & Koerber, 2021). Older adults (62-77 years) showed worse performance than young adults (19-25

years) on the Faux Pas test, particularly in Faux Pas understanding. Most older adults recognized the Faux Pas but could not understand it (Wang & Su, 2006). Thus, age impacts ToM measurements.

1.2.4 MASC

A prominent ToM measure that fulfills the ToM requirements that individuals need to recognize a mental state and distinguish their own mental state from that of the target (Quesque & Rossetti, 2020) is the Movie for the Assessment of Social Cognition (MASC; Dziobek et al., 2006). This task is an valid and reliable test (Fossati et al., 2018), which shows high reliability, good internal consistency (Cronbach's alpha $\alpha \geq .82$), and high ecological validity (Benito-Ruiz et al., 2022). This test relies on participants' responses to social interactions in multiple video clips. It is a computerized test for measuring affective and cognitive ToM that approximates the demands of everyday life. A 15-minute video about friendship and dating where four people come together for a dinner party, is presented, and repeatedly paused to ask participants to infer the thoughts or emotions of the characters using MC questions. Each question has four possible answers, with one correct and three incorrect options that represent three types of errors: exceeding ToM, less ToM, and no ToM. The MASC offers a total score, an affective ToM score, and a cognitive ToM score. However, the test includes 15 affective and 18 cognitive ToM items (Montag et al., 2010). The remaining 12 items are not assigned to either the affective or cognitive facet. Participants are also required to answer six MC control questions about non-social/non-mental-state-related situations. These questions serve as controls for variables such as memory and comprehension of the task. A broad range of social cognition concepts such as Faux Pas, first- and second-order false beliefs or sarcasm are included in the MASC. Even subtle difficulties in social understanding are visible, which closely approximates everyday life. The characteristics of the characters are quite different from each other. Questions in the video clips refer to three different mental states: emotions (what does X feel?), thoughts (what does Y think?), and intentions (why does Z say this?). Items referring to emotions vary in valence. The conversational content was designed to capture both verbal (19 items) and non-verbal communication, including facial expressions, gestures, and body language (16 items).

The MASC requires the participant to infer the mental states (affective and cognitive) that drive a character's actions. This ToM measurement was designed to

be sensitive in detecting ToM difficulties in individuals diagnosed with ASD and attempts to minimize demands on EFs by avoiding distraction such as music or quick changes of scenes (Dziobek et al., 2006). All characters in the MASC are around the same age (mid-thirties; Dziobek et al., 2006). Several clinical studies have used the MASC to assess ToM deficits, for instance among patients with BPD (Preißler et al., 2010; Ritter et al., 2011), major depressive disorders (MDD; Wolkenstein et al., 2011), and AUD (Maurage et al., 2016).

1.3 Theory of Mind and influencing factors

1.3.1 Age

1.3.1.1 Developmental aspects of Theory of Mind. Several studies have shown that the ability to understand another person's reactions and beliefs emerges gradually. In the first year of life, human infants learn a few social-cognitive abilities, such as following adults' gazes, engaging in joint attention (Carpenter et al., 1998) and understanding adults' intentional behavior behind their gestures (Behne et al., 2005). By the age of two or three, children have visual perspective-taking skills including the understanding that other's visual perspective may sometimes differ from their own (Moll & Tomasello, 2006). Moreover, around preschool age (3-5 years) children develop facial expression recognition. In addition, around the age of four or five, children develop an understanding of first-order beliefs (MacDonald et al., 1996). By the age of five or six, children connect emotions such as surprise to a lack of knowledge and false beliefs. The ability to understand second-order false beliefs develops around 6 to 7 years (Perner & Wimmer, 1985; Wimmer & Perner, 1983). Children's emotional understanding develops continuously between the ages of 3 and 11 years. From the age of 7 onwards, children distinguish between expressed and really felt emotions. Research suggests that children's understanding is arranged hierarchically (Pons et al., 2004). A previous study demonstrated that the AToM of a 6-years-old (understanding first- and second-order false beliefs) predicted their AToM at age 10; thus, AToM remained stable over time (Devine et al., 2016). More recently, Osterhaus and Koerber (2021) examined AToM in a longitudinal study, an provided evidence that individual differences in AToM develops nonlinearly. The ability reaches a critical point when children turn seven: at this time, they gain the conceptual understanding that mental states can be recursive. Moreover, this conceptual understanding of ToM has been shown to influence the development of reasoning

skills in children at the early elementary school level. This supports the hypothesis that conceptual insight is integrated into the development of AToM (Osterhaus & Koerber, 2021). Happé (1994) conducted a longitudinal study assessing understanding of the recursive nature of mental states, including the double-bluff task, also known as the strange stories task. The task involved telling children a story about a soldier who was captured by his enemies and asked to reveal the location of his weapons. In this story, the soldier reveals the genuine location. Furthermore, in late middle childhood, around the age of nine, developmental progressions in reasoning about ambiguity and social reasoning could still be observed, with improvements in ATOM visible even during elementary school (Osterhaus & Koerber, 2021).

Wellman (2020) postulates five important developmental milestones of ToM during preschool, elementary school and early adolescence. Firstly, children learn to recognize that people have different desires. Secondly, they come to understand that individuals can hold varying beliefs about the same situation. Thirdly, they develop an awareness of the distinction between knowledge and ignorance, acknowledging that people may possess different levels of information about a situation and that some may be unaware of what others know. The fourth milestone involves understanding false beliefs, where children grasp that people can maintain beliefs that are not aligned with reality. Finally, children learn that internal states such as desires, knowledge, ignorance, and thoughts can be hidden and not always visible in a person's expressions, which involves understanding that individuals may not always reveal their true beliefs or intentions. These milestones are indicative of the development of ToM (Wellman, 2020).

It is noteworthy that numerous tasks designed to measure various cognitive processes indicate a common egocentric bias in behavior. Egocentrism is assessed in a situation that involves a conflict between other’s perspective and one’s own. In these situations the child shows a systematic bias towards their own point of view (Freidson , 1958). In a study by Surtees and Apperly (2012), children aged –6-10 years were given a task that involved making inferences about a cartoon character’s visual perspective. The results showed that all age groups experienced egocentric effects, which resulted in slower responses when their perspectives differed from that of the avatar. Older children were able to make faster judgments about the avatar’s perspective and made fewer errors compared to younger children. Similarly, another

study tested children's capacity to predict a character's behavior based on their desires and beliefs, showing that older children were more accurate and faster in their judgments (Apperly et al., 2011). Moreover, evidence for egocentrism was found in all age groups (Surtees & Apperly, 2012).

1.3.1.2 Theory of Mind over the life span. Most studies have been dedicated to ToM in early and middle childhood, while only a few studies have assessed ToM in adolescents and young adults aged 11-25 years. Overall, the development of emotion recognition assessed with the RMET test, improves independently of development in reasoning, inhibition, and language, indicating a specific development. Additional, age-related improvement in social-cognitive ToM, assessed with the Faux Pas were reported (Meinhardt-Injac et al., 2020). They align with improvements in reasoning, inhibition, and language processing.

A meta-analytic review demonstrated that ToM deficits increased in late adulthood across all task types (Stories, Eyes, Videos, false belief, and Faux Pas; Henry et al., 2013). This decrease in ToM performance among healthy older adults occurred across both verbal and visual tasks (Henry et al., 2013). Poor ToM was associated with a greater number of mistakes in a social communication task among a group of healthy adults (Krych-Appelbaum et al., 2007). Moreover, a study compared younger (19-25 years) and older (65-87 years) adults (Bailey et al., 2008) and found that impairments in ToM can significantly impact social interactions in older populations: a reduction in cognitive empathy was found to mediate a decline in social participation, which correlates with social isolation among older adults. Consequently, numerous studies have highlighted the importance of ToM in older adults, pointing to its role as a potential obstacle to social interactions in daily life.

Some studies suggest that the effects of aging may be confined to cognitive ToM, with no changes observed in affective ToM (Wang & Su, 2013). Bottiroli et al. (2016) assessed ToM with the Faux Pas and found that younger adults (aged 19 to 27) performed better than older adults (aged 60 to 82) in cognitive ToM but not affective ToM. The study by Wang and Su (2013) linked the decline in cognitive ToM performance among older adults to a reduction in EF. Another study demonstrated that the decline in the cognitive aspect of ToM performance could be due to a strategic shift in the allocation of limited cognitive resources (Hess, 2014). Other studies found a decline in both ToM facets. For instance, the comparison of healthy younger (21-34

years), middle-aged (45-59 years) and older adults (≥ 70 years); showed higher cognitive and affective ToM performance among the younger vs. the older adults (≥ 70 years; Duval et al., 2011). This is in line with a meta-analysis that found reliable effects of aging on both affective and cognitive ToM (Henry et al., 2013). Using the same task to measure affective and cognitive ToM, previous studies have shown that older individuals (aged 64 to 87 years) performed more poorly than young individuals (aged 17 to 27 years) in both cognitive and affective ToM (Baksh et al., 2018; Fischer et al., 2017).

A recent study demonstrated how aging affects cognitive and affective ToM in adults, considering the type of task (verbal or nonverbal) and controlling for sex and overall cognitive functioning (Raimo et al., 2022). The results of this study indicated that affective ToM declines earlier in adulthood (starting at age 60), when assessed with the nonverbal modality rather than the verbal modality. However, cognitive ToM decreased starting at age 70, regardless of the task's modality (verbal and nonverbal). Thus, age is a strong predictor of ToM performance (Raimo et al., 2022).

1.3.2 Sex

Sex differences in ToM are still being discussed and the results are inconsistent, however, mostly small. A previous study with a healthy sample showed that women and men showed differences in social cognition (Gur et al., 2010). Women have been found to be more precise and better at recognizing emotions and expressing themselves (Kret & De Gelder, 2012; Mestre et al., 2009). In particular, several studies have reported higher scores for women in recognizing facial expressions (Campbell et al., 2002; Kohler et al., 2010; McClure, 2000; Montagne et al., 2005), along with more pronounced affective responses (Han et al., 2008). Notably, women classified happy and sad faces faster than men did (Rahman et al., 2004). There is also substantial evidence that women show higher performance on tasks measuring affective ToM compared to men (Krach et al., 2009; Xia et al., 2012). However, a recent study by Pang et al. (2023) observed that while women exhibited higher scores on a subjective measure of empathic ability, electroencephalography (EEG) analyses did not discern any sex differences in neural empathic responses, implying potential modulation by social desirability biases.

Highlighting the inconsistency in sex differences in ToM performance, one study reported superior ToM performance in healthy men than women (Russell et al.,

2007). Other studies have shown no differences between women and men in the assessment of ToM performance (Barrett et al., 1998; Derntl et al., 2010). Recent studies have investigated both affective and cognitive ToM and revealed no significant sex differences (Di Tella, Miti, et al., 2020; Navarra-Ventura et al., 2018). Interestingly, there was one exception. Women scored higher in recognizing anger on male faces (Di Tella, Miti, et al., 2020).

Several possible explanations of women's advantage in tasks involving affective ToM have been proposed. One possible explanation is that different neural regions are involved in processing emotional information (Christov-Moore et al., 2014). Another explanation might be that there are differences in social information processing strategies (Whittle et al., 2011). In this approach, men exhibited stronger systemizing strategies, whereas women exhibited empathizing strategies (Baron-Cohen et al., 2005). These approaches might lead to the female advantage on empathic processing (Di Tella, Miti, et al., 2020).

Assessing social cognition abilities not only in healthy samples but also in clinical samples provides necessary information on sex differences concerning a broader range. Two meta-analyses of individuals with AUD yielded contradictory results. In one meta-analysis the difference between patients with AUD and HC was smaller in studies with more men in the AUD sample, suggesting stronger impairments in women (Onuoha et al., 2016). In the second meta-analysis a higher ratio of men was associated with stronger ToM impairments, suggesting stronger impairments in men (Bora & Zorlu, 2017). Other clinical studies have revealed additional results on this topic. Evidence suggests that individuals with BPD are characterized by impairments in social cognition, specifically ToM (Bora, 2021). In contrast to the sample of individuals with AUD, Németh et al. (2018) did not find an effect of sex ratio on ToM among BPD patients. However, since men are typically underrepresented in BPD groups (Skodol & Bender, 2003), the effects of sex must be interpreted with caution. Considering these uncertain results, future studies should explore sex differences in clinical samples.

1.4 Mental disorder and Theory of Mind deficits

A dimensional framework for mental health, the Research Domain Criteria (RDoC), characterizes mental disorders based on various components and units of analysis. Within this framework, potential deficits in these components may lead to

mental disorders and/or functional impairments. The RDoC framework conceptualizes psychopathology as dysfunction within specific systems, each further divided into essential functional constructs and their corresponding subconstructs (Patrick & Hajcak, 2016). ToM is described in one of the six domains (Systems for Social Processes) of the matrix as a dimension describing the differences between mental disorders. RDoC (2020) suggested ToM as a transdiagnostic factor by which mental health should be described (Carcone & Ruocco, 2017; National Institute of Mental Health, 2020).

1.4.1 Mental disorders, ToM deficits and long-term treatment outcomes

Mental illness represents a major public health concern due to its high comorbidity with physical health issues, increased mortality and morbidity rates, and the socioeconomic impact stemming from a heightened risk of unemployment and disability claims. Approximately one in five individuals have experienced a mental illness in the past 12 months, and a third of the general population will be affected by mental disorders at some point in their lives (Steel et al., 2014). Epidemiological studies in Europe (Wittchen et al., 2011) reported a combined lifetime prevalence of anxiety, mood, externalizing, and substance use disorders (SUD) of between 18.1–36.1%. For instance, depression, the most prevalent mental illness, continues to pose a significant public health challenge (Chapman & Perry, 2008). Furthermore, mental disorders are often associated with a higher likelihood of being diagnosed with a physical disorder. The high rate of comorbidity between mental and physical illnesses, particularly metabolic and cardiovascular diseases, presents even and additional challenges for public health (Barnett et al., 2012; Firth et al., 2019).

Various mental and behavioral disorders are associated with dysfunctional ToM (McLaren et al., 2022). A recent review of meta-analyses examining social cognitive processes in over 30 different clinical conditions suggested that these processes are transdiagnostic clinical markers, indicative of disease progression, and treatment response (Cotter et al., 2018). These clinical conditions span developmental, psychiatric, and neurological disorders. In this review, ToM is defined as the mental operations underlying social interactions (Cotter et al., 2018). Based on this review, the most important psychiatric disorders associated with ToM impairments are discussed below, including individuals with AUD, PD, schizophrenia, MDD, bipolar disorder (BD) and eating disorders (ED).

1.4.1.1 ToM and alcohol use disorder. AUD is a mental health condition characterized by harmful drinking patterns that lead to adverse physical, emotional, and psychosocial outcomes (Tawa et al., 2016). Individuals with AUD exhibit a lack of control over their alcohol consumption, often engaging in excessive and progressively harmful drinking behaviors, despite severe negative consequences to their physical and mental health, as well as to the well-being of their families. Despite its considerable impact on public health, AUD remains one of the most under-addressed psychiatric disorders (Rehm et al., 2009). AUD is defined by the International Classification of Disease (ICD) via the following criteria: individuals display persistent alcohol use despite experiencing negative biological, psychological, behavioral, and social effects. To receive a diagnosis, individuals must experience a minimum number of these effects within a 12-month period, with the emphasis placed on heavy drinking as the central criterion (Rehm et al., 2013).

The conceptualization of AUD as a pattern of heavy drinking over an extended period aligns with neurobiological alterations in the brain that result from alcohol consumption. These changes were partially reversible following a period of sustained abstinence (Schulte et al., 2014). Alcohol consumption significantly affects the transmission of γ-aminobutyric acid (GABA) and glutamate in the brain, enhancing GABA-A receptor signaling while suppressing N-methyl-D-aspartate (NMDA) receptor signaling (Koob, 2014). SUD are complex diseases that progress through various stages, causing disturbances in three key neurocircuits. These circuits include the basal ganglia-driven binge/intoxication stage, extended amygdala-driven withdrawal/negative affect stage, and prefrontal cortex-driven preoccupation/anticipation stage (Lanius, 2014). Each stage of these disorders presents distinct characteristics that can influence an individual's behavior and mental state. The binge/intoxication stage is marked by excessive drug-seeking behavior and reduced reward function, attributed to deficits in dopamine and opioid peptides, and increased activity in the brain's stress systems. EFs are also impaired during this stage due to disruptions in specific neuronal networks within the prefrontal cortex. In the preoccupation/anticipation stage, increased drug cue-induced incentive salience interacts with low reward and high stress system functions, leading to a combination of positive and negative reinforcement processes. These processes drive destructive drug-seeking behaviors and result in cravings (Lanius, 2014). Glutamate may play a

significant role in the craving phase, underscoring the importance of pharmacological interventions in this context.

Furthermore, the most compelling evidence of the effectiveness of psychosocial treatments was found in brief interventions, social skills training, the community reinforcement approach, behavior contracting, and motivational enhancement therapy (Miller & Wilbourne, 2002). Various therapeutic approaches aim to integrate knowledge of pathophysiological and neurocognitive processes to address AUD. Effective training should contain elements that automatically trigger control when necessary (Wiers et al., 2013). One approach focuses on the increased emotional and behavioral responsiveness to alcohol cues observed in individuals with AUD, which may contribute to relapse and increased craving. Emerging evidence suggests that cognitive bias modification (CBM) training may be effective in altering the biased cognitive processing of alcohol cues by linking them with avoidance behaviors. A recent comprehensive review indicated that CBM interventions have a small impact on cognitive bias and relapse rate, but not on reducing substance use. However, further studies are required (Boffo et al., 2019). Given that AUD arises from a complex interaction of genetic, environmental, and neurobiological factors, no singular treatment approach is universally effective.

Additionally, social skills training, another empirically supported treatment method, emphasizes enhancing the client's social support network. Social skills training is known to enhance abilities such as social cognition (Dekker et al., 2021). Within this domain, the ToM framework is considered a crucial component of social cognition skills. Several systematic reviews and meta-analyses have shown that individuals with AUD have impairments in ToM with medium to high effect sizes (Bora & Zorlu, 2017; Hanegraaf et al., 2021; Onuoha et al., 2016; Sanvicente-Vieira et al., 2017). The extent to which both the affective and cognitive components of ToM are impaired in individuals with AUD remains uncertain. Previous research on this topic has primarily focused on affective ToM, with few studies investigating cognitive ToM. Among the studies investigating only affective ToM in individuals with AUD, nearly all reported impairments (e.g. Kopera et al., 2020; Mátyássy et al., 2006; Rupp et al., 2021), except Kornreich et al. (2011). The RMET has been extensively used to assess affective ToM. Nevertheless, concerns exist that the RMET may not effectively assess affective ToM (Kittel et al., 2022; Quesque & Rossetti, 2020). Consequently, research

employing the RMET to measure affective ToM might yield inconsistent outcomes compared to studies using alternative ToM assessments. This disparity in methodologies could potentially affect the overall trends observed in the existing body of literature.

There are inconsistent conclusions regarding cognitive ToM impairments in individuals with AUD. Certain studies have found impairments in both cognitive and affective ToM (Bosco et al., 2014; Buragohain & Bhagabati, 2020; Cox et al., 2018; Schmid et al., 2022). A meta-analysis indicated that individuals with SUD exhibited both significant impairments in recognizing emotions through facial expressions (affective ToM) and inferring the cognitive mental states of others (cognitive ToM; Hanegraaf et al., 2021). Other studies investigating both aspects of ToM suggest that impairments are only in affective ToM, with no significant differences in cognitive ToM between individuals with AUD to HC (Maurage et al., 2016; Nandrino et al., 2014). Previous research has often used ToM measures that do not concurrently assess both its affective and cognitive aspects, and many of these measures lack ecological validity (Pabst et al., 2022). Therefore, there's a need for studies that simultaneously investigate both affective and cognitive ToM using a highly valid measure to better understand potential impairments in both components. Additionally, given the significant impact of social cognition on the severe interpersonal dysfunction of SUD, it is essential to conduct future research that utilizes the RDoC's 'Systems for Social Processes' domain to reach a better understanding of affective and cognitive ToM (Hanegraaf et al., 2021).

1.4.1.2 ToM and personality disorder. Personality disorders (PD) are characterized by persistent and maladaptive patterns of behavior, cognition, and inner experiences that deviate from the expectations of the sociocultural environment. These patterns are inflexible and can result in subjective distress and functional impairments (American Psychiatric Association, 2013). Individuals diagnosed with PD may encounter difficulties with cognition, emotional responses, social interactions, or impulse control. PD is distinct from personality traits in terms of severity, inflexibility, and significance of consequences. Symptoms of PD differ from personality changes in that they are typically present since childhood or adolescence. However, PD can only be officially diagnosed in individuals who are 18 years or older. The Emotional Instability PD, coded as F60.3, is characterized by a tendency to act impulsively

without regard for consequences, which may manifest as violent behavior. Mood can be unstable, and PD can be characterized by outbursts of anger or violence. This disorder has two specific subtypes: impulsive and borderline. BPD is distinguished by emotional instability, lack of impulse control, uncertainty about self-image, goals, or inner preferences; and a chronic feeling of inner emptiness that frequently leads to self-harming and suicidal behaviors (Lieb et al., 2004). Affected individuals also tend to have intense yet unstable interpersonal relationships, while simultaneously experiencing a profound fear of abandonment (Bhar et al., 2008; World-Health-Organization, 1992).

The causes of PD are complex and multifactorial and include psychosocial and biological factors. Psychosocial factors include traumatic life events, physical, sexual, or psychological abuse, neglect, interpersonal stressors, experiencing a parenting style ranging from overprotective to hostile, and family history of anxiety, depression, and suicidal behavior. The most frequent adverse event during childhood is childhood sexual abuse, which is reported by a majority of patients diagnosed with BPD (Paris et al., 1994). Lack of social support is also considered as a psychosocial risk factor for the development of PD (Renneberg & Herpertz, 2020). Beck et al. (2004) proposed the cognitive theory of BPD, which suggests that individuals with BPD exhibit maladaptive cognitive, emotional, and behavioural responses to situations owing to cognitive errors and misinterpretations. They tend to evaluate their experiences and perceptions selectively and inaccurately. The neurobehavioral model proposed by Linehan (1993) identifies three main factors that contribute to the development of personality disorders, including BPD: emotional vulnerability, an invalidating environment, and traumatic childhood experiences.

Additionally, there are specific neurobiological correlates associated with certain subtypes, such as prefrontal dysfunction, in individuals with BPD. The neurobiological correlates of BPD characteristics such as dysregulation and impulsivity are poorly understood. Neuroimaging studies have shown a dysfunctional frontolimbic network that mediates important aspects of BPD symptomatology (Soloff et al., 2003). These brain regions appear to be associated with dysfunctional serotonergic neurotransmission (Soloff et al., 2000). Dysfunctional serotonergic neurotransmission is associated with disinhibited impulsive aggression in patients with BPD. Studies have shown that the amygdala, crucial for emotional regulation, exhibits hyperactivity in

BPD patients compared to matched controls (Herpertz et al., 2001; Schmahl et al., 2003). A systematic and integrative review pointed to an expanded function of the amygdala by considering its role in evaluating emotional stimuli in the context of an individual's goals and motivations (Ruocco & Carcone, 2016).

Research has concentrated only on a few PD, particularly BPD. Until the 1990s, BPD was historically viewed as a chronic and challenging condition to treat (Bateman & Fonagy, 1999). However, the development of new approaches, including pharmacotherapy and psychotherapy, has led to increased optimism. Despite this progress, more research is needed on the use of pharmacotherapy in BPD, as current studies have not provided strong evidence for its effectiveness (Gartlehner et al., 2021; Stoffers & Lieb, 2014). Second-generation antipsychotics (SGAs), antiepileptics, and antidepressants were found to be ineffective in reducing the severity of BPD. However, SGAs showed some positive effects on comorbid symptoms. The second method is psychosocial treatment, which includes psychotherapeutic interventions. This approach was recommended in a meta-analysis as the primary treatment for BPD (Stoffers-Winterling et al., 2022). Beneficial effects have been reported for dialectical behavior therapy (DBT) and skills training (Linehan, 1993), with moderate evidence. An example of one psychotherapeutic group intervention is metacognitive training (MCT) for BPD (Allen & Fonagy, 2006). A randomized controlled trial (RCT) conducted to investigate this novel approach found that, compared to progressive muscle relaxation (PMR), MCT was more effective in reducing BPD symptoms after a six-month follow-up. Conversely, PMR was shown to be more effective in alleviating depression symptoms at the same follow-up period (Schilling et al., 2018; Vogt & Norman, 2019). MCT includes interventions targeting ToM skills, which are considered critical in addressing the disturbances in interpersonal relationships observed in patients with BPD (Frick et al., 2012), thus making them a significant focus of treatment.

There is ample evidence of impairments in ToM in patients with BPD (Bora, 2021; Hanegraaf et al., 2021; Németh et al., 2018; Richman & Unoka, 2015). As mentioned above, ToM impairments in individuals with BPD are typically marked by exceeding ToM (Sharp & Vanwoerden, 2015). A recent meta-analysis revealed that exceeding ToM is found in many mental disorders such as schizophrenia, persistent somatoform pain disorder and ASD. A recent study comparing ToM performance in

patients with BPD and PD showed that exceeding ToM was associated with the severity of personality psychopathology and symptom distress not only in patients with BPD, but also in patients with PD (Burghardt et al., 2023; Normann-Eide et al., 2020). Comparing BPD and other disorders, a meta-analysis suggested that the association between psychopathology and exceeding ToM did not differ by disorder (McLaren et al., 2022). This is in line with the ToM framework as a transdiagnostic factor (National Institute of Mental Health, 2020).

Individuals with BPD suffer not only from ToM impairments, but also from comorbid symptoms of depression and anxiety (Quenneville et al., 2020; Richman & Unoka, 2015; Zabihzadeh et al., 2017). However, findings on comorbid symptoms of depression are contradictory. A meta-analysis revealed that individuals with BPD and comorbid MDD exhibited better ToM abilities than those with only BPD (Richman & Unoka, 2015). In contrast, another study concluded that individuals with both BPD and MDD demonstrated poorer ToM than those with only BPD (Zabihzadeh et al., 2017).

A meta-analysis by Németh et al. (2018) found diverse ToM facets impairments in patients with BPD, with a high heterogeneity of effect sizes. Moderator analysis revealed that variability in effect sizes between studies was influenced by specific ToM facets. Surprisingly, in contrast to the previous meta-analysis by Richman and Unoka (2015), there was no deficit in mental state decoding, which was measured with the RMET in all analyzed studies. However, patients with BPD did significantly differ from HC in their mental state reasoning abilities, with medium to high effect sizes (Németh et al., 2018), which were measured with the Faux Pas task (Baron-Cohen et al., 1999a) or the MASC (Dziobek et al., 2006). Although previous studies have suggested a decoupling of mental state decoding and reasoning abilities, as well as of affective and cognitive ToM (Baez et al., 2015; Zabihzadeh et al., 2017), the meta-analysis by Németh et al. (2018) showed that patients with BPD significantly underperformed HC in cognitive ToM as well as affective ToM when the RMET was removed from the data on affective ToM. This suggests that the relatively intact affective ToM abilities found in BPD patients in previous studies (Fertuck et al., 2009; Frick et al., 2012; Zabihzadeh et al., 2017) could primarily be due to their emotion recognition abilities measured with the RMET. Furthermore, this meta-analysis did not find a significant difference between affective and cognitive ToM in BPD patients (Németh et al., 2018). This is in line with a study by Hillmann et al. (2021), which found that patients with BPD scored

worse on affective and cognitive ToM compared to HC when assessing ToM with a false belief cartoon task and with the faux pas (Petersen et al., 2016). More studies are needed to examine affective and cognitive ToM facets in patients with PD.

1.4.1.3 ToM and schizophrenia. Schizophrenia is a chronic psychiatric disorder characterized by disruptions in thinking and perception of reality. It can be described as a syndrome that affects an individual's emotions, thoughts, and actions. The primary criteria for diagnosing schizophrenia include the existence of positive symptoms such as hallucinations, delusions, incoherent speech, disorganized or unusual behavior, and negative symptoms such as reduced emotional expression (DSM-5 American Psychiatric Association, 2013). A meta-analysis reported deficits in ToM in individuals with schizophrenia (Bora & Pantelis, 2013; Chung et al., 2014; Fett et al., 2011; Savla et al., 2013), with a large effect size from the onset of the illness (Bora, Yucel, et al., 2009; Bora & Pantelis, 2013). Bora et al. (2009) used several ToM tasks, including false belief stories and RMET. Furthermore, results indicate that schizophrenia patients showed impaired cognitive and affective ToM compared to HC (Li et al., 2017).

These deficits have been found to be highly predictive of impaired social functioning in this population, as reported by Green et al. (2016), or social and working functioning, as reported by Fett et al. (2011). Patients with schizophrenia show impairments in effectively monitoring their own and other persons' mental states. The findings indicate that this clinical group exhibits impairments in verbal and visual mentalizing tasks, with a tendency towards greater impairments in verbal mentalizing tasks compared to visual mentalizing tasks (Chung et al., 2014). Moderating variables, such as sex, age and IQ, did not affect the mean effect size (Sprong et al., 2007).

Different symptoms corresponded with different types of ToM errors. Negative symptoms may be linked to a ToM deficit, characterized by a limited capacity to interpret behavior based on mental states. The positive symptoms of schizophrenia were associated with a tendency to overinterpret mental states (Montag et al., 2012; Peyroux et al., 2019). Evidence of abnormalities in individuals with schizophrenia has been gathered using imaging techniques. Besides the volumetric abnormalities observed in the prefrontal and temporal areas of the brain, associated with impaired ToM (Koelkebeck et al., 2013), studies employing various functional imaging methods have consistently reported atypical patterns of neural activation in regions critical for

ToM. These patterns exhibit both excessive and insufficient activation, as detailed in a study by Marjoram et al. (2006).

A comprehensive review revealed evidence of significant impairments in individuals with remitted schizophrenia and their unaffected relatives. The relatives of patients exhibited a less severe deficit, with performance levels between those of HC and individuals diagnosed with schizophrenia. Furthermore, impaired ToM has also been reported for other types of psychotic disorders (Bora, Yücel, et al., 2009) and in individuals with first-episode psychosis and schizophrenia (Cotter et al., 2018). Moreover, diminished ToM is associated with a heightened probability of developing a psychotic disorder. Thus, ToM impairments have a predictive value in forecasting the onset of psychosis (Boldrini et al., 2020). Existing evidence has proposed that ToM could play a substantial moderating role not only in the risk of developing a psychotic disorder (Debbané et al., 2016), but also in the subsequent recovery process following the initial episode of psychosis (Braehler & Schwannauer, 2012).

1.4.1.4 ToM and major depression disorder. MDD is the predominant mood disorder and is characterized by abnormalities in affect, mood, neurovegetative functions and cognition. Affected individuals display depressive symptoms, such as persistent low mood and/or reduced capacity to experience pleasure and interest. Additionally, MDD is closely linked to considerable distress and impairment of functional domains (Fava & Kendler, 2000).

Furthermore, meta-analyses have revealed that ToM is impaired during acute episodes of depression (Bora & Berk, 2016; Nestor et al., 2022; Wolkenstein et al., 2011). The meta-analyses by Bora and Berk (2016) and Nestor et al. (2022) reported that individuals with depression exhibit mild-to-moderate impairments in ToM abilities throughout their lifespan. This contrasts with the more severe ToM deficits reported in the meta-analysis by Richman and Unoka (2015), which primarily included studies that employed the RMET to measure ToM.

Overall, more severe depressive symptoms are correlated with greater impairments in ToM. This is evident across various ToM tasks that investigate both cognitive and affective components, encompassing verbal, visual, reasoning, and decoding aspects (Bora & Berk, 2016). However, no significant differences were found in effect sizes across these affective and cognitive facets or between decoding and reasoning processes (Nestor et al., 2022). Moreover, while associations between ToM

and depression were evident in clinical samples, such associations were not significant in community samples.

The meta-analysis by Nestor et al. (2022) also explored sociodemographic factors. Findings reported that neither age nor sex significantly influenced the relationship between ToM and depression, with both factors considered as continuous moderators. This result aligns with the findings of an earlier meta-analysis by Bora and Berk (2016).

The impairment of ToM in individuals with MDD can be attributed to a variety of contributing factors. Depression is frequently marked by a tendency towards negative self-absorption or self-focused attention (Mor & Winquist, 2002). This inclination towards internal introspection and negative rumination regarding one's own thoughts and feelings could potentially reduce or impair empathy and involvement in comprehending the mental and emotional states of others (Nestor et al., 2022). The generally negative perspective of individuals with MDD can lead to erroneous interpretations and identifications of others' cognitive and affective states. Another explanation is that the impaired ToM abilities reported in individuals with depression are related to neural irregularities in the brain substrates associated with ToM networks. Notably, the prefrontal lobe has been identified as a critical component for successful ToM performance (Stuss et al., 2001) and has been shown to be impaired in patients with depression (Bora et al., 2012).

1.4.1.5 ToM and bipolar disorder. BD is a chronic psychiatric condition characterized by repeated episodes of mood disturbance, including depressive symptoms, and periods of relief from mood-related symptoms (Vieta et al., 2018). This disorder encompasses bipolar I disorder, characterized by the occurrence of a manic episode along with a major depressive episode, and bipolar II disorder, characterized by the occurrence of a hypomanic episode along with a major depressive episode (McIntyre et al., 2020).

Two meta-analyses revealed an association between BD and deficits in ToM performance, highlighting impaired ToM abilities in individuals with BD (Bora et al., 2016; Gillissie et al., 2022). Bora et al. (2016) examined ToM in acute, euthymic and subsyndromal BD patients and compared them to HC. Collectively, all clinical groups reported a medium-sized ToM deficit. ToM deficits have been reported in individuals with acute bipolar diseases (Bora et al., 2016; Gillissie et al., 2022; Kerr et al., 2003)

and in those in remission periods (Bora et al., 2016; Szmulewicz et al., 2019). However, ToM dysfunction was significantly more severe during acute episodes compare to periods of remission (Bora et al., 2016). The effect size was larger when focusing solely on manic patients, suggesting a smaller effect size during depressive phases. Consistent with Bora et al. (2016), moderate effect sizes in ToM impairments were reported (de Siqueira Rotenberg et al., 2020; Samamé et al., 2012). Additionally, a recent meta-analysis demonstrated moderate-to-large effect size for ToM impairments in individuals with BD (Gillissie et al., 2022). Further, a meta-analysis and a recent study showed that both patients with BD and their first-degree relatives had significant ToM deficits (Santos et al., 2017), with a small effect size (Bora & Özerdem, 2017; Santos et al., 2017). The presence of ToM impairments in first-degree relatives supports the hypothesis that these impairments may serve as a trait marker for cognitive impairments in BD (Santos et al., 2017).

Other studies have indicated that impairments in individuals with BD are not limited to ToM; instead, EFs, attention, and verbal learning abilities are also affected, although intelligence levels appear to be unaffected (Kurtz & Gerraty, 2009). Impairments in EFs such as attention may contribute to the evidence of ToM deficits (Bora et al., 2005). Moreover, a significant percentage of individuals diagnosed with BD exhibit inadequate psychosocial functioning (Burdick et al., 2010) and poorer global functioning than controls (Watson et al., 2023). A recent study has demonstrated that social functioning impairment can be partially attributed to deficits in social cognitive abilities in individuals with BD (Popolo et al., 2020).

With regard to verbal (e.g., tasks assessing the ability to infer intentions behind indirect speech assertions) and visual ToM tasks (e.g., tasks assessing the ability to infer mental states from eye gaze), a meta-analysis found significantly worse performance in verbal tasks in euthymic BD patients (de Siqueira Rotenberg et al., 2020; Haag et al., 2016). In contrast, a meta-analysis by Bora et al. (2016) found similar ToM deficits in both verbal and visual tasks. Distinguishing between the affective and cognitive aspects of ToM, cognitive ToM was impaired compared to HC, while affective ToM was not impaired (Montag et al., 2010; Shamay-Tsoory et al., 2009; Szmulewicz et al., 2019). In contrast, a meta-analysis reported deficits in both affective and cognitive ToM. The effect size for cognitive ToM impairment was

moderately greater than that for affective ToM impairment (Bora et al., 2016). In line with this, two meta-analyses considered both facets of ToM and demonstrated small differences in affective tasks (e.g. affective Faux Pas, RMET) and medium-to-large deficits in cognitively focused tasks (Samamé et al., 2012, 2015). Another study examined individuals with BD and reported no impairments in the RMET test compared to controls (Dalkner et al., 2019). Thus, divergent results have been reported regarding affective and cognitive ToM facets. Although, these studies suggest larger effect sizes in cognitive ToM impairments.

1.4.1.6 ToM and eating disorders. Bulimia nervosa (BN) is an ED characterized by repeated episodes of binge eating followed by compensatory purging behaviors. Anorexia nervosa (AN) is a severe psychiatric illness characterized by an altered perception of body image, the rigid conviction that one is overweight, and behavioral modifications, such as food restriction and excessive exercise. Both ED predominantly affect females and are connected to substantial physical health complications (Arcelus et al., 2011).

AN is associated with impaired social functioning that precedes the onset of ED symptoms (Arcelus et al., 2013). Moreover, AN has been associated with cognitive deficits, including deficits in EFs (Fagundo et al., 2012), inefficient set-shifting, characterized by rigid approaches to changing rules and preservative thinking (Fitzpatrick et al., 2012), and social cognitive deficits, including ToM impairment (Bora & Köse, 2016; Simonsen et al., 2020). Thus, individuals with AN and BN showed impairments in ToM. However, individuals with AN exhibited a diminished level of ToM ability compared to those with BN and HC, whose scores were more similar (Simonsen et al., 2020). Moreover, individuals with obesity performed worse than controls on the ToM tests. In fact, this was demonstrated for all age groups. In addition, this systematic review yielded some evidence of an association between ToM difficulties and problematic eating behaviors, such as binge eating, food addiction and emotional eating, which are commonly observed in people with obesity (Tonelli & de Siqueira Rotenberg, 2021). Impairments in cognitive ToM and ToM decoding have been linked to impaired social functioning in various psychiatric disorders (Bora et al., 2006). These deficits may contribute to interpersonal difficulties in individuals with AN. Among the various ED, the largest number of studies were conducted on AN. Research on Binge ED and BN is inconsistent and limited, although there seem to be some

indications of deficits in ToM (Mason et al., 2021).

In general, individuals with ED exhibit both exceeding and reduced ToM abilities as well as decreased accuracy in inferring affective and cognitive mental states (Corsi et al., 2021). Consequently, a cognitive ToM deficit, particularly in terms of cognitive perspective-taking, has been specifically identified in individuals with AN. The results are mixed regarding the affective ToM (Bora & Köse, 2016). Focusing on the decoding components of ToM, emerging evidence indicates impairments in ToM in individuals with AN, with the RMET used to assess decoding aspects of ToM (Tapajóz P. de Sampaio et al., 2013). The effect sizes of ToM deficits in individuals with AN indicated substantial impairment in cognitive ToM and moderate impairments in ToM decoding. BN was distinguished by a deficit in the ability to decode mental states, while reasoning about mental states remained unaffected (Bora & Köse, 2016).

1.4.2 Changes in ToM abilities during treatment

Psychological interventions are an effective treatment for mental health problems. Psychotherapeutic treatment for patients with mental disorders is associated with positive outcomes regarding symptom severity, as well as a decrease in interpersonal problems (Liebherz & Rabung, 2014). Further, social deficits have a significant impact on the long-term outcomes of many mental disorders (Uljarević et al., 2020). Focusing on changes in social deficits in response to treatment, a study by Hayden et al. (2018) found that treatment significantly increased ToM performance in patient with different mental disorders during a 6-week inpatient treatment. Moreover, improvements in ToM were associated with lower interpersonal distress and played a crucial role in enabling effective problem-solving skills in social contexts, ultimately leading to improved social adaptation. In line with this, another study found that during the course of inpatient psychodynamic treatment, the level of uncertainty about mental states decreased, with a medium effect size in individuals with BPD (Meulemeester et al., 2018).

Specific treatment programs have been developed to ameliorate ToM deficits such as MCT (Moritz et al., 2014) or mentalization-based treatment (MBT; Fonagy & Bateman, 2008). The MCT has been developed as an eclectic add-on group intervention for patients with mental disorders. The training addresses dysfunctional thinking styles and cognitive distortions (Moritz & Lysaker, 2018). MCT was originally developed for individuals with psychosis with moderate effect sizes for symptom

reduction (Eichner & Berna, 2016) and has been expanded to treat a variety of mental disorders, such as depression (Jelinek et al., 2016), obsessive-compulsive disorder (Miegel et al., 2020), pathological gamblers (Gehlenborg et al., 2021) and BPD (Schilling et al., 2018; Vogt & Norman, 2019). MCT has a high treatment efficacy for BPD patients (Schilling et al., 2018; Vogt & Norman, 2019) in terms of significant symptom reductions (Schilling et al., 2018). MBT focuses on enhancing the mentalization capacity of patients, with the aim to improve affect regulation and interpersonal functioning (Fonagy & Bateman, 2008). As an add-on intervention, it encompasses psychoeducational elements, strategies from the third wave (especially mindfulness) and CBT techniques. MBT also has an significant impact on long-term improvement among patients with BPD (Volkert et al., 2019).

Focusing on the importance of ToM, Sharp and Vanwoerden (2015) pointed out that ToM deficits are a central feature of the etiology, maintenance and treatment of BPD. The authors proposed that ToM deficits are prognostic factors for poor long-term treatment outcomes. According to the mentalization approach, mentalization is considered as a core psychological process that operates as a common factor across various treatments in individuals with BPD. Consequently, enhancement of mentalizing capacity is regarded as an important mechanism of change in any effective treatment targeting BPD (Fonagy et al., 2015). In a RCT individuals with BPD received either manualized psychodynamic treatment or treatment by experienced community therapists. Results demonstrated that significant enhancements in mentalization abilities were reported only within the group receiving the manualized psychodynamic treatment within one year of treatment (Fischer-Kern et al., 2015). In contrast, another study found no improvements in mentalization abilities in the course of a psychodynamic treatment (Vermote et al., 2010).

A recent study compared BPD patients treated with structured clinical management (SCM) with BPD patients treated with metacognitive interpersonal therapy (MIT) and revealed that both interventions decreased difficulties in emotion regulation, but MIT had a higher effect size (Rossi et al., 2023). MIT yielded larger effects on metacognitive functions and impulsivity than SCM. In contrast, SCM showed a larger decrease in BPD symptomatology. Thus, MIT is an effective intervention to improve mentalizing abilities for BPD. Nevertheless, on the long-term MBT compared to SCM resulted in better symptom reductions in BPD patients in an eight-year follow-

up study (Bateman et al., 2021). Particularly self-incurious behavior, suicide attempts and functional outcomes (engagement in purposeful activities and less use of social care interventions) could be improved.

The effectiveness of MCT in individuals with BPD is much more studied compared to individuals with AUD. According to Caselli et al. (2018), patients with AUD who received inpatient treatment exhibited decreased weekly alcohol consumption after MCT. In patients with AUD, the distress level and desire to consume alcohol, both significant indicators of treatment effectiveness, were decreased after MCT (Caselli et al., 2016). The impact of generic psychotherapy on ToM abilities remains uncertain in the absence of a targeted metacognitive intervention. A recent study conducted by Rupp et al. (2021) provided evidence that the impairment in affective ToM continued to persist even after an extended period of abstinence-oriented alcohol dependence inpatient treatment (>2 months). In contrast, a study by Frileux et al. (2020) revealed an amelioration of the affective ToM deficits subsequent to a three-week duration of alcohol detoxification therapy. However, to date there are no longitudinal studies assessing cognitive ToM in patients with AUD. Moreover, no previous studies have used the MASC to investigate the relationship between ToM and the treatment outcomes of generic psychotherapy in patients with AUD.

Empirical evidence shows that MCT is also highly effective for depression and anxiety disorders (Normann et al., 2014). A recent review investigated the efficacy of metacognitive interventions in adults diagnosed with mental disorders (Philipp et al., 2018). In controlled trials, MCT demonstrated superior effectiveness compared with both waitlist control (WLC) and CBT: MCT and MBT were similarly effective to alternative psychological interventions and outperformed non-active interventions in the treatment of schizophrenia, OCD, anxiety disorders, PTSD, and depression.

Social-cognitive training can be also applied for healthy individuals. A recent systematic review and meta-analysis with healthy individuals across the lifespan reported high effectiveness for ToM functions. Identifying specific characteristics that contribute to the success of social-cognitive training can improve the design of future training programs targeting different populations (Roheger et al., 2022).

1.4.3 ToM and changes in symptom severity

ToM impairments are common in many mental disorders. Thus, it seems important to investigate how ToM is linked to psychotherapy. It might be, that patient

with different ToM differ in their ability to engage in psychotherapy. In line with this, among individuals with mental disorders, social cognition impairments were associated with less improvements in symptom severity during psychotherapeutic treatment (Kvarstein et al., 2020; Sondermann et al., 2020). This supports the idea that patients with higher ToM benefit more from psychotherapy than patients with less ToM abilities. One form of benefit from psychotherapy is the improvement of symptom severity. Accordingly, ToM at the beginning of treatment would be associated to changes in comorbid symptom severity.

1.4.3.1 ToM and symptoms of depression. ToM has also been identified as a significant predictor of symptom reduction and social adjustment with respect to long-term treatment outcomes. For instance, a ToM deficit after symptom remission in patients with MDD predicted a higher relapse rate and lower social functioning one year after recovering from a major depressive episode (Inoue et al., 2006; Yamada et al., 2015). A prospective study by Yamada, Inoue and Kanba (2015) reported that 42.9% of patients with ToM deficits relapsed within one year. In contrast, only 5.2% of the patients without ToM deficits relapsed within that time. Moreover, a previous study by Bora and Berk (2016) showed an association between affective and cognitive ToM and symptoms of depression. More details regarding ToM impairments in patients with MDD are described in Chapter 1.4.1.4.

Individuals with mild symptoms of depression showed different results. Few studies have indicated a positive relationship between mild depressive symptoms and better ToM abilities. For instance, individuals experiencing dysphoria had greater levels of accuracy in affective ToM compared to non-dysphoric individuals (Harkness et al., 2005). Similarly, individuals with subclinical levels of depressive symptoms were associated with higher accuracy in the assessment of affective ToM (Poletti et al., 2014). These findings suggested that there may be an enhancement of affective ToM at certain levels of low mood, or that specific degrees of depressive symptoms may be characterized by heightened social sensitivity.

The role of ToM as a predictive factor for symptom improvements has also been investigated in patients with BPD (Meulemeester et al., 2018). In the context of a psychodynamic treatment program for BPD, the level of mentalization at the beginning of the treatment did not serve as a predictor of symptom improvement. However, the results showed a strong association between changes in symptomatic distress and

mentalization. Another study by Kvarstein et al. (2020) investigated patients with BPD during the course of MCT and found that insufficient ToM were associated with poorer improvement in clinical outcomes (symptoms of depression, anxiety and somatization, interpersonal problems, and social functioning).

Symptoms of depression are a frequent comorbidity among individuals diagnosed with AUD as evidenced by multiple studies (Castillo-Carniglia et al., 2019; Li et al., 2020). ToM deficits have also been linked to these comorbid symptoms (Bora & Berk, 2016). However, there is currently a lack of research investigating whether ToM capacities of individuals with AUD are associated with the reduction of comorbid symptoms of depression during AUD treatment. Nonetheless, a longitudinal study examining individuals with depression and a HC group demonstrated that a higher level of social cognition was associated with less improvement in depressive symptoms during psychotherapeutic treatment (Sondermann et al., 2020). This suggests that specific associations might exist between ToM abilities and the reduction of comorbid symptoms. Further research should investigate the role of ToM and the improvement of depression symptoms during AUD treatment.

1.4.3.2 ToM and symptoms of anxiety. Previous studies have shown that individuals with generalized anxiety disorder (GAD) have diverse interpersonal problems (Przeworski et al., 2011; Salzer et al., 2008). Moreover, individuals with GAD are more likely to show maladaptive social cognition in terms of underestimation or overestimation of hostile behaviors (Erickson & Newman, 2007). Pathological worry, as a core symptom of GAD, was also associated with interpersonal problems (Erickson et al., 2016) and poorer social cognition skills (De Vito et al., 2019). In contrast, a study by Zainal and Newman (2018) found a more accurate reasoning ToM (using the MASC) in the worried individuals with GAD compared to the HC even after controlling for EF, social anxiety, depressive symptoms, and sex. People with GAD who were worried and not relaxed, performed better with negatively rated social stimuli. Thus, ToM performance in the GAD group was impacted by state relaxation/worry (Zainal & Newman, 2018).

However, empirical studies on ToM abilities of individuals with anxiety disorders have yielded inconsistent findings. While certain studies have demonstrated lower ToM abilities (e.g. Hezel and McNally, 2014), other studies reported no significant differences between individuals with or without anxiety disorders (Lenton-Brym et al.,

2018). Nevertheless, the majority of studies and a recent meta-analysis indicate poorer ToM functionality in individuals with anxiety disorders compared to HC (Chevalier et al., 2023; Santarelli et al., 2022; Sloover et al., 2022). Further, a negative correlation between the severity of current clinically relevant anxiety symptoms and ToM has been found in healthy younger adults (Petrovic et al., 2023).

However, it is important to acknowledge that most studies included patients with social anxiety disorder (SAD). Less is known about individuals with other anxiety disorders (Sloover et al., 2022). Similar to individuals with GAD, individuals with SAD perform worse on ToM tasks than non-socially anxious individuals (Hezel & McNally, 2014) with a tendency towards hypermentalizing and more exceeding ToM errors than individuals with depression (Washburn et al., 2016).

One possible explanation for impaired ToM in anxious individuals is proposed by Bateman and Fonagy (2019): It is suggest that anxiety inhibits high-order cognitive function such as ToM functions. Thus, this could explain why reductions of anxiety might reduce ToM impairments. Accordingly, improvements in symptoms during treatment could also lead to improved ToM or require ToM. Future studies should investigate the impact of anxiety symptoms on ToM abilities during treatment.

Research on the treatment of anxiety disorders is in line with many results from studies on the efficacy on MCT in individuals with anxiety disorders. MCT had a significant positive effect on anxiety symptomatology and improved treatment success in patients with anxiety disorders (Normann et al., 2014). Nevertheless, due to the small sample size and few active control groups; this result should be interpreted with caution.

1.4.3.3 ToM and symptoms of social functioning. Investigating the association between ToM abilities and symptoms, functioning and well-being yields insights for further development of mental health interventions. The results from a non-clinical sample revealed no association between ToM and global symptoms/psychopathology (Ballespí et al., 2021). In contrast in another study, ToM is related to general, social and role functioning and well-being including self-esteem and motivation to lifegoals. This indicates that ToM could contribute to symptom management by improving well-being and social functioning and is linked to social skills and community functioning (Couture et al., 2006). In line with this Bora et al. (2006) observed that individuals with schizophrenia and with good social functional

outcomes performed better on mental state reasoning tests. Both the capacity for reasoning about intentions and decoding mental states play integral roles in perceiving and interpreting social cues. Given the significant relevance of recognizing social cues in the development of social skills, both abilities are crucial for effective social functioning (Bora et al., 2006). Interestingly, performance in the RMET also predicted social functioning. The RMET is described in this study as a mental state decoding task that exhibits a greater predictive ability for social impairments compared to mental state reasoning tasks. Other studies have argued that the RMET is an emotion recognition task (e.g. Kittel et al., 2022). Accordingly, to this assumption the study by Bora et al. (2006) demonstrated that emotion recognition might be a better predictor of social functioning than ToM.

According to Sanvicente-Vieira et al. (2017), there is a bidirectional association between impairments in ToM and SUD, whereby substance use disrupts neurological functioning that is relevant for ToM (Fluharty et al., 2018). Concurrently, deficits in ToM may hinder social functioning and contribute to the adoption of substance use as a social coping mechanism (Lannoy et al., 2020). This study assumed that interpersonal skills, as one aspect of social functioning, may contribute to excessive drinking behavior in young individuals. This supports the idea of the study by Kornreich et al. (2002), which reported that individuals with AUD showed both more interpersonal problems and deficits in facial expression recognition, which is one component of ToM abilities. ToM abilities have repeatedly been linked to social functioning in several psychiatric samples (Bora et al., 2006; Rosello et al., 2020; Wang et al., 2015). For instance patients with BPD and higher cognitive ToM showed higher improvements in social functioning (Kvarstein et al., 2020).

1.4.3.4 ToM and symptoms of somatization. Somatization disorders are characterized by the presence of somatic symptoms that indicate underlying medical conditions, but these somatic symptoms cannot be fully explained by medical conditions. Somatic symptoms have been linked to impairments in the ToM (Chevalier et al., 2023; Subic-Wrana et al., 2010; Thamby et al., 2019; Zunhammer et al., 2015). Independent of mood states, individuals with somatoform disorders have ToM deficits (Thamby et al., 2019). Comparing high extent of somatic symptoms (HRS) with low extent of somatic symptoms in the general population, HRS showed poorer affective ToM. Regarding the cognitive ToM, the groups did not differ significantly. The

distinction between affective and cognitive ToM in individuals reporting high levels of symptoms and those reporting low levels of symptoms, may be attributed to the varying demands of tasks. Therefore, comparable measurements are required. A recent study by Chevalier (2023) confirmed the association between mentalization and internalizing symptomatology, indicating that higher levels of somatization symptoms are correlated with, lower mentalization performance on solely controlled ToM tasks. Controlled ToM tasks are described as tasks that didn't activate the automatic, implicit pole. This evidence of a negative correlation between ToM and internalizing problems (e.g., somatization symptoms) provides support for the possible impact of ToM skills on psychological symptoms among clinical groups.

1.5 The present study

1.5.1 Overview of Article 1

The first article examined several individual and situational effects on the measurement of ToM abilities within two clinical samples: inpatients from an AUD treatment unit and inpatients from a personality disorder (PD) treatment unit, including especially individuals with BPD. As outlined in Section 1.2, ToM research is currently facing certain methodological problems, such that reliable measurements are needed to investigate findings on ToM abilities in clinical groups. The MASC test is a ToM measure with high ecological validity, used in the present article. Article 1 considers the effects influencing affective and cognitive ToM measurements, such as age, sex, and test duration. The research presented in Article 1 addresses the following research questions by assessing ToM abilities.

Article 1 investigated the following research questions:

1. Does the performance of affective and cognitive ToM change over the course of the test?
2. What age differences in affective and cognitive ToM performance emerge between individuals from a PD treatment unit and those from an AUD treatment unit?
3. What sex differences in affective and cognitive ToM performance emerge between individuals from a PD treatment unit and those from an AUD treatment unit?
4. Do ToM error types differ between individuals from a PD treatment unit and those from an AUD treatment unit?

5. Are test duration, sex, age, and ToM facet (affective, cognitive) associated with ToM performance?

1.5.2 Overview of Article 2

Prior research has concluded that ToM is impacted by social groups (e.g., gender) of the target and whether there is concurrence or disparity in the mental states of the target individual and the perceiving individual. A typical interaction involves a person talking (i.e., sender) and a person listening (i.e., receiver); therefore, there are two targets per interaction. Based on the literature, it is hypothesized that ToM is enhanced if both the target and perceiving individuals share a common group identity and possess congruent perspectives. Conversely, ToM is reduced if individuals belong to the same social group but exhibit divergent mental perspectives. It is still uncertain how these effects influence ToM measures. Therefore, the second article investigated the influence of gender of interaction partners on the measurement of ToM abilities within two clinical samples, patients from an AUD treatment unit, patients from a PD treatment unit and a HC group. Investigating the effects of interaction partners on ToM performance could improve the understanding of ToM assessments and might help to resolve inconsistencies in the current research on ToM performance in distinct clinical populations. This study used two clinical samples to promote the understanding of ToM as a transdiagnostic indicator of mental disorders. The research presented in Article 2 mainly addresses the influence of patients' sex, target gender and their interaction on ToM. Investigating, whether the gender of the interaction partners (sender and receiver) influences ToM performance. The following three research questions were investigated:

1. Do items with female or male targets differ in their ToM performance in clinical and non-clinical samples within the MASC?
2. What role does the same or different gender of sender and receiver play in ToM performance in clinical and non-clinical samples within the MASC?
3. Do ingroup interactions (same sender gender, receiver gender and participant's sex: only female or only male interactions) improve or hinder ToM performance in clinical and non-clinical samples within the MASC?

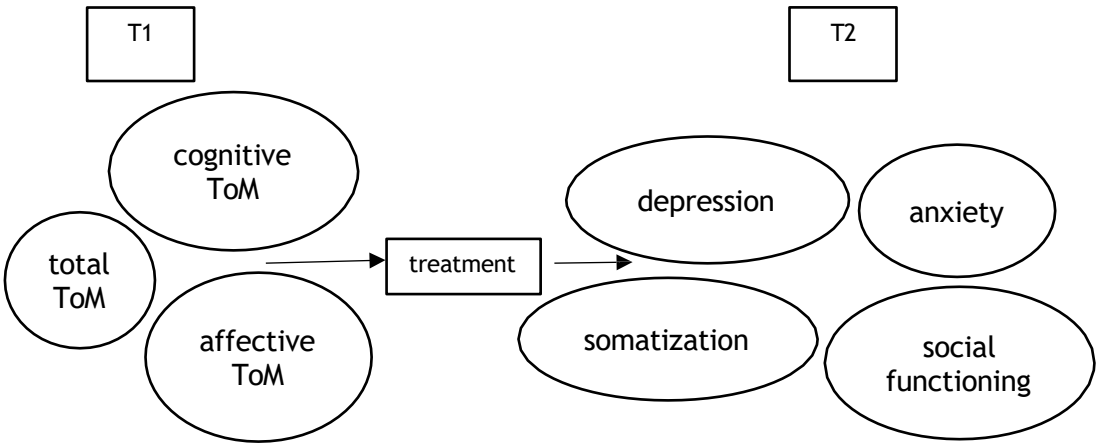
1.5.3 Overview of Article 3

There is considerable evidence indicating deficits in both affective and cognitive ToM in individuals with AUD. However, the changes in these deficits during the course of AUD treatment, and their potential relationship to co-occurring symptoms, remain unclear. The results from the third article demonstrated treatment outcomes of individuals with mental disorders, focusing on those from AUD treatment units. The article was conceptualized to investigate longitudinal data in affective and cognitive ToM. ToM abilities were measured upon admission and discharge of inpatient abstinence-oriented treatment. Data should help provide a better understanding of the possible dissociation between affective and cognitive ToM changes in individuals with AUD. Furthermore, as displayed in Figure 3, article 3 considered the associations between affective and cognitive ToM and the improvements of comorbid symptoms of depression, anxiety, somatization, and social functionality. The findings should provide a better understanding of the effects of ToM, as a prognostic value, on changes in symptoms in mental health treatments, specifically for people with chronic mental illnesses.

1. Do total, affective, and cognitive ToM abilities improve during an inpatient abstinence-oriented treatment for patients with AUD?
2. Is there a dissociation between the change of affective and cognitive ToM abilities during inpatient treatment?
3. Are affective and cognitive ToM abilities at admission related to improvements in the comorbid symptoms of depression, anxiety, somatization, and social functioning in individuals with AUD?

Figure 3

Hypothetical Model Diagram



Note. ⟶ Associations

T1, T2: measurement points; ToM: Theory of Mind

Difference of symptoms (anxiety, depression, social functioning, somatization)

2. Article 1: How Individuals and Situational Factors Influence Measures of Affective and Cognitive Theory of Mind in Psychiatric Inpatients

Reference

Knopp, M., Burghardt, J., Meyer, B., Riffer, F., & Sprung, M. (2022). How Individual and Situational Factors Influence Measures of Affective and Cognitive Theory of Mind in Psychiatric Inpatients. *Frontiers in Psychology, 13.* https://doi.org/10.3389/fpsyg.2022.855038

Abstract

Mental disorders are associated with difficulties to correctly infer the mental states of other's (theory of mind; ToM). These inferences either relate to affective states of others (affective ToM) or to their thoughts, intentions, or beliefs (cognitive ToM) and can be associated with mental disorder. The current study explores the influence of individual and situational effects on the measurement of ToM abilities within two clinical samples, to increase generalizability. We analyzed data from 229 in-patients; 103 patients treated for alcohol use disorder and 126 patients treated for a personality disorder. ToM was assessed with the Movie for the Assessment of Social Cognition (MASC). We analyzed changes in test performance over the course of the test using a logistic linear mixed effects model. Performance on the cognitive ToM items decreased over time, while performance on the affective ToM items increased over time. This difference was more pronounced among older individuals. The results show important moderators of ToM performance that might help to resolve inconsistencies in the current literature about ToM abilities in different clinical or age groups.

Magdalena Knopp (MK)'s contributions to the article: Designed and conducted the study, was responsible for methodology, formal analysis, and data curation, and took the lead in writing and revising the manuscript.

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 19 May 2022
doi: 10.3389/fpsyg.2022.855038

How Individual and Situational Factors Influence Measures of Affective and Cognitive Theory of Mind in Psychiatric Inpatients

Magdalena Knopp[1,2†], Juliane Burghardt[1†], Bernhard Meyer[1], Friedrich Riffer[3,4] and Manuel Sprung[1,2,3,4]*

[1] Division of Clinical Psychology, Department of Psychology and Psychodynamics, Karl Landsteiner University of Health Sciences, Krems an der Donau, Austria, [2] Department of Psychology, Faculty of Psychology and Educational Sciences, Ludwig-Maximilians-Universität München, Munich, Germany, [3] Psychiatric Rehabilitation Clinic Gars am Kamp, Psychosomatisches Zentrum Waldviertel, Gars am Kamp, Austria, [4] Psychosomatisches Zentrum Waldviertel, University Hospital for Psychosomatic Medicine Eggenburg, Eggenburg, Austria

OPEN ACCESS

Edited by:
Irene Ceccato,
University of Studies G. d'Annunzio Chieti and Pescara, Italy

Reviewed by:
Elena Cavallini,
University of Pavia, Italy
Ivana Stasevic Karlicic,
Clinic for Mental Disorders "Dr. Laza Lazarevic," Serbia
Mikhail Sorokin,
St. Petersburg V. M. Bekhterev Psychoneurological Research Institute, Russia

**Correspondence:*
Magdalena Knopp
Magdalena.knopp@kl.ac.at

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Psychology for Clinical Settings, a section of the journal Frontiers in Psychology

Received: 14 January 2022
Accepted: 12 April 2022
Published: 19 May 2022

Citation:
Knopp M, Burghardt J, Meyer B, Riffer F and Sprung M (2022) How Individual and Situational Factors Influence Measures of Affective and Cognitive Theory of Mind in Psychiatric Inpatients. Front. Psychol. 13:855038. doi: 10.3389/fpsyg.2022.855038

Mental disorders are associated with difficulties to correctly infer the mental states of other's (theory of mind; ToM). These inferences either relate to affective states of others (affective ToM) or to their thoughts, intentions, or beliefs (cognitive ToM) and can be associated with mental disorder. The current study explores the influence of individual and situational effects on the measurement of ToM abilities within two clinical samples, to increase generalizability. We analyzed data from 229 in-patients; 103 patients treated for alcohol use disorder and 126 patients treated for a personality disorder. ToM was assessed with the Movie for the Assessment of Social Cognition (MASC). We analyzed changes in test performance over the course of the test using a logistic linear mixed effects model. Performance on the cognitive ToM items decreased over time, while performance on the affective ToM items increased over time. This difference was more pronounced among older individuals. The results show important moderators of ToM performance that might help to resolve inconsistencies in the current literature about ToM abilities in different clinical or age groups.

Keywords: cognitive empathy, mentalizing, personality disorder, substance abuse, gender differences, age difference

INTRODUCTION

Many mental disorders are associated with difficulties to correctly infer the mental states of others, i.e., theory of mind (Németh et al., 2018). ToM deficits are seen in over 30 mental disorders (Cotter et al., 2018). These inferences either relate to affective states of others (affective ToM) or to their thoughts, intentions, or beliefs (cognitive ToM). To promote the understanding of ToM and its association with mental disorders, the influence of individual, and situational factors on the assessment of ToM needs to be considered.

Two meta-analyses and a systematic review provide evidence for ToM impairment in individuals with alcohol use disorder (AUD) relative to healthy controls (HC; Onuoha et al., 2016; Bora and Zorlu, 2017; Sanvicente-Vieira et al., 2017). Onuoha et al. (2016) summarized the result from 8

studies in their meta-analysis and found that AUD individuals (N = 187) displayed reduced ToM compared to HC (N = 187), with a large effect size, Hedges' g = −1.62 [95% CI = −2.28; 0.96]. A second meta-analysis summarizing the results from 12 studies reported that AUD individuals (N = 317) compared to HC individuals (N = 298) displayed ToM deficits with a medium effect size, Cohen's d = 0.58 [95% CI = 0,36–0.81] (Bora and Zorlu, 2017). Furthermore, ToM impairment is common among borderline personality disorder (BPD) patients (Németh et al., 2018) and has also been demonstrated in patients with other personality disorders, for instance patients with narcissistic personality disorder or cluster C personality disorders (Ritter et al., 2011; Marissen et al., 2012; Herpertz and Bertsch, 2014). However, only few studies have compared ToM abilities in patients with BPD and other personality disorders. One such study found no difference between BPD and other personality disorders in overall ToM performance, but rather ToM impairment was associated with the severity of psychopathology (Normann-Eide et al., 2020). Deficits in affective and cognitive ToM can show different associations with mental disorders. For instance, a study by Harari et al. (2010) found that patients with BPD only have cognitive ToM deficits, but later studies showed deficits in affective and cognitive ToM (Petersen et al., 2016). In line with this, a meta-analysis concluded that BPD have difficulties with cognitive as well as affective ToM tasks (Németh et al., 2018). In contrast to patients with BPD, individuals with AUD show deficits primarily in affective ToM, but their cognitive ToM is comparable to healthy controls (HC) (Maurage et al., 2016). Thus, BPD and AUD differ regarding cognitive and affective ToM abilities. Considering this dissociation, findings about ToM deficits should not be generalized to other mental disorders without further testing. Thus, affective and cognitive ToM abilities are linked to mental health. Further, they can be influenced by other characteristics of the individual. For instance, gender or age can be differentially associated with affective and cognitive ToM abilities. Women show more pronounced affective responses (Han et al., 2008) and perform better in tasks that assess affective ToM (Baron-Cohen, 2010). Further, they often show better emotion recognition (Hall and Matsumoto, 2004; Hoffmann et al., 2010; Donges et al., 2012). Many studies have found a decrease in ToM abilities in older adults, which was summarized in a meta-analysis by Henry et al. (2013). This meta-analysis found reliable aging effects on both affective and cognitive ToM. However, some recent studies found effects of aging were limited to the cognitive ToM and did not occur for the affective ToM (Wang and Su, 2013). A similar study found that age was associated with a decrease in cognitive ToM but not in affective ToM (Bottiroli et al., 2016).

Besides differentiating between affective and cognitive ToM, ToM deficits can further be divided into different error types: exceeding ToM, less ToM, no ToM. Multiple studies found that individuals with BPD showed an exceeding ToM, this means that patients with BPD displayed a tendency to interpret incidental actions in an intentional way (Sharp and Vanwoerden, 2015; Normann-Eide et al., 2020). In contrast, a meta-analysis by Onuoha et al. (2016) concluded that individuals with AUD typically displayed reduced ToM compared to healthy controls.

The present study aims to explore the influence of individual and situational effects on the measurement of affective and cognitive ToM abilities in two clinical samples, patients from an AUD treatment unit and patients from a personality disorder (PD) treatment unit; with a focus on mixed personality disorder and borderline personality disorder. The two AUD and PD samples are very different in their sociodemographic characteristics, especially regarding age, and gender and also show different ToM errors and a distinct pattern of affective and cognitive ToM. In order to detect general influences of individual and situational factors on the measurement of affective and cognitive ToM abilities using data from two distinct samples can increase the generalizability of results. Since the patients with PD in the present sample are younger and more often women, we assume that they will outperform individuals with AUD, who are older and more often men, on affective and cognitive ToM. Differential effects of age, gender, and test duration on affective and cognitive ToM performance are investigated. We expect a decrease in task performance over the course of the test. Based on previous studies we assume that age, gender, and test duration will influence ToM performance. Further we expect women and younger participants to outperform men and older participants. These effects of age, gender, and test duration may be different for affective and cognitive ToM.

MATERIALS AND METHODS

Participants and Procedure

We collected data from 390 patients from an inpatient psychiatric-psychosomatic clinic in Austria, which focusses on patients with chronic mentally illnesses (216 AUD, 174 PD). The inpatient stays last between 2 and 3 months, during which the patients receive intense therapy by an interdisciplinary team. Respective data was collected as part of the routine examination at admission.

Only data from outcome assessment at the time of intake were included in the analyses. To be admitted to the clinic patients need to be able to undergo therapies. This means they need to have at least conversational skills in German, show no marked cognitive deficits and show the motivation to engage in therapy. Exclusion criteria are acute psychosis, danger of suicide and intoxication (Knopp et al., 2021). Patients were either in a treatment unit for AUD or for PD (i.e., mixed personality disorder and borderline personality disorder). All AUD patients were recently detoxified. Data from 75 patients (64 AUD, 11 PD) were excluded due to missing values. To control for the possible influence of overall test compliance we excluded data from 86 participants who answered less than 4 out of 6 MASC control items correctly (49 AUD, 37 PD). We also analyzed the data without excluding participants who failed more than 2 out of 6 MASC control questions; this did not alter results in any significant way. After exclusion, the final AUD sample included 103 patients (66 men) with a mean age of 49.0 (SD = 8.9) years. The final PD sample included 126 patients (33 men) with a mean age of 35.0 (SD = 11.2) years. Sociodemographic variables are depicted in **Table 1**. Most patients were native speakers of German (AUD: 95, PD: 118).

In the AUD sample, the primary diagnosis of 101 patients was alcohol dependence syndrome (F10.2); 2 patients were diagnosed with harmful use of alcohol (F10.1). Additionally, patients in the AUD sample had a multitude of comorbid (secondary) diagnoses: 33 affective disorders (F31-38), 9 anxiety disorders (F40-42), 6 personality disorders (F60-61), 2 posttraumatic stress disorders (F43.1), 1 eating disorder (F50), 2 obesity (E66). In the PD sample, the primary diagnoses were: 51 mixed personality disorder (F61) and 46 borderline personality disorder (F60.3), 9 patients were diagnosed with other personality disorders (F60, F68), 7 with posttraumatic stress disorder and other reactions to severe stress (F43), 5 recurrent depressive disorder (F33), 5 phobic and other anxiety disorders (F40-42), 2 with bipolar affective disorder (F31) and 1 with Asperger syndrome (F84.5).

Procedure

The data for this study were collected between July 2017 and May 2019 as part of the routine outcome monitoring. Basic sociodemographic data, such as age and sex, were obtained from the hospital information system. All other measures were collected using the Computer-based Health Evaluation System (Holzner et al., 2012). The admission examination spans two 1-h sessions with self-report questionnaires and the ToM measure. Patients answer the questions in a computer assessment room with eight separated cubicles; so up to eight patients complete the questionnaires and the ToM task in one assessment session.

Measures

ToM was assessed using a behavioral measure. Sociodemographic variables were assessed by self-report.

Movie for the Assessment of Social Cognition

ToM was measured with the Movie for the Assessment of Social Cognition (MASC), which relies on video sequences of people interacting with each other to assess cognitive and affective ToM performance (Dziobek et al., 2006). The MASC shows a series of video clips of social interactions, which concern friendship and dating themes. The format is highly immersive. After each video clip participants are required to infer the thoughts or feelings of the individual(s) in the particular scene. Participants respond to multiple choice (MC) questions, with four possible answers. These MC-questions include one correct and three incorrect answers. Incorrect answers represent three prototypical error types. These errors types are: (a) "exceeding ToM," i.e., a mental state is attributed when there is no mental explanation for the situation, (b) "less ToM," i.e., a present mental state is misattributed, and (c) "no ToM," i.e., a total absence of mental inference (e.g., making attributions of physical causality to social situations and mental states). Test items either require inferences about a characters' feelings/emotions (affective ToM) or their thoughts/intentions (cognitive ToM). The stimulus set includes more cognitive than affective ToM items, therefore the raw scores were transformed into "percentage correct." Additionally, in order to control variables such as memory, general compliance and understanding of the task, six control questions that make no reference to mental states are included in the MASC. The control questions require inferences about non-social/non-mental state related situations or facts and are also in a MC-format. The MASC has previously been reported to have high reliability; internal consistency was high ($\alpha \geq 0.82$ for all scores), good test-retest stability and good convergent and discriminant validity (Fossati et al., 2018). The average duration to complete the MASC was 30.5 min.

Sociodemographic Characteristics

We assessed sex, age, education level, employment, duration of sick leave, duration of symptoms, and duration of psychotherapy. Sex was measured dichotomously; "0" for women and "1" for men. Age ranged from 18 to 74 years and entered regressions as a continuous variable. For the visualization of results, we used age groups (18–34; 35–45; 46–53; 54–75) based on quartiles. Educational level was coded into three graduation levels; "1" low education level (compulsory school), "2" medium education level (middle school), and "3" high education level (high school or university degree). Employment status was coded into employed (fulltime or part time) vs. unemployed (retired, unemployed, early retirement). Duration of sick leave was coded "1" if duration of sick leave has been present less than 3 months, "2" if duration of sick leave has been present 3–6 months and "3" if duration of sick leave has been present for more than 6 months. Duration of symptoms was coded "1" if symptoms have been present less than 6 months, "2" if symptoms have been present 6–12 months and "3" if symptoms have been present for more than 12 months. Duration symptoms was measured using a single item ("How long ago did you first perceive the problems that led to the treatment?").

Data Analysis

Comparative statistics (Student's *t*-test or Pearson's Chi-squared test) for both patient groups (AUD vs. PD) were calculated, to determine their complementary clinical characteristics and to replicate previous reports of specific ToM performance patterns.

Within-subject effects were analyzed with an item-level analysis, utilizing a Logistic Linear Mixed Effects model (LME), with individual MASC item responses as dichotomous (correct vs. incorrect) dependent measure. The analysis used two LME models to test effects of test duration, age, and ToM facet on ToM performance. Both LME models included the main effects of item order (marker of test duration), ToM facet (affective vs. cognitive), age, and sex, and the interaction of item order and ToM facet. (1) The first model additionally contained the interaction effects of ToM facet with sex, item order with sex and the three-way interaction of item order, ToM facet and sex.

(2) The second model instead included interaction effects of ToM facet with age, item order with age, and the three-way interaction of ToM facet, item order, and age. Resulting *p*-values for all interaction effects within the two regression models were corrected for multiple testing using the Bonferroni method (7 tests in total: $p_{adjusted} = 0.007$). Analyses were performed using R 4.0.2[1] with the packages lme4 and lmerTest for LME modeling (glmer, family: binomial), tidyverse, magrittr, dplyr, and reshape2 for data preparation, and ggplot2 for data visualization.

[1] http://cran-r-project.org/

TABLE 1 | Sociodemographic-, illness-related characteristic and MASC test results (%) *.

Sociodemographic characteristics	PD *N* (126)	AUD *N* (103)	χ^2 (df)	*p*
Sex (men/women)	26.2/73.8	64.1/35.9	41.1 (3)	**<0.001**
Educational levelLow/medium/high	71.5/19.8/8.7	68.1/20.3/11.6	146.5 (5)	**<0.001**
Employment (yes/no)	17.7/82.3	42.2/57.8	61.0 (3)	**<0.001**
Illness-related characteristics				
Duration of sick leave <3/3–6/ > 6 months	49.5/4.6/45.9	68.2/4.5/27.3	96.4 (5)	**<0.001**
Duration of symptoms <3/3–6/ > 6 months	6.4/1.6/92	14.6/6.8/78.6	301.7 (5)	**<0.001**
MASC	***M (SD)***	***M (SD)***	***t* (df)**	***p***
ToM total (%)	69.2 (13.3)	63.8 (14.2)	-2.9 (212)	**0.004**
ToM exceeding	15.8 (7.0)	16.1 (7.7)	0.4 (210)	0.705
ToM less	13.9 (6.8)	16.7 (7.9)	2.9 (202)	**0.004**
ToM no	6.8 (5.2)	8.0 (5.0)	1.7 (222)	0.098
Affective ToM	70.1 (15.4)	63.4 (16.9)	−3.1 (209)	**0.002**
Cognitive ToM	68.5 (15.8)	64.0 (16.6)	2.1 (213)	**0.038**
Control questions	80.7 (12.4)	79.9 (12.2)	−0.5 (219)	0.645

*Bold p-values are p < 0.005. *Percentage without missing.*

RESULTS

Sociodemographic and Illness-Related Characteristic

Table 1 depicts the sociodemographic characteristic, duration of symptoms/illness, and treatment history. The ratio of men and women differed between the AUD and the PD samples. There were more men than women in the AUD sample. There were more women than men in the PD sample. In both samples the majority of patients was unemployed. The proportion of unemployed individuals was higher in the PD sample. In line with this, patients in the PD sample had a lower education level than patients in the AUD sample. The majority of both samples reported a low education level (AUD participants 68.1%; PD 71.5%). The proportion of patients with low education level was even higher in the PD sample.

Regarding the history of symptoms/illness, patients in the PD sample more frequently reported a longer period of sick leave (duration of sick leave) and a longer history of subjectively experienced symptoms (duration of symptoms) than AUD patients. Moreover, 71.0% of patients in the PD sample reported being in psychotherapy (prior to the current inpatient treatment) for more than a year, compared to 46.6% of the participants in the AUD sample.

TABLE 2 | Regression of factors influencing ToM performance.

	β	*SE*	*z*	*p*
ToM facet (cognitive vs. affective)	0.08	0.08	1.10	0.288
Item order (z-values)	0.18	0.06	3.00	**0.003**
$Sex_{patient}$ (woman = 0, man = 1)	0.29	0.11	2.60	0.010
Age (z-values in years)	−0.19	0.05	−4.00	**<0.001**
ToM facet × item order	−0.31	0.08	−4.00	**<0.001**
ToM facet × $sex_{patient}$	−0.32	0.10	−3.10	**0.002**
Item order × $sex_{patient}$	0.06	0.08	0.70	0.471
ToM facet × item order × $sex_{patient}$	−0.04	0.10	−0.30	0.732

Bold p-values are p < 0.005.

Movie for the Assessment of Social Cognition Test Results

We report the results after excluding participants that failed more than 2 of MASC control questions, however, the results are highly similar with the full sample. The MASC results for the two samples are presented in **Table 1**. On average patients

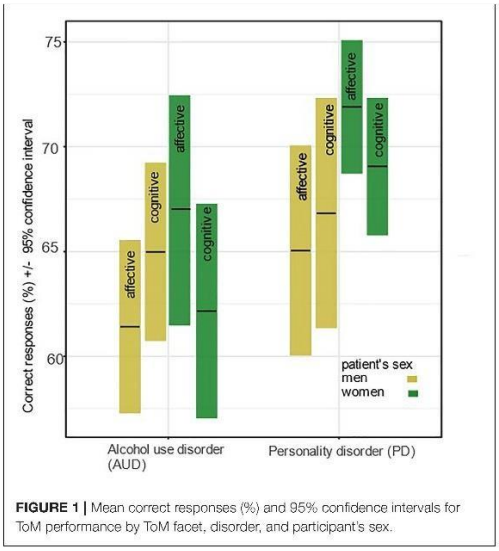


FIGURE 1 | Mean correct responses (%) and 95% confidence intervals for ToM performance by ToM facet, disorder, and participant's sex.

in the AUD sample answered 79.9% (*SD* = 12.2) of the MASC control items correctly, while in the PD sample answered 80.7% (*SD* = 12.4) of the control items correctly. Regarding ToM, items patients in the AUD sample answered *M* = 63.8 (*SD* = 14.2) of the MASC items correctly and patients in the PD sample answered *M* = 69.2 (*SD* = 13.3) of the MASC items correctly. The PD sample scored higher on the affective ToM M = 70.1 (*SD* = 15.4) compared to the AUD sample *M* = 63.4 (*SD* = 16.9), *t* (209) = –3.1, *p* = 0.002. Furthermore, the performance in the cognitive ToM facet of patients in the PD sample, *M* = 68.5 (*SD* = 15.8), was better than of patients in the AUD sample, *M* = 64.0 (*SD* = 16.6), *t* (213) = 2.1, *p* = 0.038. This difference was evident for both affective and cognitive ToM, although the difference was more pronounced for affective ToM (see **Table 1**). The most common error type among patients in the AUD sample was "less ToM," *M* = 16.7, *SD* = 7.9, which reflects misattribution of a mental state. This error type was significantly less frequent among patients in the PD sample, *t*(202) = 2.9, *p* = 0.004. Among patients in the PD "exceeding ToM" was the most common error types *M* = 15.8 (*SD* = 7), however, it was equally common among patients in the AUD sample.

Table 2 and **Figure 1** show the results of the first regression model, which tested main and interaction effects of item order (of MASC test items), the two ToM facets (affective vs. cognitive), and of participant's age and sex on ToM performance, and all interactions with sex. The analysis revealed that MASC performance decreased over the duration of the test, β = 0.18, *SE* = 0.06, *p* = 0.003. This effect was mainly due to a decrease in cognitive ToM performance. In contrast, affective ToM showed a positive slope and thus increased over time. Female participants outperformed male participants in their overall ToM performance, β = 0.29, *SE* = 0.11, *p* = 0.010, however, this effect was no longer significant after a Bonferroni correction was applied, which resulted in a critical *p*-value of *p* = 0.007. However, the interaction between ToM facet and participant sex was significant (β = –0.32, *SE* = 0.10, *p* = 0.002), suggesting that women outperformed man on the affective ToM, while their performance was comparable or lower on the cognitive ToM facet.

The second regression model tested main and interaction effects of item order, ToM facet (affective vs. cognitive), and participant sex (male vs. female) and all interactions with age. Results show a significant main effect of participant sex. These effects are visible in **Figure 2**, on the left side. There was also a significant interaction of the ToM facet with item order, β = –0.31, *SE* = 0.08, *p* < 0.001. Moreover, the decline in cognitive ToM was more pronounced for older patients, as indicated by a significant three-way interaction of ToM facet with item order and age, β = –0.15, *SE* = 0.05, *p* = 0.004. **Figure 2** illustrates this effect, showing that the slopes of the MASC test performance over the course of the testing are more pronounced among older patients. To visualize these results **Figure 2** shows four age groups. However, there was also significant main effect of age, suggesting an overall decrease in ToM with increasing age, β = –0.19, *SE* = 0.05, *p* < 0.001. The chart on the right side of **Figure 2** shows the ToM performance separately for the two

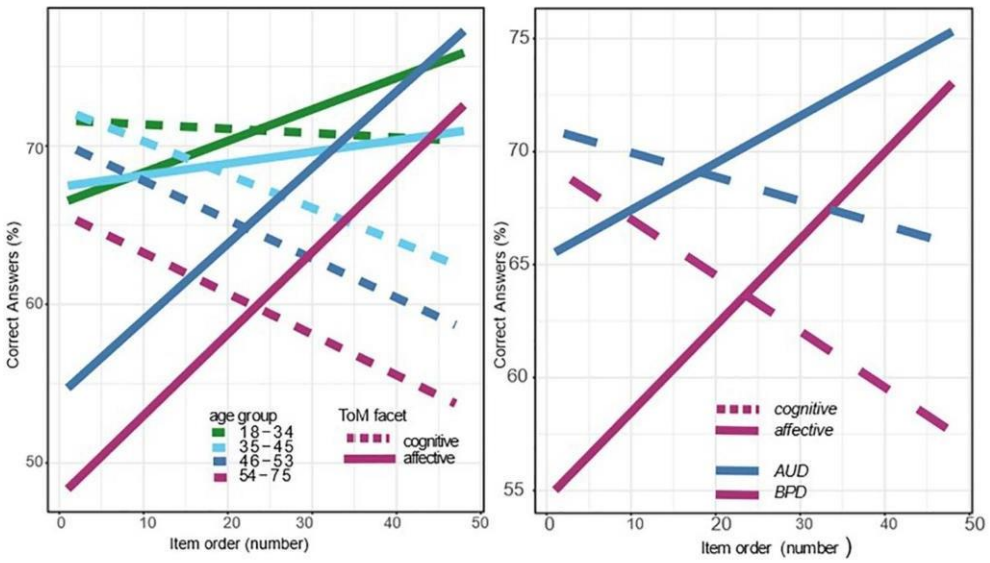


FIGURE 2 | Course of affective and cognitive ToM performance over the test duration.

samples (AUD vs. PD) and for the two ToM facets (affective vs. cognitive ToM).

DISCUSSION

The results support the notion that performance in ToM tasks, specifically the MASC, is influenced by individual and situational factors. ToM performance decreased during the course of the test. However, this progression differed between affective and cognitive ToM. While cognitive ToM decreased, affective ToM increased during the course of the test. These decreases could have resulted from a decline in task motivation, attention, or fatigue among participants. Previous studies have shown that cognitive ToM relies more strongly on working memory and attention (Gabriel et al., 2019) than affective ToM. Hence, affective ToM might be less influenced by a decrease in attention.

The pattern of decreasing cognitive ToM and increasing affective ToM was more pronounced among older participants than younger adults. Thus, the results are more in line with the studies that have shown a negative effect of aging on the cognitive ToM only (Wang and Su, 2013; Bottiroli et al., 2016) instead to those that found effects of aging on both cognitive and affective ToM to be equally strong (Henry et al., 2013). The decline in cognitive ToM performance in older adults has previously been associated with a decrease in executive functioning (Wang and Su, 2013) and a strategic reallocation of scarce cognitive resources. Therefore the motivation to engage in cognitively demanding activities is reduced (Hess, 2014). This might explain why the affective ToM shows an increase in performance over the course of the test. Participants might choose to allocate resources to attend to emotions of actors within the MASC. The results are in line with earlier studies that showed that increased task motivation enhanced older adults' performance in such a way their ToM performance level reached younger adults' level of ToM performance Zhang et al. (2018). In line with this, our findings show strong performance differences at the end of the test, when motivation is presumably low. Based on these effects it is interesting to note that in the MASC all interactions are presented by relatively young actors. Their young age might decrease the emotional closeness between the participants and the actors (Hess, 2014; Zhang et al., 2018). As a result, this could lead to lower task performance among older participants.

Further, the results showed that while women outperform men on the affective ToM, men outperform women on the cognitive ToM. This is in line with previous findings showing a higher emotional competence in women as marked by their better performance on tasks measuring the affective dimension of social cognition. Especially women showed a better decoding of non-verbal messages (Hall et al., 2000), and as well as a higher ability to recognize emotions from facial expressions (Hall and Matsumoto, 2004; Hoffmann et al., 2010; Donges et al., 2012).

There were significant differences in the overall ToM performance between the two patient samples. AUD patients had more difficulties inferring mental states than PD patients. However, despite the differences in the samples regarding their symptomatology, socio-demographic characteristics, and distinct ToM errors and pattern of affective and cognitive ToM the two samples displayed similar effects over the course of the test, which suggests that the influence of individual and situational factors may be general.

Limitation

The current study did not directly measure general cognitive abilities, executive functioning, or task motivation. This limits our understanding of the processes that underly the findings, future studies should assess motivation directly to test its effect. However, to reduce the possible influence of general cognitive abilities, executive function, or task engagement, we excluded data from participants who failed more than 2 out of 6 of MASC control questions. Thus, the results in the MASC test questions are likely to reflect genuine ToM abilities and not merely general cognitive abilities.

Since the data was collected within the standard treatment routine it was not possible to counterbalance item order. It is theoretically possible that the items in later parts of the task would be more difficult than in the beginning, which would be an alternative explanation for the effect of item order However, attributing the entire effect of item order to differences in item difficulty is complicated by the age effects. It might be possible that item difficulty increases for cognitive items and decreases for affective items. However, assuming that item difficulty increases specially for older individuals and only on cognitive items is relatively unlikely. Therefore, we assume that differences in item difficulty are not a suitable explanation for effects of item order.

Even though the two samples have a higher generalizability than a single sample, the findings are nevertheless limited and should be replicated with a broader spectrum of mental disorders.

Conclusion

The present study demonstrates how strongly performance in ToM tasks can be influenced by individual factors, i.e., sex, age, and disorder type (AUD vs. PD). The effects of age and sex should be kept in mind, when comparing results between different clinical samples, as many clinical samples have a typical distribution of age, and sexes. Further, the results illustrate the influence of situational factors, i.e., test duration, on ToM performance, since the test performance changed over the course of the test. This effect was different for affective and cognitive ToM. Cognitive ToM decreased over time, while affective ToM increased over time. Thus, the influence of individual and situational factors may explain the heterogenous evidence of ToM deficits in different mental disorders. Moreover, considering the influence of individual and situational factors can improve the validity of assessment of ToM performance.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because of the vulnerability of the study sample. Participants of this study did not agree for their data to be shared publicly,

so supporting data are not available. Requests to access the datasets should be directed to MK, magdalena.knopp@kl.ac.at.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Commission of the Karl Landsteiner University of Health Sciences (Nr: 1020/2021). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MK, JB, and MS: conceptualization. MK, JB, and BM: methodology, formal analysis, investigation, and data curation. MK and JB: writing—original draft preparation. MS and FR: writing—review and editing. All authors have read and agreed to the published version of the manuscript.

FUNDING

We are grateful to the NÖ Forschungs- und Bildungsges.m.b.H. (NFB) for the financial support as part of the endowed professorship in clinical psychology. We further acknowledge that the VAMED Institute for Gender Medicine funded the position of JB. We acknowledge support by Open Access Publishing Fund of Karl Landsteiner University of Health Sciences, Krems, Austria.

ACKNOWLEDGMENTS

We like to thank Isabel Dziobek for providing the Movie for the assessment of social cognition (MASC), and we also like to thank Isabel Dziobek, Pierre Maurage, and Ulrike Buhlmann for providing the MASC item numbers for the calculation of the affective and cognitive Theory of Mind subscales.

REFERENCES

Baron-Cohen, S. (2010). Empathizing, systemizing, and the extreme male brain theory of autism. *Prog. Brain Res.* 186, 167–175. doi: 10.1016/B978-0-444-53630-3.00011-7

Bora, E., and Zorlu, N. (2017). Social cognition in alcohol use disorder: a meta-analysis. *Addiction* 112, 40–48. doi: 10.1111/add.13486

Bottiroli, S., Cavallini, E., Ceccato, I., Vecchi, T., and Lecce, S. (2016). Theory of Mind in aging: comparing cognitive and affective components in the faux pas test. *Arch. Gerontol. Geriatr.* 62, 152–162. doi: 10.1016/j.archger.2015.09.009

Cotter, J., Granger, K., Backx, R., Hobbs, M., Looi, C. Y., and Barnett, J. H. (2018). Social cognitive dysfunction as a clinical marker: a systematic review of meta-analyses across 30 clinical conditions. *Neurosci. Biobehav. Rev.* 84, 92–99. doi: 10.1016/j.neubiorev.2017.11.014

Donges, U.-S., Kersting, A., and Suslow, T. (2012). Women's greater ability to perceive happy facial emotion automatically: gender differences in affective priming. *PLoS One* 7:e41745. doi: 10.1371/journal.pone.0041745

Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., et al. (2006). Introducing MASC: a movie for the assessment of social cognition. *J. Autism Dev. Disord.* 36, 623–636. doi: 10.1007/s10803-006-0107-0

Fossati, A., Borroni, S., Dziobek, I., Fonagy, P., and Somma, A. (2018). Thinking about assessment: further evidence of the validity of the Movie for the Assessment of Social Cognition as a measure of mentalistic abilities. *Psychoanal. Psychol.* 35, 127–141. doi: 10.1037/pap0000130

Gabriel, E. T., Oberger, R., Schmoeger, M., Deckert, M., Vockh, S., Auff, E., et al. (2019). Cognitive and affective Theory of Mind in adolescence: developmental aspects and associated neuropsychological variables. *Psychol. Res.* 85, 533–553. doi: 10.1007/s00426-019-01263-6

Hall, J. A., Carter, J. D., and Horgan, T. G. (2000). "Gender differences in nonverbal communication of emotion," in *Gender and Emotion: Social Psychological Perspectives*, ed. A. H. Fischer (Cambridge: Cambridge University Press), 97–117. doi: 10.1017/CBO9780511628191.006

Hall, J. A., and Matsumoto, D. (2004). Gender differences in judgments of multiple emotions from facial expressions. *Emotion* 4, 201–206. doi: 10.1037/1528-3542.4.2.201

Han, S., Fan, Y., and Mao, L. (2008). Gender difference in empathy for pain: an electrophysiological investigation. *Brain Res.* 1196, 85–93. doi: 10.1016/j.brainres.2007.12.062

Harari, H., Shamay-Tsoory, S. G., Ravid, M., and Levkovitz, Y. (2010). Double dissociation between cognitive and affective empathy in borderline personality disorder. *Psychiatry Res.* 175, 277–279. doi: 10.1016/j.psychres.2009.03.002

Henry, J. D., Phillips, L. H., Ruffman, T., and Bailey, P. E. (2013). A meta-analytic review of age differences in theory of mind. *Psychol. Aging* 28, 826–839. doi: 10.1037/a0030677

Herpertz, S. C., and Bertsch, K. (2014). The social-cognitive basis of personality disorders. *Curr. Opin. Psychiatry* 27, 73–77. doi: 10.1097/YCO.0000000000000026

Hess, T. M. (2014). Selective engagement of cognitive resources: motivational influences on older adults' cognitive functioning. *Perspect. Psychol. Sci.* 9, 388–407. doi: 10.1177/1745691614527465

Hoffmann, H., Kessler, H., Eppel, T., Rukavina, S., and Traue, H. (2010). Expression intensity, gender and facial emotion recognition: WOmen recognize only subtle facial emotions better than men. *Acta Psychol.* 135, 278–283. doi: 10.1016/j.actpsy.2010.07.012

Holzner, B., Giesinger, J. M., Pinggera, J., Zugal, S., Schöpf, F., Oberguggenberger, A. S., et al. (2012). The Computer-based Health Evaluation Software (CHES): a software for electronic patient-reported outcome monitoring. *BMC Med. Inform. Decis. Mak.* 12:126. doi: 10.1186/1472-6947-12-126

Knopp, M., Riffer, F., Burghardt, J., and Sprung, M. (2021). Geschlechtsspezifische Unterschiede in der psychotherapeutischen Versorgung. *Psychotherapeut* 66, 511–517. doi: 10.1007/s00278-021-00523-4

Marissen, M. A. E., Deen, M. L., and Franken, I. H. A. (2012). Disturbed emotion recognition in patients with narcissistic personality disorder. *Psychiatry Res.* 198, 269–273. doi: 10.1016/j.psychres.2011.12.042

Maurage, P., D'Hondt, F., de Timary, P., Mary, C., Franck, N., and Peyroux, E. (2016). Dissociating affective and cognitive theory of mind in recently detoxified alcohol-dependent individuals. *Alcoholism* 40, 1926–1934. doi: 10.1111/acer.13155

Németh, N., Mátrai, P., Hegyi, P., Czéh, B., Czopf, L., Hussain, A., et al. (2018). Theory of mind disturbances in borderline personality disorder: a meta-analysis. *Psychiatry Res.* 270, 143–153. doi: 10.1016/j.psychres.2018.08.049

Normann-Eide, E., Antonsen, B. T., Kvarstein, E. H., Pedersen, G., Vaskinn, A., and Wilberg, T. (2020). Are impairments in theory of mind specific to borderline personality disorder? *J. Pers. Disord.* 34, 827–841. doi: 10.1521/pedi_2019_33_417

Onuoha, R. C., Quintana, D. S., Lyvers, M., and Guastella, A. J. (2016). A meta-analysis of theory of mind in alcohol use disorders. *Alcohol Alcohol.* 51, 410–415. doi: 10.1093/alcalc/agv137

Petersen, R., Brakoulias, V., and Langdon, R. (2016). An experimental investigation of mentalization ability in borderline personality disorder. *Compr. Psychiatry* 64, 12–21. doi: 10.1016/j.comppsych.2015.10.004

Ritter, K., Dziobek, I., Preißler, S., Rüter, A., Vater, A., Fydrich, T., et al. (2011). Lack of empathy in patients with narcissistic personality disorder. *Psychiatry Res.* 187, 241–247. doi: 10.1016/j.psychres.2010.09.013

Sanvicente-Vieira, B., Romani-Sponchiado, A., Kluwe-Schiavon, B., Brietzke, E., Araujo, R. B., and Grassi-Oliveira, R. (2017). Theory of mind in substance users: a systematic minireview. *Subst. Use Misuse* 52, 127–133. doi: 10.1080/10826084.2016.1212890

Sharp, C., and Vanwoerden, S. (2015). Hypermentalizing in borderline personality disorder: a model and data. *J. Infant Child Adolesc. Psychother.* 14, 33–45. doi: 10.1080/15289168.2015.1004890

Wang, Z., and Su, Y. (2013). Age-related differences in the performance of theory of mind in older adults: a dissociation of cognitive and affective components. *Psychol. Aging* 28, 284–291. doi: 10.1037/a0030876

Zhang, X., Lecce, S., Ceccato, I., Cavallini, E., Zhang, L., and Chen, T. (2018). Plasticity in older adults' theory of mind performance: the impact of motivation. *Aging Ment. Health* 22, 1592–1599. doi: 10.1080/13607863.2017.1376313

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

3. Article 2: Effects of Interaction Partner on Theory of Mind Measurement in Clinical and Healthy Samples

Reference

Knopp, M., Burghardt, J., Meyer, B., & Sprung, M. (2023). Effects of Interaction Partner on Theory of Mind Measurement in Clinical and Healthy Samples. Manuscript submitted to Journal of social and clinical psychology.

Abstract

Introduction: Theory of Mind (ToM) is the ability to ascribe thoughts, intentions, or emotions to others. It is influenced by the social group (e.g., gender, age) of the interaction partners and the similarity of their mental states. ToM performance is facilitated if target and perceiver belong to the same group and their mental states align but is reduced if they belong to the same social group, but their mental states differ. ToM has been transdiagnostically associated with many mental disorders, but the findings in different clinical groups are inconsistent. We hypothesized that ToM scores vary with the social groups of the interaction partners and explored effects of target gender.

Methods: We analyzed 103 patients with alcohol use disorder, 126 patients with personality disorders, and 32 healthy controls. ToM was assessed with the Movie for the Assessment of Social Cognition (MASC). A logistic linear mixed effects model tested the influence of the gender of interaction partners: sender, receiver, and patient's gender on MASC test scores.

Results: In all samples, items where participant and target had the same gender resulted in lower ToM scores compared to items with different genders. In the clinical sample items with male (vs. female) targets resulted in better ToM scores.

Discussion: The results suggest that the gender of interaction partners is an important moderator of ToM performance, which may explain some of the inconsistencies in the literature on ToM abilities in different clinical groups.

Keywords: Gender, similarity, mentalizing, Personality Disorders, Alcohol Use Disorder

MK's contributions to the article: MK designed and conceptualized this study. MK curated the data and conducted the data analyzes. MK wrote the first draft. MK reviewed and edited the manuscript.

Effects of Interaction Partner on Theory of Mind Measurement in Clinical and Healthy Samples

Magdalena Knopp[1,2,*,#], Juliane Burghardt[1,#], Bernhard Meyer[1], Manuel Sprung[1,2,3]

[1] Karl Landsteiner University of Health Sciences, Clinical Psychology, Dr.-Karl-Dorrek-Straße 30, 3500 Krems an der Donau, Austria

[2] Ludwig-Maximilians-Universität München, Faculty of Psychology and Educational Sciences, Department of Psychology, Leopoldstraße 13, 80802 München, Germany

[3]University Medical Center Hamburg-Eppendorf, Department of Psychiatry and Psychotherapy, Martinistrasse 52, 20246 Hamburg, Germany

*Corresponding Author: Magdalena Knopp, Karl Landsteiner University of Health Sciences, Clinical Psychology, Dr.-Karl-Dorrek-Straße 30, 3500 Krems an der Donau, Austria, E-Mail: Magdalena.knopp@kl.ac.at

[#]These authors contributed equally to this work.

Ethics

Patient data was collected during routine outcome monitoring at the University Hospital Eggenburg. The retrospective analyses were approved by the ethics commission of the Karl Landsteiner University of Health Sciences (Nr: 1020/2021). All participants gave written informed consent to participate and agreed to the use of their data. The study complies with the ethical standards defined in the Declaration of Helsinki (World-Medical-Association, 2013). The data of the healthy control subjects were collected by Christina Andreou and her colleagues (Andreou et al., 2015). Their study was conducted in accordance with the Declaration of Helsinki and was approved by the local ethics committee.

Funding

We are grateful to the Research Promotion Agency of Lower Austria (GFF) for the financial support of this study as part of the endowed professorship in clinical psychology (endowed professor: Manuel Sprung). We further acknowledge that the VAMED Institute for Gender Medicine funded the position of Juliane Burghardt. None of the funders influenced the content or interpretation of the study.

Author Contributions

JB, MK, BM and MS designed and conceptualized this study. BM, MK and JB curated the data and conducted the data analyzes. JB and MK wrote the first draft. JB, MK and MS revised the final draft. JB, MK, BM and MS reviewed and edited the manuscript. All authors have read and agreed to the final version of the manuscript.

Acknowledgments

We are grateful to all patients at the University Hospital for Psychosomatic Medicine Eggenburg for the permission to use their data.

We further like to thank Christina Andreou and Steffen Moritz for providing data from a healthy control sample. Details about this data set are reported in Andreou et al. (2015). We would also like to thank Beate Sodian for providing helpful comments on an earlier version of this manuscript.

Declaration of conflicting interests

The authors declare no conflicts of interest.

Data Availability Statement

The datasets analyzed in this study are not readily available because of the vulnerability of the study sample. Participants of this study did not agree for their data to be shared publicly.

Abstract

Introduction: Theory of Mind (ToM) is the ability to ascribe thoughts, intentions, or emotions to others. It is influenced by the social group (e.g., gender, age) of the interaction partners and the similarity of their mental states. ToM performance is facilitated if target and perceiver belong to the same group and their mental states align but is reduced if they belong to the same social group, but their mental states differ.

ToM has been transdiagnostically associated with many mental disorders, but the findings in different clinical groups are inconsistent. We hypothesized that ToM scores vary with the social groups of the interaction partners and explored effects of target gender.

Methods: We analyzed 103 patients with alcohol use disorder, 126 patients with personality disorders, and 32 healthy controls. ToM was assessed with the Movie for the Assessment of Social Cognition (MASC). A logistic linear mixed effects model tested the influence of the gender of interaction partners: sender, receiver, and patient's gender on MASC test scores.

Results: In all samples, items where participant and target had the same gender resulted in lower ToM scores compared to items with different genders. In the clinical sample items with male (vs. female) targets resulted in better ToM scores.

Discussion: The results suggest that the gender of interaction partners is an important moderator of ToM performance, which may explain some of the inconsistencies in the literature on ToM abilities in different clinical groups.

Keywords: Gender, similarity, mentalizing, Personality Disorders, Alcohol Use Disorder

Introduction

Many mental disorders are marked by Theory of Mind (ToM) impairments (Cotter et al., 2018; McLaren et al., 2022). ToM is the ability to ascribe mental states to others, which is the basis for social functioning (Rosello et al., 2020; Sabbagh et al., 2009; Singer, 2006). However, finding the optimal ToM measure, which is necessary to improve our understanding of ToM and ToM deficits in people with mental disorders, is controversial (Quesque & Rossetti, 2020).

Alcohol use disorder (AUD) and personality disorders such as borderline personality disorder (BPD) have repeatedly been associated with ToM deficits. Multiple meta-analyses and reviews have concluded that there is consistent evidence for strong ToM deficits among individuals with AUD relative to healthy controls (HC; Bora & Zorlu, 2017; Onuoha et al., 2016; Sanvicente-Vieira et al., 2017). Notably, a meta-analysis showed that the difference between AUD and HC decreased for studies with higher proportions of men in the AUD group, suggesting stronger deficits in women (Onuoha et al., 2016). In contrast, a second meta-analysis concluded that a higher proportion of men was associated with more severe ToM deficits, thus suggesting stronger ToM deficits in men. Accordingly, it is unclear whether ToM deficits are more pronounced in men or women with AUD.

There is also substantial evidence for ToM deficits in patients with BPD (Hanegraaf et al., 2021; Németh et al., 2018) and other personality disorders (PD; Burghardt et al., 2023; Herpertz & Bertsch, 2014; Normann-Eide et al., 2020). ToM deficits were of comparable size among patients with BPD and other PD (Burghardt et al., 2023; Normann-Eide et al., 2020); however, comparatively few studies exist on other PD. In contrast to the AUD findings, there is no evidence for gender differences in ToM deficits among BPD patients (Hanegraaf et al., 2021; Németh et al., 2018). BPD samples typically include a higher proportion of female than male patients (Skodol & Bender, 2003). The opposite is typical for AUD samples (Gilbert et

al., 2019). Thus, it is possible that the type of diagnosis and gender proportion are confounded in their influence on ToM. The current study seeks to scrutinize the effects of gender and sex within ToM measurement. Sex refers to biological differences between men, women, and other individuals, while gender refers to socio-cultural differences between these groups.

When measuring ToM, participants need to infer the mental states of targets, which are identified by names or pictures. Previous experiments have shown that ToM is affected by the social group (e.g., gender, age) of the target (Ames et al., 2012; Simpson & Todd, 2017; Sudo & Farrar, 2020). Beyond ToM research, many studies have confirmed that humans use characteristics such as gender and age to categorize individuals into social groups (e.g., Klauer & Wegener, 1998; Phillips et al., 2013). This categorization affects the perception of that individual depending on the social group of the perceiver; perceiver and target can either belong to the same social group (ingroup) or to different social groups (outgroup).

Whether a person is part of an in- or outgroup influences memory, attention, and information processing concerning that target (e.g., Klauer & Wegener, 1998; Kunda & Spencer, 2003). Many studies support the notion that people prefer to interact with individuals of their ingroup relative to an outgroup (Lincoln & Miller, 1979; McDonald, 2011; McPherson & Smith-Lovin, 1987). This is partly explained by the assumption that people expect individuals from their ingroup to be more similar to themselves and therefore more likely to share mental states (e.g., information, opinions, and attitudes) than outgroup members (Chen & Kenrick, 2002; Phillips, 2003; Phillips & Loyd, 2006).

Research suggests that this expectation of similarity increases the likelihood that the perceiver (i.e., participant) will use their own mental state to infer the mental state of a similar (e.g., ingroup) target than to infer the mental state of a dissimilar target (Ames et al., 2012; Simpson & Todd, 2017; Sudo & Farrar, 2020). Thus, similarity between participant and target influences ToM. Similarity can arise from sharing surface-level (demographic) characteristics

(e.g., gender or age). Alternatively, it can arise from deep-level (psychological) similarity, which becomes more relevant in longer relationships (Harrison et al., 2002). However, despite the expectation of similarity within ingroups, ingroup members can nevertheless have opposing mental states. A meta-analytic summary of three studies found that when participants judged the mental state of an ingroup (vs. outgroup) target, the participant's own mental state interfered with that judgment (Simpson & Todd, 2017). Shared ingroup membership increased the risk that a participant would erroneously use their own mental state to infer a target's mental state. Thus, ToM was reduced by an error termed egocentric intrusion. In contrast, when the mental state of participant and ingroup target aligned, ToM performance was facilitated. Thus, ToM performance is facilitated when the perspective of the participant aligns with that of the ingroup interaction partner (i.e., target), but is reduced when their perspectives differ (Simpson & Todd, 2017). Thus, whether ToM performance is less or more accurate depends on whether the participant and target are part of the same or different social group and share the same or different mental states. Gender is a central characteristic used by humans to categorize individuals into in- and outgroups (Klauer et al., 2014). As noted before, having the same gender will lead to the expectation that the interaction partner is similar (Chen & Kenrick, 2002; Phillips, 2003; Phillips & Loyd, 2006) and should therefore increase interference between the participant's and the target's mental states (Ames et al., 2012; Simpson & Todd, 2017). Further, gender is associated with specific expectations (i.e., stereotypes). In many contexts, women are perceived as less credible sources of information than men (e.g., Armstrong & McAdams, 2009; Kang et al., 2019). Thus, ToM performance may be reduced for female targets.

Based on these previous findings, it is likely that ToM performance is influenced by characteristics of a target, particularly their gender, and by the interaction between the target's and participant's gender. These effects could help explain the variability in ToM assessed using

different measures. If the effects of gender on ToM performance are present in clinical samples, it may help explain the heterogeneity of findings in different studies of ToM among various clinical conditions. In particular, the effects could have contributed to the conflicting results regarding participant gender on ToM performance in the two meta-analyses among AUD patients (Bora & Zorlu, 2017; Onuoha et al., 2016)

The present study

This study explored whether target gender affects ToM performance and whether interactions between ingroup members result in different ToM estimates than interactions between outgroup members. These effects were analyzed in the Movie for the Assessment of Social Cognition (MASC), a prominent ToM measure that is known for high internal and ecological validity (Benito-Ruiz et al., 2022; Fossati et al., 2018). Each MASC item involves interactions between at least two individuals. A typical item involves a person talking (i.e., sender) and a person listening (i.e., receiver); thus, there are two targets per interaction. The study analyses whether the receiver gender, sender gender and participant gender influence ToM outcomes and whether trials with same-gender receiver, sender and participant show different outcomes than trials involving different genders.

The study analyzed two clinical samples and a healthy control group. Patient data were obtained from an AUD treatment unit and a personality disorder (PD) treatment unit with a focus on borderline personality disorder and mixed personality disorders. AUD and PD patients typically differ substantially regarding gender and age; thus, studying both groups allows to test effect of participant gender in clinical samples.

Method

Participants

In the current study participants refers to both patients and participants of the HC sample.

Patients

We collected data from 216 AUD and 174 PD patients treated at an inpatient psychiatric-psychosomatic clinic in Austria. The samples have already been described elsewhere (Knopp et al., 2022). The data was collected from patients in a treatment unit for AUD (detoxified and abstinent) or for PD respectively. Due to missing data 75 patients (64 AUD, 11 PD) were excluded. Additionally, we excluded data from 86 participants (49 AUD, 37 PD), who failed more than 2 out of 6 MASC control questions, to control for the influence of low test compliance. After exclusion, a total of 229 inpatients (103 patients with AUD and 126 patients with PD) were analyzed. Among AUD participants, 66 were men. Their mean age was 49.0 years (*SD* = 8.9). Among PD participants, 33 were men. Their mean age was 35.0 (*SD* = 11.2) years. The samples are described in Table 1. The participants with AUD comprised 101 patients diagnosed with alcohol dependence syndrome (F10.2) and 2 patients diagnosed with harmful use of alcohol (F10.1). The PD sample included 51 patients with mixed personality disorder (F61), 46 with borderline personality disorder (F60.3), 9 patients with other personality disorders (F60, F68), and 20 other disorders.

HC sample

We analyzed data from a sample of HC participants recruited by Andreou et al. (2015). The participants in the HC sample were approached through advertisements and word of mouth in Germany. The HC sample comprised 32 participants (62.5% females) with a mean age of 31.0 (SD = 11.4) years and 37.5% reporting a medium education level (middle school) and 62.5% a high education level (university or polytechnic degree). Further information about this data can be found elsewhere (Andreou et al., 2015). The absence of a mental disorder was determined using the Structured Clinical Interview for DSM-IV Axis II (SKID-II; Wittchen et al., 1997) and the Mini International Neuropsychiatric Interview (MINI; Sheehan, 1998). Participants were excluded if they fulfilled criteria for a current major depressive episode of

more than mild severity, a history of alcohol or drug dependence, alcohol or drug abuse during the six months previous to testing.

Procedure

The data of the clinical samples were collected using the Computer-based Health Evaluation System (Holzner et al., 2012). Basic sociodemographic data, such as age and gender were taken from the hospital information system. The study analyzed data from all patients that were admitted to the clinic between July 2017 and May 2019, if they consented to the use of their data. Data assessment took part within the routine outcome monitoring in a computer assessment room. The analyzed data was collected as part of the examination at admission, which had a duration of two hours, divided into two one-hour sessions. The HC was evaluated for symptoms of mental disorders. Afterwards they answered the MASC, measures of childhood maltreatment, mood, and cognitive abilities. More information about the data can be found elsewhere (Andreou et al., 2015).

Measures

Theory of Mind

ToM was assessed using the MASC (Dziobek et al., 2006). The MASC is a behavioral measure, which presents a series of 46 video clips of social interactions with up to 4 actors (two women, two men). After each scene, participants are asked to identify the thoughts or feelings of the individual(s) in the video clip. Participants choose between four possible answers in a multiple-choice format with one correct answer. All interactions focus on friendship and dating themes and are presented in an immersive format. Most MASC clips show two targets, which are either sender or receiver of a message. Sender was the person, who spoke during the interaction. Receiver was the person who was addressed. Two raters coded the gender of the

sender and the gender of the receiver in all items with two actors. The two raters showed a high inter-rater reliability of kappa = 0.85. Inconsistencies between raters were resolved through discussion. Three items depicted female to female interactions, 16 items depicted female to male interactions, 17 items depicted male to female interactions, and 3 items depicted male to male interactions. Ingroup interactions involve interaction partners with only males (sender, receiver and participant) or only females (sender, receiver and participant), while outgroup interactions combine male and female interaction partners. Six interactions involved all four characters, these were excluded from analysis. Furthermore, in order to control for factors such as low compliance, memory, or comprehension the MASC contains six control items. Participants with 2 or more errors were excluded. Previously, the MASC has shown good internal consistency (α =.84; Dziobek et al., 2006) and is considered to be an ecological valid task (McLaren et al., 2022).

Sociodemographic data clinical samples

We measured participant's gender, age, education level, and employment. Sociodemographic variables were assessed by self-report. Gender was assessed dichotomously; "0" for women and "1" for men. Age was measured in years. Educational level was recoded into; "1" low education (until compulsory school), "2" medium education level (until middle school), and "3" high education level (high school until university degree). Employment status was coded into unemployed (old-age pension, unemployed, early retirement) vs. employed (fulltime or part time).

Statistical analysis

The patient groups (AUD vs. PD) were compared using Student's t-test or Pearson's Chi-squared test where appropriate.

The effect of target (i.e., target and sender) gender is a within-subject effect that compares different trials among a given patient. We examined this within-subject effects using a Logistic Linear Mixed Effects model (LME), with dichotomous (correct vs. incorrect) MASC item answers as dependent variable. The first model included sender gender, receiver gender, item order and the interaction effects of sender gender and receiver gender, receiver gender and patient's gender, sender gender and patient's gender, controlling for participants gender and age. The second model contained the effect of social group, that is whether participant, sender and receiver were ingroup or outgroup members while controlling for patient's gender and age. The p-values for all interaction effects of the model were corrected for multiple testing using the Bonferroni method. The two analyses were performed on the combined clinical sample, which combined patients with AUD and PD and separately on the HC group. We analyzed the data with and without excluding patients who failed more than 2 out of 6 MASC control questions; this did not alter results in any significant way. Analyses were performed using R 3.6.0 (http://cran-r-project.org/) with the packages lme4 and lmerTest for LME modelling (glmer, family: binomial), tidyverse, magrittr, dplyr and reshape2 for data preparation, and ggplot2 for data visualization.

Results

Preliminary analysis

Table 1 displays the sample characteristics and MASC test results. The AUD and the PD samples differed in their men to women ratio. The PD sample had more women than men. In contrast, the AUD sample had more men than women. The average age of patients in the AUD sample was higher than in the PD sample. Most of the patients in the two samples were unemployed but the number of unemployed individuals was higher in the PD sample. Moreover, patients in the PD sample had on average a lower education level than patients in

the AUD sample. As reported elsewhere (Knopp et al., 2022) patients in the PD sample answered significantly more MASC items correctly than patients in the AUD sample, t (212)= -2.9, p = .004. Table 1 shows means and standard deviation of ToM MASC items and control items.

Table 1

Sociodemographic- and MASC test results in correct % of clinical samples[a]

Sociodemographic characteristics (in %)	PD *N (126)*	AUD *N (103)*	*χ2 (df)*	*p*
Participant's gender (men/women)	26.2 / 73.8	64.1 / 35.9	41.1 (3)	<.001
Educational level low/medium/high	71.5 / 19.8 / 8.7	68.1 / 20.3/ 11.6	146.5 (5)	<.001
Employment (yes/no)	17.7 / 82.3	42.2 / 57.8	61.0 (3)	<.001
MASC	*M (SD)*	*M (SD)*	*t*(df)	*p*
ToM total %	69.2 (13.3)	63.8 (14.2)	-2.9 (212)	.004
Control questions	80.7(12.4)	79.9 (12.2)	-0.5(219)	.645

Note. bold p-values are p<.005; [a] percentage without missing.

Main analyses

Figure 1 and Figure 2 show mean correct responses and 95% confidence intervals for ToM by sender and receiver gender. Figure 1 depicts ToM performance depending on sender and receiver gender and separately for male and female patients for the combined clinical samples. Figure 2 depicts ToM performance depending on sender and receiver gender separately for the two clinical samples. The first regression tested for effects of sender and receiver gender on ToM performance per item (see Table 2). The results showed that sender gender, β = .43, SE = 0.11, p < .001, and receiver gender, β = .42, SE = 0.09, p < .001, were associated with different ToM outcomes. Both male sender and male receiver were associated with better ToM performance compared to female sender and receiver. Additionally, sender and receiver gender showed a significant interaction, β = -.52, SE = 0.18, p =.004. Thus, items

in which men and women interacted were marked by higher ToM performance then items were actors with the same gender interacted. Further, the significant effect of item order showed that patients made more errors at the earlier items of the test, β = .08, *SE* = 0.03, *p* =.006. Moreover, older patients made more errors than younger patients, β = -.20, *SE* = 0.05, *p* < .001.

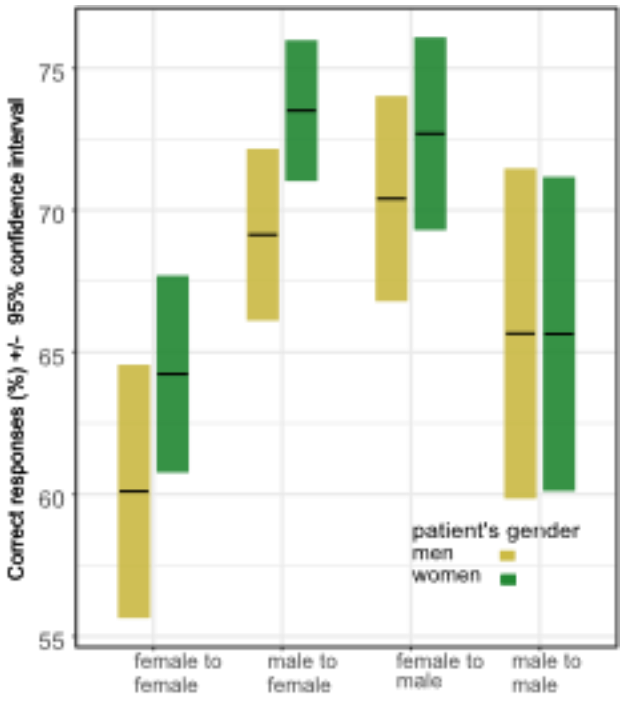


Figure 1 Mean correct responses (%) and 95% confidence intervals for ToM: item's sender to receiver gender for patient's gender

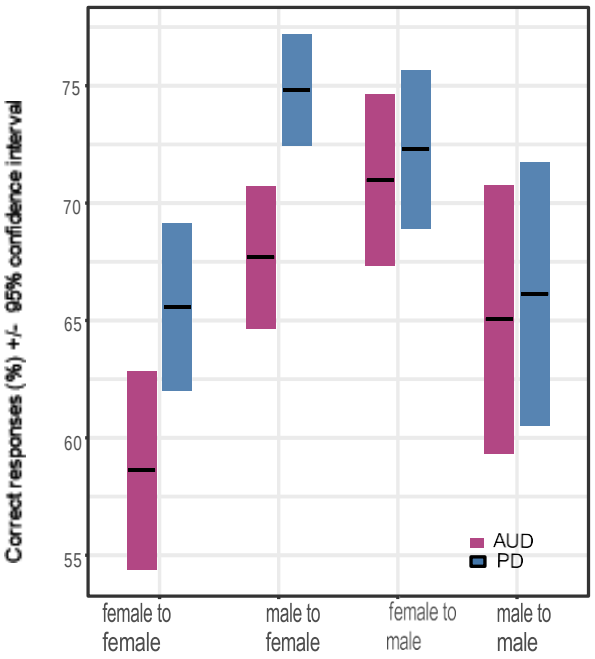


Figure 2 Mean correct responses (%) and 95% confidence intervals for ToM: item's sender to receiver gender for AUD and PD

We tested whether the results of the clinical sample would replicate in the HC. Even though some standardized coefficients were descriptively similar, none of the predictors reached significance. Table 3 shows the results of the first regression model for the HC sample.

Table 2

Regression of target and sender gender on ToM for AUD and PD

	β	*SE*	*z*	*p*
Gender receiver [a]	.43	0.11	3.98	<.001
Gender sender [a]	.42	0.09	4.49	<.001
Gender sender x Gender receiver	-.52	0.18	-2.89	.004
Gender receiver x patient's gender	-.07	0.14	-0.50	.614
Gender sender x patient's gender	.04	0.12	0.31	.755
Item order (z-values)	.08	0.03	2.73	.006
Patient's gender [b]	.07	0.13	0.54	.586
Age (z-values in years)	-.20	0.05	-4.11	<.001

Note. bold p-values are $p<.005$, [a] = female = 0, male=1, [b] = woman = 0, man=1

Table 3

Regression of target and sender gender on ToM for HC

	β	*SE*	*z*	*p*
Gender receiver [a]	.28	0.32	0.88	.380
Gender sender [a]	.32	0.28	1.14	.253
Gender sender x Gender receiver	-.10	0.42	-0.23	.819
Gender receiver x participant's gender	-.05	0.37	-0.13	.893
Gender sender x participant's gender	.04	0.35	0.10	.918
Item order (z-values)	.18	0.10	1.86	.063
Participant's gender [b]	.26	0.33	0.80	.422
Age (z-values in years)	-.10	0.10	-1.00	.316

Note. bold p-values are $p<.005$, [a] = female = 0, male=1, [b] = woman = 0, man=1

Table 4 depicts the results of the second regression within the clinical sample. The second regression tested effects of the social group (ingroup vs. outgroup interaction partners). As depicted in Table 4, the regression shows that ToM was better in outgroup interactions compared to ingroup interactions, β = -0.49, SE = 0.06, $p < .001$. Again, there was a significant effect of age, suggesting a decrease in ToM with increasing age, β = -0.19, SE = 0.05, $p < .001$.

Table 4

Regression of ingroup vs. outgroup trials on ToM in clinical samples

	β	*SE*	*z*	*p*
Ingroup vs. outgroup [a]	-.49	0.06	-8.86	<.001
Age (z-values in years)	-.19	0.05	-4.00	<.001
Patient's gender [b]	.18	0.09	1.88	.060

Note. bold p-values are p<.005; [a] outgroup = 0, ingroup =1, [b] woman = 0, man=1

Subsequently, we tested whether the results for the clinical sample were comparable to the HC sample. This time the results between HC and clinical samples aligned, ToM performance was higher in outgroup interactions than in ingroup interactions, β = -0.58, SE = 0.17, $p < .001$; ToM performance was lower in interactions including only males or only

females than in interactions with different combinations. Table 5 presents the results for the HC.

Table 5

Regression of ingroup vs. outgroup trials on ToM in HC

	β	*SE*	*z*	*p*
Ingroup vs. outgroup	-.58	0.17	-3.31	<.001
Age (z-values in years)	-.10	0.08	-1.17	.240
Participant's gender	-.14	0.20	-0.69	.490

Note. bold p-values are p<.005; [a] outgroup = 0, ingroup =1, [b] woman = 0, man=1

Discussion

The aim of this study was to test for effects of target and participant gender on the assessment of ToM using the MASC, which is a commonly applied ToM measure. In both the clinical and HC sample, items with interactions within an ingroup (i.e., same gender sender, receiver and participant) were marked by lower ToM performance than interactions between individuals of different genders. Items with interactions that combined male and female sender, receiver, and participant yielded higher ToM performance than items with interactions of same-gender sender, receiver, and participant. Thus, ToM performance depended on the interaction between target and participant gender. Within the clinical sample, items with female targets were characterized by lower ToM performance, suggesting that the perspectives of female targets were not equally likely to be considered as the perspectives of male targets. This result did not replicate in the HC. While it is possible that this finding is specific to people with mental disorders, it is also possible that the small HC sample did not provide enough statistical power to reach significance. A post hoc power analysis using G*Power 3.1.9.7 (Faul et al., 2007) showed that for a regression with 8 predictors the power to find a medium effect size (f^2 = 0.15) with the 32 participants of the HC was 0.21, while the clinical sample with 229 participants had a power of 0.99. Thus, the likelihood to find an existing effect was much lower

in the HC. Descriptively, associations for the HC and clinical sample were of similar size and in the same direction.

The aim of this study was to scrutinize ToM measures. The results suggest that items with male vs. female targets create substantial variation within the measure. Male and female participants will perform differently on male and female items. It is noteworthy that the number of male and female items is not balanced in the MASC. Male participants will have fewer ingroup items than female participants, since more female sender-female receiver items are presented than male sender-male receiver items. Small item sets are more likely to yield bigger standard errors. In this case ingroup measures for male participants would be less reliable than those for female participants. The results might help to explain the inconsistencies in previous findings on ToM in mental disorders. If participant gender interacts with target gender to influence ToM performance, then the male to female ratio will influence the likelihood to find ToM deficits in a patient group. Additionally, different ToM measures can have less or more male or female items. This would again influence the ToM assessment depending on the gender of participants. Further, the gender effects of ToM would depend on the ToM measure used and its ratio of male to female targets. For example, the inconsistent effect of gender on ToM among AUD patients (Bora & Zorlu, 2017; Onuoha et al., 2016) could be the result of different measures with different ratios of male to female targets.

In general, men using the MASC, will find fewer items in which they are part of the ingroup. This would create a higher likelihood to produce heterogeneous results. This confound is especially relevant because of the typically high ratio of women among many groups of mental health patients (Greenfield, 2002; Skodol & Bender, 2003).

Previous studies found evidence that ToM is influenced by similarity between participants and targets (Ames et al., 2012; Simpson & Todd, 2017). One form of similarity is being part of the same or a different social group. The previous studies found that if targets and

participants were similar, participants were more likely to use their own internal states to infer the target's internal states than if they were dissimilar (Ames et al., 2012; Simpson & Todd, 2017). For instance, if a female participant assessed the mental state of a woman conversing with another woman, ToM was lower than if a male participant assessed the mental states of a man interacting with a woman. Based on this, our results would imply that the ingroup targets in the MASC on average do not display the same mental state as the majority of participants, because ingroup interactions showed lower ToM than outgroup interactions. This reduced ToM in ingroup interactions was present in both the clinical and the HC sample. This could imply that the mental states presented in the MASC items are neither typical for clinical nor HC participants. However, the results are difficult to interpret since item gender was not systematically manipulated. The testing material (i.e., the MASC) was not counterbalanced and could potentially include confounds. For instance, the outgroup interactions in the MASC items often involve flirting, which might be more interesting for participants than other interactions, resulting in better performance.

Future measures should include an equal proportion of male and female targets to avoid interactions with participant gender. Previous research has repeatedly pointed out that it is crucial to select stimuli in a way that reflects reality (Burghardt & Bodansky, 2021; Fiedler, 2000). Biased inclusion of stimuli, for instance more female than male, could misrepresent cognition. Therefore, a lower inclusion of male / male interaction partners is problematic.

Within the clinical sample, items with male targets (sender and receiver) resulted in better ToM than items with female targets (sender and receiver). This is in line with findings suggesting that women can be perceived as less credible than men (e.g., Armstrong & McAdams, 2009; Kang et al., 2019). If women are less credible, it makes sense to pay less attention to them. However, these results did not replicate in the HC sample. It is unclear whether this non-replication was caused by the much smaller sample size or whether clinical

samples respond more strongly to gender than healthy participants. Somewhat in line with the latter interpretation are findings showing that alcohol consumption increases the effect of stereotypes (Bartholow et al., 2006; Schofield et al., 2015), making it possible that the patient group might use gender stereotypes more strongly. One of the male targets acts in a stereotypical masculine way and might be especially easy to predict.

The results among healthy individuals are limited by the small sample size of the HC group. Future studies should include larger samples of HC. The study did not apply an experimental design and can therefore not be interpreted causally. Interpretation of the findings is further complicated by the fact that no experimental studies have analyzed the effect similarity of two targets on ToM. Despite the limitations to interpret the causes of the target effects, the results are highly relevant for the MASC as a test instrument. Typically, items should be highly intercorrelated (Tang et al., 2014); still, the current analysis showed systematic variations between items based on target gender, which should be considered in future studies.

The current study demonstrates that performance on ToM tasks can be influenced by target gender, resulting in better ToM performance in interactions with male targets for clinical samples. Furthermore, ToM performance was better in items involving outgroup interactions compared to ingroup interactions. Thus, the results demonstrate the influence of interaction partners' gender. This influence may explain contradicting results in previous studies of ToM in clinical samples. The findings should inform the construction of future ToM measures.

References

Ames, D. R., Weber, E. U., & Zou, X. (2012). Mind-reading in strategic interaction: The impact of perceived similarity on projection and stereotyping. *Organizational Behavior and Human Decision Processes*, *117*(1), 96–110. https://doi.org/10.1016/j.obhdp.2011.07.007

Andreou, C., Kelm, L., Bierbrodt, J., Braun, V., Lipp, M., Yassari, A. H., & Moritz, S. (2015). Factors contributing to social cognition impairment in borderline personality disorder and schizophrenia. *Psychiatry Research*, *229*(3), 872–879. https://doi.org/10.1016/j.psychres.2015.07.057

Armstrong, C. L., & McAdams, M. J. (2009). Blogs of Information: How Gender Cues and Individual Motivations Influence Perceptions of Credibility. *Journal of Computer-Mediated Communication, 14*(3), 435–456. https://doi.org/10.1111/j.1083-6101.2009.01448.x

Bartholow, B. D., Dickter, C. L., & Sestir, M. A. (2006). Stereotype activation and control of race bias: Cognitive control of inhibition and its impairment by alcohol. *Journal of Personality and Social Psychology, 90*(2), 272–287. https://doi.org/10.1037/0022-3514.90.2.272

Benito-Ruiz, G., Luzón-Collado, C., Arrillaga-González, J., & Lahera, G. (2022). Development of an Ecologically Valid Assessment for Social Cognition Based on Real Interaction: Preliminary Results. *Behavioral Sciences*, *12*(2), Article 2. https://doi.org/10.3390/bs12020054

Bora, E., & Zorlu, N. (2017). Social cognition in alcohol use disorder: A meta-analysis. *Addiction, 112*(1), 40–48. https://doi.org/10.1111/add.13486

Burghardt, J., & Bodansky, A. N. (2021). Why Psychology Needs to Stop Striving for Novelty and How to Move Towards Theory-Driven Research. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.609802

Burghardt, J., Gradl, S., Knopp, M., & Sprung, M. (2023). Psychopathology and Theory of Mind in patients with personality disorders. *Borderline Personality Disorder and Emotion Dysregulation*, *10*(1), 18. https://doi.org/10.1186/s40479-023-00224-1

Chen, F. F., & Kenrick, D. T. (2002). Repulsion or attraction? Group membership and assumed attitude similarity. *Journal of Personality and Social Psychology*, *83*(1), 111–125. https://doi.org/10.1037/0022-3514.83.1.111

Cotter, J., Granger, K., Backx, R., Hobbs, M., Looi, C. Y., & Barnett, J. H. (2018). Social cognitive dysfunction as a clinical marker: A systematic review of meta-analyses across 30 clinical conditions. *Neuroscience & Biobehavioral Reviews*, *84*, 92–99. https://doi.org/10.1016/j.neubiorev.2017.11.014

Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., Kessler, J., Woike, J. K., Wolf, O. T., & Convit, A. (2006). Introducing MASC: A Movie for the Assessment of Social Cognition. *Journal of Autism and Developmental Disorders*, *36*(5), 623–636. https://doi.org/10.1007/s10803-006-0107-0

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fiedler, K. (2000). Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychological Review*, *107*(4), 659–676. https://doi.org/10.1037/0033-295X.107.4.659

Fossati, A., Borroni, S., Dziobek, I., Fonagy, P., & Somma, A. (2018). Thinking about assessment: Further evidence of the validity of the Movie for the Assessment of

Social Cognition as a measure of mentalistic abilities. *Psychoanalytic Psychology, 35*(1), 127–141. https://doi.org/10.1037/pap0000130

Gilbert, P. A., Pro, G., Zemore, S. E., Mulia, N., & Brown, G. (2019). Gender Differences in Use of Alcohol Treatment Services and Reasons for Nonuse in a National Sample. *Alcoholism: Clinical and Experimental Research, 43*(4), 722–731. https://doi.org/10.1111/acer.13965

Greenfield, S. F. (2002). Women and Alcohol Use Disorders. *Harvard Review of Psychiatry, 10*(2), 76–85. https://doi.org/10.1080/10673220216212

Hanegraaf, L., van Baal, S., Hohwy, J., & Verdejo-Garcia, A. (2021). A systematic review and meta-analysis of 'Systems for Social Processes' in borderline personality and substance use disorders. *Neuroscience & Biobehavioral Reviews, 127,* 572–592. https://doi.org/10.1016/j.neubiorev.2021.04.013

Harrison, D. A., Price, K. H., Gavin, J. H., & Florey, A. T. (2002). Time, Teams, and Task Performance: Changing Effects of Surface- and Deep-Level Diversity on Group Functioning. *The Academy of Management Journal, 45*(5), 1029–1045. https://doi.org/10.2307/3069328

Henry, J. D., Phillips, L. H., Ruffman, T., & Bailey, P. E. (2013). A meta-analytic review of age differences in theory of mind. *Psychology and Aging, 28*(3), 826–839. https://doi.org/10.1037/a0030677

Herpertz, S. C., & Bertsch, K. (2014). The social-cognitive basis of personality disorders. *Current Opinion in Psychiatry, 27*(1), 73–77. https://doi.org/10.1097/YCO.0000000000000026

Holzner, B., Giesinger, J. M., Pinggera, J., Zugal, S., Schöpf, F., Oberguggenberger, A. S., Gamper, E. M., Zabernigg, A., Weber, B., & Rumpold, G. (2012). The Computer-based Health Evaluation Software (CHES): A software for electronic patient-reported

outcome monitoring. *BMC Medical Informatics and Decision Making*, *12*, 126. https://doi.org/10.1186/1472-6947-12-126

Kang, J.-A., Hubbard, G., & Hong, S. (2019). Gender and credibility in branded storytelling. *Gender in Management: An International Journal*, *ahead-of-print*. https://doi.org/10.1108/GM-02-2019-0015

Klauer, K. C., Hölzenbein, F., Calanchini, J., & Sherman, J. W. (2014). How malleable is categorization by race? Evidence for competitive category use in social categorization. *Journal of Personality and Social Psychology*, *107*(1), 21–40. https://doi.org/10.1037/a0036609

Klauer, K. C., & Wegener, I. (1998). Unraveling social categorization in the "who said what?" paradigm. *Journal of Personality and Social Psychology, 75*(5), 1155–1178. https://doi.org/10.1037//0022-3514.75.5.1155

Knopp, M., Burghardt, J., Meyer, B., Riffer, F., & Sprung, M. (2022). How Individual and Situational Factors Influence Measures of Affective and Cognitive Theory of Mind in Psychiatric Inpatients. *Frontiers in Psychology*, *13*. https://www.frontiersin.org/article/10.3389/fpsyg.2022.855038

Kunda, Z., & Spencer, S. J. (2003). When do stereotypes come to mind and when do they color judgment? A goal-based theoretical framework for stereotype activation and application. *Psychological Bulletin*, *129*(4), 522–544. https://doi.org/10.1037/0033-2909.129.4.522

Lincoln, J. R., & Miller, J. (1979). Work and Friendship Ties in Organizations: A Comparative Analysis of Relation Networks. *Administrative Science Quarterly*, *24*(2), 181–199. https://doi.org/10.2307/2392493

McDonald, S. (2011). What's in the "old boys" network? Accessing social capital in gendered and racialized networks. *Social Networks*, *33*(4), 317–330. https://doi.org/10.1016/j.socnet.2011.10.002

McLaren, V., Gallagher, M., Hopwood, C. J., & Sharp, C. (2022). Hypermentalizing and Borderline Personality Disorder: A Meta-Analytic Review. *American Journal of Psychotherapy*, *75*(1), 21–31. https://doi.org/10.1176/appi.psychotherapy.20210018

McPherson, J. M., & Smith-Lovin, L. (1987). Homophily in Voluntary Organizations: Status Distance and the Composition of Face-to-Face Groups. *American Sociological Review*, *52*(3), 370–379. https://doi.org/10.2307/2095356

Németh, N., Mátrai, P., Hegyi, P., Czéh, B., Czopf, L., Hussain, A., Pammer, J., Szabó, I., Solymár, M., Kiss, L., Hartmann, P., Szilágyi, Á. L., Kiss, Z., & Simon, M. (2018). Theory of mind disturbances in borderline personality disorder: A meta-analysis. *Psychiatry Research*, *270*, 143–153. https://doi.org/10.1016/j.psychres.2018.08.049

Normann-Eide, E., Antonsen, B. T., Kvarstein, E. H., Pedersen, G., Vaskinn, A., & Wilberg, T. (2020). Are Impairments in Theory of Mind Specific to Borderline Personality Disorder? *Journal of Personality Disorders*, *34*(6), 827–841. https://doi.org/10.1521/pedi_2019_33_417

Onuoha, R. C., Quintana, D. S., Lyvers, M., & Guastella, A. J. (2016). A Meta-analysis of Theory of Mind in Alcohol Use Disorders. *Alcohol and Alcoholism*, *51*(4), 410–415. https://doi.org/10.1093/alcalc/agv137

Phillips, K. W. (2003). The effects of categorically based expectations on minority influence: The importance of congruence. *Personality & Social Psychology Bulletin*, *29*(1), 3–13. https://doi.org/10.1177/0146167202238367

Phillips, K. W., Duguid, M., Thomas-Hunt, M., & Uparna, J. (2013). *Diversity as Knowledge Exchange: The Roles of Information Processing, Expertise, and Status* (Q. M.

Roberson, Ed.). Oxford University Press. https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199736355.001.0001/oxfordhb-9780199736355-e-9

Phillips, K. W., & Loyd, D. L. (2006). When surface and deep-level diversity collide: The effects on dissenting group members. *Organizational Behavior and Human Decision Processes*, *99*(2), 143–160. https://doi.org/10.1016/j.obhdp.2005.12.001

Quesque, F., & Rossetti, Y. (2020). What Do Theory-of-Mind Tasks Actually Measure? Theory and Practice. *Perspectives on Psychological Science*, *15*(2), 384–396. https://doi.org/10.1177/1745691619896607

Rosello, B., Berenguer, C., Baixauli, I., García, R., & Miranda, A. (2020). Theory of Mind Profiles in Children With Autism Spectrum Disorder: Adaptive/Social Skills and Pragmatic Competence. *Frontiers in Psychology*, *11*, 567401. https://doi.org/10.3389/fpsyg.2020.567401

Sabbagh, M. A., Bowman, L. C., Evraire, L. E., & Ito, J. M. B. (2009). Neurodevelopmental Correlates of Theory of Mind in Preschool Children. *Child Development*, *80*(4), 1147–1162. https://doi.org/10.1111/j.1467-8624.2009.01322.x

Sanvicente-Vieira, B., Romani-Sponchiado, A., Kluwe-Schiavon, B., Brietzke, E., Araujo, R. B., & Grassi-Oliveira, R. (2017). Theory of Mind in Substance Users: A Systematic Minireview. *Substance Use & Misuse*, *52*(1), 127–133. https://doi.org/10.1080/10826084.2016.1212890

Schofield, T., Unkelbach, C., & Denson, T. (2015). Alcohol consumption increases bias to shoot at Middle Eastern but not White targets. *Group Processes & Intergroup Relations*, *20*. https://doi.org/10.1177/1368430215603461

Sheehan, D. V. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The Development and Validation of a Structured Diagnostic Psychiatric Interview for DSM-IV and ICD-10. *J Clin Psychiatry, 59*(Suppl 20), 22–33.

Simpson, A. J., & Todd, A. R. (2017). Intergroup visual perspective-taking: Shared group membership impairs self-perspective inhibition but may facilitate perspective calculation. *Cognition*, *166*, 371–381. https://doi.org/10.1016/j.cognition.2017.06.003

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research. *Neuroscience & Biobehavioral Reviews*, *30*(6), 855–863. https://doi.org/10.1016/j.neubiorev.2006.06.011

Skodol, A. E., & Bender, D. S. (2003). Why are women diagnosed borderline more than men? *The Psychiatric Quarterly*, *74*(4), 349–360. https://doi.org/10.1023/a:1026087410516

Sudo, M., & Farrar, J. (2020). Theory of mind understanding, but whose mind? Affiliation with the target is related to children's false belief performance. *Cognitive Development*, *54*, 100869. https://doi.org/10.1016/j.cogdev.2020.100869

Tang, W., Cui, Y., & Babenko, O. (2014). Internal consistency: Do we really know what it is and how to assess it? *Journal of Psychology and Behavioral Science*, *2*, 205–220.

Wittchen, H.-U., Zaudig, M., & Fydrich, T. (1997). *SKID. Strukturiertes Klinisches Interview für DSM-IV. Achse I und II. Handanweisung*. https://pure.mpg.de/pubman/faces/ViewItemOverviewPage.jsp?itemId=item_1646481

World-Medical-Association. (2013). World medical association declaration of helsinki: Ethical principles for medical research involving human subjects. *JAMA*, *310*(20), 2191–2194. https://doi.org/10.1001/jama.2013.281053

4. Article 3: Affective and cognitive Theory of Mind in patients with Alcohol Use Disorder: Associations with symptoms of depression, anxiety, and somatization

Reference

Knopp, M., Burghardt, J., Oppenauer, C., Meyer, B., Moritz, S., & Sprung, M. (2023). Affective and cognitive Theory of Mind in patients with Alcohol Use Disorder: Associations with symptoms of depression, anxiety, and somatization. *Journal of Substance Use and Addiction Treatment.* https://doi.org/10.1016/j.josat.2023.209227

Abstract

Introduction: Theory of Mind (ToM) is the ability to ascribe thoughts (cognitive ToM) and feelings (affective ToM) to others. There is ample evidence for impairments of affective and cognitive ToM in individuals with alcohol use disorder (AUD), however, evidence regarding changes of these impairments during AUD treatment and their possible relationship to comorbid symptoms is ambiguous. The present study analyzed changes in ToM during treatment and tested associations with comorbid symptoms of depression, anxiety, somatization, and social functioning. Methods: We analyzed data from 175 individuals with AUD. ToM and comorbid symptoms of depression, anxiety, somatization, and social functioning were assessed at the time of admission and at the time of discharge from an approximately 60 days long abstinence-oriented inpatient treatment. Affective and cognitive ToM were assessed using the Movie for the Assessment of Social Cognition, a measure with high ecological validity. Results: All symptoms, total and cognitive ToM improved following treatment, however affective ToM did not improve. Moreover, cognitive ToM at the beginning of treatment was associated to improved symptoms of depression and somatization, while affective ToM was not. Conclusions: Our study shows improvements in total and cognitive ToM as well as symptoms of depression, anxiety, somatization, and social functioning following long-term treatment. Furthermore, cognitive ToM was related to improvements in comorbid symptoms. This suggests that ToM may be an important treatment target in patients with AUD.

MK's contributions to the article: MK designed and conceptualized the study. MK conducted the data curation and data analyses. MK prepared the first draft, revised the manuscript, and prepared the final version.

157 (2024) 209227

Contents lists available at ScienceDirect

Journal of Substance Use and Addiction Treatment

journal homepage: www.journals.elsevier.com/journal-of-substance-use-and-addiction-treatment

Affective and cognitive Theory of Mind in patients with alcohol use disorder: Associations with symptoms of depression, anxiety, and somatization

Magdalena Knopp [a,b,*], Juliane Burghardt [a], Claudia Oppenauer [a], Bernhard Meyer [a], Steffen Moritz [c], Manuel Sprung [a,b,c,d]

[a] Karl Landsteiner University of Health Sciences, Division of Clinical Psychology, Dr.-Karl-Dorrek-Straße 30, 3500 Krems an der Donau, Austria
[b] Ludwig-Maximilians-Universität München, Faculty of Psychology and Educational Sciences, Department of Psychology, Leopoldstraße 13, 80802 München, Germany
[c] University Medical Center Hamburg-Eppendorf, Department of Psychiatry and Psychotherapy, Martinistrasse 52, 20246 Hamburg, Germany
[d] Psychiatric Rehabilitation Clinic Gars am Kamp, Psychosomatisches Zentrum Waldviertel, Kremserstraße 656, 3571 Gars am Kamp, Austria

ARTICLE INFO

Keywords:
Theory of mind
Social cognition
Alcohol use disorder
Substance use disorder
Treatment outcome

ABSTRACT

Introduction: Theory of Mind (ToM) is the ability to ascribe thoughts (cognitive ToM) and feelings (affective ToM) to others. Ample evidence exists for impairments of *affective and cognitive* ToM in individuals with alcohol use disorder (AUD); however, evidence regarding changes of these impairments during AUD treatment and their possible relationship to comorbid symptoms is ambiguous. The current study analyzed changes in ToM during treatment and tested associations with comorbid symptoms of depression, anxiety, somatization, and social functioning.

Methods: We analyzed data from 175 individuals with AUD. The study assessed ToM and comorbid symptoms of depression, anxiety, somatization, and social functioning at the time of admission and at the time of discharge from an approximately 60 days long abstinence-oriented inpatient treatment. We assessed affective and cognitive ToM using the Movie for the Assessment of Social Cognition, a measure with high ecological validity.

Results: All symptoms, total and cognitive ToM improved following treatment; however, affective ToM did not improve. Moreover, cognitive ToM at the beginning of treatment was associated with improved symptoms of depression and somatization, while affective ToM was not.

Conclusions: Our study shows improvements in total and cognitive ToM as well as symptoms of depression, anxiety, somatization, and social functioning following long-term treatment. Furthermore, cognitive ToM was related to improvements in comorbid symptoms. This finding suggests that ToM may be an important treatment target in patients with AUD.

1. Introduction

Alcohol use disorder (AUD) is a (chronic) mental health condition characterized by detrimental drinking patterns associated with negative emotional, physical, and psychosocial consequences (Tawa et al., 2016). Many individuals with AUD or other substance use disorders suffer from impairments in Theory of Mind (ToM), i.e. the ability to infer mental states of others (Bora & Zorlu, 2017; Hanegraaf et al., 2021; Onuoha et al., 2016; Sanvicente-Vieira et al., 2017). ToM impairments may represent a root cause for interpersonal problems and low social adjustment of individuals with AUD (Perez et al., 2022; Schmidt et al., 2016). ToM abilities can be subdivided into an affective and a cognitive component (Shamay-Tsoory & Aharon-Peretz, 2007). Affective ToM refers to inferences about emotions and feelings. Cognitive ToM refers to understanding beliefs or intentions.

Four systematic reviews and meta-analyses confirm that ToM impairments exist in individuals with AUD compared to healthy controls, with medium to high effect size (Bora & Zorlu, 2017; Hanegraaf et al., 2021; Onuoha et al., 2016; Sanvicente-Vieira et al., 2017). However, it is yet unclear whether both affective and cognitive components of ToM are

* Corresponding author at: Karl Landsteiner University of Health Sciences, Division of Clinical Psychology, Dr.-Karl-Dorrek-Straße 30, 3500 Krems an der Donau, Austria.
E-mail address: Magdalena.knopp@kl.ac.at (M. Knopp).

https://doi.org/10.1016/j.josat.2023.209227
Received 2 March 2023; Received in revised form 23 August 2023; Accepted 14 November 2023
Available online 21 November 2023

impaired in AUD. Many previous studies only assessed affective ToM and all of these studies (e.g. Kopera et al., 2020; Mátyássy et al., 2006; Rupp et al., 2021), with one exception (Kornreich et al., 2011), found that affective ToM was impaired. In the majority of these studies, the Reading-the-Mind-in-the-Eyes Test (RMET) was used to assess affective ToM. However, it has been disputed whether the RMET actually measures affective ToM or rather lower-level processes such as emotion recognition (Kittel et al., 2022; Quesque & Rossetti, 2020). The results regarding cognitive ToM impairments are more mixed. While several previous studies found impairments in cognitive as well as affective ToM (Bosco et al., 2014; Cox et al., 2018; Schmid et al., 2022; Thoma et al., 2013), some studies reported that only affective ToM was impaired in individuals with AUD, whereas cognitive ToM was comparable to healthy controls (Maurage et al., 2016; Nandrino et al., 2014). Pabst et al. (2022) pointed out that many ToM tasks used to study ToM in individuals with AUD lack ecological validity, which limits the conclusions that can been drawn from the available evidence. In contrast, the Movie for the Assessment of Social Cognition test (MASC; Dziobek et al., 2006) is a ToM measure with high ecological validity. To date only one study has used the MASC to assess the affective and cognitive ToM impairments of individuals with AUD (Maurage et al., 2016).

Even though the findings regarding cognitive ToM deficits are mixed, substantial evidence indicates that AUD patients show ToM deficits prior to treatment. Less is known about ToM changes following treatment. Since treatment typically requires abstinence, it is useful to consult studies about abstinence. One study showed that affective ToM (measured with the RMET) and facial emotion recognition can remit in individuals with AUD following a longer time (>6 months) of abstinence (Mátyássy et al., 2006). However, these findings are contradicted by the results of a meta-analysis by Bora and Zorlu (2017), which showed no ToM differences between AUD patients who were recently detoxified (<2 months) and those with a longer period of abstinence (>2 months), suggesting that longer abstinence did not lead to substantial improvements of ToM compared to shorter abstinence. In contrast to previous findings, the current study focuses on changes in ToM impairment in response to treatment. One previous study showed that treatment significantly increased ToM performance in patients with different mental disorders (Hayden et al., 2018). Regarding AUD, only two studies have addressed whether the affective ToM improves during treatment. One recent study by Rupp et al. (2021) found that impairments in affective ToM did not improve in the course of a long-term (2 months) abstinence-oriented inpatient treatment. In contrast, the study by Frileux et al. (2020) showed that emotion recognition impairments, as one aspect of affective ToM, (measured with the RMET) did improve following a 3-week long alcohol detoxification treatment. To date no studies have investigated whether cognitive ToM abilities improve during AUD treatment.

Individuals with AUD frequently suffer from comorbid symptoms of depression, anxiety, and somatization (Castillo-Carniglia et al., 2019; Hassan & Ali, 2011; Li et al., 2020; Oliveira et al., 2018). ToM deficits have been linked to more severe symptoms of depression, anxiety, and somatization (Bora & Berk, 2016; Chevalier et al., 2023; Plana et al., 2014; Preis et al., 2017; Thamby et al., 2019). Again, dissociations between affective and cognitive ToM are possible. For instance, stronger affective ToM deficits have been linked to stronger somatization symptoms, while the study found no association in regard to cognitive ToM (Thamby et al., 2019). To date, no studies have investigated whether improvements of comorbid symptoms of depression, anxiety, and somatization during AUD treatment are related to ToM abilities among patients with AUD. However, a study investigated deficits in social cognition (referred to as preoperational thinking) in individuals with depressive symptoms and found that impaired social cognition was associated with less improvement in the depressive symptoms of patients with depression (Sondermann et al., 2020). The impairments in cognitive ToM observed in patients with AUD are akin to preoperational thinking (i.e., a lack of the ability to see the perspective of others). Given ToM deficits in AUD patients, it seems relevant to consider how ToM relates to psychotherapy of patients with AUD. It is possible that individuals with different ToM abilities differ in their capacity to engage in psychotherapeutic processes. Patients with higher ToM abilities might show higher improvements in symptom severity. In line with this, impairments of lower-level social cognition (i.e. emotion recognition) in individuals with AUD have been linked to more treatment dropouts (Rupp et al., 2017). Among patients with other mental disorders, social cognition deficits were associated with lower improvements in symptom severity during psychotherapeutic treatment (Kvarstein et al., 2020; Sondermann et al., 2020). This suggests that patients with higher social cognition abilities benefit more from psychotherapy than patients with lower social cognition abilities. Alternatively, anxiety was suggested to inhibit high-order cognitive functions (Bateman & Fonagy, 2019), since ToM relies on high-order cognitive functions, a reduction of anxiety could reduce ToM deficits. Thus, improvements in comorbid symptoms during treatment might also lead to improved ToM or alternatively, require ToM. Accordingly, the current study seeks to test whether ToM at the beginning of treatment is associated to changes in symptom severity.

Moreover, only a few studies have investigated the association between ToM and social functioning in patients with AUD. One study showed that interpersonal abilities, as one aspect of social functioning, were related to heavy alcohol consumption (Lannoy et al., 2020). This is in line with a study by Kornreich et al. (2002), which showed that individuals with AUD (vs. community sample) had both more interpersonal problems and deficits in emotional facial expression recognition, which is one component of ToM. Further, difficulties describing feelings were associated to social functioning among psychiatric groups (Ospina et al., 2019). Social functioning has repeatedly been linked to ToM in several psychiatric samples (Bora et al., 2006; Rosello et al., 2020; Wang et al., 2015). One longitudinal study by Hayden et al. (2018), which involved patients with different mental disorders, found that improved ToM skills were linked to improvements in interpersonal problems during inpatient treatment. Therefore, impaired ToM should also be linked to deficits in social functioning among AUD patients.

1.1. The current study

The present study assessed changes in ToM and associations between ToM at admission and comorbid symptoms during an inpatient AUD treatment. The study relied on a naturalistic clinical care setting. First, we tested whether symptom severity (depression, anxiety, somatization, and social functioning) would improve during the psychotherapeutic treatment as an indicator for treatment success. Previous, studies have repeatedly shown dissociations between affective and cognitive ToM. However, these dissociations might simply be the result of measuring affective and cognitive ToM with different ToM tasks. To rule out this possibility, the current study measured affective and cognitive ToM within the same task using a measure with high ecological validity. Therefore, secondly, we collected data on total ToM, affective and cognitive ToM at the beginning and end of an inpatient abstinence-oriented treatment in a large sample of patients with AUD. Moreover, we explored whether affective and cognitive ToM would improve during treatment. Finally, we tested whether affective and cognitive ToM at admission were associated to changes in comorbid symptoms in the course of treatment for AUD. Based on previous literature, we hypothesized that better ToM at admission should be associated to stronger improvements in comorbid symptoms of depression, anxiety, somatization, and social functioning.

2. Materials and methods

2.1. Participants

Initially the study collected data from 362 patients from an Austrian

inpatient psychiatric-psychosomatic hospital. All patients were treated for AUD. We excluded 10 patients, because they reported scores below the cut-off (≥8) in the Alcohol Use Disorder Identification Test, indicating that the alcohol related problems of these patients were not clinically relevant. Moreover, we excluded data due to missing values at discharge. Data from 175 patients (61 % males) remained for further analyses. Patients' age ranged from 23 to 72 years, with an average age of 48.6 years ($SD = 8.7$) years. Regarding their education level, 26 patients (14.5 %) had a low education level (i.e., compulsory school education), 130 patients (75 %) had a medium education level (i.e., secondary school education), and 19 patients (10.5 %) had a high education level (i.e., college education). At the time of admission 34 % were unemployed, 19 % were retired and 36 % were employed. Further, 31 % were divorced, 22 % were living alone and 46 % were in a relationship. The primary diagnoses, made by a physician based on the current version of the WHO International Classification of Diseases (ICD-10; World Health Organization, 1992), were F10.2 alcohol dependence syndrome (97.2 %) or F10.1 harmful use of alcohol (2.8 %). Sixty-three percent of the patients had received a diagnosis of AUD at least two years prior to the current assessment, while 28 % of the patients had received a diagnosis during the last year. Patients with an acute psychosis, acute risk of suicidal behavior or current intoxication are not admitted to treatment in the psychiatric-psychosomatic hospital.

2.2. Procedure and materials

In our naturalistic longitudinal observational study, we retrospectively analyzed data collected from all abstinent patients treated in the AUD treatment units between June 2017 and April 2021 during routine clinical care. We collected all data reported in this study during the routine outcome monitoring at admission (T1) and discharge (T2). In cases where patients had multiple treatments in the respective treatment center the study used the data from the first treatment, with complete outcome monitoring. The study used the Computer-based Health Evaluation System, an electronic outcome monitoring software tool (CHES; Holzner et al., 2012), to implement the questionnaires and tests specified below. The assessment at admission and discharge lasted 2 h each, divided into two one-hour sessions. The study assessed ToM performance during the second one-hour session, while we investigated symptom severity in the first one-hour session.

The Movie for the Assessment of Social Cognition (MASC; Dziobek et al., 2006) shows a 15-minute movie about four people coming together for a dinner party. At multiple points throughout the video participants are given 45 multiple choice test questions regarding the four movie characters' emotions (affective ToM), and thoughts/intensions (cognitive ToM). We used all items for the total ToM. Following the approach suggested in previous MASC investigations (Buhlmann et al., 2015; Preißler et al., 2010; Ritter et al., 2011), the study team selected 15 items specially focused on affective ToM and 18 items specially focused on cognitive ToM from the 45 items to be used in the analysis of the subscales. The subscale score is determined by counting the correct responses corresponding to the associated items. The study did not assign the remaining 12 items to either the affective or cognitive facets. Additionally, six control questions are included in the MASC, to control variables such as understanding of the task, memory, and general compliance. Previously, the MASC demonstrated a high reliability, good internal consistency (Cronbach's alpha $\alpha \geq 0.82$ for all scores) and good convergent validity (Fossati et al., 2018). The internal consistency in the current study was also acceptable ($\alpha = 0.79$). The MASC also has high ecological validity (Benito-Ruiz et al., 2022).

The Alcohol Use Disorder Identification Test (AUDIT) is a 10-item instrument for identifying risky or harmful alcohol consumption as well as alcohol dependence and abuse (Barbor et al., 1992; Saunders et al., 1993). We used the German version with good psychometric properties (Dybek et al., 2006). The internal consistency in the current study was acceptable ($\alpha = 0.77$ for total AUDIT, $\alpha = 0.72$ for AUDIT-C and $\alpha = 0.71$ for alcohol-related problems).

The Patient Health Questionnaire-9 (PHQ-9) covers nine items to assess depressive symptoms such as lack of interest, depressed mood, sleep problems. The PHQ-9 is the depression module, evaluating and assigning scores to all 9 criteria outlined in the DSM-IV for depression assessment. The PHQ-9 previously showed good psychometric properties, with a good internal consistency ($\alpha = 0.86$), good test–retest reliability ($r = 0.84$), good convergent and discriminant validity and good sensitivity to change (Beard et al., 2016; Kroenke et al., 2010). The current study used the German version of the PHQ-9, with respectively good psychometric properties (Gräfe et al., 2004). Patients who scored above the cut-off of 10 on the PHQ-9, indicating a moderate level of depression (Kroenke et al., 2001). The internal consistency in the current study was good ($\alpha = 0.88$).

The Generalized Anxiety Disorder-7 (GAD-7) measures with seven questions assessing different symptoms of anxiety disorders in the last two weeks, e.g., nervousness, excessive worries, and incapability to relax. The internal consistency of the GAD-7 is excellent with $\alpha = 0.92$ and the test-retest reliability of $\alpha = 0.83$ is good (Spitzer et al., 2006). Scores to detect a moderate anxiety disorder are reported by Spitzer et al. (2006).The internal consistency in the current study was good ($\alpha = 0.89$).

The Patient Health Questionnaire (PHQ-15) measures somatic symptom severity and includes 15 items for women and 14 items for men that measure physical symptoms on a 3-point Likert scale during the last 4 weeks (Lowe et al., 2008). Cut-offs are demonstrated by Kroenke et al. (2002). In the present study, the PHQ-15 showed good reliability, with a $\alpha = 0.85$.

The World Health Organization disability assessment schedule (*WHODAS-12*) includes 12-items. The WHODAS-12 is used to assess patients' limitations in functioning and participation in social life in the last 30 days (Küçükdeveci et al., 2013). Based on the theoretical framework, the current study focussed on social functioning as assessed by three Items of the WHODAS-12, including Item 4, 10 and 11 ("How much of a problem did you have..." "joining in community activities?", "...dealing with people you do not know?" "...maintaining a friendship?"). The three items showed an acceptable internal consistency as indicated by a Cronbach's alpha $\alpha = 0.77$.

2.3. Treatment

In line with standard procedures in Austria, patients have completed a detoxification period (average duration of 10 days) in a specialized clinic prior to starting inpatient treatment in the psychiatric-psychosomatic hospital. Abstinence is a condition for admission at the inpatient said hospital. All participants received inpatient psychosocial treatment delivered by a multidisciplinary team of physicians, psychotherapists, clinical psychologists, nursing staff, physiotherapists, nutritionists, and social workers. The eclectic treatment comprised a structured program of 1–2 sessions of individual psychotherapy per week and 5 sessions of group psychotherapy. The interventions used in individual and group sessions were based on principles of cognitive behavioral therapy, skills training progressive muscle relaxation and mindfulness-based relapse prevention. In addition to psychosocial treatment many patients received psychopharmacological treatment. The average treatment duration was $M = 60.1$ days ($SD = 19.8$).

2.4. Data analysis

The study computed all analyses using IBM SPSS (Version 27), except for multilevel modeling, which we conducted in R 4.3.1 (http://cran-r-project.org/) with the packages lme4 and lmerTest for LME modeling (glmer, family: binomial), tidyverse, magrittr, dplyr and reshape2 for data preparation. We first analyzed whether significant improvements between admission and discharge occurred in the symptoms of depression, anxiety, somatization, and social functioning,

using a repeated measures analysis of variance (ANOVA) with Greenhouse-Geisser correction. Secondly, we investigated changes in ToM performance between admission and discharge by calculating two Logistic Linear Mixed Effects models (LME). The study assessed ToM changes as within-subject effects with dichotomous (correct vs. incorrect) MASC item answers as dependent variable. The first model included ToM facets (affective vs. cognitive), time (admission vs. discharge), and the interaction effects of ToM facets (affective vs. cognitive) × time (admission vs. discharge). Thereby, ToM was limited to items that can be described as affective or cognitive ToM. The second model used all items and contained only the factor time (admission vs. discharge). Finally, we calculated a series of linear regression analyses with MASC test scores at admission as the independent variable and changes between admission and discharge in the severity of depression, anxiety, somatic complains, and social functioning as the dependent variables. We conducted three separate regression analyses for the MASC total scores, the affective ToM score, and the cognitive ToM score. A p-values of $<.05$ was used as a threshold for all effects. However, to control for multiple testing we applied the Bonferroni-Holm p-value correction method. Different sample sizes resulted due to missing data for different analyses. The missing values resulted when participants did not take part in the discharge sessions, mostly for administrative and clinical reasons (e.g., scheduling conflict, sick leave of testing stuff, patients were discharged before the assessment was scheduled).

3. Results

3.1. Improvements in symptoms of depression, anxiety, somatization, and social functioning

The average severity of symptoms of depression, anxiety, somatization, and social functioning at admission and discharge is shown in Table 1. Symptoms improved significantly, indicated by an overall change in mean symptom levels between admission and discharge, $F(2.50, 322.54) = 134.61$, $p < .001$, $\eta_p^2 = 0.511$, indicating a large effect. The level of depressive symptoms at admission was moderate and improved to mild at discharge. The level of symptoms of anxiety at admission was mild and improved to minimal at discharge. The level of somatic complaints was mild at admission and improved to minimal at discharge. The level of deficits in social functioning at admission was moderate/mild and improved to minimal at discharge.

3.2. Improvement of ToM (MASC test scores) in the course of treatment

Table 2 displays the means and standard deviations of MASC test scores at admission and discharge. Table 3 presents the results of the two multilevel models. The first multilevel model tested main and interaction effects of ToM facet (affective vs. cognitive) and time (admission vs. discharge). The main effect of time and ToM facet were not significant. However, the interaction effect of ToM facet and time was significant, $\beta = 0.17$, $SE = 0.08$, $p = .036$. The regression model showed that only the cognitive ToM test scores improved significantly between admission and discharge. No evidence exists that the affective ToM improved between admission and discharge. The second regression model used all items and only tested the effect of time (admission vs. discharge) on ToM. The regression model showed a significant effect of time (admission vs. discharge), $\beta = 0.07$, $SE = 0.03$, $p = .037$, suggesting that total ToM improved during the course of the treatment.

Table 1
Symptoms of depression, anxiety, somatic symptoms, and social functioning for the AUD sample at admission and discharge.

Measures		Admission	Discharge	
	Max	*M* (*SD*)	*M* (*SD*)	*n*
Depression (PHQ-9)	27	10.77 (5.93)	5.92 (4.29)	135
Anxiety (GAD-7)	21	8.53 (5.08)	4.47 (3.91)	135
Somatic Symptoms (PHQ-15)	30	8.76 (5.61)	6.06 (4.82)	135
Social Functioning (WHODAS)	12	3.93 (2.77)	2.21 (2.38)	134

Note. PHQ-9: patient health questionnaire 9, GAD-7: generalized anxiety disorder 7, PHQ-15: patient health questionnaire 15, 3 Items (4, 10, 11) of the WHODAS 2.0: WHO Disability Assessment Schedule; Max: highest total score.

Table 2
MASC test scores for the AUD sample at admission and discharge.

		Admission	Discharge	
	Max	*M* (*SD*)	*M* (*SD*)	n
Total ToM	45	29.35 (5.88)	29.83 (6.85)	175
Affective ToM	15	9.65 (2.29)	9.43 (2.56)	175
Cognitive ToM	18	11.80 (2.69)	12.18 (3.26)	175

Note. AUD: Alcohol use disorder; ToM: Theory of Mind; MASC: Movie for the Assessment of Social Cognition, highest MASC score: max.

Table 3
Multilevel model of time and ToM facet on ToM performance.

	β	*SE*	*z*	*p*
Regression model 1				
ToM facets (cognitive vs. affective)	0.06	0.06	0.10	.318
Time (admission vs. discharge)	−0.07	0.06	−1.13	.258
Interaction of ToM facets & time	**0.17**	**0.08**	**2.09**	**.036**
Regression model 2				
Time (admission vs. discharge)	**0.07**	**0.03**	**2.09**	**.037**

Note. bold p-values are $p < .05$.

3.3. Associations between ToM and changes in symptoms of depression, anxiety, somatization, and social functioning

Table 4 presents the results of the three regression analyses with MASC test scores at admission as the independent variable and improvements of symptoms of depression, anxiety, somatization, and social functioning between admission and discharge. The regressions only show significant associations between the cognitive ToM score and improvements in symptoms of depression and somatization. The association between cognitive ToM and somatization was still significant after a Bonferroni-Holm correction, which resulted in a critical p-value of $p = .025$ however, the association between cognitive ToM and depression was no longer significant after the correction was applied. The study found no significant associations between the MASC ToM total and affective ToM scores and changes in any of the symptoms of depression, anxiety, somatization, or social functioning.

4. Discussion

The current study is the first investigating the role of affective and cognitive ToM abilities in the course of inpatient psychotherapeutic treatment of patients with AUD. In the course of treatment, symptoms of depression, anxiety, somatization, and social functioning improved. Additionally, total and cognitive ToM improved, but affective ToM abilities did not reliably change. Moreover, cognitive ToM at the beginning of treatment was associated with improvements in symptoms of depression and somatization, although the effect on symptoms of depression was no longer significant after a Bonferroni-Holm correction.

First, analyses of the current study show improvements in symptoms due to inpatient psychotherapeutic treatment. The results support previous evidence for symptoms improvement (depression, anxiety, and somatization) after a therapeutic inpatient treatment of four weeks (Haase et al., 2008) or even after two weeks (Kertz et al., 2015) for anxiety and depression symptoms.

Second, our results on ToM in the course of treatment suggest a

Table 4
Regression analyses of MASC test scores at admission and changes in symptoms of depression, anxiety, somatic symptoms, and social functioning.

	B	β	*t*	*p*	95 % Confidence interval of the difference	
					Lower limit	Upper limit
Total ToM						
Δ Depression (PHQ-9)	0.09	0.12	1.27	.207	−0.05	0.22
Δ Anxiety (GAD-7)	0.02	0.03	0.35	.727	−0.11	0.15
Δ Somatic Symptoms (PHQ-15)	0.06	0.10	1.08	.284	−0.05	0.17
Δ Social Functioning (WHODAS)	0.06	0.16	1.55	.123	−0.02	0.13
Affective ToM						
Δ Depression (PHQ-9)	0.06	0.03	0.33	.742	−0.28	0.39
Δ Anxiety (GAD-7)	−0.02	−0.01	−0.09	.928	−0.34	0.31
Δ Somatic Symptoms (PHQ-15)	0.02	0.01	0.14	.885	−0.26	0.30
Δ Social Functioning (WHODAS)	0.16	0.17	1.71	.090	−0.03	0.34
Cognitive ToM						
Δ Depression (PHQ-9)	**0.32**	**0.19**	**2.17**	**.032**	**0.03**	**0.61**
Δ Anxiety (GAD-7)	0.21	0.13	1.44	.152	−0.08	0.51
Δ Somatic Symptoms (PHQ-15)	**0.30**	**0.20**	**2.34**	**.021**	**0.05**	**0.55**
Δ Social Functioning (WHODAS)	0.12	0.15	1.57	.121	−0.03	0.28

Note. Δ difference between admission and discharge, PHQ-9 Patient Health Questionnaire-9, GAD-7 Generalized Anxiety Disorder 7, PHQ-15 Patient Health Questionnaire-15, 3 Items (4, 10, 11) WHODAS 2.0: WHO Disability Assessment Schedule; bold p-values are $p < .05$.

dissociation in the improvement of affective and cognitive ToM as only the cognitive ToM abilities improved in the course of an abstinence-based inpatient treatment, which lasted approximately 2 months. This is in line with a previous study, which showed that affective ToM abilities did not improve, even after a long-term abstinence-oriented inpatient treatment (Rupp et al., 2021). In contrast, a previous study by Frileux et al. (2020) reported an improvement in facial emotion recognition (measured with the RMET) following a 3-week long alcohol detoxification treatment. Emotion recognition is assumed to be a lower-level process related to affective ToM (Kittel et al., 2022; Quesque & Rossetti, 2020). Higher-level processes, such as affective ToM, which showed no improvement in the course of an inpatient treatment might be more resistant to change than low level-processes. Moreover, previous studies proposed that the affect processing system is impaired in patients with AUD (Le Berre, 2019). This implies impairments in a wide range of emotional abilities such as alexithymia, empathy, and affective ToM (Kornreich et al., 2002; Le Berre, 2019; Maurage et al., 2011, 2016), suggesting a broader, more comprehensive deficit in affect processing that might therefore be particularly difficult to address in AUD patients.

Third, the current study found that cognitive ToM abilities at the beginning of treatment were associated with improvements in comorbid symptoms of depression and somatization. However, since the results regarding the association between cognitive ToM and symptoms of depression were no longer significant after a Bonferroni-Holm correction, they should be considered with caution. However, they are in line with previous studies that found an association between comorbid symptoms of depression and ToM in patients with major depression (Bora & Berk, 2016; Chevalier et al., 2023). Additionally, these studies also found an association between symptoms of depression and affective ToM, not just cognitive ToM. This result could be specific to depression or due to differences between the studies such as different measures. Future studies should try to replicate the findings. Moreover, our findings complement results on the association of comorbid symptoms of somatization and ToM (Chevalier et al., 2023; Preis et al., 2017; Thamby et al., 2019). However, in contrast to the present study, which showed that cognitive ToM was linked to somatization symptoms, a previous study reported stronger affective ToM deficits were linked to stronger somatization symptoms (Thamby et al., 2019). Due to a small size (20 vs. 175), different measures (emotional content of animation tasks vs. MASC), and different clinical samples (somatoform disorders vs. AUD) the results are difficult to compare. Further studies are needed to clarify the role of affective/cognitive ToM and somatization/depression symptoms in AUD.

A previous study with patients with borderline personality disorders using the MASC also found that patients with higher cognitive ToM showed better improvement in symptoms and social functioning (Kvarstein et al., 2020). This fits well with findings in the current study and suggests that cognitive ToM abilities may be a particularly important facilitator in the treatment of patients with AUD and other emotional and behavioral disorders. This finding is also in line with the study by Sondermann et al. (2020), which reported that dysfunctional social cognition was linked to less improvement in depressive symptoms among patients with depression. Thus, it is possible that patients with higher ToM abilities benefit more from psychotherapy than patients with lower ToM abilities. Moreover, it is also possible that symptom improvements might have led to better ToM. Taken together the evidence suggests that multiple forms of social cognitive abilities are linked to improvements in symptoms. Therefore it might be useful for future studies to utilize the NIMH Research Domain Criteria (RDoC) 'Systems of Social Processes' framework to further investigate impairments in ToM and social cognition in individuals with AUD (see Hanegraaf et al., 2021). The RDoC dimensional framework emphasized ToM as a transdiagnostic factor underlying mental disorders. This is in line with the present transdiagnostic association of ToM and symptoms of depression and somatization in individuals with AUD.

It is possible that ToM specific interventions might be needed to sufficiently improve the ToM abilities of patients with AUD. Metacognitive therapy and training has been found to have promising effects for the improvement of ToM abilities in patients with various emotional and behavioral disorders, including AUD (Caselli et al., 2018; Philipp et al., 2018) and may also improve other treatment targets, such as alcohol cravings (Caselli et al., 2016). Moreover, recent evidence suggests that ToM impairments are not only present in many individuals with AUD but are also evident in individuals with subclinical levels of problematic alcohol consumption (Kumar et al., 2022). Given that impairments in ToM are a risk factor for treatment dropout and relapse in patients with AUD (Rupp et al., 2017), training of ToM abilities in subclinical populations may also serve as a preventive intervention for AUD (Kumar et al., 2022).

ToM is also impaired in individuals with other substance use disorders. ToM deficits have been found in cannabis (Roser et al., 2012), cocaine (Preller et al., 2014), opioid (Gandolphe et al., 2018), and methamphetamine users (Henry et al., 2009; Kim et al., 2011). A longitudinal study by Vonmoos et al. (2019) reported that cocaine users showed no ToM (measured with the MASC) change after one year, irrespective of whether they had increased or decreased their cocaine consumption. At the same time the HC in the referenced study showed a pronounced ToM increase, which is likely a training effect. Thus, patients might not typically show improvements without a specific treatment. Some research has suggested similarities between patients with

different substance-related disorders (Sanvicente-Vieira et al., 2017), future studies should test this assumption.

4.1. Limitations/future directions

A limitation of the current study is that it used only brief screening instruments (e.g., PHQ-9). The questionnaires are standardized and are all marked by high specificity and sensitivity to diagnose depression, generalized anxiety, and somatization. However, other measures assess symptom severity more comprehensively. Despite this, the current measures show high correlations to those other measures. For instance, the Beck-Depressions-Inventar-II and the PHQ-9 show a correlation of $r = 0.84$ (Dum et al., 2008). Another limitation of the present study is the absence at a healthy control group. It is, not possible to rule out that symptoms and ToM changed over time irrespective of treatment. Further it remains unclear whether ToM abilities improved to a healthy control level on the end of the treatment. However, the study is highly relevant to evaluate effects in a clinical care setting. Notably, ToM abilities may show higher improvements with higher session frequency or different treatments. Moreover, the treatment may have been too eclectic; a solely ToM-focused approach might have produced higher ToM improvements. Future studies should include the assessment of interpersonal problems, to address the link between ToM abilities, interpersonal problems, and symptoms of depression, anxiety, somatic complaints, and social functioning.

4.2. Conclusion

The study team conducted the current longitudinal study in a naturalistic setting to maximize the extent of clinical representativeness and enhance the external validity for clinical care conditions. Results indicate that affective ToM abilities in individuals with AUD did not improve due to an eclectic abstinence-oriented inpatient treatment, but the total ToM and cognitive ToM showed an improvement. More specific interventions may be necessary to obtain higher improvements in ToM abilities of patients with AUD. Further, in the course of treatment only the cognitive ToM abilities at the beginning of treatment were linked to improvements in symptoms of somatization and possibly depression. This finding suggests that cognitive ToM abilities may be a particularly promising treatment target in patients with AUD and we recommend incorporating interventions, such as metacognitive therapy and training into standard treatment protocols for patients with AUD.

Ethics

All participants consented to the use of their data. The analysis was approved by the ethics commission of the Karl Landsteiner University of Health Sciences (Nr: 1030/2021) and complies with the ethical standards of the Declaration of Helsinki (World-Medical-Association, 2013).

Funding

We are grateful to the Gesellschaft für Forschungsförderung Niederösterreich (GFF) [Society for Research Promotion Lower Austria] for funding this research as part of the endowed professorship in clinical psychology. The Open Access Publishing Fund of Karl Landsteiner University of Health Sciences, Krems, Austria supports the publication of this manuscript. The funders had no role in the study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

CRediT authorship contribution statement

MK and MS designed and conceptualized the study. MK conducted the data curation and data analyses. BM conducted additional analyses. JB and MS supervised the analyses of the data. MK prepared the first draft, with close supervision by MS. JB, CO, BM, SM and MS revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Declaration of competing interest

The authors declare that there are no conflicts of interest.

Data availability

The datasets presented in this article are not readily available because of the vulnerability of the study sample. Participants of this study did not agree for their data to be shared publicly, so supporting data is not available.

Acknowledgments

We would like to thank Beate Sodian for providing helpful feedback on earlier versions of this manuscript.

References

Bateman, A. W., & Fonagy, P. (2019). *Handbook of mentalizing in mental health practice*. American Psychiatric Pub.

Barbor, T. F., De La Fuente, J. R., Saunders, J., & Grant, M. (1992). *AUDIT: the alcohol use disorders identification test: guidelines for use in primary health care*. Geneva: World Health Organization.

Beard, C., Hsu, K. J., Rifkin, L. S., Busch, A. B., & Björgvinsson, T. (2016). Validation of the PHQ-9 in a psychiatric sample. *Journal of Affective Disorders, 193*, 267–273. https://doi.org/10.1016/j.jad.2015. 12.075

Benito-Ruiz, G., Luzón-Collado, C., Atrillaga-González, J., & Lahera, G. (2022). Development of an ecologically valid assessment for social cognition based on real interaction: Preliminary results. *Behavioral Sciences, 12*(2), Article 2. https://doi.org/10.3390/bs12020054

Bora, E., & Berk, M. (2016). Theory of mind in major depressive disorder: A meta-analysis. *Journal of Affective Disorders, 191*, 49–55. https://doi.org/10.1016/j.jad.2015.11.023

Bora, E., Eryavuz, A., Kayahan, B., Sungu, G., & Veznedaroglu, B. (2006). Social functioning, theory of mind and neurocognition in outpatients with schizophrenia; Mental state decoding may be a better predictor of social functioning than mental state reasoning. *Psychiatry Research, 145*(2), 95–103. https://doi.org/10.1016/j.psychres.2005.11.003

Bora, E., & Zorlu, N. (2017). Social cognition in alcohol use disorder: A meta-analysis. *Addiction, 112*(1), 40–48. https://doi.org/10.1111/add.13486

Bosco, F. M., Capozzi, F., Colle, L., Marostica, P., & Tirassa, M. (2014). Theory of mind deficit in subjects with alcohol use disorder: An analysis of mindreading processes. *Alcohol and Alcoholism, 49*(3), 299–307. https://doi.org/10.1093/alcalc/agt148

Buhlmann, U., Wacker, R., & Dziobek, I. (2015). Inferring other people's states of mind: Comparison across social anxiety, body dysmorphic, and obsessive–compulsive disorders. *Journal of Anxiety Disorders, 34*, 107–113. https://doi.org/10.1016/j.janxdis.2015.06.003

Caselli, G., Gemelli, A., Spada, M. M., & Wells, A. (2016). Experimental modification of perspective on thoughts and metacognitive beliefs in alcohol use disorder. *Psychiatry Research, 244*, 57–61. https://doi.org/10.1016/j.psychres.2016.07.029

Caselli, G., Martino, F., Spada, M. M., & Wells, A. (2018). Metacognitive therapy for alcohol use disorder: A systematic case series. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.02619

Castillo-Carniglia, A., Keyes, K. M., Hasin, D. S., & Cerdá, M. (2019). Psychiatric comorbidities in alcohol use disorder. *The Lancet. Psychiatry, 6*(12), 1068–1080. https://doi.org/10.1016/S2215-0366 (19)30222-6

Chevalier, V., Simard, V., & Achim, J. (2023). Meta-analyses of the associations of mentalization and proxy variables with anxiety and internalizing problems. *Journal of Anxiety Disorders, 95*, Article 102694. https://doi.org/10.1016/j.janxdis.2023.102694

Cox, S., Bertoux, M., Turner, J. J. D., Moss, A., Locker, K., & Riggs, K. (2018). Aspects of alcohol use disorder affecting social cognition as assessed using the Mini Social and Emotional Assessment (mini-SEA). *Drug and Alcohol Dependence, 187*, 165–170. https://doi.org/10.1016/j. drugalcdep.2018.03.004

Dum, M., Pickren, J., Sobell, L. C., & Sobell, M. B. (2008). Comparing the BDI-II and the PHQ-9 with outpatient substance abusers. *Addictive Behaviors, 33*(2), 381–387. https://doi.org/10.1016/j.addbeh. 2007.09.017

Dybek, I., Bischof, G., Grothues, J., Reinhardt, S., Meyer, C., Hapke, U., … Rumpf, H. J. (2006). The reliability and validity of the Alcohol Use Disorders Identification Test (AUDIT) in a German general practice population sample. *Journal of studies on alcohol, 67*(3), 473–481.

Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., … Convit, A. (2006). Introducing MASC: A movie for the assessment of social cognition. *Journal of Autism and Developmental Disorders, 36*(5), 623–636. https://doi.org/10.1007/s10803-006-0107-0

Fossati, A., Borroni, S., Dziobek, I., Fonagy, P., & Somma, A. (2018). Thinking about assessment: Further evidence of the validity of the movie for the assessment of social cognition as a measure of mentalistic abilities. *Psychoanalytic Psychology, 35*(1), 127–141. https://doi.org/10.1037/pap0000130

Frileux, S., Dubertret, C., & Le Strat, Y. (2020). Short-term impact of alcohol detoxification on facial emotions recognition. *L'Encéphale, 46*(5), 319–325. https://doi.org/10.1016/j.encep.2020.04.019

Gandolphe, M.-C., Lecluyse, B., Triquet, C., Brunelle, E., Duparcq, J.-P., & Nandrino, J.-L. (2018). Mind reading abilities in opiate-dependent patients: An exploratory study. *Comprehensive Psychiatry, 83*, 46–52. https://doi.org/10.1016/j.comppsych.2018.03.001

Gräfe, K., Zipfel, S., Herzog, W., & Löwe, B. (2004). Screening for psychiatric disorders with the Patient Health Questionnaire (PHQ). Results from the German validation study. *Diagnostica, 50*(4), 171–181. https://doi.org/10.1026/0012-1924.50.4.171

Haase, M., Frommer, J., Franke, G.-H., Hoffmann, T., Schulze-Muetzel, J., Jäger, S., Grabe, H.-J., Spitzer, C., & Schmitz, N. (2008). From symptom relief to interpersonal change: Treatment outcome and effectiveness in inpatient psychotherapy. *Psychotherapy Research, 18*(5), 615–624. https://doi.org/10.1080/10503300802192158

Hanegraaf, L., van Baal, S., Hohwy, J., & Verdejo-Garcia, A. (2021). A systematic review and meta-analysis of 'Systems for Social Processes' in borderline personality and substance use disorders. *Neuroscience & Biobehavioral Reviews, 127*, 572–592. https://doi.org/10.1016/j.neubiorev. 2021.04.013

Hassan, I., & Ali, R. (2011). The association between somatic symptoms, anxiety disorders and substance use. A literature review. *Psychiatric Quarterly, 82*(4), 315–328. https://doi.org/10.1007/s11126-011-9174-2

Hayden, M. C., Müllauer, P. K., Gaugeler, R., Senft, B., & Andreas, S. (2018). Improvements in mentalization predict improvements in interpersonal distress in patients with mental disorders. *Journal of Clinical Psychology, 74*(12), 2276–2286. https://doi.org/10.1002/jclp.22673

Henry, J. D., Mazur, M., & Rendell, P. G. (2009). Social-cognitive difficulties in former users of methamphetamine. *British Journal of Clinical Psychology, 48*(3), 323–327. https://doi.org/10.1111/j. 2044-8260.2009.tb00487.x

Holzner, B., Giesinger, J. M., Pinggera, J., Zugal, S., Schöpf, F., Oberguggenberger, A. S., … Rumpold, G. (2012). The Computer-based Health Evaluation Software (CHES): A software for electronic patient-reported outcome monitoring. *BMC Medical Informatics and Decision Making, 12*(1), 126. https://doi.org/10.1186/1472-6947-12-126

Kertz, S. J., Koran, J., Stevens, K. T., & Björgvinsson, T. (2015). Repetitive negative thinking predicts depression and anxiety symptom improvement during brief cognitive behavioral therapy. *Behaviour Research and Therapy, 68*, 54–63. https://doi.org/10.1016/j.brat.2015.03.006

Kim, Y.-T., Kwon, D.-H., & Chang, Y. (2011). Impairments of facial emotion recognition and theory of mind in methamphetamine abusers. *Psychiatry Research, 186*(1), 80–84. https://doi.org/10.1016/j. psychres.2010.06.027

Kittel, A. F. D., Olderbak, S., & Wilhelm, O. (2022). Sty in the mind's eye: A meta-analytic investigation of the nomological network and internal consistency of the "reading the mind in the eyes" test. *Assessment, 29*(5), 872–895. https://doi.org/10.1177/1073191121996469

Kopera, M., Zaorska, J., Trucco, E. M., Suszek, H., Kobyliński, P., Zucker, R. A., … Jakubczyk, A. (2020). Childhood trauma, alexithymia, and mental states recognition among individuals with alcohol use disorder and healthy controls. *Drug and Alcohol Dependence, 217*, Article 108301. https://doi.org/10.1016/j.drugalcdep.2020.108301

Kornreich, C., Delle-Vigne, D., Knittel, J., Nerincx, A., Campanella, S., Noel, X., Hanak, C., Verbanck, P., & Ermer, E. (2011). Impaired conditional reasoning in alcoholics: A negative impact on social interactions and risky behaviors? *Addiction (Abingdon, England), 106*(5), 951–959. https://doi.org/10.1111/j.1360-0443.2010.03346.x

Kornreich, C., Philippot, P., Foisy, M.-L., Blairy, S., Raynaud, E., Dan, B., Hess, U., Noël, X., Pelc, I., & Verbanck, P. (2002). Impaired emotional facial expression recognition is associated with interpersonal problems in alcoholism. *Alcohol and Alcoholism, 37*(4), 394–400. https://doi.org/10.1093/alcalc/37.4.394

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2002). The PHQ-15: Validity of a new measure for evaluating the severity of somatic symptoms. *Psychosomatic Medicine, 64*(2), 258–266.

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Kroenke, K., Spitzer, R. L., Williams, J. B. W., & Löwe, B. (2010). The patient health questionnaire somatic, anxiety, and depressive symptom scales: A systematic review. *General Hospital Psychiatry, 32*(4), 345–359. https://doi.org/10.1016/j.genhosppsych.2010.03.006

Küçükdeveci, A. A., Kutlay, Ş., Yıldızlar, D., Öztuna, D., Elhan, A. H., & Tennant, A. (2013). The reliability and validity of the World Health Organization Disability Assessment Schedule (WHODAS-II) in stroke. *Disability and Rehabilitation, 35*(3), 214–220. https://doi.org/10.3109/096 38288.2012.690817

Kumar, L., Skrzynski, C. J., & Creswell, K. G. (2022). Systematic review and meta-analysis on the association between theory of mind and alcohol problems in non-clinical samples. *Alcoholism: Clinical and Experimental Research*, 1–13. https://doi.org/10.1111/acer.14943

Kvarstein, E. H., Folmo, E., Antonsen, B. T., Normann-Eide, E., Pedersen, G., & Wilberg, T. (2020). Social cognition capacities as predictors of outcome in mentalization-based treatment (MBT). *Frontiers in Psychiatry, 11*. https://doi.org/10.3389/fpsyt.2020.00691

Lannoy, S., Gilles, F., Benzerouk, F., Henry, A., Oker, A., Raucher-Chéné, D., Besche-Richard, C., & Gierski, F. (2020). Disentangling the role of social cognition processes at early steps of alcohol abuse: The influence of affective theory of mind. *Addictive Behaviors, 102*, Article 106187. https://doi.org/10.1016/j.addbeh.2019.106187

Le Berre, A.-P. (2019). Emotional processing and social cognition in alcohol use disorder. *Neuropsychology, 33*(6), 808–821. https://doi.org/10.1037/neu0000572

Li, J., Wang, H., Li, M., Shen, Q., Li, X., Zhang, Y., Peng, J., Rong, X., & Peng, Y. (2020). Effect of alcohol use disorders and alcohol intake on the risk of subsequent depressive symptoms: A systematic review and meta-analysis of cohort studies. *Addiction, 115*(7), 1224–1243. https://doi.org/10.1111/add.14935

Lowe, B., Spitzer, R. L., Williams, J. B., Mussell, M., Schellberg, D., & Kroenke, K. (2008). Depression, anxiety and somatization in primary care: Syndrome overlap and functional impairment. *General Hospital Psychiatry, 30*(3), 191–199. https://doi.org/10.1016/j.genhosppsych.2008.01.001

Mátyássy, A., Kelemen, O., Sárközi, Z., Janka, Z., & Kéri, S. (2006). Recognition of complex mental states in patients with alcoholism after long-term abstinence. *Alcohol and Alcoholism, 41*(5), 512–514. https://doi.org/10.1093/alcalc/agl045

Maurage, P., D'Hondt, F., de Timary, P., Mary, C., Franck, N., & Peyroux, E. (2016). Dissociating affective and cognitive theory of mind in recently detoxified alcohol-dependent individuals. *Alcoholism: Clinical and Experimental Research, 40*(9), 1926–1934. https://doi.org/10.1111/acer.13155

Maurage, P., Grynberg, D., Noël, X., Joassin, F., Philippot, P., Hanak, C., … Campanella, S. (2011). Dissociation between affective and cognitive empathy in alcoholism: A specific deficit for the emotional dimension. *Alcoholism: Clinical and Experimental Research, 35*(9), 1662–1668. https://doi.org/10.1111/j.1530-0277.2011.01512.x

Nandrino, J.-L., Gandolphe, M.-C., Alexandre, C., Kmiecik, E., Yguel, J., & Urso, L. (2014). Cognitive and affective theory of mind abilities in alcohol-dependent patients: The role of autobiographical memory. *Drug and Alcohol Dependence, 143*, 65–73. https://doi.org/10.1016/j.drugalcdep.2014.07.010

Oliveira, L. M., Bermudez, M. B., Macedo, M. J.d. A., & Passos, I. C. (2018). Comorbid social anxiety disorder in patients with alcohol use disorder: A systematic review. *Journal of Psychiatric Research, 106*, 8–14. https://doi.org/10.1016/j.jpsychires.2018.09.008

Onuoha, R. C., Quintana, D. S., Lyvers, M., & Guastella, A. J. (2016). A meta-analysis of theory of mind in alcohol use disorders. *Alcohol and Alcoholism, 51*(4), 410–415. https://doi.org/10.1093/alcalc/agv137

Ospina, L. H., Shanahan, M., Perez-Rodriguez, M. M., Chan, C., Clari, R., & Burdick, K. E. (2019). Alexithymia predicts poorer social and everyday functioning in schizophrenia and bipolar disorder. *Psychiatry Research, 273*, 218–226. https://doi.org/10.1016/j.psychres.2019.01.033

Pabst, A., Gautier, M., & Maurage, P. (2022). Tasks and investigated components in social cognition research among adults with alcohol use disorder: A critical scoping review. *Psychology of Addictive Behaviors*. https://doi.org/10.1037/adb0000874

Perez, J., Beale, E., Overholser, J., Athey, A., & Stockmeier, C. (2022). Depression and alcohol use disorders as precursors to death by suicide. *Death Studies, 46*(3), 619–627. https://doi.org/10.1080/07481187.2020.1745954

Philipp, R., Kriston, L., Lanio, J., Kühne, F., Härter, M., Moritz, S., & Meister, R. (2018). Effectiveness of metacognitive interventions for mental disorders in adults—A systematic review and meta-analysis (METACOG). *Clinical Psychology & Psychotherapy, 26*(2), 227–240. https://doi.org/10.1002/cpp.2345

Plana, I., Lavoie, M.-A., Battaglia, M., & Achim, A. M. (2014). A meta-analysis and scoping review of social cognition performance in social phobia, posttraumatic stress disorder and other anxiety disorders. *Journal of Anxiety Disorders, 28*(2), 169–177. https://doi.org/10.1016/j.janxdis. 2013.09.005

Preis, M. A., Golm, D., Kröner-Herwig, B., & Barke, A. (2017). Examining differences in cognitive and affective theory of mind between persons with high and low extent of somatic symptoms: An experimental study. *BMC Psychiatry, 17*(1), 200. https://doi.org/10.1186/s12888-017-1360-9

Preißler, S., Dziobek, I., Ritter, K., Heekeren, H. R., & Roepke, S. (2010). Social cognition in borderline personality disorder: Evidence for disturbed recognition of the emotions, thoughts, and intentions of others. *Frontiers in Behavioral Neuroscience, 4*, 182. https://doi.org/10.3389/fnbeh.2010.00182

Preller, K. H., Hulka, L. M., Vonmoos, M., Jenni, D., Baumgartner, M. R., Seifritz, E., … Quednow, B. B. (2014). Impaired emotional empathy and related social network deficits in cocaine users. *Addiction Biology, 19*(3), 452–466. https://doi.org/10.1111/adb.12070

Quesque, F., & Rossetti, Y. (2020). What do theory-of-mind tasks actually measure? Theory and practice. *Perspectives on Psychological Science, 15*(2), 384–396. https://doi.org/10.1177/1745691619 896607

Ritter, K., Dziobek, I., Preißler, S., Rüter, A., Vater, A., Fydrich, T., … Roepke, S. (2011). Lack of empathy in patients with narcissistic personality disorder. *Psychiatry Research, 187*(1), 241–247. https://doi.org/10.1016/j.psychres.2010.09.013

Rosello, B., Berenguer, C., Baixauli, I., García, R., & Miranda, A. (2020). Theory of mind profiles in children with autism spectrum disorder: Adaptive/social skills and pragmatic competence. *Frontiers in Psychology, 11*, Article 567401. https://doi.org/10.3389/fpsyg.2020.567401

Roser, P., Lissek, S., Tegenthoff, M., Nicolas, V., Juckel, G., & Brüne, M. (2012). Alterations of theory of mind network activation in chronic cannabis users. *Schizophrenia Research, 139*(1), 19–26. https://doi.org/10.1016/j.schres.2012.05.020

Rupp, C. I., Derntl, B., Osthaus, F., Kemmler, G., & Fleischhacker, W. W. (2017). Impact of social cognition on alcohol dependence treatment outcome: Poorer facial emotion recognition predicts relapse/dropout. *Alcoholism: Clinical and Experimental Research, 41*(12), 2197–2206. https://doi.org/10.1111/acer.13522

Rupp, C. I., Junker, D., Kemmler, G., Mangweth-Matzek, B., & Derntl, B. (2021). Do social cognition deficits recover with abstinence in alcohol-dependent patients? *Alcoholism: Clinical and Experimental Research*, *45*(2), 470–479. https://doi.org/10.1111/acer.14537

Sanvicente-Vieira, B., Romani-Sponchiado, A., Kluwe-Schiavon, B., Brietzke, E., Araujo, R. B., & Grassi-Oliveira, R. (2017). Theory of mind in substance users: A systematic minireview. *Substance Use & Misuse*, *52*(1), 127–133. https://doi.org/10.1080/10826084.2016.1212890

Saunders, J. B., Aasland, O. G., Babor, T. F., De la Fuente, J. R., & Grant, M. (1993). Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction*, *88*(6), 791–804.

Schmid, F., Moreau, F., Benzerouk, F., Raucher-Chéné, D., Kaladjian, A., Gierski, F., & Henry, A. (2022). Faux pas recognition and executive processes in patients with alcohol use disorder: Toward an investigation of interindividual heterogeneity. *Archives of Clinical Neuropsychology*, *37*(3), 608–620. https://doi.org/10.1093/arclin/acab072

Schmidt, T., Roser, P., Juckel, G., Brüne, M., Suchan, B., & Thoma, P. (2016). Social cognition and social problem solving abilities in individuals with alcohol use disorder. *Journal of Clinical and Experimental Neuropsychology*, *38*(9), 974–990. https://doi.org/10.1080/13803395.2016.1180346

Shamay-Tsoory, & Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: A lesion study. *Neuropsychologia*, *45*(13), 3054–3067. https://doi.org/10.1016/j.neuropsychologia.2007.05.021

Sondermann, S., Stahl, J., Grave, U., Outzen, J., Moritz, S., & Klein, J. P. (2020). Preoperational thinking as a measure of social cognition is associated with long-term course of depressive symptoms. A longitudinal study involving patients with depression and healthy controls. *Frontiers in Psychiatry*, *11*. https://www.frontiersin.org/articles/10.3389/fpsyt.2020.00652.

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

Tawa, E. A., Hall, S. D., & Lohoff, F. W. (2016). Overview of the genetics of alcohol use disorder. *Alcohol and Alcoholism*, *51*(5), 507–514. https://doi.org/10.1093/alcalc/agw046

Thamby, A., Desai, G., Mehta, U. M., & Chaturvedi, S. K. (2019). Deficits in theory of mind and emotional awareness in somatoform disorders. *Indian Journal of Psychological Medicine*, *41*(4), 368–374. https://doi.org/10.4103/IJPSYM.IJPSYM_382_18

Thoma, P., Winter, N., Juckel, G., & Roser, P. (2013). Mental state decoding and mental state reasoning in recently detoxified alcohol-dependent individuals. *Psychiatry Research*, *205*(3), 232–240. https://doi.org/10.1016/j.psychres.2012.08.042

Vonmoos, M., Eisenegger, C., Bosch, O. G., Preller, K. H., Hulka, L. M., Baumgartner, M., … Quednow, B. B. (2019). Improvement of emotional empathy and cluster B personality disorder symptoms associated with decreased cocaine use severity. *Frontiers in Psychiatry*, *10*. https://www.frontiersin.org/articles/10.3389/fpsyt.2019.00213.

Wang, W.-H., Shih, Y.-H., Yu, H.-Y., Yen, D.-J., Lin, Y.-Y., Kwan, S.-Y., Chen, C., & Hua, M.-S. (2015). Theory of mind and social functioning in patients with temporal lobe epilepsy. *Epilepsia*, *56*(7), 1117–1123. https://doi.org/10.1111/epi.13023

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines* (Vol. 1). World Health Organization.

World-Medical-Association. (2013). World medical association declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA*, *310*(20), 2191–2194. https://doi.org/10.1001/jama.2013.281053

5. General Discussion

5.1 Summary of the Articles' Main Findings

5.1.1 Summary of Article 1

Article 1 presented findings on two clinical samples (AUD and PD), showing changes in MASC performance over the course of the test. Specifically, scores on affective ToM showed an upward trend, while cognitive ToM scores demonstrated a decline. Various factors might account for these findings. For instance, research suggested that cognitive ToM more strongly relies on working memory (Gabriel et al., 2019) compared to its affective facet. The PD group displayed superior performance on both affective and cognitive ToM relative to the AUD sample. Notably, the disparity was more salient in the domain of affective ToM. In terms of ToM errors, the AUD group predominantly exhibited a reduced ToM, whereas the PD group was more inclined towards an exceeding ToM. Overall, ToM decreased with increasing age. This decline in cognitive ToM performance was more pronounced for older individuals. Moreover, sex differences also emerged, with female participants showing a slightly better affective ToM relative to males, though their cognitive ToM performance was either equivalent or marginally inferior. Consequently, it is pivotal to account for individual variables (such as sex, age, and disorder type) and situational variables (i.e., test duration) when investigating ToM across clinical populations.

5.1.2 Summary of Article 2

Article 2 investigated the impact of interaction partners' gender on ToM, involving two clinical groups and a HC group. A distinct pattern emerged regarding the sequence of test items: a significant number of errors occurred in the early stages of the test. Results indicated better ToM performance in interactions involving male targets, both as senders and receivers, as opposed to those with females in the clinical samples, this effect was not significant in the HC. Interestingly, interactions between partners of the same gender resulted in lower ToM scores relative to mixed-gender interactions. Items involving interactions among a receiver, sender, and participant of different genders (outgroup interaction) showed a higher level of ToM performance compared to items involving interactions in which the receiver, sender and participant all had the same gender (ingroup interaction). This trend of diminished ToM in ingroup interactions was consistent across both clinical and HC groups. Thus, ToM depended

on the interaction of participant and target gender. These results highlight the impact of interaction partners on ToM performance and could help resolve inconsistencies in existing research on ToM abilities among various clinical groups. Recognizing these factors is crucial for accurate ToM assessment.

5.1.3 Summary of Article 3

Article 3 highlighted changes in cognitive and affective ToM during treatment and their association with changes of comorbid symptoms of depression, anxiety, somatization, and social functioning in individuals with AUD. Cognitive and affective ToM abilities were assessed simultaneously at two distinct measurement points. Over the course of the treatment process, cognitive ToM showed improvement, whereas affective ToM did not. Notably, the initial level of cognitive ToM was a significant predictor of improvements in symptoms of depression and somatization. However, the effect on symptoms of depression was no longer significant after a Bonferroni-Holm correction was applied. Affective ToM impairments remained unaddressed during the eclectic abstinence-oriented inpatient treatment, indicating that specialized interventions focusing on ToM may be necessary to address ToM abilities in patients with AUD.

5.2 Discussion of the findings on the measurement MASC

5.2.1 Aging aspects

Articles 1 and 2 are in line with previous findings regarding the influence of age on ToM. Age groups were categorized within four groups: 18-34, 35-45, 46-53 and 54-75. Younger patients exhibited superior ToM performance compared to their older counterparts, which is consistent with a meta-analysis that reported diminished ToM in older adults (Henry et al., 2013). In the results presented in Article 1, both individual (age) and situational (test duration) effects on ToM measurement using the MASC were investigated, revealing a distinct pattern of declining cognitive ToM and an increase in affective ToM over time, which was especially pronounced in older participants. The findings underscore how situational aspects, such as test duration, impact ToM performance, as there was a noticeable shift in performance over the course of the test. It was the first time ToM performance was assessed over the course of the MASC test. Some studies corroborate the findings of Article 1, emphasizing that the effects of aging predominantly impede cognitive ToM (Bottiroli et al., 2016; Wang & Su, 2013). Previous research associates this decline in cognitive ToM in older adults

with decreased EF (Wang & Su, 2013) and a strategic shift in resource allocation, potentially favoring affective tasks, as older individuals may prioritize discerning the emotions of the MASC's actors (Hess, 2014; Zhang et al., 2018).

In the MASC, the four main characters are approximately in their mid-thirties, which could make them more relevant to people of a similar age. Given that societal norms for acceptable behaviors have transformed over recent decades and linguistic expressions can differ across generations, it is plausible that individuals from the same age group would more readily comprehend the social interactions presented in the test (Dziobek et al., 2006). Future studies should develop an assessment with actors spanning a broader age range to account for this potential influence.

As adults age, a decrease in ToM, leads to reduced social involvement (Bailey et al., 2008). In this study younger adults (ages 19 to 25) were compared with older participants (ages 65 to 87), illuminating the role of ToM in the daily social interactions of older adults (Bailey et al., 2008). Consequently, the pivotal role of ToM in older adults' daily social interactions has been consistently highlighted in research, reinforcing the practical significance of ToM.

5.2.2 Sex aspects

In addition to exploring age aspects, Article 1 and 2 examined the influence of sex aspects on ToM performance. In Article 1, there was a trend for female participants to outperform male participants in total ToM. However, after applying the Bonferroni correction, this distinction was not significant, indicating no evidence for gender differences. Similarly, Article 2 suggested that women garnered higher ToM scores during outgroup interactions (either male to female or vice versa), but subsequent regression analysis didn't confirm any significant sex discrepancies. Consequently, both Articles 1 and 2 align with recent research that found no sex disparities in ToM abilities (Di Tella et al., 2020; Navarra-Ventura et al., 2018).

Article 1 used the MASC and reported that, when analyzing sex differences in affective and cognitive ToM, women performed better in affective ToM compared to men. However, in cognitive ToM, women's performance was either on par with or slightly below that of men. In fact, Di Tella et al. (2020) reported no distinct sex differences in either affective or cognitive ToM. Different findings could be due to different measurements (MASC vs. RMET).

One possible explanation for the sex differences regarding the affective ToM

outlined in article 1 could be that women are encouraged towards 'empathizing' strategies, which are centered around the understanding and recognition of others' emotions and thoughts. Support for the empathizing strategy in women is bolstered by findings indicating a female advantage in recognizing and articulating emotions (Kret & De Gelder, 2012; Mestre et al., 2009). Additionally, the literature has highlighted women's enhanced emotional accuracy, evidenced by their superior skills in interpreting nonverbal cues (Hall et al., 2000) and discerning emotions from facial expressions (Donges et al., 2012; Hall & Matsumoto, 2004; Hoffmann et al., 2010). However, Di Tella et al. (2020) found no sex differences in recognizing emotional facial expressions, except in the case of anger. In this instance, women demonstrated greater accuracy than men, especially when anger was displayed on male faces. The cognitive ToM findings in Article 1, which indicate similar performance between men and women on cognitive ToM tasks, align with the findings of a recent cross-sectional study by Navarra-Ventura et al. (2018). However, this contrasts with some studies, such as Russell et al. (2007), which have reported men outperforming women in cognitive ToM tasks.

5.2.3 Interaction partners

The primary aim of Article 2 was to examine specific variables influencing ToM measures. This study highlighted that ToM is affected by the social group (e.g., gender, age) of the target individual. Utilizing the MASC, the study investigated the influence of the sender's, receiver's, and participant's gender on ToM outcomes. This study hypothesized whether scenarios with the same gender sender, receiver, and participant (ingroup interaction) yielded different outcomes than mixed-gender interactions (outgroup interaction). Interestingly, interactions exclusively within an ingroup were associated with lower ToM performance compared to outgroup interactions.

These findings align with existing literature, suggesting that ingroup membership can lead individuals to erroneously project their own mental states onto others termed egocentric intrusion (Gopnik & Wellman, 1992; Simpson & Todd, 2017), which will lead to systematic errors when those perspectives diverge (Simpson & Todd, 2017). This reduction in ToM may stem from targets whose attitudes or emotions are not reciprocated by the perceiver. Correspondingly, a congruence in mental states between participants and targets appeared to enhance ToM

performance. Notably, such patterns were evident in both clinical and HC samples.

A study by Surtees and Apperly (2012) evaluated children ages 6-10 on their capability to interpret a cartoon character's visual perspective. The findings indicated pronounced egocentric tendencies, most evident when there was a disparity between the character's and the participants' perspectives. Interestingly, older children exhibited enhanced speed and precision in perspective-taking relative to younger participants. This advance in perspective-taking abilities was further substantiated by Apperly et al. (2011), who examined children's (aged 3-5) proficiency in predicting actions based on characters' beliefs and desires, with older children demonstrating a better understanding than their younger counterparts.

Moreover, as reported in Article 2, interactions featuring female targets were associated with diminished ToM performance in the clinical samples. This finding implies that female targets' perspectives might be less consistently or comprehensively acknowledged compared to male targets. This finding is consistent with previous research indicating that women may be perceived as less credible than men depending on the context (e.g., Armstrong & McAdams, 2009; Kang et al., 2019).

Overall, the results emphasize the influence of interaction partners' gender, a variable that might explain the heterogeneity of outcomes demonstrated in ToM studies. Thus, considering the target's gender is essential when assessing ToM performance to enhance the validity of the MASC.

5.3 Discussion of the findings on affective and cognitive ToM:

5.3.1 Affective and cognitive ToM – cross-sectional

In Article 1, the trajectory of ToM performance over the course of the MASC test was analyzed by examining the effects of item order. This revealed a significant shift in ToM performance as the test continued: while cognitive ToM declined, affective ToM improved. The decline could be explained by decreasing participant motivation, reduced attention, or fatigue, among others. Previous studies have highlighted that cognitive ToM depends more strongly on working memory compared to affective ToM (Gabriel et al., 2019). Consequently, any decrease in working memory might have less impact on affective ToM, which in contrast to cognitive ToM requires quick and automatic processing of cues, such as recognizing emotional expressions (Baron-Cohen et al., 2001). The upward trend in affective ToM performance could be due to its reliance on more automatic processes, which demand fewer cognitive resources

(Rothman et al., 2009). Given the MASC's 30-minute duration, it's likely that cognitive resources diminished over time, leading to the described patterns in ToM performance.

Neuroscientific studies have indicated that mechanisms linked to affective ToM, like decoding facial expressions, engage areas such as the superior temporal region and inferior frontal area (Dapretto et al., 2006; Malhi et al., 2008). In contrast, mental state reasoning, associated with cognitive ToM, are more cognitively challenging and involve the medial prefrontal cortex and temporoparietal junction (Amodio & Frith, 2006; Saxe & Wexler, 2005). The established link between the dorsolateral prefrontal cortex, EFs and cognitive ToM (Kalbe et al., 2010) further supports this distinction. This reported link by Kalbe et al. (2010) confirms the differentiation between cognitive and affective ToM, especially concerning EFs. It suggests that EFs are more crucial for understanding cognitive ToM compared than affective ToM.

5.3.2 Affective and cognitive ToM – longitudinal

Article 3 is the first to examine whether affective and cognitive ToM abilities improve during AUD treatment. Although a previous study reported enhanced ToM performance following treatment across diverse mental disorders, it did not differentiate between ToM facets (Hayden et al., 2018). In the realm of AUD, the literature is limited. The results indicate that after treatment, cognitive ToM demonstrated significant improvement, whereas affective ToM showed no significant changes.

Two previous studies focused exclusively on potential improvement in affective ToM during treatment in AUD patients. Rupp et al. (2021) reported no significant changes in affective ToM during a 2-month inpatient treatment, whereas Frileux et al. (2020) found enhancements in emotion recognition—a component of affective ToM (Kittel et al., 2022; Quesque & Rossetti, 2020)—following a 3-week alcohol detox. Emotion recognition may respond more readily to intervention than more complex processes such as overall affective ToM. The distinction between emotion recognition and affective ToM might account for the varied outcomes reported in affective ToM during AUD treatment (Frileux et al., 2020; Rupp et al., 2021). Previous studies indicate a broader deficit in emotional abilities, including affective ToM, among AUD patients, posing a challenge to recovery and potentially explaining the disparate trajectories of ToM facets during treatment (Kornreich et al., 2002; Le Berre, 2019;

Maurage et al., 2016; Maurage, Grynberg, Noël, Joassin, Philippot, et al., 2011).

A recent fMRI study revealed that healthy adults with a family history of AUD showed marked differences in affective ToM processing compared to counterparts without such a history. Those with a familial predisposition to AUD displayed reduced accuracy in affective ToM tasks and exhibited distinct neural activity patterns. Notably, these neural disparities persisted after controlling for depressive symptoms, anxiety, and childhood trauma. These findings suggest that affective ToM difficulties may be detectable in first-degree relatives of AUD patients, indicating potential markers of vulnerability preceding AUD onset (Schmid et al., 2023). These neurobehavioral conditions could also be reflected in the outcomes of Article 3, which reported that affective ToM impairments are more pronounced than cognitive ones in AUD. This suggests that emotional difficulties are central features of this disorder, corroborating prior research (Le Berre, 2019; Maurage et al., 2016).

In other mental disorders, such as psychopathy, individuals have exhibited impairments in affective ToM while retaining cognitive ToM abilities (Shamay-Tsoory et al., 2010). A cognitive model proposed by Shamay-Tsoory et al. (2010) suggests that the cognitive ToM serves as a prerequisite for affective ToM. This may lead to a more rapid improvement in cognitive ToM during treatment. In accordance with this model, affective ToM may require more specific training to be improved. Overall, affective, and cognitive ToM focused interventions may be crucial for enhancing ToM abilities in AUD patients. MCT has shown promise for both enhancing ToM and reducing alcohol cravings (Caselli et al., 2018; Philipp et al., 2018). Additionally, ToM deficits are present both in AUD patients and subclinical populations, making ToM training a potential preventive strategy for AUD (Kumar et al., 2022).

In conclusion, both Article 1 and Article 3 are based on the notion that ToM is not a monolithic ability. Instead, it encompasses distinct facets, specifically cognitive and affective ToM (Maleki et al., 2020; Shamay-Tsoory et al., 2007; Van den Stock et al., 2023; Zabihzadeh et al., 2017). Firstly, divergent effects were reported in affective versus cognitive ToM during the test period. Secondly, Article 1 and 2 highlighted differing aging effects on performance on these two ToM facets. Lastly, in the longitudinal study in Article 3, distinct changes in affective and cognitive ToM were investigated at two measurement points. These findings are in line with a study by Shamay-Tsoory et al. (2007), which suggests that cognitive and affective ToM abilities

are dissociable and are underpinned by separate neuroanatomical substrates.

5.3.3 Affective and cognitive ToM in different mental disorders

Article 1 showed significant differences in overall ToM performance between two patient samples: AUD and PD. Specifically, AUD patients faced greater difficulties in deducing mental states than patients diagnosed with PD. Additionally, the PD sample demonstrated superior performance in both affective and cognitive ToM compared to the AUD group. This distinction was evident for both facets, with a more pronounced difference reported for affective ToM. This implies that research should not generalize findings of ToM deficits to other mental disorders. Differentiating between affective and cognitive ToM in patients with BPD, which is the most common PD in the current sample, a meta-analysis by Németh et al. (2018) found no significant difference between affective and cognitive ToM in BPD patients, which was replicated by more recent studies (Hillmann et al., 2021).

Article 1 highlighted more pronounced challenges in affective ToM among AUD patients compared to PD patients, a finding consistent with Kopera et al. (2020) and Rupp et al. (2021), who also reported affective ToM impairments in AUD patients. Notably, despite more pronounced deficits in AUD relative to PD in both affective and cognitive ToM, a significant overlap in dysfunction was observed between the two groups, which makes sense considering the frequent comorbidity between AUD and BPD. Accordingly, characteristic ToM dysfunction are maybe shared across these disorders (Hanegraaf et al., 2021), which is postulated to be a significant contributor to the interpersonal difficulties frequently seen in both BPD and SUD (Hanegraaf et al., 2021).

The pattern of stronger deficits in the affective facet in AUD patients differs from other patient groups: Patients with schizophrenia (Li et al., 2017), MDD (Bora et al., 2016), and BD (Bora et al. (2016) exhibit deficits in both domains compared to HC. Even though newer studies, which often contain better measures, question these results again and for instances implied that the affective ToM was not impaired in BD (Szmulewicz et al., 2019). Due to these discrepancies, it's important to avoid generalizing ToM deficits across all mental disorders without further exploration of both affective and cognitive ToM.

5.3.4 ToM errors

Errors in ToM (exceeding, reduced and no ToM) can occur irrespective of the

emergence of deficits in its affective and cognitive aspects. In Article 1, the predominant error identified within the AUD sample was reduced ToM, suggesting a deficit to ascribe mental states to others. This finding is congruent with a meta-analysis by Onuoha et al. (2016), which concluded that individuals with AUD generally display a reduced ToM compared to HC groups. Considering the frequent co-occurrence of depressive symptoms and AUD, it is unsurprising that reduced ToM is also common in individuals diagnosed with depression (De Coninck et al., 2021; Safiye et al., 2023; Wolkenstein et al., 2011). Broadening the scope to other SUD, it seems that the no-ToM error is less common in SUD (Eidenmueller et al., 2021; Vaskinn et al., 2020) than other identified error types (reduced and exceeding ToM). Moreover, similar to the AUD sample, individuals with ED also display both types of errors: reduced and exceeding ToM (Corsi et al., 2021).

In contrast to the AUD sample, reduced ToM errors were rarer in the PD group, PD participants primarily exhibited errors linked to exceeding ToM, although these errors were noted within the AUD group as well. Several studies underscore that individuals diagnosed with BPD frequently manifest exceeding ToM tendencies, (Burghardt et al., 2023; McLaren et al., 2022; Normann-Eide et al., 2020). This finding aligns with patterns demonstrated in schizophrenia patients, where a marked inclination toward exceeding ToM is frequent (Montag et al., 2012; Peyroux et al., 2019). Overall, variations in ToM represent a significant clinical dimension, and close examination of these discrepancies is essential for understanding and intervening in the associated social difficulties (Livingston et al., 2019).

5.4 Discussion of the findings on changes in ToM and the associations between theories about ToM

Two primary theories explain our understanding of other minds: Theory–Theory and Simulation Theory. The Theory–Theory model proposes that our comprehension of others' mental states is rooted in a collection of concepts, accompanied by rules that dictate their relationships (Gopnik, 2003; Gopnik & Wellman, 1992, 2012). Central to this theory is a causal framework that delineates how various mental states contribute to observable behaviors. By harnessing this framework and combining it with specific information about an individual, we are able to explain and predict their actions and thoughts (Premack & Woodruff, 1978). In contrast, Simulation Theory posits a different mechanism. It argues that to understand another's mind, individuals

employ their own mental processes as a template. They identify the other person's initial mental states and by simulating these states within their own minds, they anticipated the subsequent actions and thoughts (Apperly, 2008).

Further supporting this perspective, Harris (1992) emphasized, in conjunction with Simulation Theory the significant influence of social environments in shaping ToM. Within this framework, experiential learning is critical to improving perspective-taking skills, which enables individuals to hone both their simulation skills and their ToM (Harris, 1992). Article 3 found changes in cognitive ToM over the course of treatment and indicated that treatment/inpatient setting, as one example of social environment, can benefit cognitive ToM.

In contrast, modularity theories demonstrated the dissociation between impaired ToM and general intellectual development (Leslie & Thaiss, 1992). However, they may not fully explain certain developmental phenomena (Sodian et al., 2012). Both theory-based knowledge (for predicting actions) and simulation (for understanding beliefs) are essential, as highlighted by Sodian (2005). The conclusions from Article 3 support this notion by underscoring the crucial role of experience in shaping our understanding of ToM abilities (Brüne & Brüne-Cohrs, 2006).

Simulations become more complex as multiple mental states require adjustment. Harris (1992) proposed that to simulate another person's mental state, which differs from one's own, it is necessary to set aside one's personal state. The person then tries to emulate the other's state, drawing on available information and intentions. In cases of false belief tasks, participants must overcome their own perspective as well as disregard the actual situation to simulate the mental state of someone with an incorrect belief, requiring the adjustment of two distinct mental states. Improved ToM performance is often observed when these mental states align. Research by Ames et al. (2012) and Simpson & Todd (2017) indicated that belonging to the same social group such as having the same gender, can impact participants' understanding of each other's mental states. As mentioned above, the tendency to use one's own perspective when interpreting another's is termed egocentric intrusion or simulation (Simpson & Todd, 2017). This is in line with simulation theory, which suggests that children are prone to project their current mental state onto others (Gopnik & Wellman, 1992). The congruence or disparity of these mental states is thought to affect ToM performance. This implies that, within the context of ingroup interactions as presented in Article 2,

when a misalignment of mental states occurs, the presence of egocentric simulation leads to more errors.

Simulation Theory and Theory–Theory offer distinct explanations for findings in ToM research. However, recent studies now advocate for a hybrid approach, incorporating evidence supporting both (Currie & Ravenscroft, 2002; Nichols & Stich, 2003). For instance, intuitive judgments often suggest simulation, while consistent prediction errors indicate a theoretical underpinning. To summarize, the Simulation Theory posits that we understand others' mental states by simulating their thought processes. In contrast, the Theory-Theory contends that we use an independent knowledge system to rationally model others' mental states. Cognitive ToM appears to stem from a process anchored in theoretical understandings of the mind in line with the Theory-Theory. Conversely, simulation likely underlies affective ToM. This view is reinforced by research from Adolphs et al. (2000) and Kalbe et al. (2007).

5.5 Discussion of the findings on Theory of Mind deficits in clinical samples

5.5.1 Theory of Mind deficits within the RDoC framework

The RDoC initiative by the National Institute of Mental Health (2020) aims to enhance our understanding of mental disorders by examining their foundational psychological, neural, and biological mechanisms. This framework seeks to refine research approaches by conceptualizing mental disorders as multi-layered syndromes, frequently associated with disrupted brain circuitry (Insel et al., 2010). One aspect of the RDoC framework is the emphasis on the importance of examining impaired functioning, such as cognitive difficulties and impaired motivational processes, which are critical to understanding the disorders, rather than just as markers of a syndrome (Cuthbert, 2020). Considering the substantial influence of social processes on the pronounced interpersonal dysfunction reported in AUD, integrating the RDoC's 'Systems for Social Processes' domain into future research initiatives will be pivotal for a detailed comprehension of both affective and cognitive ToM (Hanegraaf et al., 2021). Consistently, all three articles are congruent with the RDoC framework, which identifies ToM as a central transdiagnostic element in mental disorders.

In both Article 1 and Article 2, we also examined two distinct patient cohorts: those with AUD and those with PD, each displaying unique sociodemographic characteristics, particularly regarding sex and age. Individuals with AUD tend to be

older and are more often men, while those with PD are typically younger and more often women. Article 1 demonstrated that the AUD patients faced heightened difficulties in recognizing mental states compared to the PD patients, which could be the result of the underlying symptoms or be connected to the different gender and age distributions, since previous studies showed that women and younger patients typically outperform men and older patients (Krach et al., 2009; Wang & Su, 2013).

In Article 1, a comparison between the two patient cohorts revealed sociodemographic differences. However, despite disparities in symptoms and sociodemographic characteristics, the affective and cognitive ToM patterns of AUD and PD were found to be equivalent throughout the assessment. This suggests that certain individual and situational factors have a consistent influence on ToM within these mental disorders. This perspective supports the generalizability of results across diverse patient groups, further aligning with the principles of the RDoC framework. These findings support the idea that ToM impairments might be common across diverse mental disorders, reflecting the RDoC's perspective of characterizing mental health issues by their inherent challenges rather than via strict categorical boundaries. For instance, 'Systems for Social Processes' is a domain that covers these inherent challenges, with ToM serving as a crucial component of this domain. Thus, the findings of Article 3 are in line with the RDoC framework. This is particularly evident when observing symptoms of depression, anxiety, and somatization among AUD individuals.

Article 3 revealed a positive association between heightened cognitive ToM skills at treatment onset and subsequent reduction in comorbid symptoms, a pattern consistent with earlier studies with individuals with different mental disorders (Kvarstein et al., 2020; Sondermann et al., 2020). This underlines the central role of ToM during mental health recovery in various mental disorders. Given that deficits in ToM have been linked to treatment discontinuation and relapse in AUD patients (Rupp et al., 2017), enhancing ToM abilities in individuals showing early signs of AUD may prevent the initiation or worsening of the disorder (Kumar et al. 2022). Moreover, ToM impairments elevate interpersonal difficulties and diminish social connectedness (Verdejo-Garcia, 2020). Given the association between alcohol consumption and social support, it is reasonable to infer that social cognition impairments are likely to predispose individuals to problematic drinking behavior and could hinder sustained

abstinence (Robinson et al., 2018). Therefore, the RDoC framework's integration of ToM as a crucial element across a spectrum of mental disorders provides a valuable perspective for understanding and addressing the complex relations of cognitive functions and mental health challenges.

5.5.2 Symptoms – ToM deficits

5.5.2.1 Symptoms of depression. Article 3 revealed that individuals with AUD concurrently exhibited symptoms of depression, aligning with prior research reporting a prevalent comorbidity of depression within this population (Castillo-Carniglia et al., 2019; Hassan & Ali, 2011; Li et al., 2020; Oliveira et al., 2018). Moreover, the analysis presented in Article 3 corroborated the effectiveness of inpatient psychotherapeutic interventions in symptom reduction, echoing findings from studies which have demonstrated symptom reduction following two to four weeks of treatment (Haase et al., 2008; Kertz et al., 2015).

Despite the growing body of evidence of ToM deficits, current research provides limited insight into the potential relationship between ToM abilities and the reduction of comorbid depression symptoms during AUD treatment. Addressing this gap, Article 3 presented findings on the relationship between ToM and symptom reduction in individuals with AUD. The Article 3 revealed that ToM performance at the onset of psychotherapeutic treatment was related to reduction in symptoms. This finding cannot be interpreted in terms of causality. Nevertheless, findings of the Article 3 align with a study by Sondermann et al. (2020), which linked poor social cognition with limited relief in depression symptoms. The findings of Article 3 can be explained in two ways: Either those with better ToM benefit more from psychotherapy or people who respond well to treatment and therefore improve their depression improve their ToM. For instance, a reduction of depression symptoms might lead to increased EFs, which in turn improves ToM performance.

Previous research has confirmed the efficacy of MCT in treating depression disorders (Normann et al., 2014). Given that Article 3 highlighted comorbid symptoms of depression in AUD individuals, it implies that MCT could potentially address both ToM abilities and these comorbid symptoms in the AUD population. Thus, integrating MCT into treatment strategies for AUD could offer multiple benefits.

5.5.2.2 Symptoms of somatization. Article 3 provided additional insights into the relationship between ToM and comorbid symptoms of somatization, as previously

investigated in the research conducted by Chevalier et al (2023), Preis et al. (2017), and Thamby et al. (2019). Whereas Article 3 delineated a link between cognitive ToM and somatization symptoms, Thamby et al. (2019) had identified more pronounced deficits in affective ToM as being correlated with increased severity of somatization symptoms. In contrast, studies by Gabriel et al. (2019) and Rothman et al. (2009) introduced a dichotomy, associating affective ToM with the automatic dimension and cognitive ToM with a controlled dimension. This dichotomy has been shown to moderate the relationship between ToM and somatization symptoms, suggesting that higher ability in tasks requiring controlled processing, characteristic of cognitive ToM, is associated with fewer somatization symptoms. This perspective aligns with the findings in Article 3, which emphasize a correlation between cognitive ToM and somatization symptoms.

While ToM deficits may present challenges within psychotherapeutic interactions (Kvarstein et al., 2020), findings from Article 3 suggest that therapeutic progression and symptom improvement remain feasible among patients who remain in treatment. Taken together the findings underscores the significance of therapists' engagement with patients' abilities to comprehend others' cognitive perspectives for achieving effective treatment outcomes.

5.5.3 Theory of Mind and mental disorders

Previously the most pronounced ToM impairments were evident in patients with neurodegenerative disorders, while substantial deficits were likewise demonstrated in individuals with psychotic disorders (Cotter et al., 2018). The systematic review by Cotter et al. (2018) recommended repeated assessments of social cognition to explore its potential role as a longitudinal predictor of specific clinical and functional outcomes. Article 3 follows this recommendation by conducting serial assessments of ToM and confirming its predictive value for symptom reduction in individuals with mental disorders.

Direct comparison of effect sizes across various clinical conditions is methodologically complex due to inherent differences in study designs and measurement tool. Across distinct disorders, ToM deficits appeared consistent, indicating that irrespective of the diagnostic category, patients encounter similar challenges with these social cognitive domains. In contrast, within the psychiatric population, the degree of these deficits has demonstrated greater variability, likely

reflecting divergent disease trajectories and clinical severities (Cotter et al., 2018). For example, the systematic review by Cotter et al. (2018) and findings from Article 3 both indicate that patients with AUD manifested more pronounced ToM impairments compared to those with PD. Moreover, the systematic review by Cotter et al. (2018) revealed that individuals with schizophrenia exhibited the most pronounced deficits. Effect sizes for individuals with MDD, BP, AUD, and ED were similar. Interestingly, individuals with BPD showed effect sizes that were almost equal to HC (Cotter et al., 2018).

5.6 Limitation and Implication for Future Research

The present studies of affective and cognitive ToM assessed with the MASC have yielded significant insights; however, certain inherent limitations must be addressed. A principal limitation is the studies' exclusion of assessments related to general cognitive abilities. Due to feasibility in the clinical context time constrains did not allow the assessment of cognitive abilities. However, the naturalistic clinical context of the research provides valuable insights in terms of treatment targets. To mitigate the influence of general cognitive abilities, additional measures such as exclusion based on performance on the MASC control questions were employed. The exclusion of participants who incorrectly answered the MASC control questions did not yield different results compared to the full sample, suggesting that cognitive functions are not the driving force behind the change in MASC outcome. Nonetheless, future research should include comprehensive investigations of cognitive functioning to discern their relationship with ToM abilities.

Moreover, the studies did not assess EFs and thus cannot test whether EF influenced the reduction in symptoms over the course of the MASC assessment or in response to the treatment. Nevertheless, existing studies on EFs present inconsistent findings. Earlier research described ToM as a domain-specific process, distinct from other cognitive domains, particularly EFs (Fodor, 1992; Leslie, 1994). Yet more recent investigations assume that while foundational ToM skills such as gaze direction detection and joint attention are domain-specific, advanced ToM skills may be influenced by broader cognitive processes like EFs (Leslie et al., 2005; Stone & Gerrans, 2006). Clinical studies further emphasize that EFs may serve as predictors of ToM performance (Baez et al., 2015; Németh et al., 2020) (Sabbagh et al., 2006; Zelazo et al., 2002). EFs may underpin the development of sophisticated belief

reasoning in children (Apperly et al., 2009). In differentiating between affective and cognitive ToM, a pathway becomes evident (Huang et al., 2023). EF are connected to both cognitive and affective facets of ToM through the process of emotion recognition. In contrast, both Bottiroli et al. (2016) and Fischer et al. (2017) reported no significant associations between affective ToM and EFs, though they did identify associations within the cognitive ToM. Contrary to a study by Di Tella et al. (2020), which found no association between either dimension of ToM and EFs. Accordingly, we cannot ascertain whether findings in the presented studies of this dissertation are due to impairments in ToM or due to impairments in EFs. For a more comprehensive understanding of the relationship between ToM and EFs, future research should explore these constructs across various age groups within the adult population.

Another limitation in the studies encompassing this dissertation is the absence of a waiting list control group in Article 3. This absence hinders strong conclusions regarding whether the reported changes in symptoms and cognitive ToM abilities were directly attributable to the therapeutic intervention. Furthermore, it is unclear whether the post-intervention cognitive ToM abilities of participants were comparable to those of individuals on a treatment waiting list or HC. Nevertheless, for future studies, comparing treatment data with individuals on treatment waiting lists would yield a more comprehensive understanding of the treatment impact. Additionally, the studies, lacking an experimental design, cannot be interpreted in terms of causality. Despite the inability to interpret the causes of the findings, the studies are highly relevant for ToM research in a naturalistic setting.

According to Quesque and Rossetti (2020), the fundamental components of ToM encompass the representation of others' mental states and the ability to discern these from one’s own. These two core elements are assessed using the MASC, a tool consistently employed across all three studies presented in this dissertation. While many other ToM assessment tools are criticized for lacking ecological validity, as noted by Pabst et al. (2022), the MASC (Dziobek et al., 2006) stands out as both ecologically valid and reliable (Benito-Ruiz et al., 2022; Fossati et al., 2018). Nevertheless, to further ensure the understanding of the MASC’s measurements, it is necessary that future investigations maintain the methodological rigor reflected in Articles 1 and 2. Additionally, for robust validation of ToM research, it is essential that future investigations undertake comparative analyses employing validated

instruments, especially using both the MASC and the Faux Pas. Psychometric test reliability and validity are core prerequisites to ensure consistency in research outcomes and ensure a more profound understanding of social cognitive abilities. Thus, the absence of a universally accepted gold-standard measure remains a crucial gap in the field. Improvement of ToM assessment methods will also facilitate more effective treatments and interventions for individuals with mental health disorders.

5.7 Conclusion

The relationship between ToM and diverse mental disorders has been increasingly investigated (Cotter et al., 2018; McLaren et al., 2022). Cotter et al. (2018) emphasized the role of social cognitive processes as transdiagnostic indicators across various clinical presentations, underscoring their importance in discerning disease progression, and treatment efficacy. This dissertation expanded this understanding for the field of social cognition by providing data on the effects that influence ToM assessment.

Across all three studies, the MASC assessment was used to measure ToM performance, to assess affective and cognitive ToM. By differentiating between these two facets of ToM within sizable clinical samples, the present studies not only enrich our understanding of ToM assessment. To address existing inconsistencies in ToM findings, the current dissertation aimed to investigate the impact of individual and situational variables on ToM measurement through three empirical articles.

From a clinical perspective, assessing ToM in patients with mental disorders can also support intervention strategies. A comprehensive understanding of ToM is crucial to develop targeted and personalized interventions with regard to patients' strengths (Bosco et al., 2014). Evidence supports that integrative therapies can improve treatment outcomes across various psychiatric disorders (Peyroux & Franck, 2014). The findings of the studies in this dissertation emphasizes the role of ToM from a clinical perspective and highlight the need for research on affective and cognitive ToM in psychological and psychotherapeutic treatment of mental disorders. Finally, deficits in ToM significantly impair individuals' social functioning, and have a negative impact on interpersonal relationships, and social engagement (Wang et al., 2015). Considering these implications, ToM is not only an essential element for social functioning and skills but also for mental health and health-related quality of life (Bodden et al., 2010).

References

Adegboye, D., Sterr, A., Lin, J.-P., & Owen, T. J. (2017). Theory of mind, emotional and social functioning, and motor severity in children and adolescents with dystonic cerebral palsy. *European Journal of Paediatric Neurology*, *21*(3), 549–556. https://doi.org/10.1016/j.ejpn.2017.01.013

Adolphs, R., Damasio, H., Tranel, D., Cooper, G., & Damasio, A. (2000). A Role for Somatosensory Cortices in the Visual Recognition of Emotion as Revealed by Three-Dimensional Lesion Mapping. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *20*, 2683–2690. https://doi.org/10.1523/JNEUROSCI.20-07-02683.2000

Allen, J. G., & Fonagy, P. (Eds.). (2006). *The Handbook of Mentalization-Based Treatment*. John Wiley & Sons. https://doi.org/10.1002/9780470712986

Alvarez, J. A., & Emory, E. (2006). Executive Function and the Frontal Lobes: A Meta-Analytic Review. *Neuropsychology Review*, *16*(1), 17–42. https://doi.org/10.1007/s11065-006-9002-x

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (American Psychiatric Association, Ed.; 5th ed). American Psychiatric Association.

Ames, D. R., Weber, E. U., & Zou, X. (2012). Mind-reading in strategic interaction: The impact of perceived similarity on projection and stereotyping. *Organizational Behavior and Human Decision Processes*, *117*(1), 96–110. https://doi.org/10.1016/j.obhdp.2011.07.007

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*(4), 268–277. https://doi.org/10.1038/nrn1884

Andreou, C., Kelm, L., Bierbrodt, J., Braun, V., Lipp, M., Yassari, A. H., & Moritz, S. (2015). Factors contributing to social cognition impairment in borderline personality disorder and schizophrenia. *Psychiatry Research*, *229*(3), 872–879. https://doi.org/10.1016/j.psychres.2015.07.057

Andreou, M., & Skrimpa, V. (2020). Theory of Mind Deficits and Neurophysiological Operations in Autism Spectrum Disorders: A Review. *Brain Sciences*, *10*(6), Article 6. https://doi.org/10.3390/brainsci10060393

Apperly, I. A. (2012). What is “theory of mind”? Concepts, cognitive processes and individual differences. *Quarterly Journal of Experimental Psychology*, *65*(5), 825–839. https://doi.org/10.1080/17470218.2012.676055

Apperly I. A. (2008). Beyond Simulation-Theory and Theory-Theory: Why social cognitive neuroscience should use its own concepts to study "theory of mind". *Cognition, 107*(1), 266–283. https://doi.org/10.1016/j.cognition.2007.07.019

Apperly, I. A., Samson, D., & Humphreys, G. W. (2009). Studies of adults can inform accounts of theory of mind development. *Developmental Psychology*, *45*(1), 190–201. https://doi.org/10.1037/a0014098

Apperly, I. A., Warren, F., Andrews, B., Grant, J., & Todd, S. (2011). Error patterns in the belief reasoning of 3 to 5 year olds recur in reaction times from 6 years to adulthood: Evidence for developmental continuity in theory of mind. *Child Development*.

Arcelus, J., Haslam, M., Farrow, C., & Meyer, C. (2013). The role of interpersonal functioning in the maintenance of eating psychopathology: A systematic review and testable model. *Clinical Psychology Review*, *33*(1), 156–167. https://doi.org/10.1016/j.cpr.2012.10.009

Arcelus, J., Mitchell, A. J., Wales, J., & Nielsen, S. (2011). Mortality Rates in Patients

With Anorexia Nervosa and Other Eating Disorders: A Meta-analysis of 36 Studies. *Archives of General Psychiatry*, *68*(7), 724–731. https://doi.org/10.1001/archgenpsychiatry.2011.74

Armstrong, C. L., & McAdams, M. J. (2009). Blogs of Information: How Gender Cues and Individual Motivations Influence Perceptions of Credibility. *Journal of Computer-Mediated Communication*, *14*(3), 435–456. https://doi.org/10.1111/j.1083-6101.2009.01448.x

Baez, S., Marengo, J., Perez, A., Huepe, D., Font, F. G., Rial, V., Gonzalez-Gadea, M. L., Manes, F., & Ibanez, A. (2015). Theory of mind and its relationship with executive functions and emotion recognition in borderline personality disorder. *Journal of Neuropsychology*, *9*(2), 203–218. https://doi.org/10.1111/jnp.12046

Bailey, P. E., Henry, J. D., & Von Hippel, W. (2008). Empathy and social functioning in late adulthood. *Aging & Mental Health*, *12*(4), 499–503. https://doi.org/10.1080/13607860802224243

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, *14*(3), 110–118. https://doi.org/10.1016/j.tics.2009.12.006

Baksh, R. A., Abrahams, S., Auyeung, B., & MacPherson, S. E. (2018). The Edinburgh Social Cognition Test (ESCoT): Examining the effects of age on a new measure of theory of mind and social norm understanding. *PLOS ONE*, *13*(4), e0195818. https://doi.org/10.1371/journal.pone.0195818

Ballespí, S., Vives, J., Sharp, C., Chanes, L., & Barrantes-Vidal, N. (2021). Self and Other Mentalizing Polarities and Dimensions of Mental Health: Association With Types of Symptoms, Functioning and Well-Being. *Frontiers in Psychology*, *12*, Article 566254. https://doi.org/10.3389/fpsyg.2021.566254

Barnett, K., Mercer, S. W., Norbury, M., Watt, G., Wyke, S., & Guthrie, B. (2012). Epidemiology of multimorbidity and implications for health care, research, and medical education: A cross-sectional study. *Lancet (London, England)*, *380*(9836), 37–43. https://doi.org/10.1016/S0140-6736(12)60240-2

Baron-Cohen, S. (2010). Empathizing, systemizing, and the extreme male brain theory of autism. In *Progress in Brain Research* (Vol. 186, pp. 167–175). Elsevier. https://doi.org/10.1016/B978-0-444-53630-3.00011-7

Baron-Cohen, S., Knickmeyer, R. C., & Belmonte, M. K. (2005). Sex Differences in the Brain: Implications for Explaining Autism. *Science*, *310*(5749), 819–823. https://doi.org/10.1126/science.1115455

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind" ? *Cognition*, *21*(1), 37–46. https://doi.org/10.1016/0010-0277(85)90022-8

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999a). A new test of social sensitivity: Detection of faux pas in normal children and children with Asperger syndrome: *Journal of Autism and Developmental Disorders*, *29*(5), 407–418. https://doi.org/10.1023/a:1023035012436

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999b). 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. https://doi.org/10.1023/a:1023035012436

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child*

Psychology and Psychiatry, and Allied Disciplines, *42*(2), 241–251. https://doi.org/10.1111/1469-7610.00715

Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is There a “Language of the Eyes”? Evidence from Normal Adults, and Adults with Autism or Asperger Syndrome. *Visual Cognition*, *4*(3), 311–331. https://doi.org/10.1080/713756761

Barrett, L. F., Robin, L., Pietromonaco, P. R., & Eyssell, K. M. (1998). Are Women the “More Emotional” Sex? Evidence From Emotional Experiences in Social Context. *Cognition & Emotion*, *12*(4), 555–578. https://doi.org/10.1080/026999398379565

Bateman, A., Constantinou, M. P., Fonagy, P., & Holzer, S. (2021). Eight-year prospective follow-up of mentalization-based treatment versus structured clinical management for people with borderline personality disorder. *Personality Disorders, 12*(4), 291–299. https://doi.org/10.1037/per0000422

Bateman, A., & Fonagy, P. (1999). Effectiveness of Partial Hospitalization in the Treatment of Borderline Personality Disorder: A Randomized Controlled Trial. *American Journal of Psychiatry*, *156*(10), 1563–1569. https://doi.org/10.1176/ajp.156.10.1563

Bateman, A. W., & Fonagy, P. (2019). *Handbook of Mentalizing in Mental Health Practice*. American Psychiatric Pub.

Beck, A. T., Freeman, A., & Davis, D. D. (2004). *Cognitive therapy of personality disorders, 2nd ed* (pp. 3–439). Guilford Press.

Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science*, *8*(6), 492–499. https://doi.org/10.1111/j.1467-7687.2005.00440.x

Benito-Ruiz, G., Luzón-Collado, C., Arrillaga-González, J., & Lahera, G. (2022).

Development of an Ecologically Valid Assessment for Social Cognition Based on Real Interaction: Preliminary Results. *Behavioral Sciences*, *12*(2), Article 2. https://doi.org/10.3390/bs12020054

Bhar, S. S., Brown, G. K., & Beck, A. T. (2008). Dysfunctional Beliefs and Psychopathology in Borderline Personality Disorder. *Journal of Personality Disorders*, *22*(2), 165–177. https://doi.org/10.1521/pedi.2008.22.2.165

Bodden, M. E., Mollenhauer, B., Trenkwalder, C., Cabanel, N., Eggert, K. M., Unger, M. M., Oertel, W. H., Kessler, J., Dodel, R., & Kalbe, E. (2010). Affective and cognitive theory of mind in patients with parkinson's disease. *Parkinsonism & Related Disorders*, *16*(7), 466–470. https://doi.org/10.1016/j.parkreldis.2010.04.014

Boffo, M., Zerhouni, O., Gronau, Q. F., van Beek, R. J. J., Nikolaou, K., Marsman, M., & Wiers, R. W. (2019). Cognitive Bias Modification for Behavior Change in Alcohol and Smoking Addiction: Bayesian Meta-Analysis of Individual Participant Data. *Neuropsychology Review*, *29*(1), 52–78. https://doi.org/10.1007/s11065-018-9386-4

Boldrini, T., Pontillo, M., Tanzilli, A., Giovanardi, G., Di Cicilia, G., Salcuni, S., Vicari, S., & Lingiardi, V. (2020). An attachment perspective on the risk for psychosis: Clinical correlates and the predictive value of attachment patterns and mentalization. *Schizophrenia Research*, *222*, 209–217. https://doi.org/10.1016/j.schres.2020.05.052

Bora, E. (2021). A meta-analysis of theory of mind and 'mentalization' in borderline personality disorder: A true neuro-social-cognitive or meta-social-cognitive impairment? *Psychological Medicine*, *51*(15), 2541–2551. https://doi.org/10.1017/S0033291721003718

Bora, E., Bartholomeusz, C., & Pantelis, C. (2016). Meta-analysis of Theory of Mind (ToM) impairment in bipolar disorder. *Psychological Medicine*, *46*(2), 253–264. https://doi.org/10.1017/S0033291715001993

Bora, E., & Berk, M. (2016). Theory of mind in major depressive disorder: A meta-analysis. *Journal of Affective Disorders*, *191*, 49–55. https://doi.org/10.1016/j.jad.2015.11.023

Bora, E., Eryavuz, A., Kayahan, B., Sungu, G., & Veznedaroglu, B. (2006). Social functioning, theory of mind and neurocognition in outpatients with schizophrenia; mental state decoding may be a better predictor of social functioning than mental state reasoning. *Psychiatry Research*, *145*(2), 95–103. https://doi.org/10.1016/j.psychres.2005.11.003

Bora, E., Fornito, A., Pantelis, C., & Yücel, M. (2012). Gray matter abnormalities in Major Depressive Disorder: A meta-analysis of voxel based morphometry studies. *Journal of Affective Disorders*, *138*(1), 9–18. https://doi.org/10.1016/j.jad.2011.03.049

Bora, E., & Köse, S. (2016). Meta-analysis of theory of mind in anorexia nervosa and bulimia nervosa: A specific İmpairment of cognitive perspective taking in anorexia nervosa? *International Journal of Eating Disorders*, *49*(8), 739–740. https://doi.org/10.1002/eat.22572

Bora, E., & Özerdem, A. (2017). Social cognition in first-degree relatives of patients with bipolar disorder: A meta-analysis. *European Neuropsychopharmacology*, *27*(4), 293–300. https://doi.org/10.1016/j.euroneuro.2017.02.009

Bora, E., & Pantelis, C. (2013). Theory of mind impairments in first-episode psychosis, individuals at ultra-high risk for psychosis and in first-degree relatives of schizophrenia: Systematic review and meta-analysis. *Schizophrenia Research*,

144(1), 31–36. https://doi.org/10.1016/j.schres.2012.12.013

Bora, E., Vahip, S., Gonul, A. S., Akdeniz, F., Alkan, M., Ogut, M., & Eryavuz, A. (2005). Evidence for theory of mind deficits in euthymic patients with bipolar disorder. *Acta Psychiatrica Scandinavica*, *112*(2), 110–116. https://doi.org/10.1111/j.1600-0447.2005.00570.x

Bora, E., & Zorlu, N. (2017). Social cognition in alcohol use disorder: A meta-analysis. *Addiction*, *112*(1), 40–48. https://doi.org/10.1111/add.13486

Bora, E., Yücel, M., & Pantelis, C. (2009). Theory of mind impairment: A distinct trait-marker for schizophrenia spectrum disorders and bipolar disorder? *Acta Psychiatrica Scandinavica*, *120*(4), 253–264. https://doi.org/10.1111/j.1600-0447.2009.01414.x

Bora, E., Yucel, M., & Pantelis, C. (2009). Theory of mind impairment in schizophrenia: Meta-analysis. *Schizophrenia Research*, *109*(1–3), 1–9. https://doi.org/10.1016/j.schres.2008.12.020

Bosco, F. M., Capozzi, F., Colle, L., Marostica, P., & Tirassa, M. (2014). Theory of Mind Deficit in Subjects with Alcohol Use Disorder: An Analysis of Mindreading Processes. *Alcohol and Alcoholism*, *49*(3), 299–307. https://doi.org/10.1093/alcalc/agt148

Bottiroli, S., Cavallini, E., Ceccato, I., Vecchi, T., & Lecce, S. (2016). Theory of Mind in aging: Comparing cognitive and affective components in the faux pas test. *Archives of Gerontology and Geriatrics*, *62*, 152–162. https://doi.org/10.1016/j.archger.2015.09.009

Braehler, C., & Schwannauer, M. (2012). Recovering an emerging self: Exploring reflective function in recovery from adolescent-onset psychosis. *Psychology and Psychotherapy: Theory, Research and Practice*, *85*(1), 48–67.

https://doi.org/10.1111/j.2044-8341.2011.02018.x

Brothers, L., & Ring, B. (1992). A Neuroethological Framework for the Representation of Minds. *Journal of Cognitive Neuroscience*, *4*(2), 107–118. https://doi.org/10.1162/jocn.1992.4.2.107

Brüne, M. (2005). "Theory of Mind" in Schizophrenia: A Review of the Literature. *Schizophrenia Bulletin*, *31*(1), 21–42. https://doi.org/10.1093/schbul/sbi002

Brüne, M., & Brüne-Cohrs, U. (2006). Theory of mind—Evolution, ontogeny, brain mechanisms and psychopathology. *Neuroscience & Biobehavioral Reviews*, *30*(4), 437–455. https://doi.org/10.1016/j.neubiorev.2005.08.001

Buhlmann, U., Wacker, R., & Dziobek, I. (2015). Inferring other people's states of mind: Comparison across social anxiety, body dysmorphic, and obsessive–compulsive disorders. *Journal of Anxiety Disorders*, *34*, 107–113. https://doi.org/10.1016/j.janxdis.2015.06.003

Buragohain, J., & Bhagabati, D. (2020). Social Cognition in Persons with Alcohol Dependence Syndrome (ADS). *IOSR Journal Of Humanities And Social Science (IOSR-JHSS)*, *1*(25), 29–36. https://doi.org/10.9790/0837-2501012936

Burdick, K. E., Goldberg, J. F., & Harrow, M. (2010). Neurocognitive dysfunction and psychosocial outcome in patients with bipolar I disorder at 15-year follow-up. *Acta Psychiatrica Scandinavica*, *122*(6), 499–506. https://doi.org/10.1111/j.1600-0447.2010.01590.x

Burghardt, J., Gradl, S., Knopp, M., & Sprung, M. (2023). Psychopathology and Theory of Mind in patients with personality disorders. *Borderline Personality Disorder and Emotion Dysregulation*, *10*(1), 18. https://doi.org/10.1186/s40479-023-00224-1

Burke, B. L., Arkowitz, H., & Menchola, M. (2003). The efficacy of motivational

interviewing: A meta-analysis of controlled clinical trials. *Journal of Consulting and Clinical Psychology*, *71*, 843–861. https://doi.org/10.1037/0022-006X.71.5.843

Campbell, R., Elgar, K., Kuntsi, J., Akers, R., Terstegge, J., Coleman, M., & Skuse, D. (2002). The classification of 'fear' from faces is associated with face recognition skill in women. *Neuropsychologia*, *40*(6), 575–584. https://doi.org/10.1016/S0028-3932(01)00164-6

Carcone, D., & Ruocco, A. C. (2017). Six Years of Research on the National Institute of Mental Health's Research Domain Criteria (RDoC) Initiative: A Systematic Review. *Frontiers in Cellular Neuroscience*, *11*, 46. https://doi.org/10.3389/fncel.2017.00046

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social Cognition, Joint Attention, and Communicative Competence from 9 to 15 Months of Age. *Monographs of the Society for Research in Child Development*, *63*(4), 1–174. https://doi.org/10.2307/1166214

Caselli, G., Gemelli, A., Spada, M. M., & Wells, A. (2016). Experimental modification of perspective on thoughts and metacognitive beliefs in alcohol use disorder. *Psychiatry Research*, *244*, 57–61. https://doi.org/10.1016/j.psychres.2016.07.029

Caselli, G., Martino, F., Spada, M. M., & Wells, A. (2018). Metacognitive Therapy for Alcohol Use Disorder: A Systematic Case Series. *Frontiers in Psychology*, *9*, Article 2619. https://doi.org/10.3389/fpsyg.2018.02619

Castillo-Carniglia, A., Keyes, K. M., Hasin, D. S., & Cerdá, M. (2019). Psychiatric comorbidities in alcohol use disorder. *The Lancet. Psychiatry*, *6*(12), 1068–1080. https://doi.org/10.1016/S2215-0366(19)30222-6

Chapman, D. P., & Perry, G. S. (2008). Depression as a major component of public health for older adults. *Preventing Chronic Disease*, *5*(1), A22. http://www.cdc.gov/pcd/issues/2008/jan/07_0150.htm.

Chapman, E., Baron-Cohen, S., Auyeung, B., Knickmeyer, R., Taylor, K., & Hackett, G. (2006). Fetal testosterone and empathy: Evidence from the Empathy Quotient (EQ) and the “Reading the Mind in the Eyes” Test. *Social Neuroscience*, *1*(2), 135–148. https://doi.org/10.1080/17470910600992239

Chevalier, V., Simard, V., & Achim, J. (2023). Meta-analyses of the associations of mentalization and proxy variables with anxiety and internalizing problems. *Journal of Anxiety Disorders*, *95*, Article 102694. https://doi.org/10.1016/j.janxdis.2023.102694

Christov-Moore, L., Simpson, E. A., Coudé, G., Grigaityte, K., Iacoboni, M., & Ferrari, P. F. (2014). Empathy: Gender effects in brain and behavior. *Neuroscience and Biobehavioral Reviews*, *46*(4), 604–627. https://doi.org/10.1016/j.neubiorev.2014.09.001

Chung, Y. S., Barch, D., & Strube, M. (2014). A Meta-Analysis of Mentalizing Impairments in Adults With Schizophrenia and Autism Spectrum Disorder. *Schizophrenia Bulletin*, *40*(3), 602–616. https://doi.org/10.1093/schbul/sbt048

Chung, Y. S., Mathews, J. R., & Barch, D. M. (2011). The effect of context processing on different aspects of social cognition in schizophrenia. *Schizophrenia Bulletin*, *37*(5), 1048–1056. https://doi.org/10.1093/schbul/sbq012

Corsi, E., Cardi, V., Sowden, S., Coll, M.-P., Cascino, G., Ricca, V., Treasure, J., Bird, G., & Monteleone, A. M. (2021). Socio-cognitive processing in people with eating disorders: Computerized tests of mentalizing, empathy and imitation skills. *International Journal of Eating Disorders*, *54*(8), 1509–1518.

https://doi.org/10.1002/eat.23556

Cotter, J., Granger, K., Backx, R., Hobbs, M., Looi, C. Y., & Barnett, J. H. (2018). Social cognitive dysfunction as a clinical marker: A systematic review of meta-analyses across 30 clinical conditions. *Neuroscience & Biobehavioral Reviews*, *84*, 92–99. https://doi.org/10.1016/j.neubiorev.2017.11.014

Couture, S. M., Penn, D. L., & Roberts, D. L. (2006). The Functional Significance of Social Cognition in Schizophrenia: A Review. *Schizophrenia Bulletin*, *32*(1), 44–63. https://doi.org/10.1093/schbul/sbl029

Cox, S., Bertoux, M., Turner, J. J. D., Moss, A., Locker, K., & Riggs, K. (2018). Aspects of alcohol use disorder affecting social cognition as assessed using the Mini Social and Emotional Assessment (mini-SEA). *Drug and Alcohol Dependence*, *187*, 165–170. https://doi.org/10.1016/j.drugalcdep.2018.03.004

Currie, G., & Ravenscroft, I. (2002). *Recreative Minds: Imagination in Philosophy and Psychology*. Oxford University Press: https://doi.org/10.1093/acprof:oso/9780198238089.001.0001

Cuthbert, B. N. (2020). The role of RDoC in future classification of mental disorders. *Dialogues in Clinical Neuroscience*, *22*(1), 81–85. https://doi.org/10.31887/DCNS.2020.22.1/bcuthbert

Dalkner, N., Bengesser, S. A., Birner, A., Fellendorf, F. T., Hamm, C., Platzer, M., Pilz, R., Queissner, R., Rieger, A., Weber, B., Kapfhammer, H. P., Weiss, E. M., & Reininghaus, E. Z. (2019). The relationship between “Eyes Reading” ability and verbal memory in bipolar disorder. *Psychiatry Research*, *273*, 42–51. https://doi.org/10.1016/j.psychres.2019.01.015

Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y., & Iacoboni, M. (2006). Understanding emotions in others: Mirror neuron

dysfunction in children with autism spectrum disorders. *Nature Neuroscience*, *9*(1), 28-30. https://doi.org/10.1038/nn1611

De Coninck, D., Matthijs, K., & Luyten, P. (2021). Depression in Belgian first-year university students: A longitudinal study of self-definition, interpersonal relatedness, mentalizing, and integration. *Journal of Clinical Psychology*, *77*(7), 1715–1731. https://doi.org/10.1002/jclp.23149

de Siqueira Rotenberg, L., Beraldi, G. H., Okawa Belizario, G., & Lafer, B. (2020). Impaired social cognition in bipolar disorder: A meta-analysis of Theory of Mind in euthymic patients. *The Australian and New Zealand Journal of Psychiatry*, *54*(8), 783–796. https://doi.org/10.1177/0004867420924109

de Villiers, P. A. (2005). The Role of Language in Theory-of-Mind Development: What Deaf Children Tell Us. In *Why language matters for theory of mind* (pp. 266–297). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195159912.003.0013

De Vito, A., Calamia, M., Greening, S., & Roye, S. (2019). The association of anxiety, depression, and worry symptoms on cognitive performance in older adults. *Aging, Neuropsychology, and Cognition*, *26*(2), 161–173. https://doi.org/10.1080/13825585.2017.1416057

Debbané, M., Salaminios, G., Luyten, P., Badoud, D., Armando, M., Solida Tozzi, A., Fonagy, P., & Brent, B. K. (2016). Attachment, Neurobiology, and Mentalizing along the Psychosis Continuum. *Frontiers in Human Neuroscience*, *10*, Article 406. https://doi.org/10.3389/fnhum.2016.00406

Dekker, T. E. G., van der Heijden, H. S., Schirmbeck, F., van Amelsvoort, T., Bartels-Velthuis, A. A., Simons, C. J. P., de Haan, L., & Vermeulen, J. M. (2021). The association between smoking behaviour, social cognition and social functioning

in patients with a non-affective psychotic disorder: A prospective follow-up study. *Schizophrenia Research: Cognition*, *26*, 100206. https://doi.org/10.1016/j.scog.2021.100206

Dennett, D. (1978). Beliefs about beliefs [P&W, SR&B]. *Behavioral and Brain Sciences*, *1*(4), 568–570. https://doi.org/10.1017/S0140525X00076664

Derntl, B., Finkelmeyer, A., Eickhoff, S., Kellermann, T., Falkenberg, D. I., Schneider, F., & Habel, U. (2010). Multidimensional assessment of empathic abilities: Neural correlates and gender differences. *Psychoneuroendocrinology*, *35*(1), 67–82. https://doi.org/10.1016/j.psyneuen.2009.10.006

Devine, R. T., White, N., Ensor, R., & Hughes, C. (2016). Theory of mind in middle childhood: Longitudinal associations with executive function and social competence. *Developmental Psychology*, *52*(5), 758–771. https://doi.org/10.1037/dev0000105

Di Tella, M., Adenzato, M., Catmur, C., Miti, F., Castelli, L., & Ardito, R. B. (2020). The role of alexithymia in social cognition: Evidence from a non-clinical population. *Journal of Affective Disorders*, *273*, 482–492. https://doi.org/10.1016/j.jad.2020.05.012

Di Tella, M., Ardito, R. B., Dutto, F., & Adenzato, M. (2020). On the (lack of) association between theory of mind and executive functions: A study in a non-clinical adult sample. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-74476-0

Di Tella, M., Miti, F., Ardito, R. B., & Adenzato, M. (2020). Social cognition and sex: Are men and women really different? *Personality and Individual Differences*, *162*, 110045. https://doi.org/10.1016/j.paid.2020.110045

Donges, U.-S., Kersting, A., & Suslow, T. (2012). Women's Greater Ability to Perceive Happy Facial Emotion Automatically: Gender Differences in Affective Priming.

Plos one, *7*(7), e41745. https://doi.org/10.1371/journal.pone.0041745

Duval, C., Piolino, P., Bejanin, A., Eustache, F., & Desgranges, B. (2011). Age effects on different components of theory of mind. *Consciousness and Cognition*, *20*(3), 627–642. https://doi.org/10.1016/j.concog.2010.10.025

Dvash, J., & Shamay-Tsoory, S. G. (2014). Theory of Mind and Empathy as Multidimensional Constructs: Neurological Foundations. *Topics in Language Disorders*, *34*(4), 282. https://doi.org/10.1097/TLD.0000000000000040

Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., Kessler, J., Woike, J., Wolf, O., & Convit, A. (2006). Introducing MASC: A Movie for the Assessment of Social Cognition. *Journal of Autism and Developmental Disorders*, *36*, 623–636. https://doi.org/10.1007/s10803-006-0107-0

Ebert, S. (2020). Theory of mind, language, and reading: Developmental relations from early childhood to early adolescence. *Journal of Experimental Child Psychology*, *191*, 104739. https://doi.org/10.1016/j.jecp.2019.104739

Eichner, C., & Berna, F. (2016). Acceptance and Efficacy of Metacognitive Training (MCT) on Positive Symptoms and Delusions in Patients With Schizophrenia: A Meta-analysis Taking Into Account Important Moderators. *Schizophrenia Bulletin*, *42*(4), 952–962. https://doi.org/10.1093/schbul/sbv225

Eidenmueller, K., Grimm, F., Hermann, D., Frischknecht, U., Montag, C., Dziobek, I., Kiefer, F., & Bekier, N. K. (2021). Exploring Influences on Theory of Mind Impairment in Opioid Dependent Patients. *Frontiers in Psychiatry*, *12*, 721690. https://doi.org/10.3389/fpsyt.2021.721690

Erickson, T. M., & Newman, M. G. (2007). Interpersonal and Emotional Processes in Generalized Anxiety Disorder Analogues During Social Interaction Tasks. *Behavior Therapy*, *38*(4), 364–377. https://doi.org/10.1016/j.beth.2006.10.005

Erickson, T. M., Newman, M. G., Siebert, E. C., Carlile, J. A., Scarsella, G. M., & Abelson, J. L. (2016). Does Worrying Mean Caring Too Much? Interpersonal Prototypicality of Dimensional Worry Controlling for Social Anxiety and Depressive Symptoms. *Behavior Therapy*, *47*(1), 14–28. https://doi.org/10.1016/j.beth.2015.08.003

Fagundo, A. B., Torre, R. de la, Jiménez-Murcia, S., Agüera, Z., Granero, R., Tárrega, S., Botella, C., Baños, R., Fernández-Real, J. M., Rodríguez, R., Forcano, L., Frühbeck, G., Gómez-Ambrosi, J., Tinahones, F. J., Fernández-García, J. C., Casanueva, F. F., & Fernández-Aranda, F. (2012). Executive Functions Profile in Extreme Eating/Weight Conditions: From Anorexia Nervosa to Obesity. *Plos One*, *7*(8), e43382. https://doi.org/10.1371/journal.pone.0043382

Fava, M., & Kendler, K. S. (2000). Major Depressive Disorder. *Neuron*, *28*(2), 335–341. https://doi.org/10.1016/S0896-6273(00)00112-4

Felisberti, F., & King, R. (2017). Mind-Reading in Altruists and Psychopaths. In *Neuroscience and Social Science: The Missing Link* (pp. 121–140). https://doi.org/10.1007/978-3-319-68421-5_6

Fett, A.-K. J., Viechtbauer, W., Dominguez, M.-G., Penn, D. L., van Os, J., & Krabbendam, L. (2011). The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: A meta-analysis. *Neuroscience & Biobehavioral Reviews*, *35*(3), 573–588. https://doi.org/10.1016/j.neubiorev.2010.07.001

Fink, E., Begeer, S., Peterson, C. C., Slaughter, V., & de Rosnay, M. (2015). Friendlessness and theory of mind: A prospective longitudinal study. *British Journal of Developmental Psychology*, *33*(1), 1–17. https://doi.org/10.1111/bjdp.12060

Firth, J., Siddiqi, N., Koyanagi, A., Siskind, D., Rosenbaum, S., Galletly, C., Allan, S., Caneo, C., Carney, R., Carvalho, A. F., Chatterton, M. L., Correll, C. U., Curtis, J., Gaughran, F., Heald, A., Hoare, E., Jackson, S. E., Kisely, S., Lovell, K., … Stubbs, B. (2019). The Lancet Psychiatry Commission: A blueprint for protecting physical health in people with mental illness. *The Lancet Psychiatry*, *6*(8), 675–712. https://doi.org/10.1016/S2215-0366(19)30132-4

Fischer, A. L., O'Rourke, N., & Loken Thornton, W. (2017). Age Differences in Cognitive and Affective Theory of Mind: Concurrent Contributions of Neurocognitive Performance, Sex, and Pulse Pressure. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *72*(1), 71–81. https://doi.org/10.1093/geronb/gbw088

Fischer-Kern, M., Doering, S., Taubner, S., Hörz, S., Zimmermann, J., Rentrop, M., Schuster, P., Buchheim, P., & Buchheim, A. (2015). Transference-focused psychotherapy for borderline personality disorder: Change in reflective function. *British Journal of Psychiatry*, *207*(2), 173–174. https://doi.org/10.1192/bjp.bp.113.143842

Fitzpatrick, K. K., Darcy, A., Colborn, D., Gudorf, C., & Lock, J. (2012). Set-shifting among adolescents with anorexia nervosa. *International Journal of Eating Disorders*, *45*(7), 909–912. https://doi.org/10.1002/eat.22027

Flavell, J. H., Flavell, E. R., & Green, F. L. (1983). Development of the appearance-reality distinction. *Cognitive Psychology*, *15*(1), 95–120. https://doi.org/10.1016/0010-0285(83)90005-1

Fluharty, M. E., Heron, J., & Munafò, M. R. (2018). Longitudinal associations of social cognition and substance use in childhood and early adolescence: Findings from the Avon Longitudinal Study of Parents and Children. *European Child &*

Adolescent Psychiatry, *27*(6), 739–752. https://doi.org/10.1007/s00787-017-1068-x

Fodor. J. A. (1992). A theory of the child's theory of mind. *Cognition*, *44*(3), 283–296. https://doi.org/10.1016/0010-0277(92)90004-2

Fodor, J. A. (1983). *The modularity of mind.* The MIT Press eBooks. https://doi.org/10.7551/mitpress/4737.001.0001

Fonagy, P., & Bateman, A. (2008). The Development of Borderline Personality Disorder-A Mentalizing Model. *Journal of Personality Disorders*, *22*(1), 4–21. https://doi.org/10.1521/pedi.2008.22.1.4

Fonagy, P., Luyten, P., & Bateman, A. (2015). Translation: Mentalizing as treatment target in borderline personality disorder. *Personality Disorders: Theory, Research, and Treatment*, *6*, 380–392. https://doi.org/10.1037/per0000113

Fossati, A., Borroni, S., Dziobek, I., Fonagy, P., & Somma, A. (2018). Thinking about assessment: Further evidence of the validity of the Movie for the Assessment of Social Cognition as a measure of mentalistic abilities. *Psychoanalytic Psychology*, *35*(1), 127–141. https://doi.org/10.1037/pap0000130

Freidson, E. (1958). The child's conception of space, by J. Piaget & B. Inhelder. *The American Journal of Sociology, 63*(4), 450–451. https://www.jstor.org/stable/2774171

Frick, C., Lang, S., Kotchoubey, B., Sieswerda, S., Dinu-Biringer, R., Berger, M., Veser, S., Essig, M., & Barnow, S. (2012). Hypersensitivity in Borderline Personality Disorder during Mindreading. *Plos One*, *7*(8), e41650. https://doi.org/10.1371/journal.pone.0041650

Frileux, S., Dubertret, C., & Le Strat, Y. (2020). Short-term impact of alcohol detoxification on facial emotions recognition. *L'Encéphale*, *46*(5), 319–325.

https://doi.org/10.1016/j.encep.2020.04.019

Frith, U., Morton, J., & Leslie, A. M. (1991). The cognitive basis of a biological disorder: Autism. *Trends in Neurosciences*, *14*(10), 433–438. https://doi.org/10.1016/0166-2236(91)90041-R

Gabriel, E. T., Oberger, R., Schmoeger, M., Deckert, M., Vockh, S., Auff, E., & Willinger, U. (2019). Cognitive and affective Theory of Mind in adolescence: Developmental aspects and associated neuropsychological variables. *Psychological Research*, *85*(2), 533–553. https://doi.org/10.1007/s00426-019-01263-6

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *2*(12). 493–501. https://doi.org/10.1016/s1364-6613(98)01262-5

Gartlehner, G., Crotty, K., Kennedy, S., Edlund, M. J., Ali, R., Siddiqui, M., Fortman, R., Wines, R., Persad, E., & Viswanathan, M. (2021). Pharmacological Treatments for Borderline Personality Disorder: A Systematic Review and Meta-Analysis. *Cns Drugs*, *35*(10), 1053–1067. https://doi.org/10.1007/s40263-021-00855-4

Gehlenborg, J., Bücker, L., Berthold, M., Miegel, F., & Moritz, S. (2021). Feasibility, Acceptance, and Safety of Metacognitive Training for Problem and Pathological Gamblers (Gambling-MCT): A Pilot Study. *Journal of Gambling Studies*, *37*(2), 663–687. https://doi.org/10.1007/s10899-020-09975-w

Gillissie, E. S., Lui, L. M. W., Ceban, F., Miskowiak, K., Gok, S., Cao, B., Teopiz, K. M., Ho, R., Lee, Y., Rosenblat, J. D., & McIntyre, R. S. (2022). Deficits of social cognition in bipolar disorder: Systematic review and meta-analysis. *Bipolar Disorders*, *24*(2), 137–148. https://doi.org/10.1111/bdi.13163

Gopnik, A. (2003). The Theory Theory as an Alternative to the Innateness Hypothesis. In L. M. Antony & N. Hornstein (Eds.), *Chomsky and His Critics* (pp. 238–254). Blackwell Publishing Ltd. https://doi.org/10.1002/9780470690024.ch10

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*(1), 26–37. https://doi.org/10.2307/1130386

Gopnik, A., & Wellman, H. M. (1992). Why the Child's Theory of Mind Really Is a Theory. *Mind & Language*, *7*(1–2), 145–171. https://doi.org/10.1111/j.1468-0017.1992.tb00202.x

Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms and the theory theory. *Psychological Bulletin*, *138*(6), 1085–1108. https://doi.org/10.1037/a0028044

Green, M. F. (2016). Impact of Cognitive and Social Cognitive Impairment on Functional Outcomes in Patients With Schizophrenia. *The Journal of Clinical Psychiatry*, *77*(2), 8569. https://doi.org/10.4088/JCP.14074su1c.02

Gur, R. C., Richard, J., Hughett, P., Calkins, M. E., Macy, L., Bilker, W. B., Brensinger, C., & Gur, R. E. (2010). A cognitive neuroscience based computerized battery for efficient measurement of individual differences: Standardization and initial construct validation. *Journal of Neuroscience Methods*, *187*(2), 254–262. https://doi.org/10.1016/j.jneumeth.2009.11.017

Haag, S., Haffner, P., Quinlivan, E., Brüne, M., & Stamm, T. (2016). No differences in visual theory of mind abilities between euthymic bipolar patients and healthy controls. *International Journal of Bipolar Disorders*, *4*(1), 20. https://doi.org/10.1186/s40345-016-0061-5

Haase, M., Frommer, J., Franke, G.-H., Hoffmann, T., Schulze-Muetzel, J., Jäger, S., Grabe, H.-J., Spitzer, C., & Schmitz, N. (2008). From symptom relief to interpersonal change: Treatment outcome and effectiveness in inpatient psychotherapy. *Psychotherapy Research*, *18*(5), 615–624. https://doi.org/10.1080/10503300802192158

Hall, J. A., Carter, J. D., & Horgan, T. G. (2000). Gender differences in nonverbal communication of emotion. In A. H. Fischer (Ed.), *Gender and Emotion: Social Psychological Perspectives* (pp. 97–117). Cambridge University Press. https://doi.org/10.1017/CBO9780511628191.006

Hall, J. A., & Matsumoto, D. (2004). Gender Differences in Judgments of Multiple Emotions From Facial Expressions. *Emotion*, *4*(2), 201–206. https://doi.org/10.1037/1528-3542.4.2.201

Han, S., Fan, Y., & Mao, L. (2008). Gender difference in empathy for pain: An electrophysiological investigation. *Brain Research*, *1196*, 85–93. https://doi.org/10.1016/j.brainres.2007.12.062

Hanegraaf, L., van Baal, S., Hohwy, J., & Verdejo-Garcia, A. (2021). A systematic review and meta-analysis of 'Systems for Social Processes' in borderline personality and substance use disorders. *Neuroscience & Biobehavioral Reviews*, *127*, 572–592. https://doi.org/10.1016/j.neubiorev.2021.04.013

Happé, F. G. E. (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*(2), 129–154. https://doi.org/10.1007/BF02172093

Harari, H., Shamay-Tsoory, S. G., Ravid, M., & Levkovitz, Y. (2010). Double dissociation between cognitive and affective empathy in borderline personality

disorder. *Psychiatry Research*, *175*(3), 277–279. https://doi.org/10.1016/j.psychres.2009.03.002

Harbers, M., Van den Bosch, K., & Meyer, J.-J. (2012). Modeling agents with a theory of mind: Theory–theory versus simulation theory. *Web Intelligence and Agent Systems: An International Journal*, *10*(3), 331–343. https://doi.org/10.3233/WIA-2012-0250

Harkness, K., Sabbagh, M., Jacobson, J., Chowdrey, N., & Chen, T. (2005). Enhanced accuracy of mental state decoding in dysphoric college students. *Cognition & Emotion*, *19*, 999–1025. https://doi.org/10.1080/02699930541000110

Harris, P. L. (1992). From simulation to folk psychology: The case for development. *Mind & Language*, *7*, 120–144. https://doi.org/10.1111/j.1468-0017.1992.tb00201.x

Hassan, I., & Ali, R. (2011). The Association Between Somatic Symptoms, Anxiety Disorders and Substance Use. A Literature Review. *Psychiatric Quarterly*, *82*(4), 315–328. https://doi.org/10.1007/s11126-011-9174-2

Hayden, M. C., Müllauer, P. K., Gaugeler, R., Senft, B., & Andreas, S. (2018). Improvements in mentalization predict improvements in interpersonal distress in patients with mental disorders. *Journal of Clinical Psychology*, *74*(12), 2276–2286. https://doi.org/10.1002/jclp.22673

Healey, M. L., & Grossman, M. (2018). Cognitive and Affective Perspective-Taking: Evidence for Shared and Dissociable Anatomical Substrates. *Frontiers in Neurology*, *9*, 491. https://doi.org/10.3389/fneur.2018.00491

Henry, J. D., Phillips, L. H., Ruffman, T., & Bailey, P. E. (2013). A meta-analytic review of age differences in theory of mind. *Psychology and Aging*, *28*(3), 826–839. https://doi.org/10.1037/a0030677

Herpertz, S. C., Dietrich, T. M., Wenning, B., Krings, T., Erberich, S. G., Willmes, K., Thron, A., & Sass, H. (2001). Evidence of abnormal amygdala functioning in borderline personality disorder: A functional MRI study. *Biological Psychiatry*, *50*(4), 292–298. https://doi.org/10.1016/S0006-3223(01)01075-7

Hess, T. M. (2014). Selective engagement of cognitive resources: Motivational influences on older adults' cognitive functioning. *Perspectives on Psychological Science*, *9*(4), 388–407. https://doi.org/10.1177/1745691614527465

Hezel, D. M., & McNally, R. J. (2014). Theory of Mind Impairments in Social Anxiety Disorder. *Behavior Therapy*, *45*(4), 530–540. https://doi.org/10.1016/j.beth.2014.02.010

Hillmann, K., Neukel, C., Krauch, M., Spohn, A., Schnell, K., Herpertz, S. C., & Bertsch, K. (2021). Cognitive and Affective Theory of Mind in Female Patients With Borderline Personality Disorder. *Journal of Personality Disorders*, *35*(5), 672–690. https://doi.org/10.1521/pedi_2020_34_490

Hoffmann, H., Kessler, H., Eppel, T., Rukavina, S., & Traue, H. (2010). Expression intensity, gender and facial emotion recognition: Women recognize only subtle facial emotions better than men. *Acta Psychologica*, *135*, 278–283. https://doi.org/10.1016/j.actpsy.2010.07.012

Huang, Y.-L., Chen, T.-T., Dziobek, I., & Tseng, H.-H. (2023). Mentalizing in a movie for the assessment of social cognition (MASC): The validation in a taiwanese sample. *BMC Psychology*, *11*(1), 287. https://doi.org/10.1186/s40359-023-01321-0

Imuta, K., Henry, J. D., Slaughter, V., Selcuk, B., & Ruffman, T. (2016). Theory of mind and prosocial behavior in childhood: A meta-analytic review. *Developmental Psychology*, *52*, 1192–1205. https://doi.org/10.1037/dev0000140

Inoue, Y., Yamada, K., & Kanba, S. (2006). Deficit in theory of mind is a risk for relapse of major depression. *Journal of Affective Disorders*, *95*(1), 125–127. https://doi.org/10.1016/j.jad.2006.04.018

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research Domain Criteria (RDoC): Toward a New Classification Framework for Research on Mental Disorders. *American Journal of Psychiatry*, *167*(7), 748–751. https://doi.org/10.1176/appi.ajp.2010.09091379

Jelinek, L., Hauschildt, M., Wittekind, C. E., Schneider, B. C., Kriston, L., & Moritz, S. (2016). Efficacy of Metacognitive Training for Depression: A Randomized Controlled Trial. *Psychotherapy and Psychosomatics*, *85*(4), 231–234. https://doi.org/10.1159/000443699

Jenkins, J. M., & Astington, J. W. (1996). Cognitive factors and family structure associated with theory of mind development in young children. *Developmental Psychology*, *32*(1), 70–78. https://doi.org/10.1037/0012-1649.32.1.70

Kalbe, E., Grabenhorst, F., Brand, M., Kessler, J., Hilker, R., & Markowitsch, H. J. (2007). Elevated emotional reactivity in affective but not cognitive components of theory of mind: A psychophysiological study. *Journal of Neuropsychology*, *1*(1), 27–38. https://doi.org/10.1348/174866407X180792

Kalbe, E., Schlegel, M., Sack, A. T., Nowak, D. A., Dafotakis, M., Bangard, C., Brand, M., Shamay-Tsoory, S., Onur, O. A., & Kessler, J. (2010). Dissociating cognitive from affective theory of mind: A TMS study. *Cortex*, *46*(6), 769–780. https://doi.org/10.1016/j.cortex.2009.07.010

Kaltefleiter, L. J., Sodian, B., Kristen-Antonow, S., Grosse Wiesmann, C., & Schuwerk, T. (2021). Does syntax play a role in Theory of Mind development before the

age of 3 years? *Infant Behavior and Development*, *64*, 101575. https://doi.org/10.1016/j.infbeh.2021.101575

Kang, J.-A., Hubbard, G., & Hong, S. (2019). Gender and credibility in branded storytelling. *Gender in Management, 34(8), 702-714*. https://doi.org/10.1108/GM-02-2019-0015

Kerr, N., Dunbar, R. I. M., & Bentall, R. P. (2003). Theory of mind deficits in bipolar affective disorder. *Journal of Affective Disorders*, *73*(3), 253–259. https://doi.org/10.1016/S0165-0327(02)00008-3

Kertz, S. J., Koran, J., Stevens, K. T., & Björgvinsson, T. (2015). Repetitive negative thinking predicts depression and anxiety symptom improvement during brief cognitive behavioral therapy. *Behaviour Research and Therapy*, *68*, 54–63. https://doi.org/10.1016/j.brat.2015.03.006

Kittel, A. F. D., Olderbak, S., & Wilhelm, O. (2022). Sty in the Mind's Eye: A Meta-Analytic Investigation of the Nomological Network and Internal Consistency of the "Reading the Mind in the Eyes" Test. *Assessment*, *29*(5), 872–895. https://doi.org/10.1177/1073191121996469

Koelkebeck, K., Hirao, K., Miyata, J., Kawada, R., Saze, T., Dannlowski, U., Ubukata, S., Ohrmann, P., Bauer, J., Pedersen, A., Fukuyama, H., Sawamoto, N., Takahashi, H., & Murai, T. (2013). Impact of gray matter reductions on theory of mind abilities in patients with schizophrenia. *Social Neuroscience*, *8*(6), 631–639. https://doi.org/10.1080/17470919.2013.837094

Kohler, C. G., Walker, J. B., Martin, E. A., Healey, K. M., & Moberg, P. J. (2010). Facial Emotion Perception in Schizophrenia: A Meta-analytic Review. *Schizophrenia Bulletin*, *36*(5), 1009–1019. https://doi.org/10.1093/schbul/sbn192

Koob, G. F. (2014). Neurocircuitry of alcohol addiction: synthesis from animal models.

In J. W. Swanson & M. Matharu (Eds.), *Handbook of Clinical Neurology* (125, pp. 33–54). Elsevier. https://doi.org/10.1016/B978-0-444-62619-6.00003-3

Kopera, M., Zaorska, J., Trucco, E. M., Suszek, H., Kobyliński, P., Zucker, R. A., Nowakowska, M., Wojnar, M., & Jakubczyk, A. (2020). Childhood trauma, alexithymia, and mental states recognition among individuals with alcohol use disorder and healthy controls. *Drug and Alcohol Dependence*, *217*, 108301. https://doi.org/10.1016/j.drugalcdep.2020.108301

Kornreich, C., Delle-Vigne, D., Knittel, J., Nerincx, A., Campanella, S., Noel, X., Hanak, C., Verbanck, P., & Ermer, E. (2011). Impaired conditional reasoning in alcoholics: A negative impact on social interactions and risky behaviors? *Addiction*, *106*(5), 951–959. https://doi.org/10.1111/j.1360-0443.2010.03346.x

Kornreich, C., Philippot, P., Foisy, M.-L., Blairy, S., Raynaud, E., Dan, B., Hess, U., Noël, X., Pelc, I., & Verbanck, P. (2002). Impaired emotional facial expression recognition is associated with interpersonal problems in alcoholism. *Alcohol and Alcoholism*, *37*(4), 394–400. https://doi.org/10.1093/alcalc/37.4.394

Krach, S., Blümel, I., Marjoram, D., Lataster, T., Krabbendam, L., Weber, J., van Os, J., & Kircher, T. (2009). Are women better mindreaders? Sex differences in neural correlates of mentalizing detected with functional MRI. *BMC Neuroscience*, *10*(1), 9. https://doi.org/10.1186/1471-2202-10-9

Kret, M. E., & De Gelder, B. (2012). A review on sex differences in processing emotional signals. *Neuropsychologia*, *50*(7), 1211–1221. https://doi.org/10.1016/j.neuropsychologia.2011.12.022

Krych-Appelbaum, M., Law, J. B., Jones, D., Barnacz, A., Johnson, A., & Keenan, J. P. (2007). “I think I know what you mean”: The role of theory of mind in collaborative communication. *Interaction Studies*, *8*(2), 267–280.

https://doi.org/10.1075/is.8.2.05kry

Kumar, L., Skrzynski, C. J., & Creswell, K. G. (2022). Systematic review and meta-analysis on the association between theory of mind and alcohol problems in non-clinical samples. *Alcoholism: Clinical and Experimental Research*, 1–13. https://doi.org/10.1111/acer.14943

Kurtz, M. M., & Gerraty, R. T. (2009). A meta-analytic investigation of neurocognitive deficits in bipolar illness: Profile and effects of clinical state. *Neuropsychology*, *23*, 551–562. https://doi.org/10.1037/a0016277

Kvarstein, E. H., Folmo, E., Antonsen, B. T., Normann-Eide, E., Pedersen, G., & Wilberg, T. (2020). Social Cognition Capacities as Predictors of Outcome in Mentalization-Based Treatment (MBT). *Frontiers in Psychiatry*, *11,* 691. https://doi.org/10.3389/fpsyt.2020.00691

Lanius, U. F. (Ed.). (2014). *Neurobiology and treatment of traumatic dissociation: Toward an embodied self*. Springer.

Lannoy, S., Gilles, F., Benzerouk, F., Henry, A., Oker, A., Raucher-Chéné, D., Besche-Richard, C., & Gierski, F. (2020). Disentangling the role of social cognition processes at early steps of alcohol abuse: The influence of affective theory of mind. *Addictive Behaviors*, *102*, 106187. https://doi.org/10.1016/j.addbeh.2019.106187

Le Berre, A.-P. (2019). Emotional Processing and Social Cognition in Alcohol Use Disorder. *Neuropsychology*, *33*(6), 808–821. https://doi.org/10.1037/neu0000572

Lenton-Brym, A. P., Moscovitch, D. A., Vidovic, V., Nilsen, E., & Friedman, O. (2018). Theory of mind ability in high socially anxious individuals. *Anxiety, Stress, & Coping*, *31*(5), 487–499. https://doi.org/10.1080/10615806.2018.1483021

Leslie, A. M. (1994). Pretending and believing: Issues in the theory of ToMM. *Cognition*, *50*(1–3), 211–238. https://doi.org/10.1016/0010-0277(94)90029-9

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in 'theory of mind.' *Trends in Cognitive Sciences*, *8*(12), 528–533. https://doi.org/10.1016/j.tics.2004.10.001

Leslie, A. M., German, T. P., & Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cognitive Psychology*, *50*(1), 45–85. https://doi.org/10.1016/j.cogpsych.2004.06.002

Leslie, A. M., & Thaiss, L. (1992). Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition*, *43*(3), 225–251. https://doi.org/10.1016/0010-0277(92)90013-8

Li, D., Li, X., Yu, F., Chen, X., Zhang, L., Li, D., Wei, Q., Zhang, Q., Zhu, C., & Wang, K. (2017). Comparing the ability of cognitive and affective Theory of Mind in adolescent onset schizophrenia. *Neuropsychiatric Disease and Treatment*, *Volume 13*, 937–945. https://doi.org/10.2147/NDT.S128116

Li, J., Wang, H., Li, M., Shen, Q., Li, X., Zhang, Y., Peng, J., Rong, X., & Peng, Y. (2020). Effect of alcohol use disorders and alcohol intake on the risk of subsequent depressive symptoms: A systematic review and meta-analysis of cohort studies. *Addiction*, *115*(7), 1224–1243. https://doi.org/10.1111/add.14935

Lieb, K., Zanarini, M. C., Schmahl, C., Linehan, M. M., & Bohus, M. (2004). Borderline personality disorder. *The Lancet*, *364*(9432), 453–461. https://doi.org/10.1016/S0140-6736(04)16770-6

Liebherz, S., & Rabung, S. (2014). Do Patients' Symptoms and Interpersonal Problems Improve in Psychotherapeutic Hospital Treatment in Germany? - A Systematic

Review and Meta-Analysis. *PLOS ONE*, *9*(8), e105329. https://doi.org/10.1371/journal.pone.0105329

Linehan, M. M. (1993). *Skills training manual for treating borderline personality disorder*. Guilford Press.

Livingston, L. A., Carr, B., & Shah, P. (2019). Recent Advances and New Directions in Measuring Theory of Mind in Autistic Adults. *Journal of Autism and Developmental Disorders*, *49*(4), 1738–1744. https://doi.org/10.1007/s10803-018-3823-3

Longobardi, E., Spataro, P., & Rossi-Arnaud, C. (2019). Direct and Indirect Associations of Empathy, Theory of Mind, and Language with Prosocial Behavior: Gender Differences in Primary School Children. *The Journal of Genetic Psychology*, *180*(6), 266–279. https://doi.org/10.1080/00221325.2019.1653817

Luke, N., & Banerjee, R. (2013). Differentiated associations between childhood maltreatment experiences and social understanding: A meta-analysis and systematic review. *Developmental Review*, *33*(1), 1–28. https://doi.org/10.1016/j.dr.2012.10.001

MacDonald, P. M., Kirkpatrick, S. W., & Sullivan, L. A. (1996). Production of Facial Expressions of Emotion in Preschool Children. *Perceptual and Motor Skills*, *82*(1), 76–78. https://doi.org/10.2466/pms.1996.82.1.76

Maleki, G., Zabihzadeh, A., Richman, M. J., Demetrovics, Z., & Mohammadnejad, F. (2020). Decoding and reasoning mental states in major depression and social anxiety disorder. *BMC Psychiatry*, *20*(1), 463. https://doi.org/10.1186/s12888-020-02873-w

Malhi, G. S., Lagopoulos, J., Das, P., Moss, K., Berk, M., & Coulston, C. M. (2008). A

functional MRI study of Theory of Mind in euthymic bipolar disorder patients. *Bipolar Disorders*, *10*(8), 943–956. https://doi.org/10.1111/j.1399-5618.2008.00643.x

Marjoram, D., Job, D. E., Whalley, H. C., Gountouna, V.-E., McIntosh, A. M., Simonotto, E., Cunningham-Owens, D., Johnstone, E. C., & Lawrie, S. (2006). A visual joke fMRI investigation into Theory of Mind and enhanced risk of schizophrenia. *NeuroImage*, *31*(4), 1850–1858. https://doi.org/10.1016/j.neuroimage.2006.02.011

Marvin, R. S., Greenberg, M. T., & Mossler, D. G. (1976). The Early Development of Conceptual Perspective Taking: Distinguishing among Multiple Perspectives. *Child Development*, *47*(2), 511–514. https://doi.org/10.2307/1128810

Mason, T. B., Lesser, E. L., Dolgon-Krutolow, A. R., Wonderlich, S. A., & Smith, K. E. (2021). An updated transdiagnostic review of social cognition and eating disorder psychopathology. *Journal of Psychiatric Research*, *143*, 602–627. https://doi.org/10.1016/j.jpsychires.2020.11.019

Mátyássy, A., Kelemen, O., Sárközi, Z., Janka, Z., & Kéri, S. (2006). Recognition of complex mental states in patients with alcoholism after long-term abstinence. *Alcohol and Alcoholism*, *41*(5), 512–514. https://doi.org/10.1093/alcalc/agl045

Maurage, P., D'Hondt, F., de Timary, P., Mary, C., Franck, N., & Peyroux, E. (2016). Dissociating Affective and Cognitive Theory of Mind in Recently Detoxified Alcohol-Dependent Individuals. *Alcoholism: Clinical and Experimental Research*, *40*(9), 1926–1934. https://doi.org/10.1111/acer.13155

Maurage, P., Grynberg, D., Noël, X., Joassin, F., Hanak, C., Verbanck, P., Luminet, O., de Timary, P., Campanella, S., & Philippot, P. (2011). The "Reading the Mind in the Eyes" test as a new way to explore complex emotions decoding in

alcohol dependence. *Psychiatry Research*, *190*(2–3), 375–378. https://doi.org/10.1016/j.psychres.2011.06.015

Maurage, P., Grynberg, D., Noël, X., Joassin, F., Philippot, P., Hanak, C., Verbanck, P., Luminet, O., de Timary, P., & Campanella, S. (2011). Dissociation Between Affective and Cognitive Empathy in Alcoholism: A Specific Deficit for the Emotional Dimension. *Alcoholism: Clinical and Experimental Research*, *35*(9), 1662–1668. https://doi.org/10.1111/j.1530-0277.2011.01512.x

McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological Bulletin*, *126*, 424–453. https://doi.org/10.1037/0033-2909.126.3.424

McIntyre, R. S., Berk, M., Brietzke, E., Goldstein, B. I., López-Jaramillo, C., Kessing, L. V., Malhi, G. S., Nierenberg, A. A., Rosenblat, J. D., Majeed, A., Vieta, E., Vinberg, M., Young, A. H., & Mansur, R. B. (2020). Bipolar disorders. *The Lancet*, *396*(10265), 1841–1856. https://doi.org/10.1016/S0140-6736(20)31544-0

McLaren, V., Gallagher, M., Hopwood, C. J., & Sharp, C. (2022). Hypermentalizing and Borderline Personality Disorder: A Meta-Analytic Review. *American Journal of Psychotherapy*, *75*(1), 21–31. https://doi.org/10.1176/appi.psychotherapy.20210018

Meinhardt-Injac, B., Daum, M. M., & Meinhardt, G. (2020). Theory of mind development from adolescence to adulthood: Testing the two-component model. *British Journal of Developmental Psychology*, *38*(2), 289–303. https://doi.org/10.1111/bjdp.12320

Mestre, M. V., Samper, P., Frías, M. D., & Tur, A. M. (2009). Are Women More

Empathetic than Men? A Longitudinal Study in Adolescence. *The Spanish Journal of Psychology*, *12*(1), 76–83. https://doi.org/10.1017/S1138741600001499

Meulemeester, C. D., Vansteelandt, K., Luyten, P., & Lowyck, B. (2018). Mentalizing as a Mechanism of Change in the Treatment of Patients With Borderline Personality Disorder: A Parallel Process Growth Modeling Approach. *Personality Disorders: Theory, Research and Treatment*. *9*(1), 22–29 https://doi.org/10.1037/per0000256

Miegel, F., Demiralay, C., Moritz, S., Wirtz, J., Hottenrott, B., & Jelinek, L. (2020). Metacognitive Training for Obsessive-Compulsive Disorder: A study protocol for a randomized controlled trial. *BMC Psychiatry*, *20*(1), 350. https://doi.org/10.1186/s12888-020-02648-3

Miller, W. R., & Wilbourne, P. L. (2002). Mesa Grande: A methodological analysis of clinical trials of treatments for alcohol use disorders. *Addiction*, *97*(3), 265–277. https://doi.org/10.1046/j.1360-0443.2002.00019.x

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and Theory of Mind: Meta-Analysis of the Relation Between Language Ability and False-belief Understanding. *Child Development*, *78*(2), 622–646. https://doi.org/10.1111/j.1467-8624.2007.01018.x

Moll, H., & Tomasello, M. (2006). Level 1 perspective-taking at 24 months of age. *British Journal of Developmental Psychology*, *24*(3), 603–613. https://doi.org/10.1348/026151005X55370

Montag, C., Ehrlich, A., Neuhaus, K., Dziobek, I., Heekeren, H. R., Heinz, A., & Gallinat, J. (2010). Theory of mind impairments in euthymic bipolar patients. *Journal of Affective Disorders*, *123*(1), 264–269.

https://doi.org/10.1016/j.jad.2009.08.017

Montag, C., Neuhaus, K., Lehmann, A., Krüger, K., Dziobek, I., Heekeren, H. R., Heinz, A., & Gallinat, J. (2012). Subtle deficits of cognitive theory of mind in unaffected first-degree relatives of schizophrenia patients. *European Archives of Psychiatry and Clinical Neuroscience*, *262*(3), 217–226. https://doi.org/10.1007/s00406-011-0250-2

Montagne, B., Kessels, R. P. C., Frigerio, E., de Haan, E. H. F., & Perrett, D. I. (2005). Sex differences in the perception of affective facial expressions: Do men really lack emotional sensitivity? *Cognitive Processing*, *6*(2), 136–141. https://doi.org/10.1007/s10339-005-0050-6

Mor, N., & Winquist, J. (2002). Self-focused attention and negative affect: A meta-analysis. *Psychological Bulletin*, *128*, 638–662. https://doi.org/10.1037/0033-2909.128.4.638

Moritz, S., Andreou, C., Schneider, B. C., Wittekind, C. E., Menon, M., Balzan, R. P., & Woodward, T. S. (2014). Sowing the seeds of doubt: A narrative review on metacognitive training in schizophrenia. *Clinical Psychology Review*, *34*(4), 358–366. https://doi.org/10.1016/j.cpr.2014.04.004

Moritz, S., & Lysaker, P. H. (2018). Metacognition – What did James H. Flavell really say and the implications for the conceptualization and design of metacognitive interventions. *Schizophrenia Research*, *201*, 20–26. https://doi.org/10.1016/j.schres.2018.06.001

Nandrino, J.-L., Gandolphe, M.-C., Alexandre, C., Kmiecik, E., Yguel, J., & Urso, L. (2014). Cognitive and affective theory of mind abilities in alcohol-dependent patients: The role of autobiographical memory. *Drug and Alcohol Dependence*, *143*, 65–73. https://doi.org/10.1016/j.drugalcdep.2014.07.010

National Institutes of Mental Health. (2020, September 25). *NIMH » About RDoC*. Research Domain Criteria (RDoC). https://www.nimh.nih.gov/research/research-funded-by-nimh/rdoc/about-rdoc.shtml

Navarra-Ventura, G., Fernandez-Gonzalo, S., Turon, M., Pousa, E., Palao, D., Cardoner, N., & Jodar, M. (2018). Gender Differences in Social Cognition: A Cross-Sectional Pilot Study of Recently Diagnosed Patients with Schizophrenia and Healthy Subjects. *The Canadian Journal of Psychiatry*, *63*(8), 538–546. https://doi.org/10.1177/0706743717746661

Németh, N., Mátrai, P., Hegyi, P., Czéh, B., Czopf, L., Hussain, A., Pammer, J., Szabó, I., Solymár, M., Kiss, L., Hartmann, P., Szilágyi, Á. L., Kiss, Z., & Simon, M. (2018). Theory of mind disturbances in borderline personality disorder: A meta-analysis. *Psychiatry Research*, *270*, 143–153. https://doi.org/10.1016/j.psychres.2018.08.049

Németh, N., Péterfalvi, Á., Czéh, B., Tényi, T., & Simon, M. (2020). Examining the Relationship Between Executive Functions and Mentalizing Abilities of Patients With Borderline Personality Disorder. *Frontiers in Psychology*, *11*, 1583. https://doi.org/10.3389/fpsyg.2020.01583

Nestor, B. A., Sutherland, S., & Garber, J. (2022). Theory of mind performance in depression: A meta-analysis. *Journal of Affective Disorders*, *303*, 233–244. https://doi.org/10.1016/j.jad.2022.02.028

Nichols, S., & Stich, S. P. (2003). *Mindreading: An integrated account of pretence, self-awareness, and understanding other minds.* Clarendon Press/Oxford University Press. https://doi.org/10.1093/0198236107.001.0001

Njomboro, P., Deb, S., & Humphreys, G. W. (2008). Dissociation between Decoding

and Reasoning about Mental States in Patients with Theory of Mind Reasoning Impairments. *Journal of Cognitive Neuroscience*, *20*(9), 1557–1564. https://doi.org/10.1162/jocn.2008.20118

Normann, N., van Emmerik, A. A. P., & Morina, N. (2014). The Efficacy of Metacognitive Therapy for Anxiety and Depression: A Meta-Analytic Review. *Depression and Anxiety*, *31*(5), 402–411. https://doi.org/10.1002/da.22273

Normann-Eide, E., Antonsen, B. T., Kvarstein, E. H., Pedersen, G., Vaskinn, A., & Wilberg, T. (2020). Are Impairments in Theory of Mind Specific to Borderline Personality Disorder? *Journal of Personality Disorders*, *34*(6), 827–841. https://doi.org/10.1521/pedi_2019_33_417

Oakley, B. F. M. (2016). Theory of mind is not theory of emotion: A cautionary note on the Reading the Mind in the Eyes Test. *Journal of Abnormal Psychology*, *125*(6), 818–823. https://doi.org/10.1037/abn0000182

Olderbak, S., & Wilhelm, O. (2020). Overarching Principles for the Organization of Socioemotional Constructs. *Current Directions in Psychological Science*, *29*(1), 63–70. https://doi.org/10.1177/0963721419884317

Oliveira, L. M., Bermudez, M. B., Macedo, M. J. de A., & Passos, I. C. (2018). Comorbid social anxiety disorder in patients with alcohol use disorder: A systematic review. *Journal of Psychiatric Research*, *106*, 8–14. https://doi.org/10.1016/j.jpsychires.2018.09.008

Onuoha, R. C., Quintana, D. S., Lyvers, M., & Guastella, A. J. (2016). A Meta-analysis of Theory of Mind in Alcohol Use Disorders. *Alcohol and Alcoholism*, *51*(4), 410–415. https://doi.org/10.1093/alcalc/agv137

Osterhaus, C., & Bosacki, S. L. (2022). Looking for the lighthouse: A systematic review of advanced theory-of-mind tests beyond preschool. *Developmental Review*,

64, 101021. https://doi.org/10.1016/j.dr.2022.101021

Osterhaus, C., & Koerber, S. (2021). The Development of Advanced Theory of Mind in Middle Childhood: A Longitudinal Study From Age 5 to 10 Years. *Child Development*, *92*(5), 1872–1888. https://doi.org/10.1111/cdev.13627

Osterhaus, C., Koerber, S., & Sodian, B. (2016). Scaling of Advanced Theory-of-Mind Tasks. *Child Development*, *87*(6), 1971–1991. https://doi.org/10.1111/cdev.12566

Pabst, A., Gautier, M., & Maurage, P. (2022). Tasks and investigated components in social cognition research among adults with alcohol use disorder: A critical scoping review. *Psychology of Addictive Behaviors*, *38*(8), 999–1011. https://doi.org/10.1037/adb0000874

Pang, C., Li, W., Zhou, Y., Gao, T., & Han, S. (2023). Are women more empathetic than men? Questionnaire and EEG estimations of sex/gender differences in empathic ability. *Social Cognitive and Affective Neuroscience*, *18*(1). https://doi.org/10.1093/scan/nsad008

Paris, J., Zweig-Frank, H., & Guzder, J. (1994). Psychological risk factors for borderline personality disorder in female patients. *Comprehensive Psychiatry*, *35*(4), 301–305. https://doi.org/10.1016/0010-440X(94)90023-X

Patrick, C. J., & Hajcak, G. (2016). RDoC: Translating promise into progress. *Psychophysiology*, *53*(3), 415–424. https://doi.org/10.1111/psyp.12612

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, *5*(2), 125–137. https://doi.org/10.1111/j.2044-835X.1987.tb01048.x

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that…" attribution of

second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, *39*(3), 437–471. https://doi.org/10.1016/0022-0965(85)90051-7

Petersen, R., Brakoulias, V., & Langdon, R. (2016). An experimental investigation of mentalization ability in borderline personality disorder. *Comprehensive Psychiatry*, *64*, 12–21. https://doi.org/10.1016/j.comppsych.2015.10.004

Petrovic, S. A., Kaurin, N., Knezevic, J., & Maric, N. P. (2023). Theory of Mind in Typical Adults: Sex-Differences and Its Associations with Anxiety and Depression Symptoms. *Archives of Clinical Neuropsychology*, *38*(6), 913–921. https://doi.org/10.1093/arclin/acad007

Peyroux, E., & Franck, N. (2014). Rc2s: A Cognitive Remediation Program to Improve Social Cognition in Schizophrenia and Related Disorders. *Frontiers in Human Neuroscience*, *8*, Article 400. https://doi.org/10.3389/fnhum.2014.00400

Peyroux, E., Prost, Z., Danset-Alexandre, C., Brenugat-Herne, L., Carteau-Martin, I., Gaudelus, B., Jantac, C., Attali, D., Amado, I., Graux, J., Houy-Durand, E., Plasse, J., & Franck, N. (2019). From “under” to “over” social cognition in schizophrenia: Is there distinct profiles of impairments according to negative and positive symptoms? *Schizophrenia Research: Cognition*, *15*, 21–29. https://doi.org/10.1016/j.scog.2018.10.001

Philipp, R., Kriston, L., Lanio, J., Kühne, F., Härter, M., Moritz, S., & Meister, R. (2018). Effectiveness of metacognitive interventions for mental disorders in adults—A systematic review and meta-analysis (Metacog). *Clinical Psychology & Psychotherapy*, *26*(2), 227–240. https://doi.org/10.1002/cpp.2345

Poletti, M., Sonnoli, A., & Bonuccelli, U. (2014). Mild Depressive Symptoms are Associated With Enhanced Affective Theory of Mind in Nonclinical Adult

Women. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *26*(2), E63–E64. https://doi.org/10.1176/appi.neuropsych.13060123

Pons, F., Harris, P. L., & de Rosnay, M. (2004). Emotion comprehension between 3 and 11 years: Developmental periods and hierarchical organization. *European Journal of Developmental Psychology*, *1*(2), 127–152. https://doi.org/10.1080/17405620344000022

Popolo, R., Borsella, I., Meschini, L., Pianella, U., Zampaglione, G., Vinci, G., Salvatore, G., Dimaggio, G., Vecchiotti, R., Huling, K., & Lysaker, P. H. (2020). Cognitive theory of mind in bipolar disorder: Comparisons with healthy controls and associations with function. *Psychiatry Research*, *290*, 113030. https://doi.org/10.1016/j.psychres.2020.113030

Preis, M. A., Golm, D., Kröner-Herwig, B., & Barke, A. (2017). Examining differences in cognitive and affective theory of mind between persons with high and low extent of somatic symptoms: An experimental study. *BMC Psychiatry*, *17*(1), 200. https://doi.org/10.1186/s12888-017-1360-9

Preißler, S., Dziobek, I., Ritter, K., Heekeren, H. R., & Roepke, S. (2010). Social Cognition in Borderline Personality Disorder: Evidence for Disturbed Recognition of the Emotions, Thoughts, and Intentions of others. *Frontiers in Behavioral Neuroscience*, *4*, 182–182. https://doi.org/10.3389/fnbeh.2010.00182

Premack, D., & Woodruff, G. (1978). Does a chimpanzee have a theory of mind. *Behavioral and Brain Sciences*, *1*, 515–526. https://doi.org/10.1017/S0140525X00076512

Przeworski, A., Newman, M. G., Pincus, A. L., Kasoff, M. B., Yamasaki, A. S., Castonguay, L. G., & Berlin, K. S. (2011). Interpersonal Pathoplasticity in

Individuals with Generalized Anxiety Disorder. *Journal of Abnormal Psychology*, *120*(2), 286–298. https://doi.org/10.1037/a0023334

Quenneville, A. F., Kalogeropoulou, E., Küng, A.-L., Hasler, R., Nicastro, R., Prada, P., & Perroud, N. (2020). Childhood maltreatment, anxiety disorders and outcome in borderline personality disorder. *Psychiatry Research*, *284*, 112688. https://doi.org/10.1016/j.psychres.2019.112688

Quesque, F., & Rossetti, Y. (2020). What Do Theory-of-Mind Tasks Actually Measure? Theory and Practice. *Perspectives on Psychological Science*, *15*(2), 384–396. https://doi.org/10.1177/1745691619896607

Rahman, Q., Wilson, G. D., & Abrahams, S. (2004). Sex, sexual orientation, and identification of positive and negative facial affect. *Brain and Cognition*, *54*(3), 179–185. https://doi.org/10.1016/j.bandc.2004.01.002

Raimo, S., Cropano, M., Roldán-Tapia, M. D., Ammendola, L., Malangone, D., & Santangelo, G. (2022). Cognitive and Affective Theory of Mind across Adulthood. *Brain Sciences*, *12*(7), 899. https://doi.org/10.3390/brainsci12070899

Ray, L. A., Bujarski, S., Grodin, E., Hartwell, E., Green, R., Venegas, A., Lim, A. C., Gillis, A., & Miotto, K. (2019). State-of-the-art behavioral and pharmacological treatments for alcohol use disorder. *The American Journal of Drug and Alcohol Abuse*, *45*(2), 124–140. https://doi.org/10.1080/00952990.2018.1528265

Rehm, J., Marmet, S., Anderson, P., Gual, A., Kraus, L., Nutt, D. J., Room, R., Samokhvalov, A. V., Scafato, E., Trapencieris, M., Wiers, R. W., & Gmel, G. (2013). Defining Substance Use Disorders: Do We Really Need More Than Heavy Use? *Alcohol and Alcoholism*, *48*(6), 633–640. https://doi.org/10.1093/alcalc/agt127

Rehm, J., Mathers, C., Popova, S., Thavorncharoensap, M., Teerawattananon, Y., & Patra, J. (2009). Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders. *The Lancet*, *373*(9682), 2223–2233. https://doi.org/10.1016/S0140-6736(09)60746-7

Renneberg, B., & Herpertz, S. C. (2020). *Persönlichkeitsstörungen*. Hogrefe Verlag GmbH & Company KG.

Richman, M. J., & Unoka, Z. (2015). Mental state decoding impairment in major depression and borderline personality disorder: Meta-analysis. *British Journal of Psychiatry*, *207*(6), 483–489. https://doi.org/10.1192/bjp.bp.114.152108

Ritter, K., Dziobek, I., Preißler, S., Rüter, A., Vater, A., Fydrich, T., Lammers, C.-H., Heekeren, H. R., & Roepke, S. (2011). Lack of empathy in patients with narcissistic personality disorder. *Psychiatry Research*, *187*(1), 241–247. https://doi.org/10.1016/j.psychres.2010.09.013

Robinson, C. S. H., Fokas, K., & Witkiewitz, K. (2018). Relationship Between Empathic Processing and Drinking Behavior in Project MATCH. *Addictive Behaviors*, *77*, 180–186. https://doi.org/10.1016/j.addbeh.2017.10.001

Roheger, M., Hranovska, K., Martin, A. K., & Meinzer, M. (2022). A systematic review and meta-analysis of social cognition training success across the healthy lifespan. *Scientific Reports*, *12*(1), 3544. https://doi.org/10.1038/s41598-022-07420-z

Rosello, B., Berenguer, C., Baixauli, I., García, R., & Miranda, A. (2020). Theory of Mind Profiles in Children With Autism Spectrum Disorder: Adaptive/Social Skills and Pragmatic Competence. *Frontiers in Psychology*, *11*, 567401. https://doi.org/10.3389/fpsyg.2020.567401

Rossi, R., Corbo, D., & Magni, L. (2023). Metacognitive interpersonal therapy in

borderline personality disorder: Clinical and neuroimaging outcomes from the climamithe study—A randomized clinical trial. *Personality Disorders: Theory, Research and Treatment*, *14*(4), 452–466. https://doi.org/10.1037/per0000621

Rothman, A. J., Sheeran, P., & Wood, W. (2009). Reflective and Automatic Processes in the Initiation and Maintenance of Dietary Change. *Annals of Behavioral Medicine*, *38*(S1), 4–17. https://doi.org/10.1007/s12160-009-9118-3

Ruocco, A. C., & Carcone, D. (2016). A Neurobiological Model of Borderline Personality Disorder: Systematic and Integrative Review. *Harvard Review of Psychiatry*, *24*(5), 311–329. https://doi.org/10.1097/HRP.0000000000000123

Rupp, C. I., Derntl, B., Osthaus, F., Kemmler, G., & Fleischhacker, W. W. (2017). Impact of Social Cognition on Alcohol Dependence Treatment Outcome: Poorer Facial Emotion Recognition Predicts Relapse/Dropout. *Alcoholism: Clinical and Experimental Research*, *41*(12), 2197–2206. https://doi.org/10.1111/acer.13522

Rupp, C. I., Junker, D., Kemmler, G., Mangweth-Matzek, B., & Derntl, B. (2021). Do Social Cognition Deficits Recover with Abstinence in Alcohol-Dependent Patients? *Alcoholism: Clinical and Experimental Research*, *45*(2), 470–479. https://doi.org/10.1111/acer.14537

Russell, T. A., Tchanturia, K., Rahman, Q., & Schmidt, U. (2007). Sex differences in theory of mind: A male advantage on Happé's "cartoon" task. *Cognition and Emotion*, *21*(7), 1554–1564. https://doi.org/10.1080/02699930601117096

Sabbagh, M. A., Bowman, L. C., Evraire, L. E., & Ito, J. M. B. (2009). Neurodevelopmental Correlates of Theory of Mind in Preschool Children. *Child Development*, *80*(4), 1147–1162. https://doi.org/10.1111/j.1467-8624.2009.01322.x

Sabbagh, M. A., Moses, L. J., & Shiverick, S. (2006). Executive Functioning and Preschoolers' Understanding of False Beliefs, False Photographs, and False Signs. *Child Development*, *77*(4), 1034–1049. https://doi.org/10.1111/j.1467-8624.2006.00917.x

Safiye, T., Gutić, M., Dubljanin, J., Stojanović, T. M., Dubljanin, D., Kovačević, A., Zlatanović, M., Demirović, D. H., Nenezić, N., & Milidrag, A. (2023). Mentalizing, Resilience, and Mental Health Status among Healthcare Workers during the COVID-19 Pandemic: A Cross-Sectional Study. *International Journal of Environmental Research and Public Health*, *20*(8), 5594. https://doi.org/10.3390/ijerph20085594

Salzer, S., Pincus, A., Hoyer, J., Kreische, R., Leichsenring, F., & Leibing, E. (2008). Interpersonal Subtypes Within Generalized Anxiety Disorder. *Journal of Personality Assessment*, *90*, 292–299. https://doi.org/10.1080/00223890701885076

Samamé, C., Martino, D. J., & Strejilevich, S. A. (2012). Social cognition in euthymic bipolar disorder: Systematic review and meta-analytic approach. *Acta Psychiatrica Scandinavica*, *125*(4), 266–280. https://doi.org/10.1111/j.1600-0447.2011.01808.x

Samamé, C., Martino, D. J., & Strejilevich, S. A. (2015). An individual task meta-analysis of social cognition in euthymic bipolar disorders. *Journal of Affective Disorders*, *173*, 146–153. https://doi.org/10.1016/j.jad.2014.10.055

Santarelli, G., Innocenti, M., Faggi, V., Miglietta, V., Colpizzi, I., Galassi, F., Castellini, G., & Ricca, V. (2022). Preliminary evidence for a Theory of Mind impairment in patients with Anxiety Disorders. *European Psychiatry*, *65*(1), 185–186. https://doi.org/10.1192/j.eurpsy.2022.489

Santos, J. M., Pousa, E., Soto, E., Comes, A., Roura, P., Arrufat, F. X., & Obiols, J. E. (2017). Theory of Mind in Euthymic Bipolar Patients and First-Degree Relatives. *The Journal of Nervous and Mental Disease*, *205*(3), 207–212. https://doi.org/10.1097/NMD.0000000000000595

Sanvicente-Vieira, B., Romani-Sponchiado, A., Kluwe-Schiavon, B., Brietzke, E., Araujo, R. B., & Grassi-Oliveira, R. (2017). Theory of Mind in Substance Users: A Systematic Minireview. *Substance Use & Misuse*, *52*(1), 127–133. https://doi.org/10.1080/10826084.2016.1212890

Savla, G. N., Vella, L., Armstrong, C. C., Penn, D. L., & Twamley, E. W. (2013). Deficits in domains of social cognition in schizophrenia: A meta-analysis of the empirical evidence. *Schizophrenia Bulletin*, *39*(5), 979–992. https://doi.org/10.1093/schbul/sbs080

Saxe, R., & Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia*, *43*(10), 1391–1399. https://doi.org/10.1016/j.neuropsychologia.2005.02.013

Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in Cognitive Sciences*, *19*(2), 65–72. https://doi.org/10.1016/j.tics.2014.11.007

Schilling, L., Moritz, S., Kriston, L., Krieger, M., & Nagel, M. (2018). Efficacy of metacognitive training for patients with borderline personality disorder: Preliminary results. *Psychiatry Research*, *262*, 459–464. https://doi.org/10.1016/j.psychres.2017.09.024

Schmahl, C. G., Vermetten, E., Elzinga, B. M., & Douglas Bremner, J. (2003). Magnetic resonance imaging of hippocampal and amygdala volume in women with childhood abuse and borderline personality disorder. *Psychiatry Research:*

Neuroimaging, *122*(3), 193–198. https://doi.org/10.1016/S0925-4927(03)00023-4

Schmid, F., Henry, A., Benzerouk, F., Barrière, S., Portefaix, C., Gondrexon, J., Obert, A., Kaladjian, A., & Gierski, F. (2023). Neural activations during cognitive and affective theory of mind processing in healthy adults with a family history of alcohol use disorder. *Psychological Medicine*, 1–11. https://doi.org/10.1017/S0033291723002854

Schmid, F., Moreau, F., Benzerouk, F., Raucher-Chéné, D., Kaladjian, A., Gierski, F., & Henry, A. (2022). Faux Pas Recognition and Executive Processes in Patients with Alcohol Use Disorder: Toward an Investigation of Interindividual Heterogeneity. *Archives of Clinical Neuropsychology*, *37*(3), 608–620. https://doi.org/10.1093/arclin/acab072

Scholl, B. J., & Leslie, A. M. (1999). Modularity, Development and 'Theory of Mind.' *Mind & Language*, *14*(1), 131–153. https://doi.org/10.1111/1468-0017.00106

Scholl, B. J., & Leslie, A. M. (2001). Minds, Modules, and Meta-Analysis. *Child Development*, *72*(3), 696–701. https://doi.org/10.1111/1467-8624.00308

Schulte, M. H. J., Cousijn, J., den Uyl, T. E., Goudriaan, A. E., van den Brink, W., Veltman, D. J., Schilt, T., & Wiers, R. W. (2014). Recovery of neurocognitive functions following sustained abstinence after substance dependence and implications for treatment. *Clinical Psychology Review*, *34*(7), 531–550. https://doi.org/10.1016/j.cpr.2014.08.002

Schurz, M., Radua, J., Tholen, M. G., Maliske, L., Margulies, D. S., Mars, R. B., Sallet, J., & Kanske, P. (2021). Toward a hierarchical model of social cognition: A neuroimaging meta-analysis and integrative review of empathy and theory of mind. *Psychological Bulletin*, *147*, 293–327. https://doi.org/10.1037/bul0000303

Shamay-Tsoory, S. G., Harari, H., Szepsenwol, O., & Levkovitz, Y. (2009). Neuropsychological Evidence of Impaired Cognitive Empathy in Euthymic Bipolar Disorder. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *21*(1), 59–67. https://doi.org/10.1176/jnp.2009.21.1.59

Shamay-Tsoory, S. G. (2011). The Neural Bases for Empathy. *The Neuroscientist*, *17*(1), 18–24. https://doi.org/10.1177/1073858410379268

Shamay-Tsoory, S. G., Harari, H., Aharon-Peretz, J., & Levkovitz, Y. (2010). The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies. *Cortex*, *46*(5), 668–677. https://doi.org/10.1016/j.cortex.2009.04.008

Shamay-Tsoory, S. G., Shur, S., Barcai-Goodman, L., Medlovich, S., Harari, H., & Levkovitz, Y. (2007). Dissociation of cognitive from affective components of theory of mind in schizophrenia. *Psychiatry Research*, *149*(1–3), 11–23. https://doi.org/10.1016/j.psychres.2005.10.018

Sharp, C., & Vanwoerden, S. (2015). Hypermentalizing in Borderline Personality Disorder: A Model and Data. *Journal of Infant, Child, and Adolescent Psychotherapy*, *14*, 33–45. https://doi.org/10.1080/15289168.2015.1004890

Simonsen, C. B., Jakobsen, A. G., Grøntved, S., & Kjaersdam Telléus, G. (2020). The mentalization profile in patients with eating disorders: A systematic review and meta-analysis. *Nordic Journal of Psychiatry*, *74*(5), 311–322. https://doi.org/10.1080/08039488.2019.1707869

Simpson, A. J., & Todd, A. R. (2017). Intergroup visual perspective-taking: Shared group membership impairs self-perspective inhibition but may facilitate perspective calculation. *Cognition*, *166*, 371–381. https://doi.org/10.1016/j.cognition.2017.06.003

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research. *Neuroscience & Biobehavioral Reviews*, *30*(6), 855–863. https://doi.org/10.1016/j.neubiorev.2006.06.011

Skodol, A. E., & Bender, D. S. (2003). Why are women diagnosed borderline more than men? *The Psychiatric Quarterly*, *74*(4), 349–360. https://doi.org/10.1023/a:1026087410516

Slaughter, V., Imuta, K., Peterson, C. C., & Henry, J. D. (2015). Meta-Analysis of Theory of Mind and Peer Popularity in the Preschool and Early School Years. *Child Development*, *86*(4), 1159–1174. https://doi.org/10.1111/cdev.12372

Sloover, M., van Est, L. A. C., Janssen, P. G. J., Hilbink, M., & van Ee, E. (2022). A meta-analysis of mentalizing in anxiety disorders, obsessive-compulsive and related disorders, and trauma and stressor related disorders. *Journal of Anxiety Disorders*, *92*, 102641. https://doi.org/10.1016/j.janxdis.2022.102641

Sodian, B. (2005). Theory of mind—The case for conceptual development. In W. Schneider, R. Schumann-Hengsteler, & B. Sodian (Eds.), *Young children's cognitive development: Interrelationships among executive functioning, working memory, verbal ability, and theory of mind* (pp. 95–103). Lawrence Erlbaum Associates Publishers.

Sodian, B., Perst, H., & Meinhardt, J. (2012). Entwicklung der Theory of Mind in der Kindheit. In H. Förstl (Ed.), *Theory of Mind: Neurobiologie und Psychologie sozialen Verhaltens* (pp. 61–77). Springer. https://doi.org/10.1007/978-3-642-24916-7_6

Soloff, P. H., Meltzer, C. C., Becker, C., Greer, P. J., Kelly, T. M., & Constantine, D. (2003). Impulsivity and prefrontal hypometabolism in borderline personality

disorder. *Psychiatry Research: Neuroimaging*, *123*(3), 153–163. https://doi.org/10.1016/S0925-4927(03)00064-7

Soloff, P. H., Meltzer, C. C., Greer, P. J., Constantine, D., & Kelly, T. M. (2000). A fenfluramine-activated FDG-PET study of borderline personality disorder. *Biological Psychiatry*, *47*(6), 540–547. https://doi.org/10.1016/S0006-3223(99)00202-4

Sondermann, S., Stahl, J., Grave, U., Outzen, J., Moritz, S., & Klein, J. P. (2020). Preoperational Thinking as a Measure of Social Cognition Is Associated With Long-Term Course of Depressive Symptoms. A Longitudinal Study Involving Patients With Depression and Healthy Controls. *Frontiers in Psychiatry, 11*, 652. https://doi.org/10.3389/fpsyt.2020.00652

Sprong, M., Schothorst, P., Vos, E., Hox, J., & Engeland, H. V. (2007). Theory of mind in schizophrenia: Meta-analysis. *The British Journal of Psychiatry*, *191*(1), 5–13. https://doi.org/10.1192/bjp.bp.107.035899

Steel, Z., Marnane, C., Iranpour, C., Chey, T., Jackson, J. W., Patel, V., & Silove, D. (2014). The global prevalence of common mental disorders: A systematic review and meta-analysis 1980-2013. *International Journal of Epidemiology*, *43*(2), 476–493. https://doi.org/10.1093/ije/dyu038

Stoffers, J. M., & Lieb, K. (2014). Pharmacotherapy for Borderline Personality Disorder—Current Evidence and Recent Trends. *Current Psychiatry Reports*, *17*(1), 534. https://doi.org/10.1007/s11920-014-0534-0

Stoffers-Winterling, J. M., Storebø, O. J., Kongerslev, M. T., Faltinsen, E., Todorovac, A., Jørgensen, M. S., Sales, C. P., Callesen, H. E., Ribeiro, J. P., Völlm, B. A., Lieb, K., & Simonsen, E. (2022). Psychotherapies for borderline personality disorder: A focused systematic review and meta-analysis. *The British Journal of*

Psychiatry, *221*(3), 538–552. https://doi.org/10.1192/bjp.2021.204

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal Lobe Contributions to Theory of Mind. *Journal of Cognitive Neuroscience*, *10*(5), 640–656. https://doi.org/10.1162/089892998562942

Stone, V. E., & Gerrans, P. (2006). What's domain-specific about theory of mind? *Social Neuroscience*, *1*(3–4), 309–319. https://doi.org/10.1080/17470910601029221

Stuss, D. T., Gallup, G. G., Jr, & Alexander, M. P. (2001). The frontal lobes are necessary for `theory of mind'. *Brain*, *124*(2), 279–286. https://doi.org/10.1093/brain/124.2.279

Subic-Wrana, C., Beutel, M. E., Knebel, A., & Lane, R. D. (2010). Theory of Mind and Emotional Awareness Deficits in Patients With Somatoform Disorders. *Psychosomatic Medicine*, *72*(4), 404–411. https://doi.org/10.1097/PSY.0b013e3181d35e83

Surtees, A. D. R., & Apperly, I. A. (2012). Egocentrism and Automatic Perspective Taking in Children and Adults. *Child Development*, *83*(2), 452–460. https://doi.org/10.1111/j.1467-8624.2011.01730.x

Szmulewicz, A. G., Lomastro, M. J., Valerio, M. P., Igoa, A., & Martino, D. J. (2019). Social cognition in first episode bipolar disorder patients. *Psychiatry Research*, *272*, 551–554. https://doi.org/10.1016/j.psychres.2019.01.002

Tapajóz P. de Sampaio, F., Soneira, S., Aulicino, A., Martese, G., Iturry, M., & Allegri, R. F. (2013). Theory of mind and central coherence in eating disorders: Two sides of the same coin? *Psychiatry Research*, *210*(3), 1116–1122. https://doi.org/10.1016/j.psychres.2013.08.051

Tawa, E. A., Hall, S. D., & Lohoff, F. W. (2016). Overview of the Genetics of Alcohol

Use Disorder. *Alcohol and Alcoholism*, *51*(5), 507–514. https://doi.org/10.1093/alcalc/agw046

Thamby, A., Desai, G., Mehta, U. M., & Chaturvedi, S. K. (2019). Deficits in Theory of Mind and Emotional Awareness in Somatoform Disorders. *Indian Journal of Psychological Medicine*, *41*(4), 368–374. https://doi.org/10.4103/IJPSYM.IJPSYM_382_18

Thirion-Marissiaux, A.-F., & Nader-Grosbois, N. (2008). Theory of Mind "emotion", developmental characteristics and social understanding in children and adolescents with intellectual disabilities. *Research in Developmental Disabilities*, *29*(5), 414–430. https://doi.org/10.1016/j.ridd.2007.07.001

Tonelli, H., & de Siqueira Rotenberg, L. (2021). Emotion perception and theory of mind in obesity: A systematic review on the impact of social cognitive deficits on dysfunctional eating behaviors. *Surgery for Obesity and Related Diseases*, *17*(3), 618–629. https://doi.org/10.1016/j.soard.2020.10.016

Uljarević, M., Frazier, T. W., Phillips, J. M., Jo, B., Littlefield, S., & Hardan, A. Y. (2020). Mapping the Research Domain Criteria Social Processes Constructs to the Social Responsiveness Scale. *Journal of the American Academy of Child & Adolescent Psychiatry*, *59*(11), 1252–1263.e3. https://doi.org/10.1016/j.jaac.2019.07.938

Van den Stock, J., Cerami, C., & Dodich, A. (2023). Transdiagnostic overlap in brain correlates of affective and cognitive theory of mind deficits, *Brain*, e39–e42. https://doi.org/10.1093/brain/awad023

Vaskinn, A., Hauger, L. E., & Bjørnebekk, A. (2020). Theory of mind in users of anabolic androgenic steroids. *Psychopharmacology*, *237*(10), 3191–3199. https://doi.org/10.1007/s00213-020-05603-y

Vegni, N., D'Ardia, C., & Torregiani, G. (2021). Empathy, Mentalization, and Theory of Mind in Borderline Personality Disorder: Possible Overlap With Autism Spectrum Disorders. *Frontiers in Psychology*, *12*, 626353. https://doi.org/10.3389/fpsyg.2021.626353

Vellante, M., Baron-Cohen, S., Melis, M., Marrone, M., Petretto, D. R., Masala, C., & Preti, A. (2013). The "Reading the Mind in the Eyes" test: Systematic review of psychometric properties and a validation study in Italy. *Cognitive Neuropsychiatry*, *18*(4), 326–354. https://doi.org/10.1080/13546805.2012.721728

Verdejo-Garcia, A. (Ed.). (2020). *Cognition and Addiction. A Researcher's Guide from Mechanisms Towards Interventions.* Academic Press. https://doi.org/10.1016/C2017-0-03135-X

Vermote, R., Lowyck, B., Luyten, P., Vertommen, H., Corveleyn, J., Verhaest, Y., Stroobants, R., Vandeneede, B., Vansteelandt, K., & Peuskens, J. (2010). Process and Outcome in Psychodynamic Hospitalization-Based Treatment for Patients With a Personality Disorder. *The Journal of Nervous and Mental Disease*, *198*(2), 110–115. https://doi.org/10.1097/NMD.0b013e3181cc0d59

Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happé, F., Falkai, P., Maier, W., Shah, N. J., Fink, G. R., & Zilles, K. (2001). Mind Reading: Neural Mechanisms of Theory of Mind and Self-Perspective. *NeuroImage*, *14*(1), 170–181. https://doi.org/10.1006/nimg.2001.0789

Vogt, K. S., & Norman, P. (2019). Is mentalization-based therapy effective in treating the symptoms of borderline personality disorder? A systematic review. *Psychology and Psychotherapy: Theory, Research and Practice*, *92*(4), 441–464. https://doi.org/10.1111/papt.12194

Volkert, J., Hauschild, S., & Taubner, S. (2019). Mentalization-Based Treatment for Personality Disorders: Efficacy, Effectiveness, and New Developments. *Current Psychiatry Reports*, *21*(4), 25. https://doi.org/10.1007/s11920-019-1012-5

Wang, W.-H., Shih, Y.-H., Yu, H.-Y., Yen, D.-J., Lin, Y.-Y., Kwan, S.-Y., Chen, C., & Hua, M.-S. (2015). Theory of mind and social functioning in patients with temporal lobe epilepsy. *Epilepsia*, *56*(7), 1117–1123. https://doi.org/10.1111/epi.13023

Wang, Z., & Su, Y. (2013). Age-related differences in the performance of theory of mind in older adults: A dissociation of cognitive and affective components. *Psychology and Aging*, *28*(1), 284–291. https://doi.org/10.1037/a0030876

Wang, Y., & Su, Y. (2006). Theory of Mind in old Adults: The performance on Happe's stories and Faux Pas stories. *Psychologia, 49*(4), 228–237. https://doi.org/10.2117/psysoc.2006.228

Washburn, D., Wilson, G., Roes, M., Rnic, K., & Harkness, K. L. (2016). Theory of mind in social anxiety disorder, depression, and comorbid conditions. *Journal of Anxiety Disorders*, *37*, 71–77. https://doi.org/10.1016/j.janxdis.2015.11.004

Watson, M., Filia, K., Stevens, A., Cotton, S., Nelson, B., & Ratheesh, A. (2023). A systematic review and meta-analysis of global and social functioning among people at risk of bipolar disorder. *Journal of Affective Disorders*, *321*, 290–303. https://doi.org/10.1016/j.jad.2022.10.019

Wellman, H. M. (2020). *Reading Minds: How Childhood Teaches Us to Understand People*. Oxford University Press. https://doi.org/10.1093/oso/9780190878672.001.0001

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-Analysis of Theory-of-Mind Development: The Truth about False Belief. *Child Development*, *72*(3), 655–

684. https://doi.org/10.1111/1467-8624.00304

Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-Mind Tasks. *Child Development*, *75*(2), 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

Whittle, S., Yücel, M., Yap, M. B. H., & Allen, N. B. (2011). Sex differences in the neural correlates of emotion: Evidence from neuroimaging. *Biological Psychology*, *87*(3), 319–333. https://doi.org/10.1016/j.biopsycho.2011.05.003

Wiers, R. W., Gladwin, T. E., Hofmann, W., Salemink, E., & Ridderinkhof, K. R. (2013). Cognitive Bias Modification and Cognitive Control Training in Addiction and Related Psychopathology: Mechanisms, Clinical Perspectives, and Ways Forward. *Clinical Psychological Science*, *1*(2), 192–212. https://doi.org/10.1177/2167702612466547

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128. https://doi.org/10.1016/0010-0277(83)90004-5

Wittchen, H. U., Jacobi, F., Rehm, J., Gustavsson, A., Svensson, M., Jönsson, B., Olesen, J., Allgulander, C., Alonso, J., Faravelli, C., Fratiglioni, L., Jennum, P., Lieb, R., Maercker, A., van Os, J., Preisig, M., Salvador-Carulla, L., Simon, R., & Steinhausen, H.-C. (2011). The size and burden of mental disorders and other disorders of the brain in Europe 2010. *European Neuropsychopharmacology*, *21*(9), 655–679. https://doi.org/10.1016/j.euroneuro.2011.07.018

Wolkenstein, L., Schönenberg, M., Schirm, E., & Hautzinger, M. (2011). I can see what you feel, but I can't deal with it: Impaired theory of mind in depression. *Journal of Affective Disorders*, *132*(1), 104–111. https://doi.org/10.1016/j.jad.2011.02.010

World-Health-Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines* (Vol. 1). World Health Organization.

Xia, H., Wu, N., & Su, Y. (2012). Investigating the Genetic Basis of Theory of Mind (ToM): The Role of Catechol-O-Methyltransferase (COMT) Gene Polymorphisms. *PLOS ONE*, *7*(11), e49768. https://doi.org/10.1371/journal.pone.0049768

Yamada, K., Inoue, Y., & Kanba, S. (2015). Theory of mind ability predicts prognosis of outpatients with major depressive disorder. *Psychiatry Research*, *230*(2), 604–608. https://doi.org/10.1016/j.psychres.2015.10.011

Zabihzadeh, A., Maleki, G., Hatami, A., Richman, M., Alimardani, Z., & Heidari, M. (2017). Affective and cognitive theory of mind in borderline personality disorder: The role of comorbid depression. *Psychiatry Research*, *257*, 144–149. https://doi.org/10.1016/j.psychres.2017.07.034

Zainal, N. H., & Newman, M. G. (2018). Worry amplifies theory-of-mind reasoning for negatively valenced social stimuli in generalized anxiety disorder. *Journal of Affective Disorders*, *227*, 824–833. https://doi.org/10.1016/j.jad.2017.11.084

Zelazo, P., Jacques, S., Burack, J. A., & Frye, D. (2002). The relation between theory of mind and rule use: Evidence from persons with autism-spectrum disorders. *Infant and Child Development*, *11*(2), 171–195. https://doi.org/10.1002/icd.304

Zhang, X., Lecce, S., Ceccato, I., Cavallini, E., Zhang, L., & Chen, T. (2018). Plasticity in older adults' theory of mind performance: The impact of motivation. *Aging & Mental Health*, *22*(12), 1592–1599. https://doi.org/10.1080/13607863.2017.1376313

Zunhammer, M., Halski, A., Eichhammer, P., & Busch, V. (2015). Theory of Mind and

Emotional Awareness in Chronic Somatoform Pain Patients. *Plos One*, *10*(10), e0140016. https://doi.org/10.1371/journal.pone.0140016

Appendix

Movie for the Assessment of Social Cognition (MASC)

Beschreibung

Naturalistisches, Video-basiertes Testverfahren zur Erfassung sozial-kognitiver Denkfunktionen (vgl. Theory of Mind / Mentalizierung). Im MASC sehen sich die Patienten ein 15 Minuten langes Video über vier Charaktere / Protagonisten, die sich zu einem gemeinsamen Abendessen treffen, an. Das Video wird an 46 Zeitpunkten pausiert und es werden entsprechende Fragen bezüglich der Gefühle, Gedanken und Absichten der verschiedenen Charaktere gestellt (Dziobek et al., 2006).

Anleitung und Fragen

Sie sehen jetzt gleich einen ca. 15-minütigen Film. Bitte schauen Sie aufmerksam zu und versuchen Sie, sich in die handelnden Personen hinein zu versetzen.

Die Personen werden Ihnen jetzt kurz vorgestellt.
[Vorstellung d. vier Personen mit Bild: Sandra, Michael, Brigitte, Klaus]

Die vier Personen werden im Laufe des Filmes zusammentreffen und einen Abend miteinander verbringen.

Der Film wird an einigen Stellen angehalten. Sehen Sie bitte aufmerksam zu und beantworten Sie dann die eingeblendeten Fragen durch Wählen einer der vier Antwortmöglichkeiten. Bitte entscheiden Sie sich bei jeder Frage für eine der vier Möglichkeiten, auch wenn Sie der Meinung sind, dass die Antwort nur „in etwa" zutrifft.

Es geht bei den meisten Fragen um die „geistigen Zustände" der Personen, also Emotionen, Gedanken und Absichten. Wenn nach den geistigen Zuständen gefragt wird, dann beziehen sich die Fragen immer genau auf den Moment, an dem der Film stoppt.

Die Szenen können Sie sich nur einmal ansehen. Sind Sie bereit? Dann klicken Sie

auf PLAY.

Beispielhafte Beschreibung von zwei Videoszenen

Beispiel-Szene 1:

Es läutet an Sandra's Haustür. Michael stattet Sandra eine Überraschungsbesuch ab. Michael begrüßt Sandra überschwänglich und macht ihr Komplimente und streift ihr dabei mit seiner Hand durch die Haare. Er sagt, dass sie toll aussieht und ob sie wohl einen neuen Haarschnitt hat.

Testfrage: „Was fühlt Sandra?"

Antwortmöglichkeiten:

1. Ihre Haare sehen gar nicht so gut aus
2. Sie freut sich über das Kompliment
3. Sie ist wütend, dass Michael sie so bedrängt
4. Sie fühlt sich geschmeichelt aber etwas überrumpelt

Beispiel-Szene 2

Sandra und Michael unterhalten sich über einen früheren gemeinsam verbrachten Abend. Michael sagt, dass es ein toller Abend war und dass er damals ganz neue Facetten von Sandra kennengelernt hat. Er streicht ihr dabei mit seinen Händen über ihre Schultern und Arme. Sandra sagt, dass sie auch Spaß hatte und sie sich besonders gefreut hat Michael's Freund Klaus kenngelernt zu haben. Sandra fragt ob er (Michael) und sein Freund (Klaus) nicht beide einmal zu ihr zum Essen kommen wollen. Michael macht eine abschüttelnde Bewegung und sagt, dass Sandra doch lieber zuerst einmal mit ihm Squash spielen gehen soll und dass sie danach Essen gehen können.

Testfrage: Warum sagt Michael das?

Antwortmöglichkeiten:

1. Er will Sandra mit seinem guten Spiel beeindrucken
2. Er will sich mit Sandra alleine treffen
3. Er ist ein guter Squash-Spieler
4. Er geht lieber Squash spielen als zum Essen

Auswertung / Skalierung

Zur Auswertung wird die Gesamtpunktezahl aller richtigen Theory of Mind Antworten aufsummiert. Die anderen Antworten werden einer von 3 „Fehlerarten" zugeordnet und entsprechende Skalen berechnet.

1. Gesamtpunktezahl Richtige ToM (c): z.B. Szene / Frage 1 = Antwort D
2. Übermäßige ToM (+): z.B. Szene / Frage 1 = Antwort C
3. Geringere ToM (-): z.B. Szene / Frage 1 = Antwort B
4. Keine ToM (o): z.B. Szene / Frage 1 = Antwort A

Außerdem gibt es Kontrollfragen die zu einer Kontrollskala aufsummiert werden. Fragen: 13 (c), 33 (c), 38 (a), 46 (b), 47 (a), 48 (a)

In nachfolgender Tabelle finden sich pro Item und jeweiliger Antwort (a bis d) die je zuzuordnenden Skalen.

Item	**Gesamtpunktzahl richtig ToM (c)**	**übermäßiges ToM (+)**	**geringeres ToM (-)**	**kein ToM (o)**
1.	D	C	B	A
2.	B	A	D	C
3.	A	B	D	C
4.	A	D	C	B
5.	C	B	A	D
6.	A	C	D	B
7.	C	A	B	D
8.	B	D	A	C
9.	B	C	D	A
10.	B	C	A	D
11.	C	D	B	A
12.	C	A	D	B
13.	C	D	A	B
14.	D	A	B	C
15.	D	B	C	A
16.	C	B	D	A

17.	B	A	C	D
18.	B	D	C	A
19.	A	D	B	C
20.	D	C	A	B
21.	D	B	A	C
22.	A	B	C	D
23.	D	A	C	B
24.	C	D	B	A
25.	B	C	A	D
26.	B	A	D	C
27.	B	A	C	D
28.	B	D	A	C
29.	D	B	A	C
30.	C	A	D	B
31.	D	A	B	C
32.	A	B	C	D
33.	A	D	B	C
34.	A	D	C	B
35.	C	B	D	A
36.	D	B	C	A
37.	A	C	B	D
38.	D	C	B	A
39.	C	D	A	B
40.	C	B	A	D
41.	A	C	D	B
42.	C	A	B	D
43.	D	A	C	B
44.	B	D	C	A
45.	A	B	D	C