

**Universität
Basel**

Fakultät für
Psychologie



Development of novel diagnostic tools for behavioural variant frontotemporal dementia

Inauguraldissertation zur Erlangung der Würde einer Doktorin der Philosophie vorgelegt der Fakultät für Psychologie der Universität Basel von

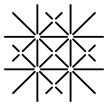
Anna Semenkova

aus Mariupol, Ukraine

Basel, 2024

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel

edoc.unibas.ch



Universität
Basel

Fakultät für
Psychologie



Genehmigt von der Fakultät für Psychologie auf Antrag von

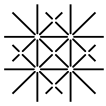
Prof. Dr. Marc Sollberger

Prof. Dr. Rui Mata

Datum des Doktoratsexamen: 04.09.2024

Prof. Dr. Rainer Greifeneder

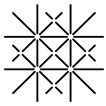
Dekan der Fakultät für Psychologie



Erklärung zur wissenschaftlichen Lauterkeit

Ich erkläre hiermit, dass die vorliegende Arbeit ohne die Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel selbstständig verfasst habe. Zu Hilfe genommene Quellen sind als solche gekennzeichnet. Die veröffentlichten oder zur Veröffentlichung in Zeitschriften eingereichten Manuskripte wurden in Zusammenarbeit mit den Koautoren erstellt und von keinem der Beteiligten an anderer Stelle publiziert, zur Publikation eingereicht, oder einer anderen Prüfungsbehörde als Qualifikationsarbeit vorgelegt. Es handelt sich dabei um folgende Manuskripte:

1. **Semenkova, A.**, Piguet, O., Johnen, A., Schroeter, M. L., Godulla, J., Linnemann, C., Mühlhauser, M., Sauer, T., Baumgartner, M., Anderl-Straub, S., Otto, M., Felbecker, A., Kressig, R. W., Berres, M., & Sollberger, M. (2023). The Behavioural Dysfunction Questionnaire discriminates behavioural variant frontotemporal dementia from Alzheimer's disease dementia and major depressive disorder. *Journal of Neurology*. <https://doi.org/10.1007/s00415-023-11666-6> **Published**
2. Jarsch, M., Piguet, O., Berres, M., Sluka, C., **Semenkova, A.**, Kressig, R.W., Monsch, A., McDonald, S., & Sollberger, M. (2023). Development of the Basel Version of the Awareness of Social Inference Test – Theory of Mind (BASIT-ToM) in Healthy Adults. *Journal of Neuropsychology*, 17, 125–145. <http://doi.org/10.1111/jnp.12290> **Published**
3. Jarsch, M., **Semenkova, A.**, Monsch, A, Kressig, R.W., & Sollberger, M. (2022). Eine Lücke, die es zu schließen gilt: Die Untersuchung sozial-kognitiver Fähigkeiten an deutschsprachigen Memory Kliniken. *Zeitschrift für Neuropsychologie*, 33, 129–137. <https://doi.org/10.1024/1016-264X/a000358> **Published**



Spezifizierung des eigenen Forschungsbeitrags zu den Manuskripten:

1. Eigener Beitrag nach CRediT¹:

- | | | |
|--|---|---|
| <input type="checkbox"/> Conceptualization | <input checked="" type="checkbox"/> Data curation | <input checked="" type="checkbox"/> Formal Analysis |
| <input type="checkbox"/> Funding acquisition | <input type="checkbox"/> Investigation | <input checked="" type="checkbox"/> Methodology |
| <input checked="" type="checkbox"/> Project administration | <input type="checkbox"/> Resources | <input type="checkbox"/> Software |
| <input type="checkbox"/> Supervision | <input type="checkbox"/> Validation | <input checked="" type="checkbox"/> Visualization |
| <input checked="" type="checkbox"/> Writing – original draft | | |
| <input checked="" type="checkbox"/> Writing – review & editing | | |

Das Manuskript wurde bisher für keine anderen Qualifikationsarbeiten eingereicht.

2. Eigener Beitrag nach CRediT¹:

- | | | |
|--|--|---|
| <input type="checkbox"/> Conceptualization | <input type="checkbox"/> Data curation | <input type="checkbox"/> Formal Analysis |
| <input type="checkbox"/> Funding acquisition | <input type="checkbox"/> Investigation | <input type="checkbox"/> Methodology |
| <input type="checkbox"/> Project administration | <input type="checkbox"/> Resources | <input type="checkbox"/> Software |
| <input type="checkbox"/> Supervision | <input type="checkbox"/> Validation | <input checked="" type="checkbox"/> Visualization |
| <input type="checkbox"/> Writing – original draft | | |
| <input checked="" type="checkbox"/> Writing – review & editing | | |

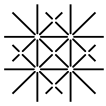
Das Manuskript wurde bereits für andere Qualifikationsarbeiten bei Jarsch M. eingereicht.

3. Eigener Beitrag nach CRediT¹:

- | | | |
|--|--|---|
| <input type="checkbox"/> Conceptualization | <input type="checkbox"/> Data curation | <input type="checkbox"/> Formal Analysis |
| <input type="checkbox"/> Funding acquisition | <input type="checkbox"/> Investigation | <input type="checkbox"/> Methodology |
| <input type="checkbox"/> Project administration | <input type="checkbox"/> Resources | <input type="checkbox"/> Software |
| <input type="checkbox"/> Supervision | <input type="checkbox"/> Validation | <input checked="" type="checkbox"/> Visualization |
| <input type="checkbox"/> Writing – original draft | | |
| <input checked="" type="checkbox"/> Writing – review & editing | | |

Das Manuskript wurde bereits für andere Qualifikationsarbeiten bei Jarsch M. eingereicht.

¹ <https://casrai.org/credit/>



Open-Science Aspekte der Manuskripte:

1. Preregistration: ja nein
Open-Access-Publikation: ja nein
Open-Access-Data/Analyse: ja nein
Ort/URL der Daten und Analysen: N/A

2. Preregistration: ja nein
Open-Access-Publikation: ja nein
Open-Access-Data/Analyse: ja nein
Ort/URL der Daten und Analysen: N/A

3. Preregistration: ja nein
Open-Access-Publikation: ja nein
Open-Access-Data/Analyse: ja nein
Ort/URL der Daten und Analysen: N/A

Ort, Datum

Basel, 25.06.2024

Signatur

A. Semenkova

Vorname Nachname

Anna Semenkova

Acknowledgements

I would like to sincerely thank my first supervisor Prof. Dr. Marc Sollberger who trusted from the beginning in my skills to complete this great project. Thanks to his comprehension and guidance, his encouragement, support, and strong involvement in the projects, I learned a lot during these three years and was happy to be a part of his team.

I also want to thank my second supervisor Prof. Dr. Rui Mata for his valuable inputs, the extensive expertise he shared and the guidance throughout my PhD.

And of course, I want to thank our research team and the team of the Memory Clinic Basel who made the collection of the data for these projects possible and created the best team environment in which this project could be carried out.

I am also sincerely grateful to all my friends who encouraged me in difficult times and always kept me motivated to become a better me. In Switzerland and Germany, far away from my home, they became my second family and have always been there for me.

And last but not least, I want to thank my family. All generations of my family worked very hard to enable each next generation to receive better education and to reach more. I am happy to become the first person in my family to complete a PhD program, and this at the oldest university in Switzerland. This became possible only thanks to the hard work and sacrifices of all previous generations of my family. I am incredibly grateful to them for trusting me with this honour and continuously supporting me on this journey.

Devoted to all non-academic generations behind all academics whose hard work and sacrifices made these graduations possible

Contents

Glossary of abbreviations.....	2
Abstract	3
Introduction.....	4
Theoretical background.....	5
Behavioural variant frontotemporal dementia: Clinical profile and diagnostic challenges	5
Behavioural disorders in behavioural variant frontotemporal dementia.....	7
Social cognition as a new diagnostic domain for behavioural variant frontotemporal dementia	8
Social cognition: Definition, structure and neuronal correlates.....	9
Theory of mind and its role in diagnosis of behavioural variant frontotemporal dementia	10
Neuropsychological assessment in diagnosis of behavioural variant frontotemporal dementia	11
Assessment of behavioural symptoms in patients with brain disorders.....	12
Assessment of social cognition and theory of mind in clinical routine	13
Summary	14
STUDY I: The Behavioural Dysfunction Questionnaire discriminates behavioural variant frontotemporal dementia from Alzheimer’s disease dementia and major depressive disorder	15
STUDY II: Eine Lücke, die es zu schließen gilt: Die Untersuchung sozial-kognitiver Fähigkeiten an deutschsprachigen Memory Kliniken.....	25
STUDY III: Development of the Basel Version of the Awareness of Social Inference Test – Theory of Mind (BASIT-ToM) in Healthy Adults	35
Discussion	57
Behavioural Dysfunction Questionnaire (BDQ): The first behavioural scale operationalising the current diagnostic criteria for behavioural variant frontotemporal dementia	57
Insights on the current diagnostic criteria based on the findings in the Study I	58
Clinical implication.....	58
Further research.....	59

Assessment of social cognition at the German-speaking memory clinics	60
Basel Version of the Awareness of Social Inference Test – Theory of Mind (BASIT-ToM): A reality-close assessment tool of social cognition subdomain Theory of Mind	61
Development phase	61
Validation phase.....	63
Contribution of the developed tools to diagnosis of the behavioural variant frontotemporal dementia and their critical evaluation	63
Outlook in further research and practical implications	65
Conclusion.....	66
References	67
Appendix A. Diagnostic criteria for frontotemporal dementia (Englund et al., 1994)	82
Appendix B. Diagnostic criteria for frontotemporal dementia (Neary et al., 1998)	84
Appendix C. Diagnostic criteria for behavioural variant frontotemporal dementia (Rascovsky et al., 2011).....	86
Appendix D. Study I: Supplementary materials.....	88
Appendix E. Study II: Supplementary materials.....	112
Appendix F. Study III: Supplementary materials.....	115

Glossary of abbreviations

ADD	Alzheimer disease dementia
AUC	Area under the curve
BASIT-ToM	Basel Version of The Awareness of Social Inference Test – Theory of Mind
BDQ	Behavioural Dysfunctional Questionnaire
bvAD	Behavioural variant Alzheimer disease
bvFTD	Behavioural variant frontotemporal dementia
CDR	Clinical Dementia Rating scale
DSM-5	Diagnostic and Statistical Manual, 5th edition
FBI	Frontal Behavioural Inventory
FRS	Frontotemporal Dementia Rating Scale
FTD	Frontotemporal dementia
FTDC	Frontotemporal dementia Criteria Consortium
FTLD	Frontotemporal lobar degeneration
ICD-11	International Classification of Diseases, 11th edition
MDD	Major depressive disorder
NPI	Neuropsychiatric Inventory
TASIT	The Awareness of Social Inference Test
ToM	Theory of Mind

Abstract

Behavioural variant frontotemporal dementia (bvFTD) is one of the most common early-onset dementia with the first symptoms occurring before the age of 65. Due to the lack of reliable biomarkers, neuropsychological assessment plays a crucial role in bvFTD diagnosis. In 2011 revised diagnostic criteria for bvFTD showed high sensitivity to the early bvFTD symptoms. However, bvFTD is still often misdiagnosed with other neurodegenerative and psychiatric disorders, especially at the early stages of the disease. The aim of this dissertation was to provide clinicians with new diagnostic tools that can increase diagnostic accuracy of early bvFTD. The focus lay on two diagnostic domains, (1) behavioural disorders as outlined in the current diagnostic criteria and (2) Theory of Mind, a subdomain of social cognition often impaired in bvFTD patients even before other behavioural or cognitive symptoms occur. As a result, two diagnostic instruments were developed. The first one is the Behavioural Dysfunctional Questionnaire, the first behavioural scale operationalising the current diagnostic criteria in form of a standardised informant-report questionnaire. The second one is the Basel Version of The Awareness of Social Inference Test – Theory of Mind, a video-based test assessing Theory of Mind ability close to real-life setting and fulfilling the requirements on clinical tools. The implementation of the developed tools in clinical evaluations of the patients with suspect on bvFTD can increase the diagnostic accuracy but also contribute to better understanding of symptoms' development and progression. This dissertation describes the development of these tools and discusses the possible effects of their implementation for diagnosis and therapy strategies of bvFTD.

Introduction

Worldwide, dementia is the 7th leading cause of death and one of the main diseases associated with aging (World Health Organization, 2020). With increasing life-expectancy, the relevance of research on aging related diseases, such as dementia, grows respectively. There are different neurodegenerative pathologies underlying dementia (e.g., Alzheimer's disease, Lewy body disease, frontotemporal lobar degeneration (FTLD) and others). Some of these pathologies manifest in different clinical syndromes (e.g., FTLD can be presented in the form of behavioural, language or motor dysfunction (Rascovsky & Grossman, 2013)). The interest of this dissertation is behavioural variant frontotemporal dementia (bvFTD), the most common clinical syndrome caused by FTLD. BvFTD belongs to young-onset dementia with symptoms occurring before the age of 65. It is characterized by prominent changes in personality, behaviour and social cognition, with relatively stable performance in such cognitive domains as memory or attention (Piguet et al., 2017; Rosness et al., 2016). Currently, there are no reliable biomarkers for non-genetic bvFTD cases accounting to ~60% of all bvFTD cases (Johnen & Bertoux, 2019), so an extensive neuropsychological assessment remains a gold-standard if bvFTD is suspected (Younes & Miller, 2020). However, there are severe challenges in differential diagnosis of bvFTD due to its heterogeneous cognitive profile and an overlap in behavioural symptoms with psychiatric conditions and other neurodegenerative disorders (Musa et al., 2020; Pose et al., 2013; Shinagawa et al., 2016). Especially at the early stages when appropriate symptoms' therapies would be the most efficient, bvFTD is often misdiagnosed (Shinagawa et al., 2016; Woolley et al., 2011). Therefore, development of new diagnostic tools increasing diagnosis accuracy at the early stages of the disease is necessary.

The aim of this dissertation is to provide clinicians with new bvFTD-specific neuropsychological tools sensitive for the early symptoms of the disease. The focus lies on two diagnostic domains recommended to be evaluated during a neuropsychological assessment if bvFTD is suspected, i.e., behavioural symptoms and social cognition. As result two instruments were developed: 1) Behavioural Dysfunctional Questionnaire (BDQ) assessing bvFTD-typical behavioural symptoms according to the latest diagnostic criteria for bvFTD, and 2) Basel Version of The Awareness of Social Inference Test – Theory of Mind (BASIT-ToM), a video-based instrument assessing Theory of Mind, one of the main subdomains of the social cognition.

This manuscript is organised in three parts. The first part is a general introduction into the clinical profile of bvFTD followed by a detailed presentation of the mentioned diagnostic domains, namely behavioural disorders and social cognition. The second part contains three studies introducing the new diagnostic instruments developed for the assessment of these

domains. The first study presents development and validation of the BDQ, an other-report questionnaire that showed good discriminating power between bvFTD and Alzheimer disease dementia (ADD, McKhann et al. 2011), and between bvFTD and major depressive disorder (MDD, American Psychiatric Association, 2013), respectively. The second study demonstrates the current state of social cognition assessment in German-speaking memory clinics and the existing gap in instruments fulfilling requirements of good diagnostic tools applicable in clinical routine. The last study describes the development of BASIT-ToM, a performance-based test assessing Theory of Mind, a social cognition subdomain often impaired in bvFTD. Finally, relevance and applicability of the new tools in bvFTD diagnosis and directions of further research are discussed.

Theoretical background

Behavioural variant frontotemporal dementia: Clinical profile and diagnostic challenges

BvFTD belongs to the spectrum of frontotemporal dementia (FTD), the second most common young onset dementia (after ADD) with the symptoms occurring before 65 years (Loi et al., 2021). The prevalence of FTD varies between studies and ranges from 7 to 30 per 100,000 individuals younger than 65. So, FTD accounts to about 10.2% of all young onset dementia cases (Hogan et al., 2016). FTD also encompasses two language syndromes, i.e., non-fluent primary progressive aphasia and semantic dementia (Gorno-Tempini et al., 2011), however, more than the half of all FTD cases account to bvFTD (Johnson et al., 2005).

BvFTD is characterized by early progressive changes in personality, behaviour and social conduct as well as specific cognitive deficits (Johnen & Bertoux, 2019). These symptoms correspond to the damaged brain areas. The neurodegeneration starts mostly in frontal and orbital paralimbic cortices involving anterior cortical and subcortical structures (Broe et al., 2003; Seeley et al., 2008). With the disease progression the neurodegeneration extends to temporal cortices and basal ganglia (Broe et al., 2003; Peet, Spina, et al., 2021).

To time three diagnostic criteria sets have been established in clinical practice and research (Englund et al., 1994; Neary et al., 1998; Rascovsky et al., 2011). These criteria sets are presented in Appendix A-C. Within this dissertation the current diagnostic criteria of Rascovsky et al. (2011) are used. These criteria were developed by 46 members of the International Behavioural Variant FTD Criteria Consortium (FTDC) based on the current literature and were evaluated in a sample of 137 bvFTD patients. Similarly to the previous diagnostic criteria, behavioural disorders received more attention than cognitive deficits resulting in five of six diagnostic domains (Appendix A-C). The behavioural domains are presented in more detail in the next chapter. The cognitive impairments in bvFTD are

characterised with deficits in executive functions – the higher-order cognitive functions, like planning or decision making – with relative sparing of memory and visuospatial functions. Noteworthy, though severe impairments in memory are traditionally seen as signs of ADD and therefore as an exclusion criterion for bvFTD (Piguet et al., 2011), current studies showed very heterogeneous cognitive profile of bvFTD patients including also severe memory impairments overlapping with the scores of ADD patients (Bertoux et al., 2018; Fernández-Matarrubia et al., 2017; Poos et al., 2018).

Comparing to previous diagnostic criteria, the FTDC-criteria have several advantages:

- clear operational definition (authors defined the time criterion ‘early’ and provided behavioural examples to ensure consistent understanding of criteria),
- reduced number of diagnostic features with no distinction between core and supportive features,
- flexibility in how patients can meet the criteria (not all described symptomatic domains need to be presented),
- hierarchy of diagnostic certainty (possible, probable, or definitive bvFTD).

All these aspects resulted in a greater sensitivity of the FTDC-criteria (86% and 76% for possible and probable bvFTD, respectively) compared to the Neary criteria (53%) (Rascovsky et al., 2011). Furthermore, the current diagnostic criteria demonstrated high interrater diagnostic agreement (.81 and .82 for possible and probable bvFTD, respectively) (LaMarre et al., 2013).

Even with the high sensitivity and standardisation of the new criteria, an accurate diagnosis of bvFTD remains challenging, especially at the early stages of the disease. So, about half of the bvFTD patients received first a psychiatric diagnosis (Krudop et al., 2017; Woolley et al., 2011). Oppositely, about 60% of patients who priority received a bvFTD diagnosis were misdiagnosed with other dementia syndromes (mostly ADD), psychiatric disorders or even clinically normal individuals (Shinagawa et al., 2016). There are several reasons explaining so high misdiagnosis rates:

- heterogeneous cognitive profile of bvFTD patients (e.g., memory performance can span from relatively intact (Ranasinghe et al., 2016) to severely impaired (Bertoux et al., 2018)),
- behavioural symptoms overlapping with other brain disorders (see chapter 1.2),
- lack of reliable biomarkers for non-genetic bvFTD (Johnen & Bertoux, 2019),
- lack of standardised guidelines on assessment methods and standardised tools operationalising the diagnostic criteria (Johnen & Bertoux, 2019).

The first aim of this dissertation is to provide a solution for the last described problem in bvFTD diagnosis, namely lack of standardised tools operationalising current diagnostic criteria. The focus was on the behavioural disorders as the most prominent symptoms in bvFTD. These are presented in the next chapter.

Behavioural disorders in behavioural variant frontotemporal dementia

Behavioural changes belong to criteria of already prodromal bvFTD (Barker et al., 2022). According to FTDC-criteria, typical behavioural disorders in bvFTD can be grouped into following five diagnostic domains:

- A. Early behavioural disinhibition,
- B. Early apathy or inertia,
- C. Early loss of sympathy or empathy,
- D. Early perseverative or compulsive behaviour,
- E. Hyperorality or dietary changes.

Importantly, all these behaviours need to be novel and, except for domain E, occur within the first three years after clinical onset (Rascovsky et al., 2011).

Disinhibition is one of the most salient symptoms in bvFTD and is characterised by personality change, withdrawal of social norms and impulsivity. These behavioural changes are associated with pathologies in ventromedial prefrontal and medial orbitofrontal cortices, and anterior insula (Massimo et al., 2009; Reber & Tranel, 2019) that are also involved in social cognition processes which will be described later. Another behavioural domain also partly explained with impairments in social cognition is *loss of empathy*. It partly overlaps in the areas of neuropathology with disinhibition and is associated with damages in ventromedial prefrontal, inferior frontal and insular cortices (Dermoddy et al., 2016; Shamy-Tsoory et al., 2009). Often reported symptoms are distancing, decreased social interest and lack of concern for feelings of others (Rascovsky et al., 2011). *Apathy* is one of the most common symptoms characterised by loss of motivation and goal-directed behaviour and is associated with pathology in dorsomedial prefrontal cortex and cingulate gyrus (Massimo et al., 2009; Reber & Tranel, 2019). *Stereotyped, ritualistic and obsessive-compulsive behaviour* often presented in bvFTD is associated with pathology in basal ganglia (particularly in striatum, left putamen and bilateral globus pallidus) and in lateral temporal lobe (particularly the left middle and inferior temporal gyri) (Josephs et al., 2008; Perry et al., 2012; Pickering-Brown, 2007). Lastly, *dietary changes* characterised by increased food consumption or changes in food preferences are associated with hypothalamus atrophy and changes in neuronal connection between the hypothalamus and

reward pathways in orbitofrontal cortex (Ahmed et al., 2021; Ahmed et al., 2014; Ahmed et al., 2015; Perry et al., 2014).

Though some of these behaviours can also be presented in other neurodegenerative disorders, the number and the severity of presented domains is distinctively for differential diagnosis. For example, disinhibition, stereotypic behaviour and dietary changes are shown to differentiate reliably between bvFTD and ADD as well as between bvFTD and semantic dementia (Ahmed et al., 2014; Bozeat et al., 2000; Luciano I Mariano, O'Callaghan, et al., 2020; Musa et al., 2020). Apathy is reported both in ADD and bvFTD, but bvFTD patients usually have more severe symptoms (Kumfor et al., 2018). The discrimination between bvFTD and psychiatric disorders is more challenging due to extensive overlap in behavioural symptoms (Peet, Castro-Suarez, et al., 2021). So, apathy and dietary changes can be mistaken for depression, impulsivity can be interpreted as mania and stereotyped behaviour is also a main symptom in obsessive-compulsive disorder (American Psychiatric Association, 2013). That's why assessment of only behavioural disorders is not always specific and other diagnostic domains need to be additionally evaluated to increase diagnosis accuracy. According to current literature, social cognition can be such a cognitive domain (Ducharme et al., 2020). Moreover, some of behavioural disorders in bvFTD, such as disinhibited behaviour, apathy and loss of empathy, can be caused and explained by impairments of social cognition and, consequently, the disability of bvFTD patients to estimate the mental state of others and the effects of their behaviour (Dilcher et al., 2023). The role of social cognition in bvFTD diagnosis is illustrated in the next chapter.

Social cognition as a new diagnostic domain for behavioural variant frontotemporal dementia

Social cognition assessment is not required by the current diagnostic criteria for bvFTD. However, these criteria were developed over twenty years ago, when the research on social cognition in patients with cognitive disorders just started. To date, there is extensive evidence emphasising the importance of social cognition deficits in bvFTD diagnosis (Dilcher et al., 2023; Kipps et al., 2009; Mendez, 2022). Indeed, the social cognition impairments often occur already at the early stage of the disease and have been shown to be more sensitive to early bvFTD than “traditional” cognitive tasks (Torralva et al., 2009). Moreover, social cognition deficits in bvFTD are more profound than in most other brain disorders (for overview see Dodich et al., 2021), so its assessment can increase the accuracy of differential diagnosis.

Social cognition: Definition, structure and neuronal correlates

Social cognition encompasses a range of cognitive processes enabling individuals to comprehend and engage with others. Therefore, it is a crucial ability enabling humans to function effectively in society. Although social cognition has been an interest of research in different sciences (e.g., psychology, neuroscience, sociology), there is neither a unified definition nor consistent structure of this concept among researchers. Yet, in consonance with information processing theory seeing the human brain as a computer that perceives inputs, processes these and generates outputs (Frith, 2008), social cognition can be defined as *perception, interpretation and response to social stimuli* (Adolphs, 2001; Arioli et al., 2018). *Social stimuli* are all information coming from other individuals during the social interaction, e.g., facial or vocal expressions. For a proper social interaction these stimuli first need to be *perceived*. So, a person should be able to detect face expressions, intonations etc. In the next step, the perceived information needs to be *interpreted* according to the person's explicit knowledge, e.g., the meaning of a certain facial expression. Finally, the person will *respond* to the perceived stimuli guiding own behaviour according to one's knowledge about appropriate behaviour in such situation. Therefore, social behaviour as response to social stimuli relies strongly on the ability to perceive and to interpret these stimuli as well as the knowledge of social norms (Desmarais et al., 2018; Magno, Canu, Agosta, et al., 2022).

The structure of social cognition is very inconsistent among researchers. Many concepts, such as emotion processing, empathy, perspective taking, understanding of humour or lie, moral judgment, etc. are classified as social cognition subdomains (Eddy, 2019). Considering the various structures and the complexity of all processes included in social cognition, it is difficult to identify a particular brain region responsible for it. The most named brain areas are medial frontal and prefrontal cortex and temporoparietal junction (Amodio & Frith, 2006; Bzdok et al., 2012; Magno, Canu, Filippi, et al., 2022; Van Overwalle, 2009). However, it is advisable in the context of social cognition to speak about neuronal networks instead of brain areas (Krendl & Betzel, 2022). Especially noteworthy is the "salience network" involved in identifying behaviour relevant stimuli (Seeley et al., 2007). Research of last decades showed that damage in the salience network is both needed and sufficient to cause bvFTD typical behavioural disorders (Rankin, 2021). This network involves, inter alia, anterior insula and anterior cingulate cortex (Seeley et al., 2007) that are known to coactivate in response to emotionally significant stimuli (Critchley, 2005; Toller et al., 2018) and overlap with in bvFTD often damaged brain structures (Seeley et al., 2012).

Since the aim of this dissertation is to provide clinicians with new diagnostic tools, it is reasonable to follow the Diagnostic and Statistical Manual (DSM-5, American Psychiatric Association, 2013) recommending to evaluate two subdomains of social cognition, i.e., emotion recognition (Ferretti & Papaleo, 2019) and Theory of Mind (Premack & Woodruff, 1978). This dissertation focuses only on the Theory of Mind, a more complex concept that has been often reported to be impaired in bvFTD (Henry et al., 2014).

Theory of mind and its role in diagnosis of behavioural variant frontotemporal dementia

Theory of Mind (ToM), or mentalizing, is the ability to infer about mental state of others, their feelings, beliefs and intentions (Premack & Woodruff, 1978). ToM is a complex concept, including affective (mentalizing of others' emotional state) and cognitive (mentalizing of others' thoughts and intentions) components (Hynes et al., 2006; Kalbe et al., 2010; Shamay-Tsoory & Aharon-Peretz, 2007). Considering that social interaction involves often several individuals, ToM can be additionally discriminated in orders (Perner & Wimmer, 1985; Tesar et al., 2020). Usually, first- and second-order ToM is studied. First-order ToM is what a tested person thinks about individual's A mental state. Second-order ToM involves additionally one more person and asks what a tested person believes that individual A thinks about individual's B mental state. Each next order includes one additional person (Tesar et al., 2020). It is important to note that differently to other cognitive skills, ToM can be impaired in two ways: underperformed (hypomenthalising) and overperformed (hypermentalising). By the latter, patients generate additional hypotheses about someone's mental state that cannot be supported with the cues from the social situation (Frith, 2004). This complex structure of ToM makes its assessment to a challenge in research and in diagnostic.

Neuroanatomical ToM is associated with typical brain areas involved in social cognition, such as anterior cingulate cortex, medial prefrontal cortex, frontoinsula and temporoparietal junction (Strikwerda-Brown et al., 2019). Interestingly, affective and cognitive ToM are partly dissociable. Affective ToM is more associated with activation in inferior frontal gyrus, orbitofrontal and ventromedial lobes, while cognitive ToM is more associated with dorsolateral prefrontal cortex, precuneus and regions in the temporal lobes (Healey & Grossman, 2018; Hynes et al., 2006; Kalbe et al., 2010; Molenberghs et al., 2016; Schlaffke et al., 2015; Shamay-Tsoory & Aharon-Peretz, 2007). Thus, the in ToM involved brain areas overlap strongly with the ones typically damaged in bvFTD.

Indeed, many studies showed that bvFTD patients perform poorly on different ToM tasks, like perspective taking (Strikwerda-Brown et al., 2022), faux-pas stories (Delbeuck et al., 2022; Giovagnoli et al., 2019), attribution (Van den Stock et al., 2019) or false-belief understanding

(Fernandez-Duque et al., 2009), also if comparing with tasks involving similar cognitive demand without social interpretations (Henry et al., 2014). Interestingly, low scores on ToM in healthy subjects in age of 50-60 years were predictive for decline in prefrontal function in follow-up evaluation two years later, including findings suggestive for bvFTD (Pardini et al., 2013). More important for differential diagnosis is that bvFTD patients perform significantly worse on ToM task than most other diagnostic groups, e.g., ADD (Bora et al., 2015; Buhl et al., 2013; Luciano Inácio Mariano, Caramelli, et al., 2020), Huntington's disease (Snowden et al., 2003), bipolar disorder (Baez et al., 2019; Barbosa et al., 2023), or MDD (Bertoux et al., 2012; Lichtenstein et al., 2023). Talking about discriminating between bvFTD and ADD patients by means of their ToM performance, it is important to mention that ADD patients also show ToM deficits but mostly in highly performance task (e.g., second-order cognitive ToM) or at the late stages of the disease (Setién-Suero et al., 2022). The performance on ToM tasks highly differentiated ADD from bvFTD even in cases when memory performance of bvFTD was comparable to the one in ADD (Bertoux et al., 2016).

Thus, assessment of ToM skills during neuropsychological evaluation can increase diagnostic accuracy of bvFTD at the early stages of the disease and provide first cue on the risk of development of the bvFTD later in the future. The next chapter presents the current state of neuropsychological assessment in bvFTD with focus on ToM evaluation and assessment of behavioural disorders.

Neuropsychological assessment in diagnosis of behavioural variant frontotemporal dementia

Neuropsychological assessment plays a crucial role in diagnosis of patients with brain disorders, e.g., dementia, especially in cases with absent or unreliable biomarkers, like in sporadic bvFTD. According to the definition of the National Academy of Neuropsychology, clinical neuropsychology is a field of psychology specialized on brain-behaviour relationships across the lifespan (Barth et al., 2003). It aims to transform unobservable brain functioning processes (behavioural responses) in quantitative values (scores) that can be compared to normative values in the questioned population. Comparing these scores with the normative values, neuropsychologists can answer the question of whether any or a certain brain function is impaired (Russell, 2012, pp. 69-110).

The practice of neuropsychological assessments relies strongly on the accuracy of the assessment instruments. That's why it is important that such instruments are developed with high accuracy and fulfil the psychometric requirements. Except the general psychometric criteria, like reliability and validity (Moosbrugger & Kelava, 2020), neuropsychological tools

need to fulfil additionally specific criteria, i.e., proven association between test score and questioned brain function, provided normative values for the population of interest and applicability in clinical setting (Russell, 2012). Ideally, the instruments need to avoid ceiling- and floor-effects to be able to assess both deficits and overperformance in questioned functions. Thus, the process from development of a neuropsychological tool to its use in clinical practice usually contains several phases (e.g., theory- or data-driven development of the tool, its validation in different samples as well as calculation of the normative values or discriminating cut-offs) and requires several studies with different patients' groups and healthy controls.

Next, the current state of the assessment of the described above diagnostic domains will be discussed explaining the need for new diagnostic instruments.

Assessment of behavioural symptoms in patients with brain disorders

Information on behavioural symptoms can be collected in two ways, i.e., observation of patients or anamnesis reports from patients or their close persons (informants). Both methods have their advantages and disadvantages in the context of bvFTD diagnosis.

Observation of the patients' behaviour during the clinical evaluation requires clinicians to have extensive experiences with typical symptoms of the suspected diagnosis to be able to make reliable observation and rating. These observations need to be transformed into scores to be compared with normative values. This can be done by the use of standardised scales, like the Social Observer Behavior Checklist (Rankin et al., 2021) allowing clinicians to evaluate spontaneous behaviour of the patients during cognitive assessment. Though observational methods belong to the objective assessment methods, they are limited by time and context (Rascovsky & Grossman, 2013). So, clinicians can only evaluate behaviour during the clinical evaluation when not all behaviours can be shown (e.g., eating changes or hoarding). Furthermore, clinicians cannot evaluate whether the occurred behaviour is novel or typical for the individual and its progression over time. As mentioned above, novelty, quantitative difference from the previous behaviour and prominent progression of the symptoms are important features helping to discriminate bvFTD from other disorders with similar behavioural symptoms (Johnen & Bertoux, 2019). Since these features cannot be evaluated during an observation, the anamnesis reports provide an essential source of information.

Anamnesis reports are probably the most used methods in the clinical routine to assess information on symptoms in almost all diseases including behaviour disorders. This method belongs to the subjective ones and requires awareness of and sensitivity to changes in patients' behaviours (Johnen & Bertoux, 2019). Since bvFTD patients often lack insight in their disease and symptoms (Eslinger et al., 2005; Rankin et al., 2005), informant reports are essential if

bvFTD diagnosis is suspected. The information can be collected in the form of an interview or a questionnaire. Interviews can provide more information due to direct interaction between clinicians and informants, but they also require a trained person and additional time that is often limited in clinical practice. Unlike, questionnaires can be filled out by informants in absence of a clinician and enable assessment of patients' behavioural symptoms in a standardised way within a short time. Moreover, it is easy to provide normative values for questionnaires due to their highly standardised form. Informants can provide information about the patients' behaviour over a longer period of time (usually several years), in different contexts, and comparing to previous behaviour (to ensure that behaviour is novel and not a personality trait). However, it is important to keep in mind, that information from informants can be influenced by subjective estimation (personal values and norms) and individual mental state (often informants show high stress and fatigue levels (Koyama et al., 2018; Lima-Silva et al., 2015)).

Currently, several informant-report scales on patients' behaviour exist, e.g., Clinical Dementia Rating scale (CDR, Hughes et al., 1982), Neuropsychiatric Inventory (NPI, Cummings et al., 1994), Frontotemporal Dementia Rating Scale (FRS, Mioshi et al., 2010) or Frontal Behavioural Inventory (FBI, Kertesz et al., 1997). However, most of these scales are either not focused on bvFTD-typical behaviour (e.g., NPI, CDR) or based on the outdated diagnostic criteria (e.g., FRS, FBI). This leads to the problem that despite consensus symptoms' definitions in the current diagnostic criteria and their high sensitivity to bvFTD symptoms, there are still no instruments that have operationalised these criteria in the form of a diagnostic tool.

Assessment of social cognition and theory of mind in clinical routine

Social cognition has been recognised as an important cognitive domain crucial for daily functioning and has been added as a diagnostic domain to the DSM-5 (American Psychiatric Association, 2013) and recently to the International Classification of Diseases (ICD-11, World Health Organization, 2022), two most common diagnostic manuals. However, it is still only rarely tested in clinical routine. Two main reasons can be named: 1) lack of guidelines, which subdomains and how should be tested, and 2) poor number of tools validated in clinical settings and with available normative values (Samtani et al., 2023). The current situation with social cognition assessment at the German-speaking memory clinics as well as the requirements of clinicians on social cognition tools will be illustrated and discussed in the Study II.

Talking specific about the assessment of ToM, it is worth highlighting the complexity of this concept and the variability of the paradigms (e.g., understanding of sarcasm or lie (McDonald et al., 2003), faux pas (Stone et al., 1998), or perspective taking (Samson et al., 2010)) and tasks (e.g., text stories (Happé, 1994), portrait photos (Baron-Cohen et al., 1997),

films (Dziobek et al., 2006), or questionnaires (Davis, 1980)) with which ToM is currently assessed. Interestingly, many instruments applied in ToM assessment are also applied in assessment of other social cognition skills, like empathy or emotion recognition, raising a question which brain function exactly they assess (Quesque & Rossetti, 2020). Moreover, most of these instruments were developed for research purposes and do not fulfil the requirements of practitioners. Most limitations of the existing tools are a high demand on additional cognitive skills (e.g., working memory), task settings far from real-world situations (low ecological validity), and lack of normative values and validations with different patients' groups (Eddy, 2019; Quesque et al., 2022). Finally, only very few instruments consider the complex structure of ToM (affective vs. cognitive, first- vs. second order, and hypo- vs. hypermentalising) that could be especially useful for differential diagnoses. Thus, though tens of instruments assessing ToM abilities exist, there is still the need for new instruments assessing ToM in reality-close setting and fulfilling requirements of good diagnostic tools.

Summary

Summing up, though the research on bvFTD is extensive, its differential diagnosis remains challenging especially at the early stages of the disease. The new diagnostic criteria of Rascovsky et al. (2011) showed high sensitivity to bvFTD, however there are still no instruments operationalising these criteria in the form of a standardised diagnostic tool. Moreover, the behavioural symptoms representing five of six diagnostic domains for bvFTD can be also presented in other brain disorders demanding researchers and clinicians to search for additional diagnostic domains both sensitive for the early bvFTD symptoms and able to discriminate bvFTD from other brain disorders already at the early stage. Social cognition, and particularly Theory of Mind, have been recommended as such domains. Indeed, the international diagnostic guidelines added social cognition as one of the diagnostic domains to be evaluated during a neuropsychological assessment. However, there is still a lack of diagnostic tools fulfilling specific requirements of clinical setting and applicable in German-speaking patients with brain disorders.

The goal of this dissertation is to close these two gaps providing clinicians with appropriate tools that can increase diagnosis accuracy at the early stage of bvFTD. The focus was on (1) development of a tool operationalising the current diagnostic criteria for bvFTD with the focus on behavioural symptoms, and (2) development of a tool assessing Theory of Mind close to real-life conditions and applicable in the clinical setting. The results are presented in the three following studies and discussed afterwards.

STUDY I: The Behavioural Dysfunction Questionnaire discriminates behavioural variant frontotemporal dementia from Alzheimer's disease dementia and major depressive disorder



The Behavioural Dysfunction Questionnaire discriminates behavioural variant frontotemporal dementia from Alzheimer's disease dementia and major depressive disorder

Anna Semenkova^{1,2} · Olivier Piguet³ · Andreas Johnen⁴ · Matthias L. Schroeter^{5,6} · Jannis Godulla^{5,6} · Christoph Linnemann⁷ · Markus Mühlhauser⁷ · Thomas Sauer⁷ · Markus Baumgartner⁸ · Sarah Anderl-Straub⁹ · Markus Otto^{9,10} · Ansgar Felbecker¹¹ · Reto W. Kressig¹ · Manfred Berres¹² · Marc Sollberger^{1,13}

Received: 7 December 2022 / Revised: 22 February 2023 / Accepted: 12 March 2023
© The Author(s) 2023

Abstract

Background and objectives Early-stage behavioural variant frontotemporal dementia (bvFTD) is often misdiagnosed, highlighting the need for new diagnostic instruments. Based on the revised diagnostic criteria for bvFTD, we developed the Behavioural Dysfunction Questionnaire (BDQ). In this explorative study, we aimed to determine the best scoring and analytical method for the BDQ to discriminate between bvFTD and non-bvFTD patients.

Materials and methods 34 patients with early-stage bvFTD, 56 with early-stage Alzheimer's disease dementia (ADD) and 41 with major depressive disorder (MDD) were recruited. We calculated BDQ-items with or without inclusion of a time criterion: (a) without time criterion, (b) with 10 years' time criterion (symptom presence less than 10 years), and (c) with 3 years' time criterion (symptom presentation within the first 3 years). Using these three differently calculated items, we generated six variables, i.e. 3*2 [BDQ-Global Score (BDQ-GS; domains average score); BDQ-Global Domain Score (BDQ-GDS; domains categorical score)]. Then, we performed univariate and bivariate (BDQ-GS and BDQ-GDS combined) ROC analyses.

Results Models including BDQ-GS, BDQ-GDS or both variables combined discriminated similarly between groups. In contrast, models without time criterion or with 10 years' time criterion discriminated better than models including variables with 3 years' time criterion. These models discriminated highly (AUC = 85.98–87.78) between bvFTD and MDD and bvFTD and ADD, respectively.

Conclusion BDQ-scores without any time criterion discriminated highly between early-stage bvFTD and non-bvFTD groups, which could improve the early diagnosis of bvFTD. With its standardised procedure, the BDQ is also appropriate for repeated assessments.

Keywords Behavioural variant frontotemporal dementia · Alzheimer's disease · Depressive disorder · Behavioural disorder · Questionnaire

Introduction

After Alzheimer's disease dementia (ADD), frontotemporal dementia (FTD) is the second most common younger-onset dementia, with the behavioural variant (bvFTD) as its most frequent clinical syndrome [27, 31, 35]. BvFTD is a neurodegenerative disorder associated with early progressive

changes in personality, behaviour, and social interactions [27, 35], often with only mild and nonspecific cognitive deficits in the early stages of the disease [24]. Currently, no biomarkers exist that enable a reliable early diagnosis in sporadic (i.e. non-genetic) bvFTD cases [13, 35]. Thus, diagnoses rely strongly on clinical assessment, in which one of the main diagnostic challenges is the clinical overlap of bvFTD with primary psychiatric disorders (PPD) [9, 25, 32, 36] and neurodegenerative disorders such as behavioural variant Alzheimer's disease (AD) [21, 23]. Indeed, up to 50% of bvFTD patients are first diagnosed with PPD and vice versa [15, 29], of which major depressive disorder

✉ Marc Sollberger
marc.sollberger@felixplatter.ch

Extended author information available on the last page of the article

(MDD) is probably the most common misdiagnosis [21, 29, 36]. Similarly, 7–17% of patients diagnosed with bvFTD are found to have AD pathology post-mortem [2, 12].

In 2011, the revised diagnostic criteria for bvFTD were published [26]. These criteria are based on six clinical domains including five behavioural domains (i.e. early behavioural disinhibition; early apathy/inertia; early loss of sympathy/empathy; early perseverative, stereotyped or compulsive/ritualistic behaviour; and hyperorality/dietary changes) and one cognitive domain (primary executive dysfunction). The risk of examiner-biased assessment of these domains seems quite low among bvFTD experts according to LaMarre, et al. [16], who found moderate to high inter-rater agreement ($\kappa=0.41\text{--}0.80$) between the six domains. However, the majority of clinicians, who are evaluating these domains, are not bvFTD experts, which increases the likelihood of examiner-biased assessment. Moreover, endorsement or non-endorsement of a behavioural domain is a crude approach, which does not capture the severity of the behavioural disorders. Consequently, the development of an informant questionnaire, which assesses these five behavioural domains in a standardised and quantitative way, is warranted.

To the best of our knowledge, two instruments are currently available to assess bvFTD-specific behavioural disorders, namely the Frontal Behavioral Inventory (FBI) [14] and DAPHNE (named for Disinhibition, Apathy, Perseverations, Hyperorality, Personal neglect and Loss of Empathy) [6]. The FBI, which exists both as an informant interview and as an informant questionnaire, was designed to optimise diagnostic accuracy for the Lund–Manchester criteria for frontotemporal dementia [11], but also includes items based on the authors' experience [14]. Accordingly, the FBI comprises items such as inattention or incontinence [14] that are not part of the current diagnostic criteria for bvFTD [26].

Unlike the FBI, DAPHNE [6] is based on the five behavioural domains of the diagnostic criteria for bvFTD [26]. It is an informant interview that is composed of 10 items with five possible answer categories, designed as semi-structured propositions similarly to the Clinical Dementia Rating scale [20]. DAPHNE does not reflect completely the structure of the behavioural domains of the bvFTD diagnostic criteria. First, "personal neglect" is an addition to the five behavioural domains. Second, the behavioural domains of DAPHNE are weighted differently, i.e. behavioural disinhibition represents 40%, hyperorality represents 20% and each of the other four domains represent 10% of the ten items. Third, DAPHNE does not consider the time criterion "early" as required by the diagnostic criteria for scoring four of the five behavioural domains [26].

In light of this evidence, we aimed to develop an instrument that would operationalise the diagnostic criteria for bvFTD more precisely. To meet this goal, we operationalised

the five bvFTD-behavioural domains according to the examples of behavioural disorders of the consensus paper for bvFTD [26]. We opted for an informant questionnaire rather than an interview instrument to facilitate its use in clinical practice, and named it "Behavioural Dysfunction Questionnaire" (BDQ). Information on the development of the BDQ is provided in the Methods section and Supplementary Material A.

In this study, we administered the BDQ to informants of patients with probable bvFTD [26], probable ADD [18] and major depressive disorder (MDD) [34].

Our aims were (1) to determine the best scoring method for BDQ to discriminate between bvFTD and the other two patient groups, and (2) to compare its discriminatory power with that of the FBI [14], probably the most common inventory for assessing bvFTD-specific behavioural symptoms at present.

Methods

Participants

In total, 131 patients were recruited from several Swiss and German medical centres with expertise in early diagnosis of bvFTD, ADD and/or MDD. Thirty-four patients with probable bvFTD [26], 56 patients with either probable ADD with evidence of the AD pathophysiological process ($n=49$) or probable ADD ($n=7$) [18] and 41 patients with MDD (i.e. at least moderate depressive episode according to ICD-10 [34]) were recruited. As the BDQ should primarily help discriminating early-stage bvFTD patients from other patient groups, we included only bvFTD and ADD patients with a major neurocognitive disorder at mild stage according to DSM-5 [3]. An additional inclusion criterion for all patients was availability of a reliable informant (>18 years) who has regular contact with the patient. Specific exclusion criteria for bvFTD and ADD patients were a major neurocognitive disorder at or above moderate stage according to DSM-5, history of severe depressive episode or current depressive episode according to ICD-10 and history of or current major psychiatric disorders according to ICD-10. Specific exclusion criteria for MDD patients were a neurocognitive disorder according to DSM-5 and any other major psychiatric disorders according to ICD-10. Exclusion criteria for all patient groups were history of or current drug or/and alcohol abuse as well as drug- or/and alcohol-related disorders according to ICD-10 and traumatic brain injuries, systemic disorders or brain diseases that could result in behavioural changes.

To increase diagnostic certainty, we confirmed the majority of patients' diagnoses [86% (88% bvFTD, 88% ADD (88% probable ADD with evidence of the AD pathophysiological process, 86% probable ADD), 83% MDD)] by at least one

follow-up assessment either in the institution or by a standardised phone interview (mean time period of 24 ± 11 months).

As some BDQ-items such as cursing are to some degree not necessarily pathological, we also collected BDQ-data of 414 cognitively and mentally healthy Central European individuals [52% women; age 69.21 ± 12.53 years; 14.61 ± 3.16 years of education] to determine the range of behaviour as measured by the BDQ in the general population. Inclusion and exclusion criteria for the healthy participants are outlined in Supplementary Material B.

Instruments

Behavioural Dysfunction Questionnaire (BDQ)

The BDQ is an informant questionnaire based on the five behavioural domains of the bvFTD diagnostic criteria [26]. Items of each behavioural domain are scored for their frequency or severity on a Likert-scale from 0 (none) to 5 (very often/very severe). For each endorsed item, the informant was also required to state its time of onset. To know whether endorsed items fulfil the time criterion “early” as required by Rascovsky, et al. [26], informants were asked to state both, the time of onset of the endorsed item and the time when the first symptoms appeared. The development and design of the BDQ are presented in Supplementary Material A.

Frontal Behavioral Inventory (FBI)

To test the convergent validity of the BDQ, we administered the German version of the FBI as an informant questionnaire [14]. It contains 24 items scored on a Likert-scale from 0 (none) to 3 (severe/most of the time). The total score is the sum of all items.

Statistical approach

Evaluating the need for adjustment of patients’ scores in relation to healthy subjects’ scores

First, we compared healthy subjects’ item scores (Supplementary Material C) with bvFTD patients’ item scores using Kendall-Tau [1] and adjusted patients’ items scores, based on these coefficients. Comparisons between the adjusted and not adjusted patients’ items scores using ROC analyses revealed no significant differences between these scores. As such, all subsequent analyses use unadjusted scores.

Exclusion of BDQ-items affirmed by less than 5% of bvFTD patients’ informants

To omit behavioural items not related to early-stage bvFTD, we excluded items endorsed by fewer than 5% of bvFTD

patients’ informants. In doing so, we removed 6 of the 56 items (see removed items in Supplementary Material A).

Scoring of item scores

Next, we devised three different item scoring methods (a–c):

- (a) Original item scores (i.e. scores unchanged);
- (b) Item scores adjusted according to the 10 years’ time criterion. By assuming that a patient’s behaviour that exists over 10 years is rather a personality trait than due to a neurodegenerative disease or another brain disease, we set any endorsed items with a duration longer than 10 years to zero;
- (c) Item scores of four bvFTD domains (i.e. disinhibition, apathy/inertia, loss of empathy, stereotypic behaviour) were adjusted for time criterion “early” as defined by the diagnostic criteria for bvFTD [26]. “Early” refers to symptom presentation within the first 3 years [26]. Endorsed items that did not fulfil this criterion were set to zero.

Generating two global BDQ-scores

BDQ-Global Score (BDQ-GS) was calculated as an average score of the mean domains’ scores. By taking this approach, we ensured that each domain score contributed equally to the total score.

BDQ-Global Domain Score (BDQ-GDS) represents the number of endorsed behavioural domains (0–5). According to the diagnostic criteria, a behavioural domain is endorsed if at least one behavioural feature (i.e. item) of this domain is “persistent or recurrent, rather than single or rare ...” [26]. Accordingly, we considered a domain as endorsed if an item was scored as “sometimes/moderate” or greater. As several items were also endorsed in healthy participants, we added “and above healthy subjects’ 99th percentile of this item” as an additional criterion.

By applying the above-mentioned three different scoring methods on these two global scores, we generated six variables:

- 1a. BDQ-GS without time criterion
- 1b. BDQ-GS with 10 years’ time criterion
- 1c. BDQ-GS with 3 years’ time criterion
- 2a. BDQ-GDS without time criterion
- 2b. BDQ-GDS with 10 years’ time criterion
- 2c. BDQ-GDS with 3 years’ time criterion

Data analysis

First, we derived a non-bvFTD group by combining ADD and MDD patients, in order to compare the discriminatory

power of the six variables between bvFTD and non-bvFTD patients. To test if the merging of these two patient groups is statistically meaningful, we compared the “BDQ-GS without time criterion” score between the ADD and MDD groups. A Kruskal–Wallis test, followed by pairwise Wilcoxon tests, revealed no statistically significant group differences (data not shown), allowing us to combine these participants into a non-bvFTD group.

Next, we run six univariate logistic regressions, followed by ROC analyses. As the BDQ-GS and the BDQ-GDS represent the data differently (i.e. BDQ-GS represents all items independent of domain structure, BDQ-GDS represents number of endorsed domains), we run as well three bivariate (i.e. variables 1a and 2a, variables 1b and 2b, and variables 1c and 2c) logistic regressions in an effort to best separate the groups. Using the Delong’s method [8], we compared the discriminatory power of these nine regression models. We aimed to select the regression model with the highest discriminatory power between bvFTD and non-bvFTD patients. Finally, by taking the best regression model, we aimed (1) to examine its discriminatory power between bvFTD and ADD and between bvFTD and MDD, respectively, and (2) to compare its discriminatory power with the FBI score.

Results

Analyses of covariance followed by Tukey–Kramer post hoc analyses for age and education and Chi-square test for sex showed no differences among patient groups (Table 1). BvFTD and ADD patients were cognitively more impaired than MDD patients as measured by the Montreal Cognitive Assessment [22]. In addition, bvFTD patients showed higher FBI scores than ADD and MDD patients (Table 1). Based on the Frontotemporal Dementia Rating Scale [19], an informant questionnaire, bvFTD patients experienced on average

moderate to severe functional dependence and behavioural disturbances (Table 1).

The internal consistencies of the five BDQ-domains ranged from poor ($\alpha = 0.54$; “hyperorality and dietary changes”), over acceptable ($\alpha = 0.67$; “early apathy/inertia”) to good ($\alpha = 0.76$ – 0.86 ; “early behavioural disinhibition”, “early loss of sympathy/empathy” and “early perseverative/stereotyped behaviour”). The BDQ showed an excellent overall internal consistency ($\alpha = 0.92$). Please see Supplementary Material D for more detail.

Determining the best BDQ scoring method

The nine regression models showed acceptable to excellent [17] discriminatory power between bvFTD and non-bvFTD patients (AUC ranging between 78.08 and 87.78%) (Table 2). Neither models including BDQ-GS variables nor models including BDQ-GDS variables turned out stronger. Likewise, bivariate regression models did not discriminate better than univariate regression models. To determine whether our findings were driven by single domains, we run post hoc univariate regression analyses with each behavioural domain’s mean score and found similar discriminatory accuracies (Supplementary Material E). Despite the fact that both global scores discriminated similarly, we favoured the BDQ-GS over the BDQ-GDS, as the BDQ-GS is more informative, i.e. it considers the degree of each item, whereas the BDQ-GDS is limited to the number of endorsed behavioural domains.

Regarding the time criterion, models that included the variables without time criterion or with 10 years’ time criterion tended to discriminate better than the models that included variables with the 3 years’ time criterion ($p = 0.02$ – 0.08). Post hoc analyses in patients with follow-up BDQ assessments ($n = 44$; mean time period of 16.5 ± 6.43 months) revealed that informants’ data on

Table 1 Demographic and clinical characteristics of study participants ($N = 131$) classified by diagnostic group

	bvFTD ($n = 34$)	ADD ($n = 56$)	MDD ($n = 41$)	Test (df)	Post hoc
Age (years)	64.76 ± 9.78	67.68 ± 10.97	63.32 ± 10.40	$2.17_{(2,128)}^a$	
Sex (m/f)	20/14	22/34	17/24	$3.57_{(2)}^b$	
Education (years)	13.88 ± 2.80	13.14 ± 3.53	13.31 ± 3.34	$0.53_{(2,120)}^a$	
MoCA (0–30)	18.6 ± 5.49	17.82 ± 5.10	25.21 ± 3.99	$37.84_{(2)}^{c*}$	MDD > bvFTD, ADD*
FBI (0–72)	26.69 ± 12.78	13.6 ± 8.77	12 ± 8.52	$27.53_{(2)}^{c*}$	bvFTD > ADD, MDD*
FRS (0–100%)	39.41 ± 23.79	n/a	n/a		

bvFTD behavioural variant frontotemporal dementia, *MDD* major depressive disorder, *ADD* Alzheimer’s disease dementia, *MoCA* Montreal Cognitive Assessment, *FBI* Frontal Behavioral Inventory, *FRS* Frontotemporal Dementia Rating Scale (0–2 = profound; 3–12 = very severe; 13–40 = severe; 41–79 = moderate; 80–96 = mild; 97–100 = very mild), *n/a* not applicable

* $p < 0.001$

^aAnalysis of variance

^bChi-square test

^cKruskal–Wallis test

Table 2 Area under the curves of six univariate and three bivariate logistic regression models in bvFTD and non-bvFTD patients

	Without time criterion	Ten years' time criterion	Three years' time criterion	Delong's method
BDQ-GS models	85.98 (CI 78.73–93.22)	86.11 (CI 78.43–93.79)	79.52 (CI 69.98–89.05)	Model 1a, model 1b > model 1c*
BDQ-GDS models	86.43 (CI 79.05–93.81)	86.81 (CI 79.26–94.36)	78.08 (CI 68.10–88.05)	Model 2a, model 2b > model 2c**
BDQ-GS & BDQ-GDS models	87.36 (CI 80.18–94.53)	87.78 (CI 80.49–95.07)	79.64 (CI 70.11–89.17)	Model 3a, model 3b > model 3c**

1a. BDQ-GS without time criterion, 1b. BDQ-GS with 10 years' time criterion, 1c. BDQ-GS with 3 years' time criterion, 2a. BDQ-GDS without time criterion, 2b. BDQ-GDS with 10 years' time criterion, 2c. BDQ-GDS with 3 years' time criterion, 3a. BDQ-GS without time criterion and BDQ-GDS without time criterion, 3b. BDQ-GS with 10 years' time criterion and BDQ-GDS with 10 years' time criterion, 3c. BDQ-GS with 3 years' time criterion and BDQ-GDS with 3 years' time criterion

BDQ-GS BDQ-Global Score, BDQ-GDS BDQ-Global Domain Score

* $p < 0.1$, ** $p < 0.05$

symptoms' onset, on which we based the time criteria, had low reliability. They deviated on average by 14.47 months (0.03–145.33 months) between two time points. Given these findings, plus the fact that the collection of time data turned out to be elaborate, we decided to omit the time criterion in all subsequent analyses.

In the end, we decided to select the variable "BDQ-GS without time criterion" for further analyses.

Discriminatory power of BDQ between bvFTD and ADD and between bvFTD and MDD, respectively

Similar discriminatory power between bvFTD and ADD (AUC = 87.84%) and between bvFTD and MDD (AUC = 83.43%) was observed based on the "BDQ-GS without time criterion" variable. Applying the Youden-Index, no single cut-off scores with sufficient sensitivity and specificity were identified. Therefore, we decided to identify two

cut-off scores, that is, in each analysis, the cut-off scores with at least 90% sensitivity or 90% specificity [10, 30].

When examining the discriminatory power between bvFTD and ADD, we found a score of > 1.4 to be strongly indicative for bvFTD (sensitivity 65%, specificity 91%) and score of < 0.6 to be strongly indicative for ADD (sensitivity 91%, specificity 59%). Scores between 0.6 and 1.4 were considered equivocal (Fig. 1).

When examining the discriminatory power between bvFTD and MDD, we found a score of > 1.6 to be strongly indicative for bvFTD (sensitivity 56%, specificity 90%) and score of < 0.6 to be strongly indicative for MDD (sensitivity 91%, specificity 56%). Scores between 0.6 and 1.6 were considered equivocal (Fig. 2).

Lastly, we compared the discriminatory power of the "BDQ-GS without time criterion" variable with the one of the FBI variable. We found similar discriminatory powers of the two variables between bvFTD and ADD patients, and between bvFTD and MDD patients, respectively.

Fig. 1 Cut-offs between bvFTD and ADD patients. Percentages of patients with bvFTD who were correctly classified [sensitivity, orange line with triangles] and percentages of correctly classified ADD patients [specificity, blue line with circles] in relation to the BDQ-scores. Two cut-offs with either sensitivity or specificity above 90% are highlighted by dashed lines. The solid black line represents the optimal cut-off using the Youden-Index (sensitivity = 79%; specificity = 88%). *bvFTD* behavioural variant frontotemporal dementia, *ADD* Alzheimer's disease dementia, *BDQ* Behavioural Dysfunction Questionnaire

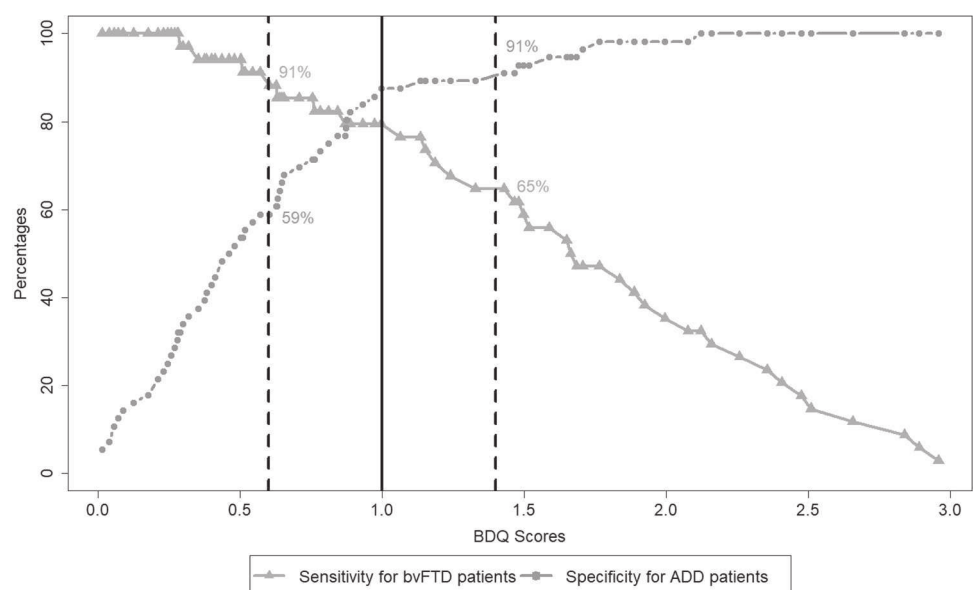
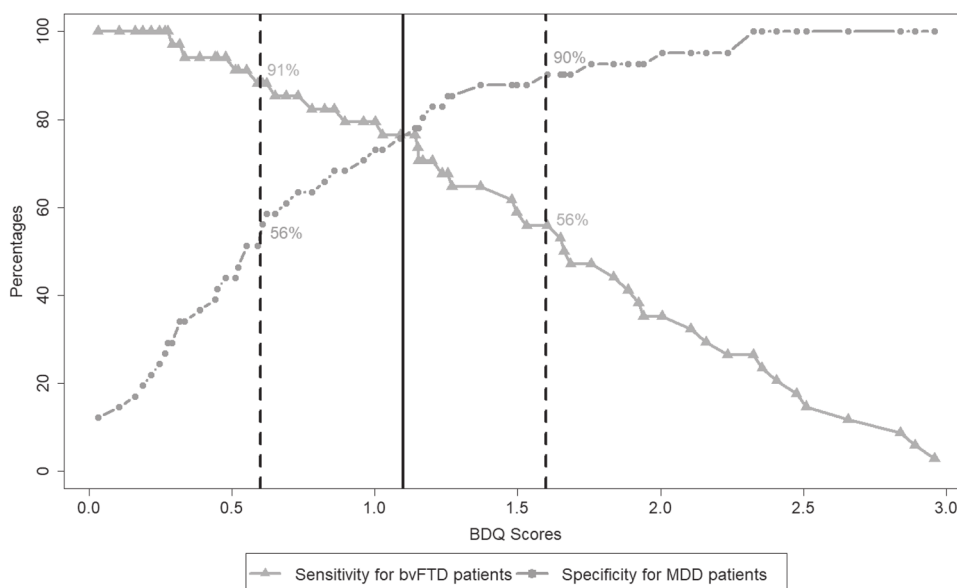


Fig. 2 Cut-offs between bvFTD and MDD patients. Percentages of patients with bvFTD who were correctly classified [sensitivity, orange line with triangles] and percentages of correctly classified MDD patients [specificity, blue line with circles] in relation to the BDQ-scores. Two cut-offs with either sensitivity or specificity above 90% are highlighted by dashed lines. The solid black line represents the optimal cut-off using the Youden-Index (sensitivity = 76%; specificity = 78%). *bvFTD* behavioural variant frontotemporal dementia, *MDD* major depressive disorder, *BDQ* Behavioural Dysfunction Questionnaire



Discussion

We developed an informant questionnaire based on the five behavioural domains of the revised diagnostic criteria for bvFTD [26]. This questionnaire, named BDQ, demonstrated excellent [17] discriminatory power between early-stage bvFTD and early-stage ADD patients (AUC = 88%) and between early-stage bvFTD and MDD patients (AUC = 83%), respectively.

Comparison of two different global scoring methods

We examined whether scoring by a global score (i.e. average score across all domains), by a global domain score (i.e. number of endorsed domains) or by both scores combined would yield different discriminatory powers. These different approaches showed similar results. Given that the global score incorporates all item scores and the global domain score incorporates the number of endorsed domains based on at least one item, these results suggest that items within a domain represent similar behaviours. Indeed, the internal consistencies of the five behavioural domains were acceptable to good apart from the domain “hyperorality and dietary changes” ($\alpha = 0.54$). In addition, given that an overall behavioural measure, independent of domain structure, as measured by the global score separates groups similarly to a behavioural pattern measure, as measured by the global domain score, suggests that each domain discriminated similarly well. Indeed, we found similar discriminatory power of the domain scores between bvFTD and non-bvFTD patients. Taken together, we favoured the global score over the global domain score for further use as it represents a more fine-grained assessment of the different behavioural domains.

By contrast, the global domain score is useful for describing the behavioural disorder pattern; information that the global score lacks.

Investigating whether to include a time criterion in the BDQ scoring

As required by the bvFTD diagnostic criteria [26], we included the time criterion “early” in the BDQ scoring. The time criterion “early” limits scoring to symptoms that appear within the first 3 years in four (i.e. disinhibition, apathy/inertia, loss of empathy and stereotypical behaviour) of the five behavioural domains. As suggested by the diagnostic criteria, we expected that inclusion of this time criterion would increase the discriminatory power between bvFTD and non-bvFTD patients. Inclusion of this strict time criterion, however, resulted in a weaker discriminatory power. Inclusion of a more lenient time criterion, namely limitation of scoring to symptoms that are present for less than 10 years (for removal of any potential personality-associated behavioural abnormalities), did not increase the discriminatory power. In light of these results, we wondered about the reliability of informants' time data on symptoms' onset and compared informants' time data at two different time points. This analysis revealed large data variability, ranging from 0.03 to 145.33 months. Our findings are consistent with a previous study which showed variance in patients' recollections of their past symptoms from one inquiry to the next [4]. These findings show the difficulty in perceiving and recalling the time of one's [4] or another person's symptom onset, likely even more so if the symptoms develop gradually and affect behaviour. Next to this, the collection of time data turned out to be time-consuming as informants often forgot

to state it. Accordingly, we decided to leave out time criteria in future BDQ scoring.

Comparison of the discriminatory power of the BDQ with existing instruments

The BDQ showed similar discriminatory power to the FBI [14]. At first look, this result seems surprising, since FBI-items represents primarily the Lund–Manchester criteria [11], which should be less sensitive for bvFTD than the current diagnostic criteria for bvFTD [26] upon which the BDQ is based. However, information on the specificity of the revised diagnostic criteria for bvFTD is lacking [26], which limits prediction on their diagnostic accuracy. Moreover, the FBI includes items like “loss of insight” or “personal neglect” that are typical for bvFTD [5, 28], but that are not (i.e. loss of insight), or not prominently (i.e. personal neglect), present in the diagnostic criteria for bvFTD. Nevertheless, the BDQ provides, in contrast to the FBI, not only a global score, but also scores of each behavioural domain, allowing quantitative representation of a patient's pattern of behavioural disturbances. In short, though BDQ and FBI discriminated patient groups similarly well, we consider the BDQ more useful for further use as it has an arranged structure of the behavioural domains and employs the current bvFTD diagnostic criteria [26]. Furthermore, the relatively large number of items lends itself to subsequent data-based weighting of items depending on the comparison group to bvFTD, which would improve the discriminatory power of the BDQ.

The second existing instrument that captures bvFTD-specific behavioural disorders is DAPHNE [6]. In one study, DAPHNE scores discriminated between bvFTD and ADD patients with AUC values between 95 and 99% [6]. These scores are higher than our AUC value (i.e. 88%). However, comparison in discriminatory power between these two instruments should be made with caution as they were used in different samples. Having said this, in the DAPHNE sample, the FBI discriminated between bvFTD and ADD patients similarly well to the DAPHNE [6]. This finding, in turn, corresponds to our finding, i.e. BDQ and FBI separated the two groups in our sample similarly. Accordingly, one may imagine that BDQ and DAPHNE discriminate similarly between bvFTD and ADD patients. The structure of the two instruments, however, differs in that DAPHNE is a semi-structured interview instrument that allows some interaction between informants and examiners, whereas BDQ is a self-administered informant questionnaire that can be completed in the absence of an examiner. Next, although the ten items of DAPHNE are based on the bvFTD diagnostic criteria [26], their compilation and structure is based on French experts' opinions in bvFTD [6]. Accordingly, unlike the BDQ, DAPHNE does not fully represent the structure

of the behavioural domains of the diagnostic criteria for bvFTD. It would be worthwhile to use both instruments in future studies in the same sample to investigate whether they discriminate differently between bvFTD and other patient groups.

Limitations

Despite the acceptable to excellent discriminatory power of the BDQ between bvFTD and the other two patient groups, a large gap of equivocal results between the two 0.9 sensitivity/specificity thresholds (0.6–1.4 for bvFTD vs. ADD and 0.6–1.6 for bvFTD vs. MDD, respectively) was present. However, it should be noted that the BDQ has only a supportive role in the diagnosis of bvFTD. It records the report of a significant other about a patient's bvFTD-specific behavioural features in daily life in a standardised way. For the final assessment of these behavioural features, the clinical impression of the examiner on the patient needs to be added. Of course, the assessment of bvFTD-specific behavioural features does not suffice to diagnose bvFTD. A comprehensive clinical assessment plus a brain MRI needs to take place, ideally complemented by further imaging techniques (e.g. FDG-PET, amyloid-PET or tau-PET) and/or laboratory tests (e.g. CSF biomarkers for AD or neurofilament light chain) to increase diagnostic certainty [7, 33].

Our study is limited by the absence of post-mortem pathological confirmations of our patients' diagnoses. Accordingly, we cannot exclude the possibility of patients' misdiagnoses. To increase diagnostic certainty, we confirmed the majority of patients' diagnoses by follow-up assessments (mean time period of 24 ± 11 months).

Conclusion

In summary, we demonstrated that the BDQ, the first instrument that operationalises the five behavioural domains of the diagnostic criteria for bvFTD [26], discriminates well between bvFTD and two non-bvFTD patient groups. Importantly, it allows a quantitative assessment of these domains that is independent of the examiner's expertise in bvFTD. This point is significant as knowledge and expertise in bvFTD is generally low outside of research institutions what likely contributes to bvFTD over- and underdiagnoses [21, 29, 36]. With its standardised approach, the BDQ would also be appropriate for assessing the severity of single and all bvFTD-specific behavioural features together. Similarly, it would be also appropriate for follow-up assessments. Last, but not least, the self-administrative format of the BDQ enables time saving behavioural disorder assessment, which is of increasing importance in clinical routine.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00415-023-11666-6>.

Acknowledgements The authors would like to thank Sabine Bruchmann, Annerose Engel, Sylvia Zimmerer, Jasmin Käch, Nils Kleine, Nina Huber, Nina Zach, Natalie Heger, Susanne Steffen, Fabienne Geiger, Lisa Rickenbacher, Jorinde Straub, Lisa Heid, Teresa Müggler, Nathalie Müller, Jonas Jeker, Valentina Memma, Aimée Winkenjo-hann, Luana Palermo and Valerie Jankowski for help in data collection.

Funding Open access funding provided by University of Basel. AS and MS are supported by the Alzheimer Forum Schweiz, Stiftung Empiris, Béatrice Ederer-Weber Stiftung and Freiwillige Akademische Gesellschaft (FAG). OP is supported in part by a National Health and Medical Research Council of Australia Leadership Fellowship (GNT2008020). MLS is supported by the German Research Foundation (DFG, SCHR 774/5-1) and the eHealthSax Initiative of the Sächsische Aufbaubank (SAB; project TelDem). Accordingly, this study is co-financed with tax revenue based on the budget approved by the Saxon state parliament.

Data availability The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflicts of interest The authors declare that they have no conflict of interest.

Ethical approval The study was approved by the Ethics Committee of Northwestern and Central Switzerland and has been performed in accordance with the ethical standards of the 1964 Declaration of Helsinki and its later amendments.

Consent to participate and publish All participants provided written informed consent.








Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abdi H (2007) The Kendall rank correlation coefficient. *Encyclopedia of Measurement and Statistics* Sage, Thousand Oaks, pp 508–510
- Alladi S, Xuereb J, Bak T, Nestor P, Knibb J, Patterson K, Hodges JR (2007) Focal cortical presentations of Alzheimer's disease. *Brain* 130:2636–2645
- Association AP (2013) *Diagnostic and statistical manual of mental disorders*, 5th edn. <https://doi.org/10.1176/appibooks9780890425596>
- Barsky AJ (2002) Forgetting, fabricating, and telescoping: the instability of the medical history. *Arch Intern Med* 162:981–984
- Bathgate D, Snowden JS, Varma A, Blackshaw A, Neary D (2001) Behaviour in frontotemporal dementia, Alzheimer's disease and vascular dementia. *Acta Neurol Scand* 103:367–378
- Boutoleau-Brettonnière C, Evrard C, Hardouin JB, Rocher L, Charriau T, Etcharry-Bouyx F, Auriacombe S, Richard-Mornas A, Lebert F, Pasquier F (2015) DAPHNE: a new tool for the assessment of the behavioral variant of frontotemporal dementia. *Dement Geriatr Cogn Disord Extra* 5:503–516
- Del Campo M, Zetterberg H, Gandy S, Onyike CU, Oliveira F, Udeh-Momoh C, Lleo A, Teunissen CE, Pijnenburg Y (2022) New developments of biofluid-based biomarkers for routine diagnosis and disease trajectories in frontotemporal dementia. *Alzheimer's Dement* 18:2292–2307
- DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44:837–845
- Ducharme S, Dols A, Laforce R, Devenney E, Kumfor F, van den Stock J, Dallaire-Théroux C, Seelaar H, Gossink F, Vijverberg E (2020) Recommendations to distinguish behavioural variant frontotemporal dementia from psychiatric disorders. *Brain* 143:1632–1650
- Ducharme S, Pearl-Dowler L, Gossink F, McCarthy J, Lai J, Dickerson BC, Chertkow H, Rapin L, Vijverberg E, Krudop W (2019) The frontotemporal dementia versus primary psychiatric disorder (FTD versus PPD) checklist: a bedside clinical tool to identify behavioral variant FTD in patients with late-onset behavioral changes. *J Alzheimers Dis* 67:113–124
- Englund B, Brun A, Gustafson L, Passant U, Mann D, Neary D, Snowden J (1994) Clinical and neuropathological criteria for frontotemporal dementia. *J Neurol Neurosurg Psychiatry* 57:416–418
- Forman MS, Farmer J, Johnson JK, Clark CM, Arnold SE, Coslett HB, Chatterjee A, Hurtig HI, Karlawish JH, Rosen HJ, Van Deerlin V, Lee VMY, Miller BL, Trojanowski JQ, Grossman M (2006) Frontotemporal dementia: clinicopathological correlations. *Ann Neurol* 59:952–962
- Johnen A, Bertoux M (2019) Psychological and cognitive markers of behavioral variant frontotemporal dementia—a clinical neuropsychologist's view on diagnostic criteria and beyond. *Front Neurol* 10:594
- Kertesz A, Davidson W, Fox H (1997) Frontal behavioral inventory: diagnostic criteria for frontal lobe dementia. *Can J Neurol Sci* 24:29–36
- Krudop WA, Dols A, Kerssens CJ, Eikelenboom P, Prins ND, Moller C, Schouws S, Rhebergen D, van Exel E, van der Flier WM, Sikkes S, Scheltens P, Stek ML, Pijnenburg YAL (2017) The pitfall of behavioral variant frontotemporal dementia mimics despite multidisciplinary application of the FTDC criteria. *J Alzheimers Dis* 60:959–975
- LaMarre AK, Rascovsky K, Bostrom A, Toofanian P, Wilkins S, Sharon JS, Perry DC, Miller ZA, Naasan G, Hagen J (2013) Interrater reliability of the new criteria for behavioral variant frontotemporal dementia. *Neurology* 80:1973–1977
- Mandrekar JN (2010) Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol* 5:1315–1316
- McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R, Mohs RC, Morris JC, Rossor MN, Scheltens P, Carrillo MC, Thies B, Weintraub S, Phelps CH (2011) The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement* 7:263–269

19. Mioshi E, Hsieh S, Savage S, Hornberger M, Hodges JR (2010) Clinical staging and disease progression in frontotemporal dementia. *Neurology* 74:1591–1597
20. Morris JC (1993) The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* 43:2412–2414
21. Musa G, Slachevsky A, Muñoz-Neira C, Méndez-Orellana C, Villagra R, González-Billault C, Ibáñez A, Hornberger M, Lillo P (2020) Alzheimer's disease or behavioral variant frontotemporal dementia? Review of key points toward an accurate clinical and neuropsychological diagnosis. *J Alzheimers Dis* 73:833–848
22. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H (2005) The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 53:695–699
23. Ossenkoppele R, Pijnenburg YA, Perry DC, Cohn-Sheehy BI, Scheltens NM, Vogel JW, Kramer JH, van der Vlies AE, La Joie R, Rosen HJ, van der Flier WM, Grinberg LT, Rozemuller AJ, Huang EJ, van Berckel BN, Miller BL, Barkhof F, Jagust WJ, Scheltens P, Seeley WW, Rabinovici GD (2015) The behavioural/dysexecutive variant of Alzheimer's disease: clinical, neuroimaging and pathological features. *Brain* 138:2732–2749
24. Piguet O, Kumfor F, Hodges J (2017) Diagnosing, monitoring and managing behavioural variant frontotemporal dementia. *Med J Aust* 207:303–308
25. Pose M, Cetkovich M, Gleichgerrcht E, Ibáñez A, Torralva T, Manes F (2013) The overlap of symptomatic dimensions between frontotemporal dementia and several psychiatric disorders that appear in late adulthood. *Int Rev Psychiatry* 25:159–167
26. Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, Van Swieten JC, Seelaar H, Dopper EG, Onyike CU (2011) Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain* 134:2456–2477
27. Seeley WW (2019) Behavioral variant frontotemporal dementia. *CONTINUUM: lifelong learning. Neurology* 25:76–100
28. Shany-Ur T, Lin N, Rosen HJ, Sollberger M, Miller BL, Rankin KP (2014) Self-awareness in neurodegenerative disease relies on neural structures mediating reward-driven attention. *Brain* 137:2368–2381
29. Shinagawa S, Catindig JA, Block NR, Miller BL, Rankin KP (2016) When a little knowledge can be dangerous: false-positive diagnosis of behavioral variant frontotemporal dementia among community clinicians. *Dement Geriatr Cogn Disord* 41:99–108
30. Thomann AE, Berres M, Goettel N, Steiner LA, Monsch AU (2020) Enhanced diagnostic accuracy for neurocognitive disorders: a revised cut-off approach for the Montreal Cognitive Assessment. *Alzheimers Res Ther* 12:1–10
31. Vieira RT, Caixeta L, Machado S, Silva AC, Nardi AE, Arias-Carrión O, Carta MG (2013) Epidemiology of early-onset dementia: a review of the literature. *Clin Pract Epidemiol Ment Health CP EMH* 9:88
32. Vijverberg EG, Gossink F, Krudop W, Sikkes S, Kerssens C, Prins N, Stek M, Scheltens P, Pijnenburg Y, Dols A (2017) The diagnostic challenge of the late-onset frontal lobe syndrome: clinical predictors for primary psychiatric disorders versus behavioral variant frontotemporal dementia. *J Clin Psychiatry* 78:e1197–e1203
33. Ward J, Ly M, Raji CA (2023) Brain PET imaging: frontotemporal dementia. *PET Clin* 18:123–133
34. WHO (2004) ICD-10: international statistical classification of diseases and related health problems: tenth revision, 2nd edn. World Health Organization. <https://apps.who.int/iris/handle/10665/42980?locale-attribute=en&>
35. Woollacott IO, Rohrer JD (2016) The clinical spectrum of sporadic and familial forms of frontotemporal dementia. *J Neurochem* 138(Suppl 1):6–31
36. Woolley JD, Khan BK, Murthy NK, Miller BL, Rankin KP (2011) The diagnostic challenge of psychiatric symptoms in neurodegenerative disease: rates of and risk factors for prior psychiatric diagnosis in patients with early neurodegenerative disease. *J Clin Psychiatry* 72:126–133

Authors and Affiliations

Anna Semenkova^{1,2}  · Olivier Piguet³  · Andreas Johnen⁴ · Matthias L. Schroeter^{5,6} · Jannis Godulla^{5,6} · Christoph Linnemann⁷ · Markus Mühlhauser⁷ · Thomas Sauer⁷  · Markus Baumgartner⁸ · Sarah Anderl-Straub⁹ · Markus Otto^{9,10} · Ansgar Felbecker¹¹  · Reto W. Kressig¹  · Manfred Berres¹²  · Marc Sollberger^{1,13} 

¹ Memory Clinic, University Department of Geriatric Medicine FELIX PLATTER, Basel, Switzerland

² Faculty of Psychology, University of Basel, Basel, Switzerland

³ School of Psychology and Brain and Mind Centre, The University of Sydney, Sydney, NSW, Australia

⁴ Clinic for Neurology, Münster University Hospital, Münster, Germany

⁵ Clinic for Cognitive Neurology, University Hospital Leipzig, Leipzig, Germany

⁶ Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

⁷ University Psychiatric Clinic, Basel, Switzerland

⁸ Memory Clinic Sonneweid, Wetzikon, Switzerland

⁹ Department of Neurology, University of Ulm, Ulm, Germany

¹⁰ Department of Neurology, University Hospital Halle, Halle, Germany

¹¹ Clinic of Neurology und Neurophysiology, Canton Hospital St. Gallen, St. Gallen, Switzerland

¹² Department of Mathematics and Technology, University of Applied Sciences Koblenz, Koblenz, Germany

¹³ Department of Neurology, University Hospital Basel, Basel, Switzerland

**STUDY II: Eine Lücke, die es zu schließen gilt: Die Untersuchung sozial-kognitiver
Fähigkeiten an deutschsprachigen Memory Kliniken**



Eine Lücke, die es zu schließen gilt: Die Untersuchung sozial-kognitiver Fähigkeiten an deutschsprachigen Memory-Kliniken

Marianne Jarsch^{1,2} , Anna Semenkova^{1,2} , Andreas U. Monsch^{1,2}, Reto W. Kressig¹ 
und Marc Sollberger^{1,3} 

¹ Memory Clinic, Universitäre Altersmedizin FELIX PLATTER und Universität Basel, Schweiz

² Fakultät für Psychologie, Universität Basel, Schweiz

³ Neurologische Klinik, Universitätsspital und Universität Basel, Schweiz

Zusammenfassung: Die Untersuchung sozial-kognitiver Fähigkeiten kann zur Diagnostik kognitiver Störungen beitragen. In diesem Zusammenhang stellt sich die Frage, inwieweit sozial-kognitive Fähigkeiten in deutschsprachigen Memory-Kliniken (MK) untersucht werden. Wir führten dazu eine Befragung in 87 deutschsprachigen MK in Deutschland, Österreich und der Schweiz durch. Es zeigte sich, dass die Untersuchung sozial-kognitiver Fähigkeiten mehrheitlich befürwortet wird, jedoch nur selten erfolgt. Die Mehrheit der MK könnte sich jedoch vorstellen, in Zukunft sozial-kognitive Diagnostik durchzuführen, vorausgesetzt, dass hierfür adäquate psychometrische Testverfahren vorliegen. Die Forderung der MK nach psychometrischen sozial-kognitiven Tests gewinnt an Bedeutung angesichts der Tatsache, dass die kürzlich publizierten ICD-11-Kriterien (International Statistical Classification of Diseases and Related Health Problems) die Untersuchung sozial-kognitiver Fähigkeiten bei der Abklärung einer kognitiven Störung empfehlen.

Schlüsselwörter: Emotionserkennung, Theory of Mind, soziale Kognition, neuropsychologische Untersuchung, neurokognitive Störung

A Gap that Needs to be Closed: The Assessment of Social-Cognitive Abilities in German-Speaking Memory Clinics

Abstract: The assessment of social-cognitive abilities can contribute to diagnosing cognitive disorders. This leads to the question of to what extent social-cognitive abilities are examined by German-speaking memory clinics (MCs). For this purpose, we surveyed 87 German-speaking MCs in Germany, Austria, and Switzerland and found that most MCs support the examination of social-cognitive abilities but do it only rarely. Most MCs could imagine conducting social-cognitive diagnostics in the future, if adequate psychometric test procedures are available. This demand by the MCs gains in importance in view of the fact that the recently published ICD-11 criteria recommend the examination of social-cognitive abilities in the assessment of a cognitive disorder.

Keywords: emotion recognition, Theory of Mind, social cognition, neuropsychological assessment, neurocognitive disorder

Soziale Kognition ist ein multidimensionales kognitives Konstrukt, welches uns befähigt, sozio-emotionale Stimuli wahrzunehmen und zu interpretieren, um unser Verhalten der jeweiligen sozialen Situation anpassen zu können (Arioli, Crespi & Canessa, 2018; Henry, von Hippel, Molenberghs, Lee & Sachdev, 2016; Pinkham et al., 2014; Sollberger, Rankin & Miller, 2010). Sozial-kognitive Fähigkeiten ermöglichen uns, adäquat mit anderen Personen zu interagieren (Kennedy & Adolphs, 2012), was von grundlegender Bedeutung ist und entscheidend zu unserer Lebensqualität beiträgt (Javed & Charles, 2018; Yagrajiah & Mula, 2019).

Sozial-kognitive Fähigkeiten sind bei den meisten kognitiven Störungsbildern beeinträchtigt (Cotter et al., 2018), sei es bei Schädel-Hirn-Traumata (Babbage et al., 2011), Entwicklungsstörungen (Fonagy & Goodyer, 2008), Temporallappenepilepsien (Bora & Meletti, 2016), neurodegenerativen (Christidi, Migliaccio, Santamaria-Garcia, Santangelo & Trojsi, 2018; Elamin, Pender, Hardiman & Abrahams, 2012; Maresca et al., 2020), vaskulären (Adams, Schweitzer, Molenberghs & Henry, 2019) oder autoimmun-entzündlichen (Cotter et al., 2016) Hirnkrankheiten oder bei psychiatrischen Krankheiten (Bora & Berk, 2016; Kohler, Walker, Martin, Healey & Moberg,

2010; Savla, Vella, Armstrong, Penn & Twamley, 2013). Entsprechend erscheint es sinnvoll, sozial-kognitive Fähigkeiten im Rahmen einer neuropsychologischen Testung zu untersuchen.

Abgesehen von der Häufigkeit des Auftretens sozial-kognitiver Störungen bei Hirnkrankheiten und Schädel-Hirn-Traumata, gilt es die Bedeutung der sozial-kognitiven Testung in der Ätiologiefindung kognitiver Störungen zu erwähnen. Die aktuell in der klinischen Routine verwendeten kognitiven Tests sind diesbezüglich nur begrenzt hilfreich. Dies liegt daran, dass diese Tests nicht alle Hirnregionen, respektive neuronalen Netzwerke, welche kognitive Prozesse repräsentieren, erfassen. Insbesondere werden neuronale Netzwerke wie das „Salience Network“ (Seeley et al., 2007) oder das „Semantic Appraisal Network“ (Yeo et al., 2011), welche zentrale Aspekte sozial-kognitiver Fähigkeiten repräsentieren (Christidi et al., 2018; Toller et al., 2018), mit in der klinischen Routine üblicherweise verwendeten Tests nicht erfasst. Diese Netzwerke wiederum sind jedoch bei bestimmten neurodegenerativen Krankheiten wie der Verhaltensvariante der frontotemporalen Demenz oder bei bestimmten psychiatrischen Krankheiten wie der Schizophrenie betroffen (Menon, 2011; Seeley, Crawford, Zhou, Miller & Greicius, 2009).

Im deutschsprachigen Raum führen Memory-Kliniken (MK) als interdisziplinäre Kompetenzzentren die Diagnostik von Personen mit Verdacht auf eine kognitive Störung durch (Diehl-Schmid, Lautenschlager & Kurz, 2011). In diesem Rahmen leistet die neuropsychologische Untersuchung, insbesondere bei klinisch komplexen Fällen, einen wesentlichen Beitrag zur Diagnosestellung (Deuschl & Maier, 2016). Die neuropsychologische Untersuchung beinhaltet die Testung mehrerer kognitiver Bereiche. Was für Bereiche und wie viele Bereiche dabei untersucht werden sollen, variiert je nach Klassifikationssystem. Das im deutschsprachigen Raum wahrscheinlich am häufigsten verwendete Klassifikationssystem ist die zehnte Revision der Internationalen Klassifikation der Krankheiten (ICD-10; World Health Organization, 2015). Gemäß ICD-10 sind „Gedächtnis, Denken, Orientierung, Auffassung, Rechnen, Lernfähigkeit, Sprache und Urteilsvermögen“ (World Health Organization, 2015), nicht jedoch soziale Kognition, zu testen. Demgegenüber empfehlen die kürzlich publizierten ICD-11-Kriterien (World Health Organization, 2022) die Testung der sozialen Kognition im Rahmen einer neuropsychologischen Untersuchung. Die fünfte Auflage des Diagnostischen und Statistischen Manuals Psychischer Störungen (DSM-5; American Psychiatric Association, 2013), ein im deutschsprachigen Raum seltener verwendetes Klassifikationssystem, setzt ebenfalls die Testung der sozialen Kognition voraus.

Es stellt sich nun die Frage, wie die Empfehlungen der ICD-11- und DSM-5-Kriterien im klinischen Alltag umge-

setzt werden können. Dazu ist zunächst zu klären, wie soziale Kognition getestet werden soll. Angesichts des multidimensionalen Konstrukts und den unterschiedlichen Definitionen der sozialen Kognition (American Psychiatric Association, 2013; Arioli et al., 2018; Henry et al., 2016; Pinkham et al., 2014; Sachdev et al., 2014; Sollberger et al., 2010) ist dies jedoch schwierig. Beispielsweise definiert DSM-5 soziale Kognition als eine kognitive Domäne, welche sich aus den Bereichen Emotionserkennung (EE) und Theory of Mind (ToM) zusammensetzt (American Psychiatric Association, 2013). Während EE die Fähigkeit beschreibt, die Emotionen anderer wahrzunehmen und zu benennen, beinhaltet ToM die Fähigkeit, Gedanken, Absichten und Emotionen anderer zu verstehen und diese vom eigenen mentalen und/oder affektiven Zustand abzugrenzen (Henry et al., 2016). Andere Autor_innen erwähnen zusätzliche Bereiche wie beispielsweise Selbstwahrnehmung (Sachdev et al., 2014), soziales Verhalten (Henry et al., 2016) oder soziale Wahrnehmung, Attributionsstil und gezielter Umgang mit wahrgenommenen Emotionen (Pinkham et al., 2014). Den verschiedenen Definitionen ist jedoch die Erwähnung der Bereiche EE und ToM gemeinsam. Dazu passend finden sich primär Testresultate zu EE und ToM in Studien zu sozialer Kognition bei Patient_innen mit Hirnkrankheiten und/oder Hirntraumata (Cotter et al., 2018; Henry, Cowan, Lee & Sachdev, 2015). Unter Berücksichtigung dieser Punkte favorisieren wir die Testung der Bereiche EE und ToM zur Abklärung der sozialen Kognition in der klinischen Routine.

Was die Testung der sozialen Kognition im deutschsprachigen Raum aktuell erschwert, ist der Mangel an hierfür adäquaten Testverfahren. Im Gegensatz zu den etablierten kognitiven Bereichen existieren im Bereich der sozialen Kognition, abgesehen von Tests wie dem „Geneva Emotion Recognition Test“ (Schlegel, Grandjean & Scherer, 2014), dem „Reading the Mind in the Eyes Test“ (Kynast et al., 2020) oder dem „Movie for the Assessment of Social Cognition“ (Dziobek et al., 2006), keine normierten und/oder validierten deutschsprachigen Tests. In Bezug auf den „Movie for the Assessment of Social Cognition“ ist zu ergänzen, dass dieser Test primär für Patient_innen mit Autismus-Spektrum-Störungen, spricht für jüngere Patient_innen mit verhältnismäßig leichten kognitiven Störungen, entwickelt worden ist (Dziobek et al., 2006). Entsprechend ist davon auszugehen, dass der Test nur begrenzt für Patient_innen höheren Alters und/oder Patient_innen mit neuropsychologischen Störungen in mehreren kognitiven Bereichen geeignet ist.

Im Rahmen dieser Studie wollten wir erfahren, i) wie deutschsprachige MK zur Testung sozial-kognitiver Fähigkeiten bei der Abklärung einer kognitiven Störung stehen,

ii) ob, und falls ja, in welcher Form sie sozial-kognitive Fähigkeiten testen und iii) was ihre Vorstellungen hinsichtlich der Testung sozial-kognitiver Fähigkeiten in der näheren Zukunft sind. Zu diesem Zweck entwickelten wir einen explorativen Onlinefragebogen und kontaktierten MK in Deutschland, Österreich und der deutschsprachigen Schweiz.

Methoden

Studienteilnehmer_innen

Wir kontaktierten 198 deutschsprachige MK (Deutschland: $n = 155$, Österreich: $n = 15$, Schweiz: $n = 28$), die im Januar 2021 in Deutschland, Österreich und der Schweiz offiziell als deutschsprachige MK ausgewiesen waren. Die jeweiligen Leitungspersonen erhielten per E-Mail einen Weblink, durch den sie auf den Onlinefragebogen weitergeleitet wurden. Falls die Leitungspersonen keine Aussagen zur sozial-kognitiven Diagnostik treffen konnten, bestand die Möglichkeit, den Fragebogen intern an die entsprechenden Fachpersonen weiterzuleiten. Um sicher zu sein, dass die Fragen von Fachpersonen beantwortet werden, erhoben wir die Berufsgruppenzugehörigkeit der Fragebogenteilnehmenden. Weitere Informationen über die Auswahl- und Teilnahmebedingungen der Studienteilnehmenden sind im elektronischen Supplement (ESM) 1 einsehbar. In Absprache mit der Ethikkommission der Nordwest- und Zentralschweiz bedurfte es zur Durchführung der Umfrage keiner Ethikbewilligung.

Expert_innenfragebogen

Der explorative Onlinefragebogen wurde mit der webbasierten Umfragesoftware LimeSurvey Version 3.25.8+210118 (LimeSurvey Project Team, 2021) erstellt.

Format der Fragen

Der Fragebogen beinhaltete die Antwortformate Einfachauswahl, Mehrfachauswahl oder offene Antworten. Die Teilnehmenden konnten zu bereits beantworteten Fragen zurückkehren, um ihre Antworten abzuändern. Es war ihnen jedoch nicht möglich, Fragen zu überspringen, mit Ausnahme der Fragen mit offenem Antwortformat. Zusätzlich konnten sie den Bearbeitungsstand des Fragebogens zwischenspeichern, um zu einem späteren Zeitpunkt fortzufahren.

Bereiche des Fragebogens

Diagnostische Relevanz verschiedener Bereiche der sozialen Kognition

Die MK wurden gebeten, die Relevanz von fünf Bereichen der sozialen Kognition (EE, ToM, gezielter Umgang mit wahrgenommenen Emotionen, soziale Wahrnehmung, Attributionsstil), welche im Rahmen der „Social Cognition Psychometric Evaluation (SCOPE) study“ (Pinkham et al., 2014) definiert worden waren, auf einer vierstufigen Skala („nicht relevant“ bis „relevant“) zu beurteilen. Zusätzlich bestand die Möglichkeit, in einem Freitextfeld weitere relevante Bereiche der sozialen Kognition zu nennen.

Relevanz sozial-kognitiver Untersuchungen in Abhängigkeit bestimmter Hirnkrankheiten oder Hirntraumata

Die Institutionen wurden gebeten, anzugeben, wie relevant sie sozial-kognitive Diagnostik bei bestimmten Hirnkrankheiten, respektive bei Schädel-Hirn-Traumata, erachten. Folgende Krankheiten standen zur Auswahl: neurodegenerative, vaskuläre und neuroinflammatorische Hirnkrankheiten, Entwicklungsstörungen und psychiatrische Krankheiten. Im Freitextfeld konnten zusätzliche Krankheitsbilder, welche hinsichtlich sozial-kognitiver Diagnostik als relevant betrachtet wurden, genannt werden.

Durchführung von Tests zur Emotionserkennung und/oder Theory of Mind

Die MK wurden gefragt, ob sie EE- und/oder ToM-Tests durchführen und, falls ja, gebeten, die Namen der Tests zu nennen. Diejenigen MK, welche sozial-kognitive Tests bejaht hatten, wurden gefragt, wie zufrieden sie mit den Tests sind. Dazu stand ihnen eine fünfstufige Skala („unzufrieden“ bis „zufrieden“) zur Verfügung. In einem Freitextfeld konnten Vor- und Nachteile der Tests genannt werden.

Wir beschränkten uns auf Tests aus den Bereichen EE und ToM, da wir diese Bereiche, wie eingangs erwähnt, als Kernbereiche der sozialen Kognition betrachten.

Mögliche Gründe gegen die Durchführung von Tests zur Emotionserkennung und/oder Theory of Mind

Diejenigen MK, welche die Durchführung sozial-kognitiver Tests verneint hatten, wurden nach den Gründen hierfür gefragt. Dafür gab es die folgenden fünf Auswahlmöglichkeiten mit der Möglichkeit von Mehrfachantworten: (1) Tests sind für Patient_innen mit Störungen in mehreren kognitiven Domänen ungeeignet, (2) Durchführung dauert zu lange, (3) zu niedrige Testsensitivität, (4) keine Therapierrelevanz der Testergebnisse und (5) Durchführung wurde bis dato nicht in Erwägung gezogen. Weitere Gründe konnten in einem Freitextfeld genannt werden.

Qualitative Beurteilung sozial-kognitiver Fähigkeiten

Institutionen, welche die Durchführung von sozial-kognitiven Tests verneint hatten, wurden gefragt, ob sie alternative Erfassungsmethoden wie Befragungen von Patient_innen und/oder Angehörigen durchführen. Weitere Erfassungsmethoden konnten in einem Freitextfeld genannt werden.

Bevorzugte Testdauer zur Messung sozial-kognitiver Fähigkeiten für Screening und ausführliche Testung

Wir erfragten, welche maximale Zeitdauer die MK für das Screening, respektive für eine ausführliche Testung, sozial-kognitiver Fähigkeiten als vertretbar erachten. Dafür stand den MK eine siebenstufige Skala („max. 10 Min.“, „max. 20 Min.“, „max. 30 Min.“, „max. 40 Min.“, „max. 50 Min.“, „max. 1h“, „über 1h“) zur Auswahl.

Wahrscheinlichkeit des zukünftigen Gebrauchs von in der klinischen Routine praktikablen sozial-kognitiven Tests mit guten psychometrischen Eigenschaften

Wir erfragten mittels einer vierstufigen Skala („unwahrscheinlich“ bis „wahrscheinlich“) die Wahrscheinlichkeit des zukünftigen Gebrauchs von in der klinischen Routine praktikablen, validen und verlässlichen sozial-kognitiven Tests.

Analysen

Die Datenanalyse wurde mit IBM SPSS Statistics für Macintosh, Version 27.0.1.0 (IBM Corp., 2007) durchgeführt. Die Daten wurden mittels Häufigkeitsanalysen ausgewertet.

Ergebnisse

Anzahl teilnehmender Memory-Kliniken und Rücklaufquote des Fragebogens

Von den insgesamt 198 angeschriebenen deutschen ($n = 155$), schweizerischen ($n = 28$) und österreichischen ($n = 15$) MK nahm knapp die Hälfte (43.9%; $n = 87$) an der Befragung teil. Die Rücklaufquote lag bei den deutschen MK bei 41.9%, bei den schweizerischen MK bei 67.9% und bei den österreichischen MK bei 20.0%. Die teilnehmenden MK teilten sich, nach Ländern, wie folgt auf: Deutschland (74.7%; $n = 65$), Schweiz (21.8%; $n = 19$) und Österreich (3.4%; $n = 3$). Die überwiegende Mehrheit (94.3%; $n = 82$) der MK füllte den Fragebogen vollständig aus. 5.7% ($n = 5$) beantworteten ihn teilweise.

Anzahl monatlicher Abklärungen pro Memory-Klinik und Berufsgruppen, welche den Fragebogen ausfüllten

Bezüglich der Anzahl monatlicher Abklärungen teilten sich die MK in absteigender Häufigkeit wie folgt auf: 5–20 Patient_innen (36.6%), 20–40 Patient_innen (25.6%), > 60 Patient_innen (13.4%), 40–60 Patient_innen (13.4%), 1–4 Patient_innen (11.0%).

Die Mehrheit (63.6%) der Personen, welche den Fragebogen ausgefüllt hatten, waren Neuropsycholog_innen, Psycholog_innen oder Psychotherapeut_innen. Die anderen Berufsgruppen, waren Psychiater_innen (17.8%), Neurolog_innen (12.7%), Ärzte/Ärztinnen in Weiterbildung (2.5%) oder andere Berufsgruppen (3.4%).

Diagnostische Relevanz verschiedener Bereiche der sozialen Kognition

Die Mehrheit der MK bewertete die diagnostische Relevanz der fünf zur Auswahl gestellten Bereiche der sozialen Kognition als „relevant“ oder „eher relevant“ (EE: 86.2%; $n = 75$, ToM: 83.9%; $n = 73$, soziale Wahrnehmung: 95.4%; $n = 83$, gezielter Umgang mit wahrgenommenen Emotionen: 77.0%; $n = 67$, Attributionsstil: 73.6%; $n = 64$; Abbildung 1).

Als zusätzliche Bereiche wurden soziale Verhaltensweisen (Sozialverhalten, psychosoziales Funktionsniveau, Compliance bei Unterstützungsangeboten, Delinquenz: 11.5%; $n = 10$) und Empathiefähigkeit (4.6%; $n = 4$) genannt.

Relevanz sozial-kognitiver Diagnostik in Abhängigkeit bestimmter Hirnkrankheiten oder Hirntraumata

Die klare Mehrheit der MK (92.7%) befürwortete sozial-kognitive Testungen bei neurodegenerativen Krankheiten. In absteigender Häufigkeit befürworteten MK sozial-kognitive Testungen bei anderen Hirnkrankheiten, respektive bei Hirntraumata, wie folgt: 72.0% bei vaskulären Hirnkrankheiten, 67.1% bei psychiatrischen Erkrankungen, 58.5% bei Schädel-Hirn-Traumata, 47.6% bei Entwicklungsstörungen und 33.0% bei neuroinflammatorischen Krankheiten. Im Freitextfeld wurden keine zusätzlichen Krankheiten genannt.

Untersuchung der Emotionserkennung

Die Minderheit (13.9%; $n = 12$) der 86 MK, welche die Frage zur EE-Untersuchung beantwortet hatten, bejahte die Durchführung von EE-Tests (Abbildung 2a). Die Hälfte da-

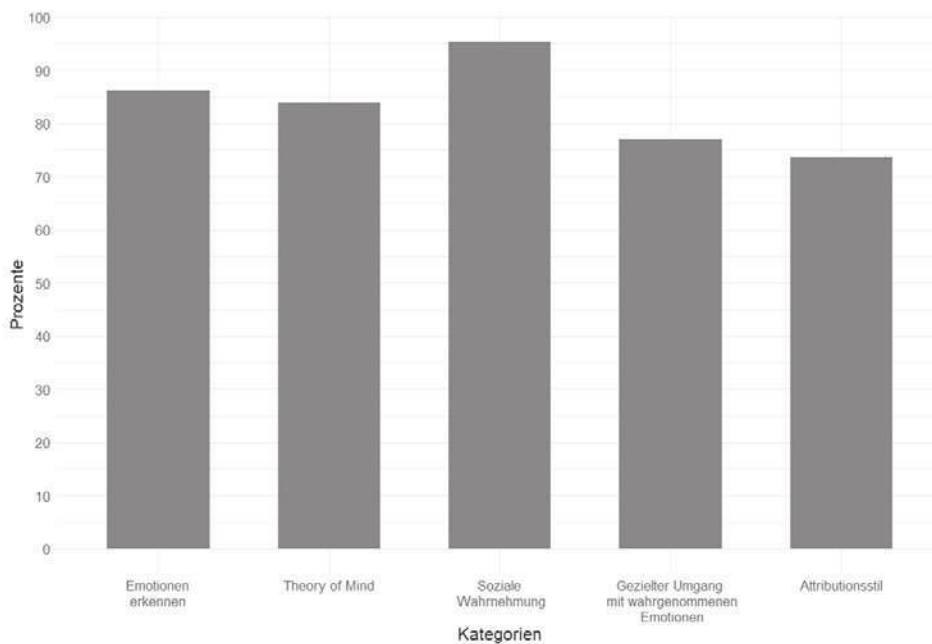


Abbildung 1. Anteile der Institutionen in Prozent, welche die von der „Social Cognition Psychometric Evaluation Study“ definierten Bereiche der sozialen Kognition als „relevant“ oder „eher relevant“ beurteilten.

von gab an, dazu statische Stimuli zu verwenden. Die andere Hälfte machte diesbezüglich keine Angaben. Fünf (83.3%) der MK, welche statische Stimuli bejaht hatten, nannten Ekman Faces (Ekman & Friesen, 1976) als Stimulusmaterial. Die sechste dieser MK machte dazu keine Angaben. Ein Viertel ($n = 3$) der MK, welche die Durchführung von EE-Tests bejaht hatten, gab an, die Tests nur in wissenschaftlichem Rahmen durchzuführen.

Was die Zufriedenheit mit den eingesetzten EE-Tests betraf, waren 58.3% ($n = 7$) „zufrieden“ oder „eher zufrieden“, 8.3% ($n = 1$) „neutral“ und 33.3% ($n = 4$) „eher unzufrieden“ oder „unzufrieden“. Befragt nach den Vor- und Nachteilen dieser Tests, wurden primär Nachteile genannt. Als nachteilig wurden fehlende Normierung (41.7%; $n = 5$), hoher Zeitaufwand (25.0%; $n = 3$), zu geringe Realitätsnähe der Tests (25.0%; $n = 3$), fehlende Validierung für spezifische Patientengruppen (8.3%; $n = 1$), fragwürdige Validität (8.3%; $n = 1$), zu hohe sprachliche Komplexität (8.3%; $n = 1$) und fehlendes Scoring (8.3%; $n = 1$) genannt. Als Vorteile wurden die Eignung zum Screening von EE-Störungen (16.7%; $n = 2$) und zur EE-Einschätzung (8.3%; $n = 1$), sowie die akzeptable Zeitdauer der Stimuli-Administration (8.3%; $n = 1$) genannt.

Die Mehrheit (86.1%; $n = 74$) der MK verneinte die Durchführung von EE-Tests. Zu den Gründen machten sechs MK keine Angaben. Von den anderen MK gaben 63.2% ($n = 43$) an, die Durchführung von EE-Tests bisher nicht erwogen zu haben und 36.8% ($n = 25$) hatten sich aufgrund folgender Gründe dagegen entschieden: Durch-

führung dauert zu lange (52.0%; $n = 13$), verfügbare EE-Tests sind für Patient_innen mit Störungen in mehreren kognitiven Domänen ungeeignet (44.0%; $n = 11$) und zu niedrige Testsensitivität (44.0%; $n = 11$). Weitere Gründe, welche jeweils nur von einer Institution genannt wurden, sind im ESM 2 aufgeführt.

Von den MK, welche die Durchführung von EE-Tests verneint hatten, bejahten 62.2% ($n = 46$) qualitative Erfassungsmethoden. Davon gaben 80.4% ($n = 37$) an, Angehörige und Patient_innen zu befragen; 10.9% ($n = 5$) gaben an, nur Angehörige zu befragen; 4.3% ($n = 2$) gaben an, nur Patient_innen zu befragen und 2.2% ($n = 1$) gaben an, EE auf der Basis der klinischen Beobachtung zu beurteilen. Eine Institution (2.2%) machte dazu keine Angaben.

Untersuchung von Theory of Mind

Die Minderheit (15.5%; $n = 13$) der 84 MK, welche die Frage zur ToM-Untersuchung beantwortet hatten, bejahte die Durchführung von ToM-Tests (Abbildung 2b). Eine dieser MK äußerte sich dabei nicht zum Stimulusmaterial. Die anderen MK gaben an, statische Stimuli in Form von Texten und/oder Bildern, respektive Bildergeschichten, einzusetzen. Folgende Tests wurden genannt: „Faux-Pas-Geschichten“ (nicht näher bezeichnet; 41.7%; $n = 5$), „Reading the Mind in the Eyes Test“ (Baron-Cohen, Wheelwright, Hill, Raste & Plumb, 2001; 25.0%; $n = 3$), „Sally & Anne Test“ (Baron-Cohen, Leslie & Frith, 1985;

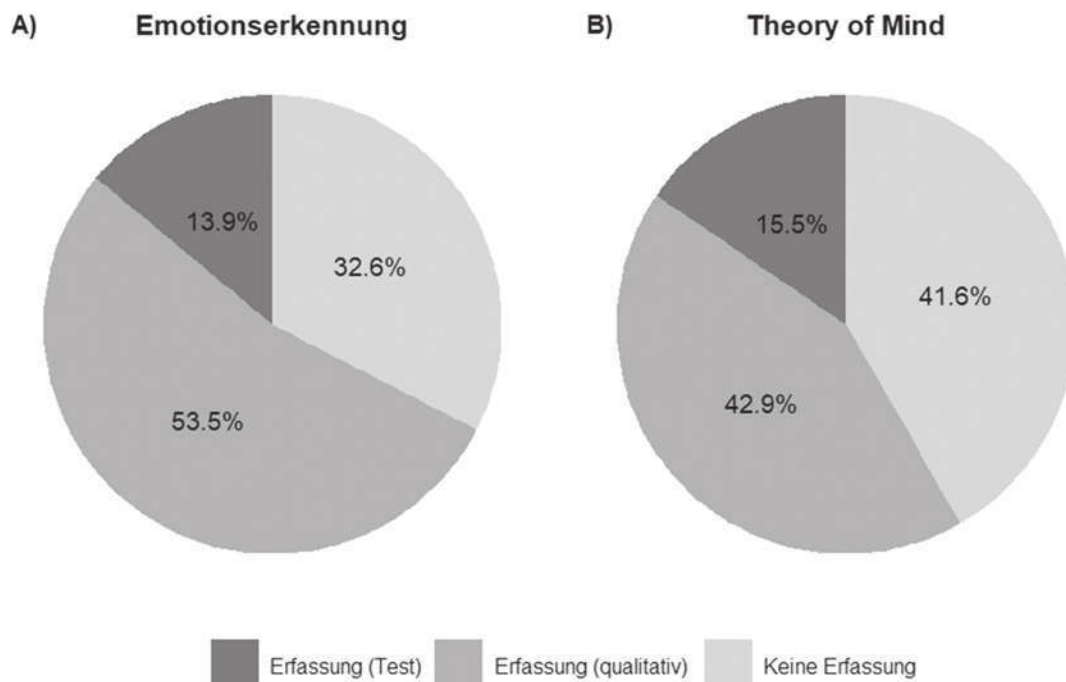


Abbildung 2. Anteile der Institutionen in Prozent, welche A) Emotionserkennung bzw. B) Theory of Mind entweder mittels Tests oder qualitativ erfassen oder welche A) Emotionserkennung bzw. B) Theory of Mind nicht erfassen.

25.0%; $n = 3$), „Geneva Social Cognition Scale“ (Martory et al., 2015; 8.3%; $n = 1$), „ToM Cartoons von Brüne“ (Brüne, 2005; 8.3%; $n = 1$) und Bildergeschichten (nicht näher bezeichnet; 8.3%; $n = 1$). Zwei (15.4%) der MK, welche die Durchführung von ToM-Tests bejaht hatten, gaben an, die Tests nur in wissenschaftlichem Rahmen einzusetzen.

Was die Zufriedenheit mit den jeweiligen ToM-Tests betraf, waren 30.8% ($n = 4$) „zufrieden“ oder „eher zufrieden“, 23.1% ($n = 3$) „neutral“ und 46.2% ($n = 6$) „unzufrieden“ oder „eher unzufrieden“. Wie bei den EE-Tests überwogen auch bei den ToM-Tests die Nachteile gegenüber den Vorteilen. Als nachteilig wurden fehlende Validierung (46.2%; $n = 6$), hoher Zeitaufwand (30.8%; $n = 4$), ungenaue Angaben zur Auswertung (15.4%; $n = 2$) und erschwerte Durchführung bei Patient_innen mit Sprach- und Leseschwierigkeiten (15.4%; $n = 2$) genannt. Nur eine (8.3%) MK nannte einen Vorteil, nämlich die Vielfältigkeit und Durchführungsdauer der Geneva Social Cognition Scale (Martory et al., 2015).

Die Mehrheit (84.5%; $n = 71$) der MK verneinte die Durchführung von ToM-Tests. Zu den Gründen machte eine MK keine Angaben. Von den anderen MK gaben 52.9% ($n = 37$) an, deren Durchführung bisher nicht erwogen zu haben, und 47.1% ($n = 33$) hatten sich aufgrund folgender Gründe dagegen entschieden: Durchführung dauert zu lange (60.6%; $n = 20$), Tests sind für Patient_innen mit Störungen in mehreren kognitiven Domänen ungeeignet (39.4%; $n = 13$), zu niedrige Testsensitivität (27.3%;

$n = 9$), Testergebnisse sind irrelevant für Therapie (15.2%; $n = 5$) und Stimuli sind eher für jüngere als für ältere Patient_innen geeignet (3.0%; $n = 1$).

Von den MK, welche die Durchführung von ToM-Tests verneint hatten, bejahten 50.7% ($n = 36$) qualitative Erfassungsmethoden. Davon gaben 77.8% ($n = 28$) an, Angehörige und Patient_innen zu befragen; 16.7% ($n = 6$) gaben an, nur Angehörige zu befragen, und je eine Institution (2.8%) gab an, nur Patient_innen zu befragen, respektive ToM auf der Basis klinischer Beobachtung zu beurteilen.

Bevorzugte Durchführungsdauer zum Screening sozial-kognitiver Störungen

Die Mehrheit (86.7%; $n = 72$) der MK favorisierte max. 10 Minuten zum Screening sozial-kognitiver Störungen. Eine Minderheit (12.0%; $n = 10$) favorisierte max. 20 Minuten und eine MK (1.2%) favorisierte max. 30 Minuten.

Bevorzugte Durchführungsdauer zur ausführlichen Testung sozial-kognitiver Fähigkeiten

Ein gutes Drittel der MK (38.6%; $n = 32$) favorisierte eine Durchführungsdauer von max. 20 Minuten zur ausführlichen Testung sozial-kognitiver Fähigkeiten. Annähernd

ein Drittel (30.1%; $n = 25$) favorisierte max. 30 Minuten. Die anderen Optionen zur Durchführungsdauer wurden nur selten gewählt: max. 10 Minuten (10.8%; $n = 9$), max. 1 Stunde (10.8%; $n = 9$), max. 40 Minuten (7.2%; $n = 6$), max. 50 Minuten (1.2%; $n = 1$), und > 1 Stunde (1.2%; $n = 1$).

Wahrscheinlichkeit des zukünftigen Gebrauchs von in der klinischen Routine praktikablen sozial-kognitiven Tests mit guten psychometrischen Eigenschaften

Die Mehrheit der MK (89.0%; $n = 73$) schätzte es als „wahrscheinlich“ oder „eher wahrscheinlich“ ein, zukünftig in der klinischen Routine praktikable, sozial-kognitive Testverfahren mit guten psychometrischen Eigenschaften einzusetzen. Nur 11.0% ($n = 9$) der MK schätzten die Wahrscheinlichkeit hierfür als „unwahrscheinlich“ oder „eher unwahrscheinlich“ ein.

Diskussion

Unsere Studie zeigt, dass deutschsprachige MK die Testung sozial-kognitiver Fähigkeiten mehrheitlich als wichtig erachten. Jedoch führt aktuell nur eine Minderheit (knapp 15%) sozial-kognitive Tests durch. Die Zufriedenheit mit diesen Tests ist dabei begrenzt. Eine klare Mehrheit der befragten MK könnte sich jedoch vorstellen, in Zukunft soziale Kognition zu testen, sofern hierfür Tests vorliegen, welche den Qualitätsstandards neuropsychologischer Tests genügen.

Unsere Befragung ergab, dass ein wahrscheinlich wesentlicher Grund für die aktuell kaum stattfindende Testung der sozialen Kognition in der fehlenden Verfügbarkeit hierfür adäquater Tests liegt. Ein dabei oft genannter Kritikpunkt war die zu lange Testdauer. Weitere Kritikpunkte, respektive Gründe für den Verzicht auf die Testung, waren die fehlende Normierung und/oder Validierung der Tests, der fehlende Realitätsbezug der Tests, die zu niedrige Testsensitivität und die fehlende Testeignung für Patient_innen mit Störungen in mehreren kognitiven Bereichen. Die erwähnten Punkte weisen darauf hin, dass die verfügbaren sozial-kognitiven Tests nicht den Qualitätsstandards neuropsychologischer Tests entsprechen. Die von den MK eingesetzten Tests wie „Ekman Faces“ (Ekman & Friesen, 1976), nicht näher bezeichnete „Faux-Pas-Tests“, „Reading the Mind in the Eyes Test“ (Baron-Cohen et al., 2001) oder der „Sally & Anne Test“ (Baron-Cohen et al., 1985) sind zwar in der Forschung etablierte Tests (Cotter et al., 2018; Henry et al., 2015),

jedoch sind deren deutschsprachige Versionen nur teilweise validiert und/oder normiert (Kynast et al., 2020).

Die aktuell kaum stattfindende Testung der sozialen Kognition in deutschsprachigen MK steht sicher auch in Zusammenhang mit ICD-10 (World Health Organization, 2015), dem im deutschsprachigen Raum mehrheitlich verwendeten Klassifikationssystem für kognitive Störungen. Gemäß ICD-10 ist soziale Kognition bei der Abklärung einer kognitiven Störung nicht zu testen (World Health Organization, 2015). Die 2013 publizierten DSM-5-Kriterien und die kürzlich publizierten ICD-11-Kriterien setzen demgegenüber die Testung der sozialen Kognition voraus (American Psychiatric Association, 2013; World Health Organization, 2022). Entsprechend wichtig und dringlich ist die Entwicklung von Tests, welche die Qualitätsstandards neuropsychologischer Tests erfüllen und den Anforderungen der klinischen Routine genügen.

In diesem Zusammenhang ist eine in Entwicklung befindliche deutschsprachige sozial-kognitive Testbatterie zu erwähnen. Diese Testbatterie (Basel Version of the Awareness of Social Inference Test [BASIT]) setzt sich aus einem EE-Test (BASIT-Emotion Recognition [BASIT-ER]; Jarsch et al., 2022) und einem ToM-Test (BASIT-Theory of Mind [BASIT-ToM]; Jarsch et al., in Revision) zusammen. Beide Tests beinhalten kurze, realitätsnahe Filmszenen, in welchen Personen miteinander interagieren. Im Rahmen einer präliminären Validierungsstudie bei 240 kognitiv und psychisch gesunden Personen unterschiedlichen Alters wurden Szenen ohne Boden- und Deckeneffekte für EE, respektive für ToM, ausgewählt (Jarsch et al., 2022; Jarsch et al., in Revision). Die Validierung der BASIT-Testbatterie sollte in nächster Zeit beginnen.

Die Mehrheit der befragten MK erachtete alle fünf von der „Social Cognition Psychometric Evaluation study“ (Pinkham et al., 2014) genannten Bereiche der sozialen Kognition als mindestens „eher relevant“. Die Mehrheit der MK zeigte sich auch motiviert, in Zukunft soziale Kognition zu untersuchen, sofern hierfür psychometrische Tests vorliegen. Zur ausführlichen Testung sozial-kognitiver Funktionen wäre je ein Drittel der MK bereit, max. 20 Minuten, respektive max. 30 Minuten, einzusetzen. In diesen genannten Zeitspannen wäre es jedoch unmöglich, sämtliche oben gewählten Bereiche der sozialen Kognition zu testen. Es wäre dann Aufgabe der MK, festzulegen – in Abhängigkeit des klinischen Bildes und der Fragestellung –, welche Bereiche der sozialen Kognition getestet werden sollen. Alternativ könnte sich die Testung auf die beiden Kernbereiche der sozialen Kognition, EE und ToM, beschränken. Wie in der Einführung erwähnt, favorisieren wir dieses Vorgehen.

Mehr als die Hälfte der MK, welche soziale Kognition nicht testen, evaluieren soziale Kognition mittels Befragung von Angehörigen und/oder Patient_innen. Zählt

man die Zahl der MK, welche soziale Kognition in Form von Tests und/oder Befragungen abklären, zusammen, ergibt sich daraus die Mehrheit der an der Studie teilgenommenen MK. Dieses Resultat sollte jedoch nicht generalisiert werden, da unsere Daten nur auf knapp der Hälfte (43.9%) der angeschriebenen deutschsprachigen MK beruhen. Es ist anzunehmen, dass diejenigen MK, welche den Fragebogen nicht ausgefüllt haben, an der Evaluation der sozialen Kognition kaum bis nicht interessiert sind. Es ist deshalb eher davon auszugehen, dass die meisten deutschsprachigen MK sozial-kognitive Fähigkeiten aktuell nicht abklären.

Bezüglich unserer Resultate gilt es einschränkend zu erwähnen, dass sie primär (75%) auf den Angaben von deutschen MK beruhen. Während MK der deutschsprachigen Schweiz noch zu 22% der Daten beitrugen, waren es bei den österreichischen MK nur 3%. Der sehr geringe Anteil der österreichischen MK am Datensatz ist in erster Linie durch deren sehr niedrige Rücklaufquote (20%) begründet. Die Gründe hierfür kennen wir leider nicht.

Zusammenfassend lässt sich festhalten, dass die Testung sozial-kognitiver Fähigkeiten noch kein fester Bestandteil der neuropsychologischen Diagnostik in deutschsprachigen MK darstellt. Dieser Umstand ist wahrscheinlich primär Ausdruck der Vorgabe des ICD-10-Klassifikationssystems (World Health Organization, 2015), welches die Testung der sozialen Kognition nicht voraussetzt. Die kürzlich publizierten ICD-11-Kriterien, respektive die bereits seit mehreren Jahren publizierten DSM-5-Kriterien, setzen demgegenüber die Testung der sozialen Kognition voraus (American Psychiatric Association, 2013; World Health Organization, 2022). Die Umsetzung der Vorgaben dieser beiden Klassifikationssysteme bedingt die Verfügbarkeit von deutschsprachigen Tests, welche den Qualitätsstandards neuropsychologischer Tests entsprechen und den Anforderungen der klinischen Routine genügen. Die aktuell verfügbaren Tests erfüllen diese Anforderungen jedoch kaum bis nicht. Entsprechend dringlich ist die Entwicklung adäquater deutschsprachiger sozial-kognitiver Tests.

Elektronisches Supplement

Das elektronische Supplement (ESM) ist mit der Onlineversion dieses Artikels verfügbar unter <https://doi.org/10.1024/1016-264X/a000358>

ESM 1. Zusätzliche Informationen zu den Auswahl- und Teilnahmebedingungen der Studienteilnehmenden

ESM 2. Gründe, keine Tests zur Emotionserkennung durchzuführen, die jeweils von nur einer Institution genannt wurden

Literatur

- Adams, A. G., Schweitzer, D., Molenberghs, P. & Henry, J. D. (2019). A meta-analytic review of social cognitive function following stroke. *Neuroscience & Biobehavioral Reviews*, *102*, 400–416. <https://doi.org/10.1016/j.neubiorev.2019.03.011>
- American Psychiatric Association. (2013). *Diagnostisches und Statistisches Manual Psychischer Störungen (DSM-5)*. Göttingen: Hogrefe.
- Arioli, M., Crespi, C. & Canessa, N. (2018). Social cognition through the lens of cognitive and clinical neuroscience. *BioMed Research International*, *2018*, 4283427. <https://doi.org/10.1155/2018/4283427>
- Babbage, D. R., Yim, J., Zupan, B., Neumann, D., Tomita, M. R. & Willer, B. (2011). Meta-analysis of facial affect recognition difficulties after traumatic brain injury. *Neuropsychology*, *25*, 277–285. <https://doi.org/10.1037/a0021908>
- Baron-Cohen, S., Leslie, A. M. & Frith, U. (1985). Does the autistic child have a theory of mind. *Cognition*, *21*, 37–46. [https://doi.org/10.1016/0010-0277\(85\)90022-8](https://doi.org/10.1016/0010-0277(85)90022-8)
- 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*, *42*, 241–251. <https://doi.org/10.1111/1469-7610.00715>
- 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. & Meletti, S. (2016). Social cognition in temporal lobe epilepsy: A systematic review and meta-analysis. *Epilepsy & Behavior*, *60*, 50–57. <https://doi.org/10.1016/j.yebeh.2016.04.024>
- Brüne, M. (2005). Emotion recognition, 'theory of mind', and social behavior in schizophrenia. *Psychiatry Research*, *133*, 135–147. <https://doi.org/10.1016/j.psychres.2004.10.007>
- Christidi, F., Migliaccio, R., Santamaria-Garcia, H., Santangelo, G. & Trojsi, F. (2018). Social cognition dysfunctions in neurodegenerative diseases: Neuroanatomical correlates and clinical implications. *Behavioural Neurology*, *2018*, 1849794. <https://doi.org/10.1155/2018/1849794>
- Cotter, J., Firth, J., Enzinger, C., Kontopantelis, E., Yung, A. R., Elliott, R. et al. (2016). Social cognition in multiple sclerosis: A systematic review and meta-analysis. *Neurology*, *87*, 1727–1736. <https://doi.org/10.1212/WNL.0000000000003236>
- 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 and Biobehavioral Reviews*, *84*, 92–99. <https://doi.org/10.1016/j.neubiorev.2017.11.014>
- Deuschl, G. & Maier, W. (2016). S3-Leitlinie „Demenzen“ 2016. In Deutsche Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde e.V. (DGPPN) & Deutsche Gesellschaft für Neurologie e.V. (DGN) (Hrsg.), *Leitlinien für Diagnostik und Therapie in der Neurologie*. Verfügbar unter <https://www.awmf.org/leitlinien/detail/ll/038-013.html>
- Diehl-Schmid, J., Lautenschlager, N. T. & Kurz, A. (2011). Gedächtnisprechstunden (Memory-Kliniken). In H. Förstl (Hrsg.), *Demenzen in Theorie und Praxis* (S. 419–436). Berlin: Springer.
- 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. *Journal of Autism and Developmental Disorders*, *36*, 623–636. <https://doi.org/10.1007/s10803-006-0107-0>
- Ekman, P. & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.
- Elamin, M., Pender, N., Hardiman, O. & Abrahams, S. (2012). Social cognition in neurodegenerative disorders: A systematic review. *Journal of Neurology, Neurosurgery and Psychiatry*, *83*, 1071–1079. <https://doi.org/10.1136/jnnp-2012-302817>

- Fonagy, P. & Goodyer, I. (Eds.). (2008). *Social cognition and developmental psychopathology*. Oxford: Oxford University Press.
- Henry, J.D., Cowan, D.G., Lee, T. & Sachdev, P.S. (2015). Recent trends in testing social cognition. *Current Opinion in Psychiatry*, 28, 133–140. <https://doi.org/10.1097/YCO.0000000000000139>
- Henry, J.D., von Hippel, W., Molenberghs, P., Lee, T. & Sachdev, P.S. (2016). Clinical assessment of social cognitive function in neurological disorders. *Nature Reviews Neurology*, 12, 28–39. <https://doi.org/10.1038/nrneurol.2015.229>
- Jarsch, M., Bartsch, K., Berres, M., Ryf, I., Jurisic, K., Sluka, C. et al. (2022). The Basel Version of the Awareness of Social Inference Test-Emotion Recognition (BASIT-ER): Preliminary validation analyses in healthy adults. *Neuropsychology*, 36, 175–184. <https://doi.org/10.1037/neu0000786>
- Jarsch, M., Piguot, O., Berres, M., Sluka, C., Kressig, R.W., Monsch, A.U. et al. (in Revision). The Basel Version of the Awareness of Social Inference Test – Theory of Mind (BASIT-ToM): Preliminary Validation Analyses in Healthy Adults.
- Javed, A. & Charles, A. (2018). The importance of social cognition in improving functional outcomes in schizophrenia. *Frontiers in Psychiatry*, 9, 157. <https://doi.org/10.3389/fpsy.2018.00157>
- Kennedy, D.P. & Adolphs, R. (2012). The social brain in psychiatric and neurological disorders. *Trends in Cognitive Sciences*, 16, 559–572. <https://doi.org/10.1016/j.tics.2012.09.006>
- 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, 1009–1019. <https://doi.org/10.1093/schbul/sbn192>
- Kynast, J., Polyakova, M., Quinque, E.M., Hinz, A., Villringer, A. & Schroeter, M.L. (2020). Age- and sex-specific standard scores for the Reading the Mind in the Eyes Test. *Frontiers in Aging Neuroscience*, 12, 607107. <https://doi.org/10.3389/fnagi.2020.607107>
- LimeSurvey Project Team. (2021). *LimeSurvey: An open source survey tool*. LimeSurvey GmbH. Available from <https://limesurvey.org>
- Maresca, G., Maggio, M.G., Latella, D., Naro, A., Portaro, S. & Calabro, R.S. (2020). Understanding the role of social cognition in neurodegenerative Disease: A scoping review on an overlooked problem. *Journal of Clinical Neuroscience*, 77, 17–24. <https://doi.org/10.1016/j.jocn.2020.05.013>
- Martory, M.D., Pegna, A.J., Sheybani, L., Metral, M., Bernasconi Pertusio, F. & Annoni, J.M. (2015). Assessment of social cognition and theory of mind: Initial validation of the Geneva Social Cognition Scale. *European Neurology*, 74, 288–295. <https://doi.org/10.1159/000442412>
- Menon, V. (2011). Large-scale brain networks and psychopathology: A unifying triple network model. *Trends in Cognitive Sciences*, 15, 483–506. <https://doi.org/10.1016/j.tics.2011.08.003>
- Pinkham, A.E., Penn, D.L., Green, M.F., Buck, B., Healey, K. & Harvey, P.D. (2014). The Social Cognition Psychometric Evaluation study: Results of the expert survey and RAND panel. *Schizophrenia Bulletin*, 40, 813–823. <https://doi.org/10.1093/schbul/sbt081>
- Sachdev, P.S., Blacker, D., Blazer, D.G., Ganguli, M., Jeste, D.V., Paulsen, J.S. et al. (2014). Classifying neurocognitive disorders: The DSM-5 approach. *Nature Reviews Neurology*, 10, 634–642. <https://doi.org/10.1038/nrneurol.2014.181>
- 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, 979–992. <https://doi.org/10.1093/schbul/sbs080>
- Schlegel, K., Grandjean, D. & Scherer, K.R. (2014). Introducing the Geneva Emotion Recognition Test: An example of Rasch-based test development. *Psychological Assessment*, 26, 666–672. <https://doi.org/10.1037/a0035246>
- Seeley, W.W., Crawford, R.K., Zhou, J., Miller, B.L. & Greicius, M.D. (2009). Neurodegenerative diseases target large-scale human brain networks. *Neuron*, 62, 42–52. <https://doi.org/10.1016/j.neuron.2009.03.024>
- Seeley, W.W., Menon, V., Schatzberg, A.F., Keller, J., Glover, G.H., Kenna, H. et al. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience*, 27, 2349–2356. <https://doi.org/10.1523/JNEUROSCI.5587-06.2007>
- Sollberger, M., Rankin, K.P. & Miller, B.L. (2010). Social cognition. *Continuum (Minneapolis)*, 16(4 Behavioral Neurology), 69–85. <https://doi.org/10.1212/01.CON.0000368261.15544.7c>
- Toller, G., Brown, J., Sollberger, M., Shdo, S.M., Bouvet, L., Sukhanov, P. et al. (2018). Individual differences in socioemotional sensitivity are an index of salience network function. *Cortex*, 103, 211–223. <https://doi.org/10.1016/j.cortex.2018.02.012>
- World Health Organization. (2015). *Internationale Klassifikation psychischer Störungen (ICD-10)*, hrsg. von H. Dilling, W. Mombour, M.H. Schmidt & E. Schulte-Markwort. Göttingen: Hogrefe.
- World Health Organisation. (2022). *Internationale Klassifikation psychischer Störungen (ICD-11)*. Entwurfsfassung verfügbar unter https://www.bfarm.de/DE/Kodiersysteme/Klassifikation/ICD/ICD-11/uebersetzung/_node.html
- Yeo, B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M. et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106, 1125–1165. <https://doi.org/10.1152/jn.00338.2011>
- Yogarajah, M. & Mula, M. (2019). Social cognition, psychiatric comorbidities, and quality of life in adults with epilepsy. *Epilepsy & Behavior*, 100(Pt B), 106321. <https://doi.org/10.1016/j.yebeh.2019.05.017>

Historie

Eingereicht: 23. August 2021

Akzeptiert: 06. Mai 2022

Interessenkonflikt

Es bestehen keine Interessenkonflikte.

Finanzielle Unterstützung

MJ, AS und MS werden von der Velux Stiftung (Grant 1225), dem Alzheimer Forum Schweiz, der Stiftung für Demenzforschung Basel, dem Forschungs-Pool Medizin des Universitätsspitals Basel, der Sonnweid Stiftung, Merz Pharma (Schweiz), Novartis und AC Immune und Axovant Sciences finanziell unterstützt.

ORCID

Marianne Jarsch

 <https://orcid.org/0000-0003-1871-8435>

Anna Semenkova

 <https://orcid.org/0000-0002-6034-5911>

Reto W. Kressig

 <https://orcid.org/0000-0002-3215-7681>

Marc Sollberger

 <https://orcid.org/0000-0001-8863-3731>

PD Dr. med. Marc Sollberger

Memory Clinic

Universitäre Altersmedizin FELIX PLATTER

Burgfelderstr. 101

4055 Basel










Schweiz

marc.sollberger@felixplatter.ch

**STUDY III: Development of the Basel Version of the Awareness of Social Inference Test
– Theory of Mind (BASIT-ToM) in Healthy Adults**

RESEARCH ARTICLE

Development of the Basel Version of the Awareness of Social Inference Test – Theory of Mind (BASIT-ToM) in healthy adults

Marianne Jarsch^{1,2}  | Olivier Piguet³  | Manfred Berres⁴  |
Constantin Sluka⁵  | Anna Semenkova^{1,2}  | Reto W. Kressig¹  |
Andreas U. Monsch^{1,2}  | Skye McDonald⁶  | Marc Sollberger^{1,7} 

¹Memory Clinic, University Department of Geriatric Medicine Felix Platter, Basel, Switzerland

²Faculty of Psychology, University of Basel, Basel, Switzerland

³School of Psychology and Brain and Mind Centre, The University of Sydney, Sydney, New South Wales, Australia

⁴Faculty of Mathematics and Technology, University of Applied Sciences Koblenz, Koblenz, Germany

⁵Department of Clinical Research, University of Basel and University Hospital Basel, Basel, Switzerland

⁶University of New South Wales Psychology, Sydney, New South Wales, Australia

⁷Department of Neurology, University Hospital Basel and University of Basel, Basel, Switzerland

Correspondence

Marc Sollberger, Memory Clinic, University Department of Geriatric Medicine Felix Platter, Burgfelderstrasse 101, 4055 Basel, Switzerland.
Email: marc.sollberger@felixplatter.ch

Funding information

National Health and Medical Research Council of Australia, Grant/Award Number: GNT20008020; AC Immune and Axovant Sciences; Novartis; Merz Pharma (Schweiz); Sonnweid Stiftung; Research-Pool Medicine of the University Hospital Basel; Stiftung für Demenzforschung Basel; Alzheimer Forum Schweiz; Velux Foundation, Grant/Award Number: 1225

Abstract

Impairments of Theory of Mind (ToM) abilities occur in a wide range of brain disorders. Therefore, reliable and ecologically valid examination of these abilities is a crucial part of any comprehensive neuropsychological assessment. An established and ecologically valid, English-language test identifying deficits in ToM abilities is “The Awareness of Social Inference Test – Social Inference Minimal (TASIT-SIM)”. However, no comparable German-language ToM test currently exists. In this study, we aimed to develop the first German-language adaption of TASIT-SIM in healthy adults. We selected 13 scenes [four scenes per message type (i.e., honesty, simple sarcasm, paradoxical sarcasm) and one practice scene] out of the 30 TASIT-SIM scenes. In collaboration with a film institute, we filmed each scene at three different intensities. These intensity version scenes were then administered to 240 healthy adults, equally distributed in sex and age, ranging from 35 to 92 years. By applying Rasch analysis, we selected intensity versions that showed neither floor nor ceiling effects in the majority of ToM questions in participants whose ToM abilities were in the medium range. In conclusion, we have developed the first German-language adaption of TASIT-SIM, i.e., the “Basel

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Journal of Neuropsychology* published by John Wiley & Sons Ltd on behalf of The British Psychological Society.

Version of the Awareness of Social Inference Test – Theory of Mind (BASIT-ToM)”. The BASIT-ToM incorporates the strengths of TASIT-SIM, while overcoming its limitations such as inconsistencies in cinematic realization and ceiling effects in healthy participants. Next, the BASIT-ToM needs to be validated in healthy people and clinical populations.

KEYWORDS

dynamic stimuli, healthy subjects, social cognition, test development, theory of mind

BACKGROUND

Theory of Mind (ToM) – the adoption of cognitive (cognitive ToM) and affective (affective ToM) perspectives of others (Henry et al., 2016) – is a critical social cognitive skill that enables us to interact successfully with others (Kennedy & Adolphs, 2012). Adoption of another person's perspective requires both the ability in interpreting the other person's linguistic and paralinguistic characteristics (e.g., facial expressions, gestures and/or prosody) and distinguishing them from one's own (Henry et al., 2016). While cognitive ToM enables people to draw inferences about cognitive mental states such as thoughts and beliefs of another person, affective ToM enables people to infer affective mental states such as emotions from another person (Henry et al., 2016).

Apart from the distinction between cognitive and affective ToM, ToM can also be subdivided into so-called orders. In general, first- and second-order ToM tasks are assessed in clinical studies (Amodio & Frith, 2006; Canty et al., 2015; Castelli et al., 2011; Wang et al., 2018). First-order ToM enables a person to draw social inferences about another person (e.g., first-order cognitive ToM: Person A thinks that person B wants to quit his job; Meijering et al., 2011). Forming a second-order ToM, however, is cognitively more complex as it requires higher-level-reasoning (Canty et al., 2017) and enables a person to mentalize social inferences that another person draws about another person (e.g., second-order affective ToM: Person A thinks that person B feels sorry for person C as person B seems to perceive person C as sad; Meijering et al., 2011).

Lastly, apart from content and degree of cognitive complexity within the concept of ToM, ToM can be conceptualized as a continuum that ranges from no mentalization through to hypomentalization, adequate mentalization and finally hypermentalization (Abu-Akel & Bailey, 2000). Whilst hypomentalizing can be seen as reduced mentalization, hypermentalization can be described as a kind of overattribution, in which additional mental hypotheses are generated, that cannot be directly deduced from the social situation (Abu-Akel & Bailey, 2000; Frith, 2004). In other words, people with hypermentalization mentalize, but in an excessive and inadequate way. When examining ToM in patients with neuropsychological disorders, one needs to be aware of the continuum of ToM, as disturbances in ToM vary depending on the brain disorders. For example, patients with Asperger's syndrome, euthymic bipolar disorder, behavioural variant frontotemporal dementia or schizophrenia with negative symptoms tend to hypomentalize (Bora et al., 2015; Canty et al., 2017; Dziobek et al., 2006; Montag et al., 2010), whereas patients with borderline personality disorder or patients with schizophrenia with positive symptoms tend to hypermentalize (Canty et al., 2017; Sharp et al., 2011).

Deficits in ToM occur in a wide range of brain disorders, including developmental, neurological and psychiatric disorders (for review see Cotter et al., 2018). Given the high occurrence and clinical relevance of ToM deficits, the Diagnostic and Statistical Manual for Mental Disorders, Fifth Edition (DSM-5; American Psychiatric Association, 2013) requires the assessment of ToM as a subdomain of social cognition for any comprehensive diagnosis of a neurocognitive disorder. To meet this goal, sensitive

and reliable tests with high construct validity that assess different aspects of ToM within a practicable timeframe in clinical populations are needed. In addition, performance on tests of ToM should ideally predict the examined person's ability to adopt cognitive and affective perspectives of others in real life. In this regard, dynamic, realistic and socially relevant stimuli (e.g., film scenes) may meet this demand more than static stimuli do.

In terms of tests that contain dynamic, socially relevant stimuli validated for German-speaking patients, only the Movie for the Assessment of Social Cognition (MASC; Dziobek et al., 2006) is available, to the best of our knowledge. In this test, a fifteen-minute long film is shown about four adolescents who meet for a cooking and games evening. The film is interrupted 45 times. Each time, participants are asked questions about the interactions that have just been shown, which relate to emotions, thoughts and intentions of the protagonists. It has been shown that the MASC is a reliable, ecologically valid test which has proved sensitive to detect ToM deficits in different brain disorders such as Asperger's syndrome (Dziobek et al., 2006), personality disorder (Sharp et al., 2011) or schizophrenia (Martinez et al., 2017). However, given its quite challenging cognitive design (Dziobek et al., 2006; Pottgen et al., 2013), the test seems little suitable for patients with deficits in multiple cognitive domains such as patients with neurodegenerative diseases. In addition, the topic of the protagonists' interactions seem more of interest for young adults (for whom the test was primarily developed) than for other age groups, which limits its relevance in the general population. Lastly, test application takes a long time in subjects with cognitive disorders, e.g., 30–45 min in patients with mild to moderate traumatic brain injuries [J. Quinting, personal communication, October 29, 2021; (Quinting et al., 2020)], which precludes its regular use in clinical settings. Therefore, a German-language test based on dynamic, multimodal stimuli in social interactions that examines ToM in a less cognitively demanding way for the general adult population in a timely fashion is needed.

We decided to develop a German-language adaption of “The Awareness of Social Inference Test – Social Inference Minimal” (TASIT-SIM), an established and ecologically valid, English-language test identifying deficits in ToM abilities (McDonald et al., 2003). TASIT-SIM consists of two test sets of 15 brief film scenes showing simple but realistic day-to-day social interactions. By use of four questions per scene, ToM abilities can be assessed based on the comprehension of intentions, beliefs, meanings and emotions in either sarcastic or honest conversational messages. TASIT-SIM was designed as a criterion-based test that can identify clear deficits in ToM abilities in adults (McDonald et al., 2003). Indeed, TASIT-SIM can discriminate between healthy individuals and patients with different brain disorders including neurodegenerative diseases (e.g., behavioural variant frontotemporal dementia (Kumfor et al., 2017), semantic variant primary progressive aphasia (Rankin et al., 2009), progressive supranuclear palsy (Ghosh et al., 2012)), multiple sclerosis (Genova & McDonald, 2020), traumatic brain injuries (McDonald et al., 2017) and psychiatric disorders such as schizophrenia or bipolar disorder (Quidé et al., 2020). It has adequate reliability and there is evidence for construct and ecological validity (for review see McDonald, 2012; McDonald et al., 2004).

TASIT-SIM, however, contains limitations that need to be addressed (see also Jarsch et al., 2022). First, because of its duration, the test does not lend itself to be easily used in clinical settings (Henry et al., 2014; Westerhof-Evers et al., 2014). Second, the social interactions often appear unrealistic due to the high intensity of paralinguistic features and somewhat outdated appearances of actors, potentially biasing participants' capacity for ToM judgements. Third, the high intensity of the paralinguistic features result in near ceiling scores in the majority of healthy people (McDonald et al., 2006, 2015), preventing the detection of subtle ToM deficits. Finally, questions after each scene were not explicitly developed to evaluate different ToM concepts (i.e., first-/second-order ToM, cognitive/affective ToM, hypermentalization), but to evaluate overall ToM ability, which hampers the evaluation of these different aspects of ToM.

In this study, we created an adapted German-language version of TASIT-SIM with professional actors in collaboration with a film institute and administered it to 240 healthy participants to select the scenes that showed neither floor nor ceiling effects in the majority of ToM questions in participants

whose ToM abilities were in the medium range. The adapted German-language version was named the “Basel Version of the Awareness of Social Inference Test-Theory of Mind (BASIT-ToM)”.

METHODS

The Awareness of Social Inference Test (TASIT) – Social Inference Minimal (SIM) and Social Inference Enriched (SIE)

TASIT – Social Inference Minimal (SIM) and TASIT – Social Inference Enriched (SIE) are parts of TASIT (McDonald et al., 2003). TASIT comprises three parts, i.e., TASIT – Emotion Evaluation Test (TASIT-EET), TASIT-SIM and TASIT-SIE. TASIT-EET examines the ability to recognize basic emotions in social situations, whereas TASIT-SIM and TASIT-SIE examine the ability to make ToM judgements in social interactions (McDonald et al., 2003). The adapted German-language version of TASIT-EET, called the “Basel Version of the Awareness of Social Inference Test – Emotion Recognition” has been published previously (Jarsch et al., 2022).

TASIT-SIM examines the ability to make ToM judgements based on linguistic and paralinguistic cues (e.g., gestures, facial expressions, prosody) in short, realistic, everyday dialogue interactions with either honest or sarcastic content and minimal contextual information (i.e., no information provided other than that presented in the short interaction). In honest exchanges, the paralinguistic features of the speaker are congruent with the literal message allowing the meaning to be inferred directly. In sarcastic statements, the paralinguistic features are incongruent with the spoken text. If this incongruity cannot be perceived, statements might be perceived as honest, meaningless, or bizarre (McDonald et al., 2003). In the SIM, sarcasm is subdivided into simple sarcasm (i.e., the literal message could be misinterpreted as honest) and paradoxical sarcasm (i.e., the literal message only appears meaningful if the sarcastic meaning is recognized).

TASIT-SIE comprises a set of scenes showing dialogue interactions with either sarcastic messages or lies. As in the SIM, viewers have to perceive and interpret paralinguistic features of the speaker/s to identify the meaning of the literal message. In contrast to the SIM, scenes of the SIE provide additional information to the viewer about the true belief of the speaker or the true state of affairs by presenting visual or auditory/verbal cues before or after the interaction scene.

Two parallel test versions, i.e., form A and form B, exist for both SI-forms. The SIM comprises 15 scenes [five scenes per three different message types (i.e., honesty, simple sarcasm, and paradoxical sarcasm)]. The SIE comprises 16 scenes [eight scenes per two different message types (i.e., sarcasm, lie)]. The scenes last 15–60 s and are administered in a quasi-randomized order.

In both SI-forms, ToM ability is assessed with four questions: one about the belief (thinking question), one about the intention (doing question), one about the message (saying question) and the last one about the emotion (feeling question) of the main actor who behaves either honestly or sarcastically, or who lies. Participants can answer each question with “yes”, “no”, or “do not know”.

The SIM and the SIE were developed to differentiate between cognitively healthy people “with average social skills” and individuals with clear ToM deficits (McDonald et al., 2003). Accordingly, actors were asked to act each scene in an exaggerated fashion such that the majority of healthy people should be able to answer the four questions correctly (McDonald et al., 2003).

Development of the BASIT-ToM scenes

We decided to choose only one of the SI-forms as they differ only in their richness of contextual information, and both assess ToM similarly. Indeed, factor analysis conducted on data collected in patients with acquired brain injuries and healthy individuals demonstrated the presence of a single factor underlying both SI-forms (Honan et al., 2016). As the BASIT-ToM is aimed for use in patients with cognitive deficits, we

favoured SIM over SIE, as it likely requires less cognitive capacities than SIE given its lower amount of contextual information. SIE comprises additional camera shots on the true state of affair by mean of prologues or epilogues, whereas SIM does not. Arguably, however, by leaving out SIE, which contains more contextual information than SIM, we may not assess fully the brain regions associated with contextual adjustment in ToM (Lavoie et al., 2016). Accordingly, one may wish to consider adapting the SIE scenes in a comparable way to the SIM scenes depending on the patient groups to be investigated and the research questions.

We named the German-language adaption of TASIT-SIM, the BASIT-ToM. In the following, we present the development of the BASIT-ToM for which we took a similar approach as for the BASIT-ER (Jarsch et al., 2022).

Selection of the scenes from TASIT-SIM

One test form was developed, rather than two, given negligible practice effects on TASIT-SIM scores (McDonald et al., 2003). First, we selected 13 film scenes (4×3 message types for test scenes and one practice scene) from the 30 TASIT-SIM scenes (15 form A scenes and 15 form B scenes; McDonald et al., 2003) based on 12 evaluation criteria that focus on different aspects of cinematographic quality, target group fit and cinematic feasibility (see Appendix A1). Three raters (MJ, MS, and a master's degree psychology student) rated each scene independently, and then agreed on a joint rating and selected the 13 scenes in consensus (see Appendix B: Tables B1.1–B1.3). The selected TASIT-SIM scenes are depicted in the Appendix B: Table B2.

Conceptualization of the BASIT-ToM scenes

We developed the conceptualization of the scenes and the cinematographic realization in collaboration with the film production company East End Film GmbH (Germany; <https://www.eastendfilm.de>) and adapted TASIT-SIM as follows:

Screenplay

First, the transcription of TASIT-SIM scenes (Westerhof-Evers et al., 2014) was adapted to the German language without major changes of content. In three scenes (i.e., A9, A10, A13), we deleted text passages that either lengthened the scene without providing additional information or made it too difficult to understand. In two scenes (i.e., A11, B15), we supplemented the script with short sentences to achieve a more realistic conversation. Second, as in the BASIT-ER (Jarsch et al., 2022), we omitted naming the actors to exclude any potential bias associated with naming. All scenes were cast with one woman and one man. BASIT-ToM screenplays and description of the main changes in the text are presented in the Appendix C.

Cast and acting of the actors

We cast the scenes with four female and four male, middle-aged, professional German-speaking actors. Each actor portrayed a message type (i.e., honesty, simple sarcasm or paradoxical sarcasm) only once to avoid any actor-specific associations to a message type. Each pair of actors played only one scene together to omit any potential biases due to varying role relationships between actors. Actors' sex were equally distributed across the different message types.

To make the scenes appear as realistic as possible, both actors in the scene acted in a realistic way. This type of acting differs from TASIT-SIM in which the conversation partners act in a neutral way (i.e., the conversation partner does not seem to perceive the main actor's sarcasm). In real life, however, one would not expect another person to react in a neutral way to someone who is very sarcastic. Indeed, the conversation partner's neutral behaviour may confuse a viewer with preserved ToM capacity. The distribution of the actors to each of the 13 scenes is presented in the Appendix D.

Realism of scenes and camera work

We placed emphasis on a high degree of realism on the film scenes through realistic set designs either at office or at home backgrounds. All scenes were filmed as medium shots. Medium shots focus on the person in the scene while still showing some environment. Scenes were shot in 4K-resolution.

Three intensity levels of the portrayed paralinguistic cues

Actors were asked to portray the paralinguistic cues at three different levels of intensity (i.e., low, medium, and high). Text and set design remained identical at the three intensity versions. Background information was written for each scene and each intensity to help actors portraying the message types. After the shooting, MJ, MS and six master's degree psychology students selected the best scenes for each intensity by consensus. One exemplary scene (scene#4 of the message type paradoxical sarcasm) at low, medium and high levels of intensity, can be found at <https://figshare.com/s/5ca4ff99790fcc5866db>.

Background information for each BASIT-ToM scene is described in the Appendix E.

Adaption of the questions of the BASIT-ToM

In TASIT-SIM, four kinds of questions were originally asked tapping thoughts, intentions, feelings and meanings of the speakers. While belief and intention questions, in particular, were originally designed to tap first-order and second-order ToM respectively, they varied in the extent to which this was successful. This was especially the case for the intention questions, which often asked for understanding of motives rather than beliefs. Other question types also tapped ToM to varying degrees (S. McDonald, personal communication, November 15, 2021; McDonald et al., 2003).

We decided to categorise existing questions as to whether they explicitly referred to (1) affective or (2) cognitive ToM and whether this was (3) first-order or (4) second-order. We also reworded several first-order ToM questions and created additional second-order ToM questions to get a more balanced representation of first- and second-order questions. We were able to reword these questions because conversation partners act, in contrast to TASIT-SIM, in a realistic way in the BASIT-ToM scenes (see above "Cast and acting of the actors"). The generation of new second-order ToM questions resulted in a more balanced proportion of first- and second-order ToM questions [i.e., 65% first-order ToM questions (first-order cognitive ToM: $n = 17$, first-order affective ToM: $n = 17$), 35% second-order questions (second-order cognitive ToM: $n = 8$, second-order affective ToM: $n = 10$)]. First- and second-order ToM and cognitive and affective ToM were partly balanced within and between the three message types, i.e., honesty: 19% first-order cognitive ToM, 13% second-order cognitive ToM, 44% first-order affective ToM, 25% second-order affective ToM; simple sarcasm: 38% first-order cognitive ToM, 25% second-order cognitive ToM, 25% first-order affective ToM, 12% second-order affective ToM; paradoxical sarcasm: 38% first-order cognitive ToM, 6% second-order cognitive ToM, 31% first-order affective ToM, 25% second-order affective ToM. Importantly, the distribution of ToM types may undergo further refinement, depending on the results of the planned validation study. We will also consider combining simple and paradoxical sarcasm scores that may result in a more balanced distribution of ToM types within and between message types. Taken together, we assume that separate analyses of ToM types will be feasible in the final BASIT-ToM version. This is critical with regard to the future use of the BASIT-ToM in clinical samples. Scores reflecting different ToM abilities, which partly reflect different neuroanatomic substrates (Corradi-Dell'Acqua et al., 2020; Fortier et al., 2018; Poletti et al., 2012; Ryan et al., 2017), will likely better discriminate between brain diseases (Lancaster et al., 2019; Poletti et al., 2012; Rossetto et al., 2018) than an overall ToM score.

Unlike other ToM tests (e.g., MASC (Dziobek et al., 2006), Virtual Assessment of Mentalising Ability (VAMA; Canty et al., 2015)), the BASIT-ToM contains multiple mentalizing questions per scene. Using normalization, we will be able to assess a participant's mentalization ability, ranging from no mentalization, to reduced mentalization (i.e., hypomentalization) to adequate mentalization. This approach, however, says nothing about hypermentalization (i.e., overattribution of another person's cognitive and/

or affective mental states), which is why we created a hypermentalization question (hyperToM) for each scene. The hypermentalization question may be useful to discriminate between different clinical populations with behavioural disorders. For example, individuals diagnosed with schizophrenia who experience positive symptoms tend to hypermentalize (Canty et al., 2017), whereas patients with behavioural variant frontotemporal dementia, another syndrome associated with behavioural disorders, but due to neurodegeneration (Rascovsky et al., 2011), hypomentalize or are even unable to mentalize (Bora et al., 2015).

In TASIT-SIM, participants could answer each question with “yes”, “no”, or “do not know” (McDonald et al., 2003). We reduced the three response alternatives to a forced-choice paradigm, namely “yes” and “no”, as we considered the risk of people with cognitive disorders choosing the “do not know” option due to uncertainty or little motivation to be greater than the problem of guessing probability of 0.5 per question. Moreover, we reworded some questions to get a more balanced proportion of “yes” and “no” answers per scene to avoid any content-independent tendency towards “yes” or “no” answers, respectively. For illustration, you find the adaption of a first-order, cognitive TASIT-SIM question into a second-order, affective BASIT-ToM question, as well as a hypermentalization question in Table 1.

All questions of each BASIT-ToM scene and the respective TASIT-SIM scene with representation of ToM types, ToM orders and correct answers, as well as the hypermentalization questions are found in the Appendix F. The main differences between BASIT-ToM and TASIT-SIM scenes are described in the Appendix G.

Selection of the BASIT-ToM intensity version scenes for use in clinical populations

Participants

Next, we administered the BASIT-ToM intensity version scenes to 240 cognitively and mentally healthy Central European subjects (50% women) with mother language (Swiss)-German to select scenes neither showing floor nor ceiling effects. We opted for a large age range to make the results applicable to the general adult population. We defined five age groups (35–44, 45–54, 55–64, 65–74, >75) to achieve an even distribution of sex by age. Each group consisted of 48 participants (50% women; at maximum three participants were at the same age). Participants were included if they met the following inclusion criteria: 35 years of age or older, total education of seven years or greater, German and/or Swiss German as first language, and self-report of good health. Exclusion criteria were conditions with potential negative influence on the test results, including signs of depressive mood (i.e., scoring ≥ 10 points on the Beck Depression Inventory (Beck et al., 1961) for individuals below 65 years, or ≥ 5 on the Geriatric Depression Scale (Yesavage & Sheikh, 1986) for individuals aged ≥ 65 years), cognitive deficits [i.e., Montreal Cognitive Assessment (Nasreddine

TABLE 1 Example of an adaption of a first-order, cognitive TASIT-SIM question into a second-order, affective BASIT-ToM question as well as an example of a hypermentalization question

	Question	ToM type	ToM order	Correct answer
TASIT-SIM	Does Michael think she took it easy on the weekend?	cogn	1st	no
BASIT-ToM	Does he think it was okay for her working all weekend?	aff	2nd	yes
	Is he asking her about the report to provoke her?	hyperToM		no

Abbreviations: aff, affective ToM; BASIT-ToM, Basel Version of the Awareness of Social Inference Test-Theory of Mind; cogn, cognitive ToM; hyperToM, hypermentalization; TASIT-SIM, The Awareness of Social Inference Test-Social Inference Minimal; ToM, Theory of Mind.

et al., 2005) score below the demographically-adjusted, fifth percentile for cognitively healthy individuals (Thomann et al., 2018)], systemic or brain diseases, psychiatric disorders according to the ICD-10 criteria, traumatic brain injury, chronic pain, history or current regular intake of any psychoactive drugs (except benzodiazepines for sleep), general anaesthesia within the last three months and severe sensory and/or motor deficits. The study was approved by the local ethics committee and all participants provided written informed consent.

Application of the BASIT-ToM scenes

BASIT-ToM scenes and the test paradigm were displayed on a monitor with a diagonal size of 24 inches, a 16:10 aspect ratio and a 1920:1200 display resolution using Python 2.7 (Peirce, 2007) and PsychoPy 1.84.2 package (Peirce, 2009). Details of the programming and data storage are described in the Appendix G.

Before the test started, participants were provided with standardized instructions by the examiner. Then, they read the test instructions on the computer monitor. Next, one practice scene was shown to familiarize the participant with the procedure and to clarify any potential questions. The 12 test scenes were then administered in a pseudo-randomized order (not the same message type twice in a row). Each message type [i.e., honesty (H), simple sarcasm (sS), paradoxical sarcasm (pS)] was shown in four scenes (3×4). Of the four scenes of each message type, participants watched one scene at low intensity, another scene at medium intensity, another scene at high intensity, and another scene at either low, medium, or high intensity. For example, participant#1 watched H#1 at low intensity, H#2 at medium intensity, H#3 at high intensity, and H#4 at low intensity, whereas participant#2 watched H#1 at medium intensity, H#2 at high intensity, H#3 at low intensity, and H#4 at medium intensity. As there were three intensity levels per scene, each intensity version scene was watched by 80 participants. Each participant watched 13 intensity version scenes, i.e., one practice scene at a given intensity +12 test scenes at given intensities [$3(\text{message types}) \times 4(\text{intensity versions})$].

By taking the approach that each participant watches only one intensity per scene, we avoided potential biases on responses arising from watching a scene that you have seen and rated previously at a different intensity.

Following each scene, participants were required to answer the four ToM questions and the hyper-ToM question with “yes” or “no.” There was no time limit for answering the questions.

If participants were uncertain about a scene's content, they could rewatch the respective scene. They could rewatch a scene as many times as they wanted before answering. We noted the number of times a scene was watched. By noting it, we learned, in addition to the participants' ToM responses, about potential difficulties in the understanding of the respective scene content. We considered this information critical given the fact that we had developed new scenes, albeit based on TASIT-SIM templates.

Similarly, participants could change their answers within a question block. An exemplary representation of the BASIT-ToM computer-based application process is depicted in Appendix H.

Data analysis

The aim of the data analysis of the 39 intensity version scenes [$3(\text{message types}) \times 4(\text{scenes}) \times 3(\text{intensities}) = 36$ intensity versions + 3 practice scene intensity versions] was to select one intensity version per scene that resulted in neither floor nor ceiling performance in participants whose ToM abilities were in the medium range for as many of the respective four ToM questions as possible. We prioritized the four ToM questions over the hyperToM question in our analysis approach as we aimed primarily for scenes that are adequate to detect decreased ToM ability. To achieve this type of scene selection, we took the following four-steps approach:

Step 1

We conducted Rasch analysis (Rasch, 1960) by applying the R ltm package (Rizopoulos, 2006) with RStudio (R Studio Team, 2016) in R version 3.5.1 (R Core Team, 2020) to analyze the relations between the estimated participants' ToM ability parameters and the probabilities of correct answering the four ToM questions of the respective intensity version.

In each of the three message type analyses, we included the respective twelve intensity versions (4 scenes per message type \times 3 intensities). In the practice scene analysis, we included the respective three intensity versions. Separate analyses were carried out for each message type intensity version scene and each practice scene intensity version. In order to check the model fit of the data, we conducted the ltm package parametric Bootstrap goodness-of-fit test using Pearson's χ^2 , based on 201 data sets (the original data set plus 200 simulated datasets). The scenes that showed model fit were selected for further analysis.

Step 2

Based on the graphical assessment of the item characteristic curve (ICC) plots of the intensity versions showing model fit, we aimed to choose one intensity per message type scene and one intensity for the practice scene that predicted a correct answering probability between 0.5 and 0.8 for the majority of the respective four ToM questions in participants with a medium ToM ability around 0 [-0.5, 0.5].

Step 3

Next, we examined the 95% confidence intervals (CIs) of the estimated difficulties of the selected intensity versions in relation to the required difficulty interval. CIs were calculated as estimate $\pm 1.96 \times$ standard error of estimate. The required difficulty was calculated according to the scene selection criteria, mentioned above at step 2 (i.e., probability between 0.5 and 0.8 in participants with a medium ToM ability around 0). This step allowed us to check whether the CI of the estimated difficulties of the selected scenes covered the required difficulty interval. In addition, we evaluated whether systematic differences in difficulty were present between first- and second-order ToM questions and between cognitive and affective ToM questions, respectively. For this, we checked whether the estimated difficulty intervals of these questions overlapped (i.e., equal difficulties of the questions) or not (i.e., different difficulties of the questions).

Step 4

After completing scene selection, we ran separate Rasch analyses with each of the selected intensity versions by additionally including the respective hyperToM question to evaluate the probability of correct answering the hyperToM questions.

Distribution of demographic variables between the selected intensity version scenes

As participants were randomly assigned to each intensity version scene and given the potential influence of demographic variables on ToM performance, we compared the mean scores and standard deviations of age and years of education as well as the sex ratio across the selected intensity version scenes.

Transparency and openness

In accordance with the Transparency and Openness Promotion (TOP; Nosek et al., 2015) and the Journal Article Reporting Standards-Quantitative for non-experimental designs (JARS-Quant; Appelbaum et al., 2018), we have reported in detail how we developed the BASIT-ToM stimulus material. The data that support the findings of this study are openly available in “figshare” at <https://figshare.com/s/d7194f6f2f18082e623c>.

RESULTS

Selection of the scenes

Step 1: Selection of intensity versions with an acceptable model fit

The parametric Bootstrap goodness-of-fit test showed an acceptable fit of the Rasch model ($p > .05$) for the analyses of the four ToM questions for 22 of the 39 intensity versions. ICC of the 22 intensity versions with an acceptable model fit are shown in Appendix J.

Step 2: Selection of intensity versions based on response probabilities

We selected 10 intensity versions [i.e., 3 intensity versions per message type (i.e., 3×3) and 1 intensity version for the future practice scene] from the 22 intensity versions that showed model fit of the Rasch model and predicted probabilities of a correct response between 0.5 and 0.8 for as many of the respective four ToM questions as possible in participants with a medium ToM ability around 0 [−0.5, 0.5] (see the H scenes in Figure 1a–c, the pS scenes in Figure 2a–d and the sS scenes in Figure 3a–c). All selected intensity versions are shown in Table 2, and the ToM types and orders of the questions associated with these intensity versions are shown in Table 3.

The number of ToM questions that met the required response probabilities varied between the selected intensity versions. In 8 (80%) of the 10 selected intensity version scenes at least half of the questions met the required response probabilities, i.e., in 4 scenes (i.e., H#2_low, pS#1_low, pS#2_low, practice(pS)_low), the response probabilities were met by all four ToM questions; in 2 scenes (i.e., sS#1_low, sS#3_medium), they were met by three questions and in another 2 scenes (i.e., H#1_low, pS#4_low), they were met by two questions. In the H#4_medium and sS#2_low scenes, only one question met the required response probability. The response probabilities of the questions that were outside the required range were predicted higher (>0.8) apart from the affective second-order question in sS#1_low, which was predicted lower (<0.5).

Step 3: Evaluation of the estimated difficulties of the ten selected intensity versions

Next, we examined how well the 95% CI of the estimated difficulties of the four ToM questions of the selected intensity versions covered the required range of difficulty [i.e., $-1.386, 0$ (minus the logits of $0.8 = -1.386$, minus the logits of $0.5 = 0$); see the H scenes in Figure 4, the pS scenes in Figure 5 and the sS scenes in Figure 6]. The estimated difficulties were considered within the required range of difficulty when at least half of the CI covered the required range. In two of the 10 selected scenes (20%; i.e., H#2_low, pS#1_low), all four questions were within the required range. In another two scenes (i.e., pS#2_low, practice(pS)_low), three of the four questions were within the required range, whereas one question was slightly above the required difficulty (too difficult). In three scenes (i.e., pS#4_low, sS#1_low, sS#3_low), two questions were within the required range. In pS#4_low, the other two questions were below

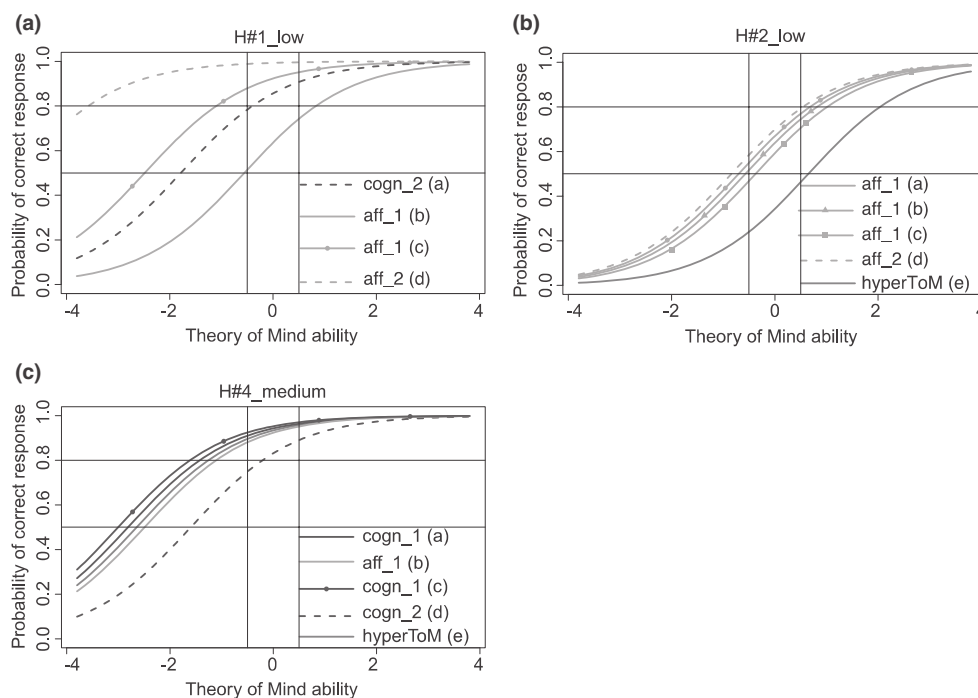


FIGURE 1 Item characteristic curves (ICC) of the theory of mind (ToM) and hypermentalization questions (hyperToM) of the three selected BASIT-ToM scenes for message type honesty (i.e., a: Scene H#1_low, b: Scene H#2_low, c: Scene H#4_medium). ICC of the hyperToM question is only depicted in presence of a model fit. Scene code consists of a scene number (scene), a letter (message type, H = honesty), and the portrayed intensity (i.e., low or medium). Question code consists of abbreviations of type (i.e., cogn = cognitive ToM, aff = affective ToM, hyperToM = hypermentalization), order (i.e., 1 = first-order ToM, 2 = second-order ToM) of ToM, and question's identification letter (a-d = ToM questions, e = hypermentalization question). Line with circles = ToM question of same type and order as another ToM question in the same panel; line with squares = ToM question of same type and order as two other questions in the same panel. BASIT-ToM, Basel Version of the Awareness of Social Inference Test – Theory of Mind

the required difficulty (too easy), whereas in sS#1_low and sS#3_low the other two questions were above the required difficulty (too difficult). In one scene (i.e., H#1_low), one question was within the required range, whereas the other three questions were below the required difficulty (too easy). In two scenes (i.e., H#4_medium, sS#2_low), all questions were below the required difficulty (too easy).

Notably, we found no systematic differences in terms of difficulty between first- and second-order ToM questions and between cognitive and affective ToM questions, respectively (Figures 4–6).

Step 4: Evaluation of the HyperToM questions in terms of model fit and response probabilities

Finally, we checked whether the model fit of each of the 10 selected intensity versions persisted after inclusion of the hyperToM question in the model and if yes, how well the 95% CI of the estimated difficulties of the hyperToM questions covered the required range of difficulty (i.e., $-1.386, 0$). In six scenes (60%; i.e., H#2_low, H#4_medium, pS#1_low, pS#4_low, sS#1_low, sS#2_low), there was still a model fit. In two [i.e., pS#1_low (Figure 2a), sS#2_low (Figure 3b)] of these scenes, the response probability of the hyperToM question was within the required range of 0.5–0.8 (Figures 5 and 6); in three scenes [i.e., H#4_medium (Figure 1c), pS#4_low (Figure 2c), sS#1_low (Figure 3a)], it was above 0.8 [too easy, Figures 4–6] and in one scene [i.e., H#2_low (Figure 1b)], it was below 0.5 [too difficult,

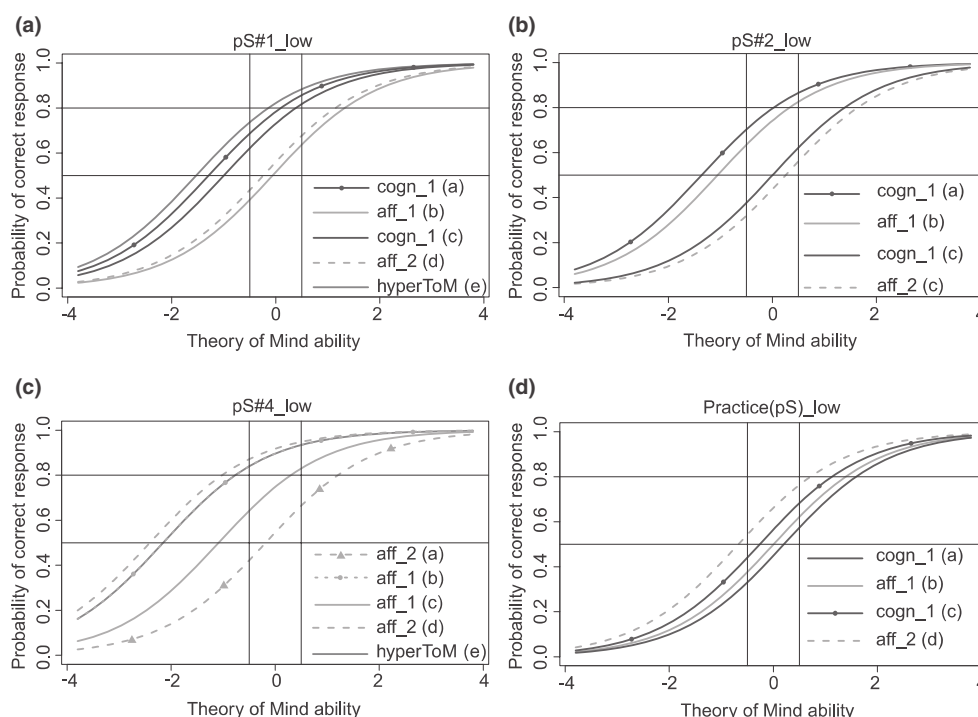


FIGURE 2 Item characteristic curves (ICC) of the theory of mind (ToM) and hypermentalization questions (hyperToM) of the four selected BASIT-ToM scenes for message type paradoxical sarcasm [i.e., a: Scene pS#1_low, b: Scene pS#2_low, c: Scene pS#4_low, d: Practice (pS)_low]. ICC of the hyperToM question is only depicted in presence of a model fit. Scene code consists of a scene number (scene), a letter (message type, pS = paradoxical sarcasm), and the portrayed intensity (i.e., low). Question code consists of abbreviations of type (i.e., cogn = cognitive ToM, aff = affective ToM, hyperToM = hypermentalization) and order (i.e., 1 = first-order ToM, 2 = second-order ToM) of ToM, and question's identification letter (a–d = ToM questions, e = hypermentalization question). Line with circles or line with triangles = ToM question of same type and order as another ToM question in the same panel. Practice(pS_low) will be used as a test scene for paradoxical sarcasm in BASIT-ToM, whereas PS#4_low will be used as practice scene. BASIT-ToM, Basel Version of the Awareness of Social Inference Test – Theory of Mind

Figure 4]. A model fit was no longer present in four scenes (i.e., H#1_low, pS#2_low, practice(pS)_low, sS#3_medium).

Number of times the selected intensity version scenes were watched

The majority of participants (83%) watched the selected intensity version scenes once only. Selected intensity version scenes were watched more than once as follows: twice: 16%, three times: 1%, and four times: <1%. None of the selected intensity version scenes was watched only once by all 80 participants. Percentages of scenes watched once only ranged from 58% (pS#2_low) to 91% (H#1_low, sS#1_low, sS#3_medium). For more details, please see Table K.1 in the Appendix K.

Distribution of demographic variables between the selected intensity version scenes

We found similar distributions of the demographic variables across the selected intensity version scenes: For years of age, means ranged from 56.58 to 62.59 with standard deviations ranging from 13.34 to

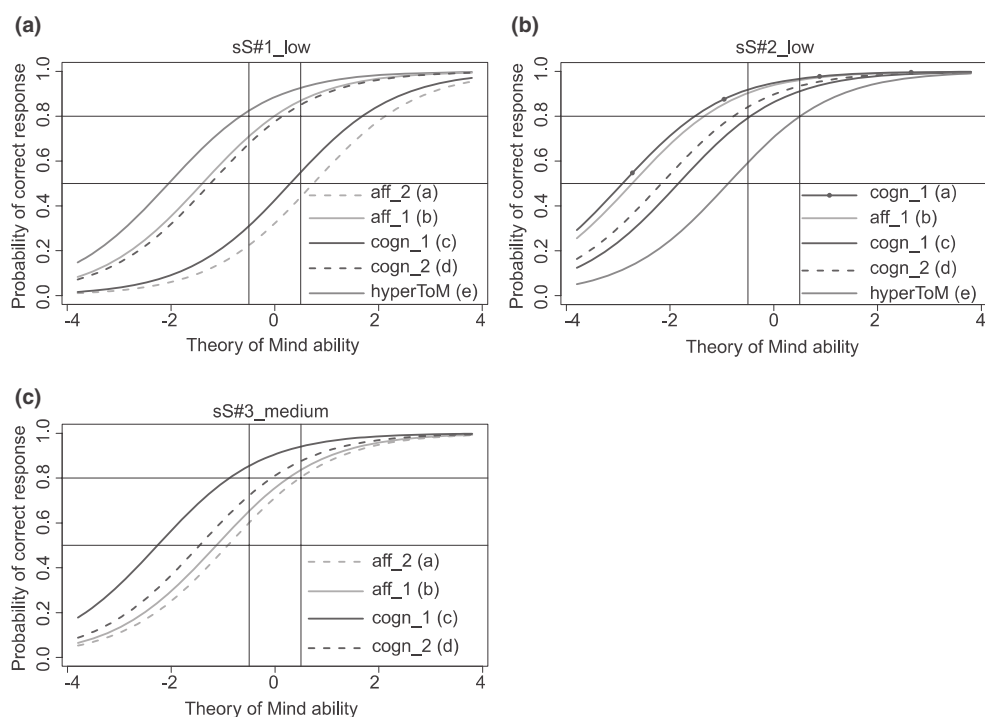


FIGURE 3 Item characteristic curves (ICC) of the theory of mind (ToM) and hypermentalization questions (hyperToM) of the three selected BASIT-ToM scenes for message type simple sarcasm (i.e., a: Scene sS#1_low, b: Scene sS#2_low, c: Scene sS#3_medium). ICC of the hyperToM question is only depicted in presence of a model fit. Scene code consists of a scene number (scene), a letter (message type, sS = simple sarcasm) and the portrayed intensity (i.e., low or medium). Question code consists of abbreviations of type (i.e., cogn = cognitive ToM, aff = affective ToM, hyperToM = hypermentalization) and order (i.e., 1 = first-order ToM, 2 = second-order ToM) of ToM, and question's identification letter (a–d = ToM questions, e = hypermentalization question). Line with circles = ToM question of same type and order as another ToM question in the same panel. BASIT-ToM, Basel Version of the Awareness of Social Inference Test – Theory of Mind

TABLE 2 Selected intensity version scenes for the BASIT-ToM for further use in clinical populations

Message type	Intensity version				
	Scene 1	Scene 2	Scene 3	Scene 4	Scene 5
Honesty	Low	Low	–	Medium	n/a
Simple sarcasm	Low	Low	Medium	–	n/a
Paradoxical sarcasm	Low	Low	–	Low ^a	Practice (pS)_low ^b

Abbreviations: BASIT-ToM, Basel Version of the Awareness of Social Inference Test-Theory of Mind; n/a, not available; pS, paradoxical sarcasm.

^apS#4_low (scene #4 containing paradoxical sarcasm at low intensity) will be used as practice scene in BASIT-ToM.

^bPractice(pS)_low will be used as a test scene for paradoxical sarcasm in BASIT-ToM.

16.34; for years of education, means ranged from 15.09 to 16.06 with standard deviations ranging from 2.52 to 3.15 and for sex, percentage of females ranged from 41% to 59%.

DISCUSSION

In this study, we developed the BASIT-ToM, the first German-language TASIT-SIM adaption and administered it to a sample of 240 healthy adults with a large age range. We adapted TASIT-SIM to address

TABLE 3 Types and orders of theory of mind of each question of the ten selected scenes

Scene	ToM type and order				
	a	b	c	d	e
H#1	cogn_2 nd	aff_1 st	aff_1 st	aff_2 nd	Hyper
H#2	aff_1 st	aff_1 st	aff_1 st	aff_2 nd	Hyper
H#4	cogn_1 st	aff_1 st	cogn_1 st	cogn_2 nd	Hyper
sS#1	aff_2 nd	aff_1 st	cogn_1 st	cogn_2 nd	Hyper
sS#2	cogn_1 st	aff_1 st	cogn_1 st	cogn_2 nd	Hyper
sS#3	aff_2 nd	aff_1 st	cogn_1 st	cogn_2 nd	Hyper
pS#1	cogn_1 st	aff_1 st	cogn_1 st	aff_2 nd	Hyper
pS#2	cogn_1 st	aff_1 st	cogn_1 st	aff_2 nd	Hyper
pS#4	aff_2 nd	aff_1 st	aff_1 st	aff_2 nd	Hyper
Practice (pS)	cogn_1 st	aff_1 st	cogn_1 st	aff_2 nd	Hyper

Note: ToM, Theory of Mind. Scene code consists of letters (i.e., H, honesty, pS, paradoxical sarcasm, sS, simple sarcasm) and the respective scene numbers. Question's identification letter = a–e (a–d = ToM questions, e = hypermentalization question); aff_1st, affective ToM first-order; aff_2nd, affective ToM second-order; cogn_1st, cognitive ToM first-order; cogn_2nd, cognitive ToM second-order; hyper, hypermentalization question.

some methodological and practical limitations of the original test, such as inconsistencies in cinematic realization, ceiling effects in healthy participants, and the long-time of administration. In addition, we created a hypermentalization question for each of the BASIT-ToM scenes.

We identified 10 scenes [1 practice item and 3 scenes apiece for the three message types (i.e., honesty, simple sarcasm and paradoxical sarcasm) that will comprise the BASIT-ToM]. These scenes showed in general neither floor nor ceiling effects in the majority of ToM questions. As evidence of validity, we found that gradual changes in expressed intensities of either honest or sarcastic message types related to difficulties in ToM as measured by participants' correct answering probabilities.

Regarding the hypermentalization question, we were able to reliably analyse its response probabilities in six of the 10 scenes. These questions were in general easy to answer by the healthy participants, which is unsurprising given the fact that these participants are cognitively and mentally healthy. Future studies will show whether patients who hypermentalize such as schizophrenic patients with positive symptoms (Canty et al., 2017) or patients with borderline personality disorder (Sharp et al., 2011) will fail on the hypermentalization questions. In this regard, it is worth mentioning that BASIT-ToM contains less contextual information than those tests [i.e., MASC (Dziobek et al., 2006), VAMA (Canty et al., 2015)] that were able to detect hypermentalization in these two patient groups. It will be interesting to investigate whether these patients do also hypermentalize when having watched BASIT-ToM scenes, which contain little contextual information. If so, then the BASIT-ToM would allow assessing not only decrease (hypomentalization) or loss (no mentalization) of ToM but also excessive ToM (hypermentalization). This is of high clinical value in the assessment of patients with neuropsychological disorders, because as mentioned above, some patient groups hypermentalize (Canty et al., 2017; Sharp et al., 2011), whereas for example patients with behavioural variant frontotemporal dementia, Asperger's syndrome, euthymic bipolar disorder, or schizophrenia with negative symptoms hypomentalize or do not mentalize at all (Bora et al., 2015; Canty et al., 2017; Dziobek et al., 2006; Montag et al., 2010).

Notably, we found no systematic differences in difficulty either between cognitive and affective questions or between first- and second-order ToM questions. The absence of a difference in difficulty between first- and second-order ToM questions may seem surprising as higher order ToM questions tend to be cognitively more demanding as shown among others by recruitment of more ToM-associated brain regions in high-order than low-order ToM questions in cognitively healthy subjects (Lewis et al., 2017). Clinically, this seems to show up primarily in the healthy subjects' reaction times, but not

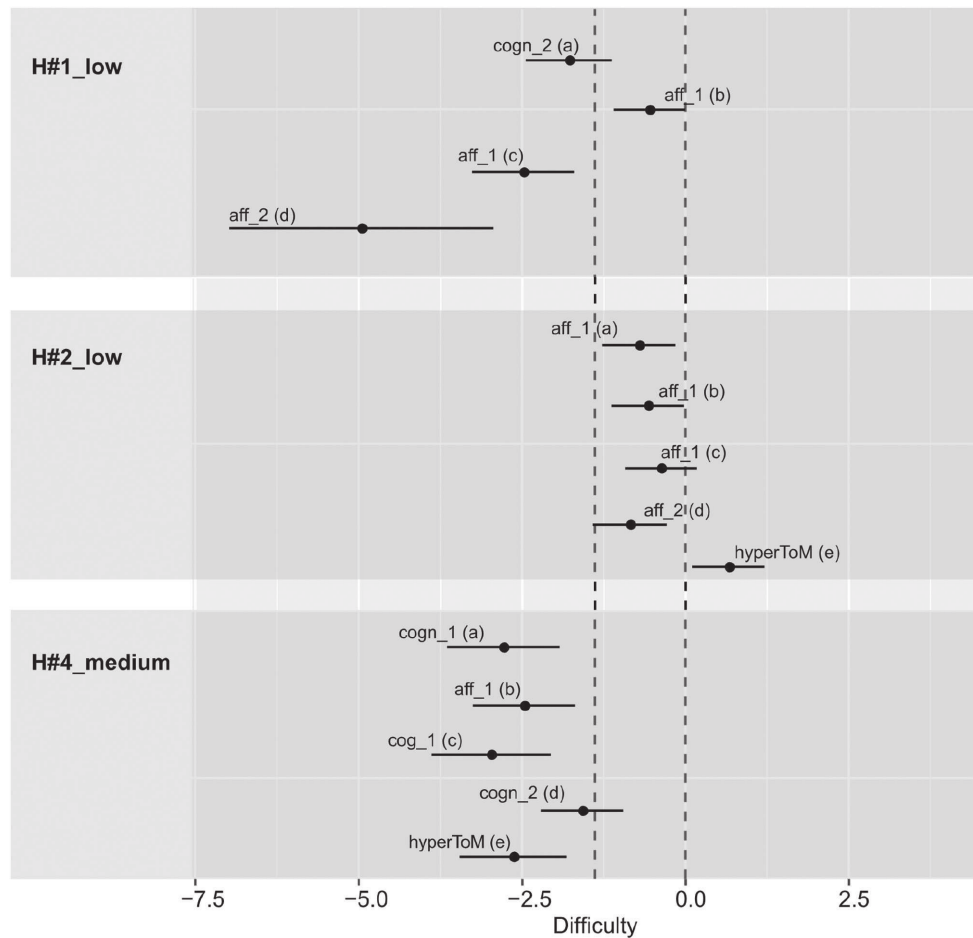


FIGURE 4 95% confidence intervals (CI) of the estimated difficulties of the theory of mind (ToM) and hypermentalization questions (hyperToM) of the three selected BASIT-ToM scenes for message type honesty (i.e., a: Scene H#1_low, b: Scene H#2_low, c: Scene H#4_medium). CI of the hyperToM question is only depicted in presence of a model fit. Scene code consists of a scene number (scene), a letter (message type, H = honesty) and the portrayed intensity (i.e., low or medium). Question code consists of abbreviations of the type (i.e., cogn = cognitive ToM, aff = affective ToM, hyperToM = hypermentalization) and order (i.e., 1 = first-order ToM, 2 = second-order ToM) of ToM, and question's identification letter (a-d = ToM questions, e = hypermentalization question). The dashed lines illustrate the required range of difficulty $[-1.386, 0]$. BASIT-ToM, Basal Version of the Awareness of Social Inference Test – Theory of Mind

in response accuracy (Lewis et al., 2017). In our paradigm, however, we measured the accuracy of responses rather than reaction times. Nevertheless, for individuals with cognitive deficits, for whom the test is designed for, it can be assumed that the higher cognitive load of the second-order ToM questions will be reflected in the test scores.

Participants had the option to rewatch the scenes in case they had not understood the content of the scene. We set this option to get an idea about participants' understanding of the scene play. Notably, most participants did not use this option, which demonstrates the quality of the scene play and contents. Arguably, one of the selected intensity versions was watched at least twice by 42% of the participants. The majority of this scene's ToM questions, however, were within the required range of difficulty, indicating adequate understanding of the scene's content by the participants. We will reanalyse the scene's capacity in measuring ToM in the subsequent validation study with patients with neuropsychological disorders. As we speak of patients with neuropsychological disorders, who make up the target

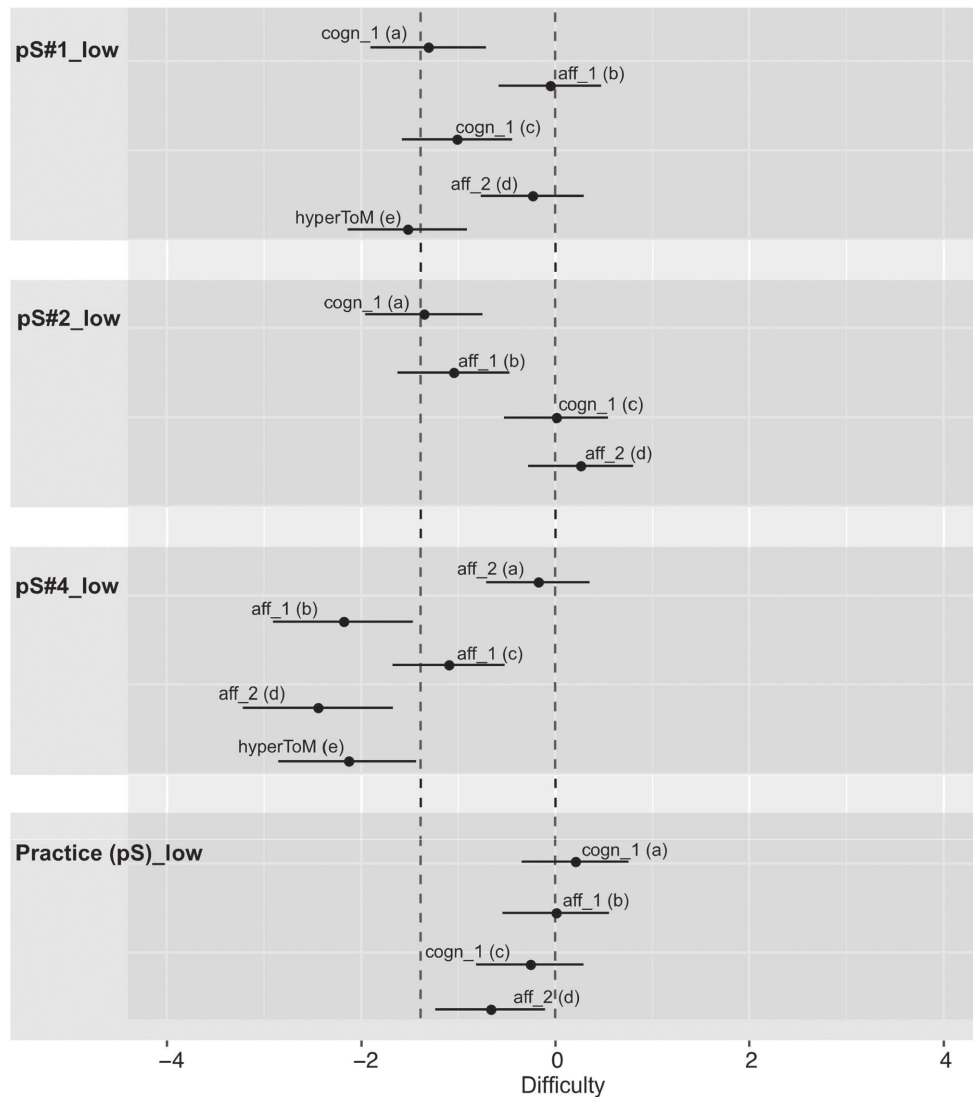


FIGURE 5 95% confidence intervals (CI) of the estimated difficulties of the theory of mind (ToM) and hypermentalization questions (hyperToM) of the four selected BASIT-ToM scenes for message type paradoxical sarcasm [i.e., a: Scene pS#1_low, b: Scene pS#2_low, c: Scene pS#4_low, d: Practice (pS)_low]. CI of the hyperToM question is only depicted in presence of a model fit. Scene code consists of a scene number (scene), a letter (message type, pS = paradoxical sarcasm) and the portrayed intensity (i.e., low). Question code consists of abbreviations of the type (i.e., cogn = cognitive ToM, aff = affective ToM, hyperToM = hypermentalization) and order (i.e., 1 = first-order ToM, 2 = second-order ToM) of ToM, and question's identification letter (a–d = ToM questions, e = hypermentalization question). Practice(pS_low) will be used as a test scene for paradoxical sarcasm in BASIT-ToM, whereas pS#4_low will be used as practice scene. The dashed lines illustrate the required range of difficulty $[-1.386, 0]$. BASIT-ToM, Basel Version of the Awareness of Social Inference Test – Theory of Mind

population for whom we designed the BASIT-ToM, the opportunity to rewatch the scene before answering the questions to be sure about the scene content is critical. This way, we minimize the risk that low ToM test scores are due to cognitive dysfunction other than dysfunction in mentalizing. We agree that no “rewatch button” exists in real life. However, in real life, you are experiencing the unfolding of the scene, which likely facilitates ToM ability.

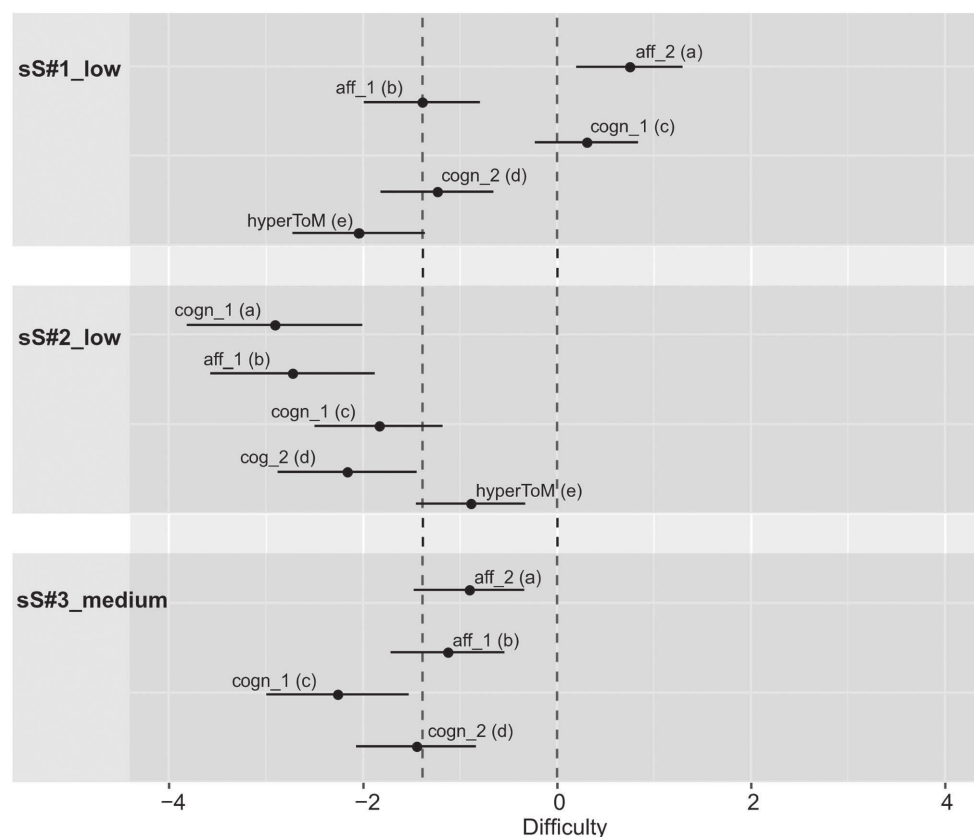


FIGURE 6 95% confidence intervals (CI) of the estimated difficulties of the theory of mind (ToM) and hypermentalization questions (hyperToM) of the three selected BASIT-ToM scenes for message type simple sarcasm (i.e., a: Scene sS#1_low, b: Scene sS#2_low, c: Scene sS#3_medium). CI of the hyperToM question is only depicted in presence of a model fit. Scene code consists of a scene number (scene), a letter (message type, sS = simple sarcasm), and the portrayed intensity (i.e., low or medium). Question code consists of abbreviations of the type (i.e., cogn = cognitive ToM, aff = affective ToM, hyperToM = hypermentalization) and order (i.e., 1 = first-order ToM, 2 = second-order ToM) of ToM, and question's identification letter (a-d = ToM questions, e = hypermentalization question). The dashed lines illustrate the required range of difficulty [-1.386, 0]. BASIT-ToM, Basel Version of the Awareness of Social Inference Test – Theory of Mind

Our study contains some limitations. First, in six of the selected scenes (i.e., H#1_low, H#4_medium, pS#4_low, sS#1_low, sS#2_low, sS#3_medium), some ToM questions were too easy to answer for participants with medium ToM abilities. Thus, these questions may not capture subtle deficits in ToM abilities. We will examine this in the subsequent validation study in a clinical sample and, if needed, remove scenes with too little discriminative power between patients with known ToM deficits such as patients with behavioural variant frontotemporal dementia (Bora et al., 2015; Henry et al., 2014) and healthy individuals.

Second, in four scenes (i.e., H#1_low, pS#2_low, practice(pS)_low, sS#3_medium), the inclusion of the hyperToM question in the data analysis resulted in a loss of model fit of the Rasch model. Likewise, these four questions seemed to be unclear to the participants, showing percentages of correct answers between 11% and 50%. Interestingly, all four questions referred to protagonists' affects ("Is he in love with her?", "Can they not stand Simone?"), whereas the hyperToM questions with an adequate model fit referred to protagonists' intentions. The scenes' paralinguistic features and contextual information were probably insufficient to generate meaningful affective hyperToM questions. Based on these findings, we decided to limit ourselves to cognitive hyperToM questions

in the BASIT-ToM and, accordingly, will replace these four affective hyperToM questions by cognitive hyperToM questions.

Third, first-, and second-order ToM and cognitive and affective ToM were only partly balanced within and between the three message types. Importantly, however, the distribution of ToM types will undergo further refinement if needed, depending on the results of the planned validation study. We will also consider combining the simple and paradoxical sarcasm scores that may result in a more balanced distribution of ToM types within and between message types.

Finally, the large age range of our sample may be considered a limitation, given the evidence of an association between age and ToM (Henry et al., 2016). Similarly, sex and years of education may also influence performance. We therefore compared the distribution of the demographic variables between the selected intensity version scenes and found similar distributions. Accordingly, the influence of demographic variables on test performance appears unlikely. A related point, however, is that given the relatively high educational level of the participants, generalization of our findings to the general population needs to be done with caution.

Similar to the procedure used with the already published BASIT-ER (Jarsch et al., 2022), we plan to validate the BASIT-ToM both in a healthy population and in a clinical sample to examine its reliability, construct and ecological validity, sensitivity in detecting ToM deficits, and the potential influence of cognitive deficits on test performance. Thereby, we will score both, honest and sarcastic exchanges, and the different types and orders of ToM separately to evaluate whether this scoring approach improves discrimination between patients with different brain disorders compared to an overall ToM score. Likewise, we will evaluate the utility of hypermentalization questions to discriminate between patient groups. Based on the results of the validation study, we will set the final version of the BASIT-ToM for later use in clinical settings.

AUTHOR CONTRIBUTIONS

Marianne Jarsch: Conceptualization; data curation; formal analysis; methodology; visualization; writing – original draft; writing – review and editing. **Olivier Piguet:** Writing – review and editing. **Manfred Berres:** Formal analysis; methodology; writing – review and editing. **Constantin Sluka:** Data curation; resources; software; writing – review and editing. **Anna Semenkova:** Formal analysis; writing – review and editing. **Reto W. Kressig:** Writing – review and editing. **Andreas U. Monsch:** Writing – review and editing. **Skye McDonald:** Writing – review and editing. **Marc Sollberger:** Conceptualization; formal analysis; funding acquisition; investigation; methodology; project administration; resources; supervision; validation; writing – original draft; writing – review and editing.

ACKNOWLEDGEMENTS

MJ, AS and MS are supported by the Velux Foundation grant 1225, the Alzheimer Forum Schweiz, the Stiftung für Demenzforschung Basel, the Research-Pool Medicine of the University Hospital Basel, the Sonnweid Stiftung, Merz Pharma (Schweiz), Novartis, AC Immune and Axovant Sciences. OP is supported in part by a National Health and Medical Research Council of Australia Leadership Fellowship (GNT20008020). The authors would like to thank the film production company East End Film GmbH (Germany) and the actors for their professionalism and dedication for the film project. Lastly, the authors would like to thank Ricarda Aucher, Anna-Lisa Krieg, Valentina Arnold, Julia Desax, Nerissa Dettling, Riane Baur, Vera Nina Looser, Anne-Catherine Rueegg, Teresa Müggler, Nathalie Müller, Vaisnavi Varatharajan, Sophie von Achenbach, Andreas Gassilloud and Nina Huber for help in participant recruitment, data collection or data preparation. Open access funding provided by Universität Basel.








CONFLICT OF INTEREST

All authors report no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in “figshare” at <https://figshare.com/s/d7194f6f2f18082e623c>. One exemplary test scene of the BASIT-ToM at three intensity levels of the portrayed paralinguistic cues can be found in “figshare” at <https://figshare.com/s/5ca4ff99790fcc5866db>.

ORCID

Marianne Jarsch  <https://orcid.org/0000-0003-1871-8435>
Olivier Piguet  <https://orcid.org/0000-0002-6696-1440>
Manfred Berres  <https://orcid.org/0000-0003-0409-8554>
Constantin Sluka  <https://orcid.org/0000-0001-7672-4986>
Anna Semenková  <https://orcid.org/0000-0002-6034-5911>
Reto W. Kressig  <https://orcid.org/0000-0002-3215-7681>
Andreas U. Monsch  <https://orcid.org/0000-0003-1692-1560>
Skeye McDonald  <https://orcid.org/0000-0003-0723-6094>
Marc Sollberger  <https://orcid.org/0000-0001-8863-3731>

REFERENCES

- Abu-Akel, A., & Bailey, A. L. (2000). Correspondence. *Psychological Medicine*, *30*(3), 735–738.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5 (Vol. 5)*. American Psychiatric Association.
- Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*(4), 268–277.
- Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA publications and communications board task force report. *American Psychologist*, *73*(1), 3–25.
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*(6), 561–571.
- Bora, E., Walterfang, M., & Velakoulis, D. (2015). Theory of mind in behavioural-variant frontotemporal dementia and Alzheimer's disease: A meta-analysis. *Journal of Neurology, Neurosurgery & Psychiatry*, *86*(7), 714–719.
- Canty, A. L., Neumann, D. L., Fleming, J., & Shum, D. H. (2015). Evaluation of a newly developed measure of theory of mind: The virtual assessment of mentalising ability. *Neuropsychological Rehabilitation*, *27*(5), 834–870.
- Canty, A. L., Neumann, D. L., & Shum, D. H. (2017). Using virtual reality to assess theory of mind subprocesses and error types in early and chronic schizophrenia. *Schizophrenia Research: Cognition*, *10*, 15–19.
- Castelli, I., Pini, A., Alberoni, M., Liverta-Sempio, O., Baglio, F., Massaro, D., Marchetti, A., & Nemni, R. (2011). Mapping levels of theory of mind in Alzheimer's disease: A preliminary study. *Aging & Mental Health*, *15*(2), 157–168.
- Corradi-Dell'Acqua, C., Ronchi, R., Thomasson, M., Bernati, T., Saj, A., & Vuilleumier, P. (2020). Deficits in cognitive and affective theory of mind relate to dissociated lesion patterns in prefrontal and insular cortex. *Cortex*, *128*, 218–233.
- 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.
- 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.
- Fortier, J., Besnard, J., & Allain, P. (2018). Theory of mind, empathy and emotion perception in cortical and subcortical neurodegenerative diseases. *Revue Neurologique*, *174*(4), 237–246.
- Frith, C. D. (2004). Schizophrenia and theory of mind. *Psychological Medicine*, *34*(3), 385–389.
- Genova, H. M., & McDonald, S. (2020). Social cognition in individuals with progressive multiple sclerosis: A pilot study using TASIT-S. *Journal of the International Neuropsychological Society*, *26*(5), 539–544.
- Ghosh, B. C., Calder, A. J., Peers, P. V., Lawrence, A. D., Acosta-Cabronero, J., Pereira, J. M., Hodges, J. R., & Rowe, J. B. (2012). Social cognitive deficits and their neural correlates in progressive supranuclear palsy. *Brain*, *135*(7), 2089–2102.
- Henry, J. D., Phillips, L. H., & Von Hippel, C. (2014). A meta-analytic review of theory of mind difficulties in behavioural-variant frontotemporal dementia. *Neuropsychologia*, *56*, 53–62.
- Henry, J. D., Von Hippel, W., Molenberghs, P., Lee, T., & Sachdev, P. S. (2016). Clinical assessment of social cognitive function in neurological disorders. *Nature Reviews Neurology*, *12*(1), 28–39.

- Honan, C. A., McDonald, S., Sufani, C., Hine, D. W., & Kumfor, F. (2016). The awareness of social inference test: Development of a shortened version for use in adults with acquired brain injury. *The Clinical Neuropsychologist, 30*(2), 243–264. <https://doi.org/10.1080/13854046.2015.1136691>
- Jarsch, M., Bartsch, K., Berres, M., Ryf, I., Jurisic, K., Sluka, C., Kressig, R. W., Monsch, A. U., McDonald, S., Piguet, O., & Sollberger, M. (2022). The Basel Version of the Awareness of Social Inference Test-Emotion Recognition (BASIT-ER): Preliminary validation analyses in healthy adults. *Neuropsychology, 36*(2), 175–184. <https://doi.org/10.1037/neu0000786>
- Kennedy, D. P., & Adolphs, R. (2012). The social brain in psychiatric and neurological disorders. *Trends in Cognitive Sciences, 16*(11), 559–572. <https://doi.org/10.1016/j.tics.2012.09.006>
- Kumfor, F., Honan, C., McDonald, S., Hazelton, J. L., Hodges, J. R., & Piguet, O. (2017). Assessing the “social brain” in dementia: Applying TASIT-S. *Cortex, 93*, 166–177.
- Lancaster, K., Stone, E. M., & Genova, H. M. (2019). Cognitive but not affective theory of mind deficits in progressive MS. *Journal of the International Neuropsychological Society, 25*(8), 896–900.
- Lavoie, M. A., Vistoli, D., Sutliff, S., Jackson, P. L., & Achim, A. M. (2016). Social representations and contextual adjustments as two distinct components of the theory of mind brain network: Evidence from the REMICS task. *Cortex, 81*, 176–191. <https://doi.org/10.1016/j.cortex.2016.04.017>
- Lewis, P. A., Birch, A., Hall, A., & Dunbar, R. I. M. (2017). Higher order intentionality tasks are cognitively more demanding. *Social Cognitive and Affective Neuroscience, 12*(7), 1063–1071. <https://doi.org/10.1093/scan/nsx034>
- Martinez, G., Alexandre, C., Mam-Lam-Fook, C., Bendjemaa, N., Gaillard, R., Garel, P., Dziobek, I., Amado, I., & Krebs, M. O. (2017). Phenotypic continuum between autism and schizophrenia: Evidence from the movie for the assessment of social cognition (MASC). *Schizophrenia Research, 185*, 161–166. <https://doi.org/10.1016/j.schres.2017.01.012>
- McDonald, S. (2012). New frontiers in neuropsychological assessment: Assessing social perception using a standardised instrument, the awareness of social inference test. *Australian Psychologist, 47*(1), 39–48.
- McDonald, S., Bornhofen, C., Shum, D., Long, E., Saunders, C., & Neulinger, K. (2006). Reliability and validity of the awareness of social inference test (TASIT): A clinical test of social perception. *Disability and Rehabilitation, 28*(24), 1529–1542.
- McDonald, S., Fisher, A., Flanagan, S., & Honan, C. A. (2017). Impaired perception of sincerity after severe traumatic brain injury. *Journal of Neuropsychology, 11*(2), 291–304.
- McDonald, S., Fisher, A., Togher, L., Tate, R., Rushby, J., English, T., Kelly, M., Mathersul, D., Froreich, F., & Francis, H. (2015). Adolescent performance on the awareness of social inference test: TASIT. *Brain Impairment, 16*(1), 3–18.
- McDonald, S., Flanagan, S., Martin, I., & Saunders, C. (2004). The ecological validity of TASIT: A test of social perception. *Neuropsychological Rehabilitation, 14*(3), 285–302.
- McDonald, S., Flanagan, S., Rollins, J., & Kinch, J. (2003). TASIT: A new clinical tool for assessing social perception after traumatic brain injury. *The Journal of Head Trauma Rehabilitation, 18*(3), 219–238.
- Meijering, B., Van Rijn, H., Taatgen, N., & Verbrugge, R. (2011). *I do know what you think I think: Second-order theory of mind in strategic games is not that difficult*. Paper presented at the Proceedings of the Annual Meeting of the Cognitive Science Society.
- 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*(1–3), 264–269. <https://doi.org/10.1016/j.jad.2009.08.017>
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695–699.
- Nosek, B., Alter, G., & Banks, G. (2015). SCIENTIFIC STANDARDS. Promoting an open research culture. *Science, 348*(6242), 1422–1425.
- Peirce, J. W. (2007). PsychoPy—Psychophysics software in python. *Journal of Neuroscience Methods, 162*(1–2), 8–13.
- Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics, 2*, 10.
- Poletti, M., Enrici, I., & Adenzato, M. (2012). Cognitive and affective theory of mind in neurodegenerative diseases: Neuropsychological, neuroanatomical and neurochemical levels. *Neuroscience & Biobehavioral Reviews, 36*(9), 2147–2164.
- Pottgen, J., Dziobek, I., Reh, S., Heesen, C., & Gold, S. M. (2013). Impaired social cognition in multiple sclerosis. *Journal of Neurology, Neurosurgery, and Psychiatry, 84*(5), 523–528. <https://doi.org/10.1136/jnnp-2012-304157>
- Quidé, Y., Wilhelmi, C., & Green, M. J. (2020). Structural brain morphometry associated with theory of mind in bipolar disorder and schizophrenia. *PsyCh Journal, 9*(2), 234–246.
- Quinting, J., Jonas, K., Wendt, C., & Stenneken, P. (2020). Social cognition in cognitive-communication disorders following traumatic brain injury: A video-based assessment. *Aphasia Und Verwandte Gebiete, 48*(2), 6–18.
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- R Studio Team. (2016). *RStudio: Integrated development environment for R (1.1.463)*. RStudio, Inc.
- Rankin, K. P., Salazar, A., Gorno-Tempini, M. L., Sollberger, M., Wilson, S. M., Pavlic, D., Stanley, C. M., Glenn, S., Weiner, M. W., & Miller, B. L. (2009). Detecting sarcasm from paralinguistic cues: Anatomic and cognitive correlates in neurodegenerative disease. *NeuroImage, 47*(4), 2005–2015.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Paedagogische Institut.

- Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., van Swieten, J. C., Seelaar, H., Dopper, E. G., Onyike, C. U., Hillis, A. E., Josephs, K. A., Boeve, B. F., Kertesz, A., Seeley, W. W., Rankin, K. P., Johnson, J. K., Gorno-Tempini, M. L., Rosen, H., ... Miller, B. L. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*, *134*(9), 2456–2477.
- Rizopoulos, D. (2006). Ltm: An R package for latent variable modeling and item response analysis. *Journal of Statistical Software*, *17*, 1–25.
- Rossetto, F., Castelli, I., Baglio, F., Massaro, D., Alberoni, M., Nemni, R., Shamay-Tsoory, S., & Marchetti, A. (2018). Cognitive and affective theory of mind in mild cognitive impairment and Parkinson's disease: Preliminary evidence from the italian version of the yoni task. *Developmental Neuropsychology*, *43*(8), 764–780.
- Ryan, N. P., Catroppa, C., Beare, R., Silk, T. J., Hearps, S. J., Beauchamp, M. H., Yeates, K. O., & Anderson, V. A. (2017). Uncovering the neuroanatomical correlates of cognitive, affective and conative theory of mind in paediatric traumatic brain injury: A neural systems perspective. *Social Cognitive and Affective Neuroscience*, *12*(9), 1414–1427.
- Sharp, C., Pane, H., Ha, C., Venta, A., Patel, A. B., Sturek, J., & Fonagy, P. (2011). Theory of mind and emotion regulation difficulties in adolescents with borderline traits. *Journal of the American Academy of Child and Adolescent Psychiatry*, *50*(6), 563–573. <https://doi.org/10.1016/j.jaac.2011.01.017>
- Thomann, A. E., Goettel, N., Monsch, R. J., Berres, M., Jahn, T., Steiner, L. A., & Monsch, A. U. (2018). The Montreal cognitive assessment: Normative data from a German-speaking cohort and comparison with international normative samples. *Journal of Alzheimer's Disease*, *64*(2), 643–655.
- Wang, Y. Y., Wang, Y., Zou, Y. M., Ni, K., Tian, X., Sun, H. W., Lui, S. S. Y., Cheung, E. F. C., Suckling, J., & Chan, R. C. K. (2018). Theory of mind impairment and its clinical correlates in patients with schizophrenia, major depressive disorder and bipolar disorder. *Schizophrenia Research*, *197*, 349–356.
- Westerhof-Evers, H. J., Visser-Keizer, A. C., McDonald, S., & Spikman, J. M. (2014). Performance of healthy subjects on an ecologically valid test for social cognition: The short, Dutch version of the awareness of social inference test (TASIT). *Journal of Clinical and Experimental Neuropsychology*, *36*(10), 1031–1041.
- Yesavage, J. A., & Sheikh, J. I. (1986). 9/geriatric depression scale (GDS) recent evidence and development of a shorter version. *Clinical Gerontologist*, *5*(1–2), 165–173.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix A–K

How to cite this article: Jarsch, M., Piguet, O., Berres, M., Sluka, C., Semenкова, A., Kressig, R. W., Monsch, A. U., McDonald, S., & Sollberger, M. (2023). Development of the Basel Version of the Awareness of Social Inference Test – Theory of Mind (BASIT-ToM) in healthy adults. *Journal of Neuropsychology*, *17*, 125–145. <https://doi.org/10.1111/jnp.12290>

Discussion

The goal of this dissertation was to develop new neuropsychological tools that could increase diagnostic accuracy of bvFTD at the early stage of the disease. The focus was on two diagnostic domains, i.e., behavioural disorders as described in the current diagnostic criteria for bvFTD, and social cognition, particularly its subdomain Theory of Mind (ToM), that is not included in the current diagnostic criteria but is recommended by the latest literature to be assessed if bvFTD diagnosis is suspected. As a result, two diagnostic tools were developed. The first instrument is an other-report questionnaire, the Behavioural Dysfunction Questionnaire (BDQ), operationalising current diagnostic criteria for bvFTD. In the validation study, it showed good discriminatory power between bvFTD and two other brain disorders, i.e., ADD and MDD, that can often be misdiagnosed with bvFTD. Before the development of new social cognition assessment tools, the current situation of the social cognition assessment in German-speaking memory clinics was investigated. A survey of 87 memory clinics revealed the willingness of clinicians to assess social cognition in patients with brain disorders. However, it is still only rarely done during clinical evaluations due to the lack of appropriate diagnostic tools applicable in German-speaking areas and fulfilling the criteria of good neuropsychological tools. To close this gap, the second instrument, the Basel Version of the Awareness of Social Inference Test – Theory of Mind (BASIT-ToM), was developed. It is a video-based performance test assessing ToM ability close to real-life setting.

In the following, the application of the developed instruments in bvFTD diagnosis, their strengths, and limitations as well as further research directions are discussed.

Behavioural Dysfunction Questionnaire (BDQ): The first behavioural scale operationalising the current diagnostic criteria for behavioural variant frontotemporal dementia

The aim of the BDQ was to develop the first behavioural scale operationalising the current diagnostic criteria for bvFTD as exactly as possible. Therefore, BDQ relies strongly on the structure and the examples provided in FTDC-criteria (Rascovsky et al., 2011). The validation study (Study I) was performed in a sample of patients with bvFTD and two its most common misdiagnoses, i.e., ADD and MDD. The aim was to investigate the discriminatory power of the BDQ between bvFTD and two named diseases, respectively. Thanks to the strong conformity with the diagnostic criteria, the findings of the performed validation study can also provide new insights on the current diagnostic criteria and its discriminatory ability.

Insights on the current diagnostic criteria based on the findings in the Study I

BDQ showed good discriminatory power between bvFTD and ADD (AUC=88%) and between bvFTD and MDD (AUC=83%), respectively. However, the sensitivity, based on the Youden-index optimal cut-off scores, was lower than expected. Adjusting the cut-off scores to reach sensitivity of at least 90% led to decrease of the specificity up to 56%. This finding points to the overlap in behavioural symptoms between investigated diseases, especially between bvFTD and MDD patients that is also demonstrated with misdiagnosis rate of bvFTD (Woolley et al., 2011). The current diagnostic criteria were developed with the aim to increase the sensitivity to the early bvFTD symptoms while the specificity was not in focus. Later studies showed heterogeneous findings regarding specificity of the diagnostic criteria especially for possible bvFTD based solely on the evaluation of the behavioural and cognitive changes without biomarkers such as brain images. So, Harris et al. (2013) reported for possible bvFTD the specificity of 82% at the sensitivity of 95% while Vijverberg et al. (2016) reported the specificity of only 27% at the sensitivity of 85%. The findings of the BDQ validation study lie in between the findings of these two studies. However, this inconsistency in reported specificity of the current diagnostic criteria with the tendency to the low values highlights the demand on revising the current criteria in order to increase the diagnosis accuracy of bvFTD.

Another important finding was that inclusion of the time criterion “early” (the presentation of the behaviours within the three years after symptoms’ onset) highlighted in the FTDC-criteria for four of the five behavioural domains did not improve the diagnostic accuracy of the bvFTD. In turn, the inclusion of this time criterion resulted in lower discriminatory power of the BDQ. The reasons for that are discussed in the Study I. Importantly, this finding queries the relevance of the time aspect in the diagnostic criteria in general and provides a basis for more research on this topic. Possibly, it can be enough to ensure that a behaviour is novel without restriction to the exact time of the first occurrence.

Clinical implication

While bvFTD typical behavioural symptoms can be present also in other brain disorders, an important discriminatory feature is whether these symptoms progress over time. That’s why repeated clinical examination is advisable. So, the behavioural symptoms usually worsen in bvFTD, improve in psychiatric disorders and stay stable in other neurodegenerative disorders (Reus et al., 2018). To be able to evaluate symptoms’ development between different assessments, standardized assessment tools are necessary. The BDQ is the first tool allowing the comparison of the behavioural symptoms according to the FTDC-criteria over time and among different clinicians that can change between assessment points.

Another strength of the BDQ designed in form of an informant questionnaire is the possibility to evaluate different behavioural symptoms over the course of and in different contexts in a time-efficient way. However, it is important to keep in mind that the informants' evaluations can depend on their stress level, personality and sensitivity to patients' behavioural changes (Rascovsky & Grossman, 2013). That's why it would be useful to assess this information from several informants to compare different perspectives. In this context, it can be welcomed to provide a digital version of the questionnaire that can be filled out by informants at any time and at any place increasing their willingness to provide information.

Lastly, though initially developed for assessment of the early bvFTD behavioural symptoms, BDQ can be also used in evaluation of the behavioural disorders in other diseases due to the normative values for each behavioural domain and the total score developed gender-specific in a sample of 414 healthy controls. Indeed, it has been already successfully implemented in the clinical routine of the Memory Clinic Basel and is occasionally used in other clinical institutions.

Further research

To increase discriminatory power of the BDQ, a second validation study is planned. It intends to investigate whether the discriminatory power of the BDQ can be increased if adjusting the total score depending on the differential diagnoses since the discriminatory power of different behaviours' symptoms (items) might differ depending on the comparison disease. Such an approach is reasonable for the clinical purposes since clinicians usually have two or three suspected diagnoses and a more precise discriminating score can be helpful in such situations. The second study will be conducted in an extended sample of the patients with bvFTD (estimated N = 50), ADD (estimated N = 120) and MDD (estimated N = 80) diagnoses. The two main aims of the study are: 1) to replicate the results of the first validation study, and 2) to increase the discriminatory power of the BDQ adjusting the total score by the use of a data-driven approach. For that, the items with the best discriminatory power between bvFTD and other differential diagnosis (ADD or MDD) will be identified and only these will be included in the final BDQ score. It is expected that exclusion of the overlapping behavioural symptoms will increase the diagnosis accuracy. The planned sample size is considered as sufficient to ensure the generalisability of the calculated cut-offs to the general clinical population. The generalisability of the results will be additionally evaluated with cross-validation analysis. Next, similar studies with other diagnoses versus bvFTD (e.g., behavioural variant Alzheimer disease (bvAD) (Ossenkoppele et al., 2022) or bipolar disorder (American

Psychiatric Association, 2013)) can be performed. The results of these studies can be used as a basis for eventually revision of the current diagnostic criteria to increase their specificity.

Lastly, filling out BDQ by patients themselves and their informants can be used as a tool to evaluate patients' insight into their disease and symptoms. The reliability and validity of this method need to be evaluated in the future studies.

Assessment of social cognition at the German-speaking memory clinics

Social cognition has been acknowledged by diagnostic guidelines as a cognitive domain that needs to be evaluated together with other brain functions in diagnosis of brain disorders. The aim of the Study II was to evaluate whether and how these recommendations have been implemented in the clinical routine in the German-speaking area. For that an online survey of 87 memory clinics in Germany, Switzerland and Austria was conducted. The results of this survey showed that clinicians agree about the importance of social cognition in brain disorders and show willingness to assess this. However, it is still only rarely measured in the clinical routine. Similar results showed a survey with French neuropsychologists (Quesque et al., 2022). In this study, about 93% of the participants agreed that social cognition has the same importance in a neuropsychological evaluation as other cognitive domains, however only 8% confirmed that they evaluate social cognition on a daily basis (Quesque et al., 2022). This situation can be explained from several perspectives, e.g., lack of guidelines what exactly and how should be assessed (Samtani et al., 2023), poor training of clinicians on social cognition assessment (Quesque et al., 2022) or lack of appropriate tools (Eddy, 2019).

The last reason was closely investigated in the Study II assessing the position of practitioners to the current social cognition tools and their requirements on the tools to use these in the clinical evaluations. Indeed, many clinicians confirmed the lack of diagnostic tools evaluating social cognition that is also often discussed in literature. Though tens of social cognition tools are available, most of them were developed for research purposes and do not fulfil the requirements for diagnostic instruments. So, many instruments lack psychometric evaluations, proven associations with questioned brain functions and normative values (Eddy, 2019). Moreover, many of the instruments are not validated in patients with brain disorders. The last is especially important for the tools assessing ToM since many existing tools demand intact skills in other cognitive domains, like memory or language comprehension (Magno, Canu, Agosta, et al., 2022) that are often impaired in patients with brain disorders and can result in low performance on typical ToM tests. Also, stimuli often used in social cognition assessment tasks are questioned. Most of them are unimodal (either visual or auditory), often using static pictures or even pure texts (Henry et al., 2015). Such stimuli are far from the real-

life social interactions and result in low ecological validity – a psychometric parameter referring to the extent to which test scores reflect real-life skills (Osborne-Crowley, 2020). That is especially important in clinical evaluation aiming to identify the impairments in daily functionalities. To increase ecological validity the test setting should be as close to the real-life situations as possible. Current research recommends using dynamic, multimodal (visual and auditory) and context-embedded stimuli in social cognition assessment (Khosdelazad et al., 2020; Osborne-Crowley, 2020).

Summing up, not only current literature and diagnostic guidelines but also practitioners agree on the importance of social cognition in brain disorders. However, good diagnostic tools fulfilling requirements of clinical practice are still lacking. Development of such instruments can promote evaluation of social cognition during clinical assessments. In diagnosis of bvFTD this can play a crucial role, since social cognition deficits in the patients often occur even before other cognitive or behavioural symptoms (Torralva et al., 2009). But social cognition also plays an important role in other disorders. A systematic review of 31 meta-analyses showed that social cognition deficits vary a lot between different brain disorders (Cotter et al., 2018) and can in this way provide valuable information for differential diagnosis.

Following the demand of practitioners of new diagnostic tools and the intent to promote social cognition assessment in the clinical routine, a new assessment tool for Theory of Mind, one of the main social cognition subdomains was developed, namely the BASIT-ToM. Its strengths and applicability in clinical routine, and particularly in bvFTD diagnosis, are discussed in the next chapter.

Basel Version of the Awareness of Social Inference Test – Theory of Mind (BASIT-ToM): A reality-close assessment tool of social cognition subdomain Theory of Mind

The aim of the BASIT-ToM was to overcome the described above limitations of the existing tools addressing specific needs of the clinical practice. The integration of the BASIT-ToM in clinical routine consists of two phases: 1) test development, including stimuli selection, and 2) test validation in a sample of healthy adults and patients with bvFTD and other brain disorders (ADD, bvAD and MDD). Within this dissertation the first phase, i.e., test development, was performed.

Development phase

The conducted development phase aimed to overcome several often criticised points on the existing tools for ToM assessment as discussed in the Study III. One of the main aims was to ensure the ecological validity of the stimuli. For that, short video scenes were selected as test stimuli. Such stimuli allow perception of information on different modalities (visual and

auditory) and in dynamic, as recommended by the latest literature (Osborne-Crowley, 2020). Moreover, they additionally provide situational context that is especially important for the diagnosis of disorders involving frontal lobe damages, such as bvFTD (Ibañez & Manes, 2012). For higher ecological validity the literature recommends using interactive tasks or at least open answer format (Osborne-Crowley, 2020). However, such an approach is hardly applicable in clinical routine due to high costs, time effort and difficulties in standardisation and calculation of the normative values that is crucial for diagnosis. The used approach with videos displaying different daily situations and following yes/no questions to the played scenes is a good balance between ecological validity and tool's standardisation.

Another challenge associated with increasing the ecological validity is that use of reality-close stimuli (e.g., videos) increases the cognitive demand of the task. It is often criticized in the literature that ToM tasks strongly rely on other cognitive skills, like memory or language comprehension (Magno, Canu, Agosta, et al., 2022). However, any attempt to simplify the tasks in order to decrease cognitive demand has the consequence of decreasing the ecological validity. Since the primary aim of the BASIT-ToM is to identify the early signs of the impaired social cognition, it is to assume that the targeted patients' population has sufficient or only slightly declined other cognitive skills up to that time. Therefore, high cognitive demand of the stimuli should not be a problem in the targeted patients' population. Nevertheless, the BASIT-ToM Control Task was developed additionally to the main test. It contains one video scene similar to the ones of the BASIT-ToM followed by four multiple-choice questions about scene context and content. The aim of this task is to ensure that a tested person is able to understand language and the situational context of the videos and can keep this information in mind long enough to answer the following after scenes questions. Besides that, the sincere scenes can be seen as additional control tasks since sincere communication is usually well understood also by patients with reduced ToM skills (Kumfor et al., 2017; Rankin et al., 2009; Shany-Ur et al., 2012).

Noteworthy is the diversity of the BASIT-ToM questions assessing different components of the ToM, i.e., affective vs cognitive ToM, first- and second-order, and hypermentalising. If the importance of the assessment of both the cognitive and the affective components of the ToM separately was also acknowledged by other instruments (Henry et al., 2015), only very few tools can additionally examine the hypermentalising (Eddy, 2019). Implementation of all ToM components in one test enables their better comparison due to identical stimuli and test setting. Such differential assessment of the different ToM components is beneficial both for more accurate differential diagnosis and tracking of the progressions of different diseases (Gregory et al., 2002; Kumfor et al., 2014; Poletti et al., 2012; Torralva et al., 2015).

To address additional requirements of the clinical practice, i.e., psychometric characteristics, normative values, and feasibility in patients with brain disorders, the next validation phase is planned.

Validation phase

Currently, a validation study in a sample of 210 healthy German-speaking adults is being conducted. The aims of this study are (1) to provide the psychometric characteristics of the BASIT-ToM (i.e., convergent, divergent and ecological validity as well as test-retest reliability), and (2) to calculate normative values depending on gender, age and years of education, the demographic characteristics often shown to influence ToM performance (Białecka-Pikul et al., 2017; Henry et al., 2013; Li et al., 2013).

Recently, another validation study with patients has started. It includes patients with bvFTD, ADD, bvAD and MDD. The aims of this study are (1) to examine the sensitivity of the BASIT-ToM to reduced ToM abilities in patients with bvFTD if compare with healthy controls, (2) to investigate its discriminatory power between different patients' groups, and (3) to ensure the feasibility of the test in patients with brain disorders, including once with reduced cognitive skills. The BASIT-ToM was developed based on the Social Inference – Minimal subtest of The Awareness of Social Inference Test (TASIT, McDonald et al., 2003) which has proven good psychometric characteristics (McDonald et al., 2006; McDonald et al., 2004). Different language versions of the TASIT showed good sensitivity to ToM impairments, in patients with different brain disorders (e.g., traumatic brain injuries (McDonald & Flanagan, 2004), schizophrenia (Sparks et al., 2010), autism spectrum disorder (Mathersul et al., 2013) and also bvFTD (Buhl et al., 2013; Kipps et al., 2009; Kumfor et al., 2014)). Therefore, it is to assume that the BASIT-ToM will also show good psychometric characteristics and sensitivity to ToM impairments in the running validation studies.

After conducting the planned validation studies, the BASIT-ToM will be ready for clinical use to evaluate ToM abilities in patients with brain disorders and particularly in ones with bvFTD diagnosis.

Contribution of the developed tools to diagnosis of the behavioural variant frontotemporal dementia and their critical evaluation

The presented tools were developed with the primary aim to increase diagnosis accuracy of the bvFTD at the early stage. The instruments were developed in accordance with good psychometric practice and considering the needs of clinical routine. Based on the results of the Study I, the BDQ has been implemented since April 2022 in the clinical routine of the Memory clinic Basel. The first evaluation of psychometric characteristics of the BASIT-ToM as well as

calculation of the normative values are planned for the beginning of the year 2025. After that, BASIT-ToM can also be implemented in clinical evaluation of the patients with brain disorders. Indeed, it was already requested by several institutions in Switzerland and Germany for use with research or clinical purposes.

The implementation of the developed tools in the clinical routine will:

- increase diagnosis accuracy of the bvFTD at the early stage,
- build a fundament for establishment of new valid biomarkers for the bvFTD,
- allow evaluation of the changes in behavioural symptoms and social cognition, namely its subdomain ToM, on a standardised way that could be useful to estimate both disease progression but also success of the applied therapies,
- promote examination of the behavioural disorders and ToM in bvFTD and other brain disorders.

Additionally, broad use of the standardised tools in assessment of the behavioural disorders and ToM could provide more data and promote extensive research on these domains in bvFTD but also in other disorders. This can result in revision of the current diagnostic criteria for bvFTD in order to increase their specificity. The promoted research could also extend our knowledge and understanding of the behavioural disorders and ToM in different diseases, including their causes and consequences.

However, it is important to remember that though neuropsychological assessment provides clinicians with valuable and fundamental information on patients' symptoms, biomarkers, such as brain imaging, are certainly needed to increase diagnosis accuracy (Dodich et al., 2018). Another information to keep in mind is that a single test can never fully represent the brain's functioning and pathology. That's why a group of tests integrated in test batteries are needed for more elaborate representations of brain functions (Russell, 2012). A single diagnostic instrument also has its limitations. So, the BDQ is a subjective method relying strongly on informants' perception and interpretation of patients' behaviours. And the BASIT-ToM provides information about the patients' performance only in the certain time point without knowing their previous or typical performance on ToM. Lastly, the calculated cut-offs and normative values should be always interpreted carefully since they are calculated in a sample, and the entire population can never be sampled. Due to usually strict selection of study participants during screening, the sample populations tend to be overnormal and not average (Russell, 2012, pp. 96-97). Moreover, the motivation of the study's participants is difficult to control and can also influence the reliability of the data with which the cut-offs or normative values are developed. That's why further replication and cross-validations studies with developed instruments are needed to ensure the generalisability of the results.

Outlook in further research and practical implications

As mentioned above, the developed instruments can provide a good basis for further research on behavioural disorders and social cognition impairments in bvFTD and other brain disorders by enabling a standardised and reality-close assessment of these symptoms and promoting their assessment in clinical routine. The first already planned studies with the BDQ and BASIT-ToM were described above and aim mainly instruments' validations verifying their ability to identify bvFTD patients at the early stage and to classify them correctly if comparing with other patients' groups. Currently, only four patients' groups are planned, i.e., bvFTD, ADD, bvAD and MDD. But research with other patients' groups is also important, especially the ones symptomatically close to bvFTD, such as late-onset schizophrenia or late-life attention deficit hyperactivity disorder (Pose et al., 2013).

Next, longitudinal studies with the patients fulfilling criteria for prodromal bvFTD (Barker et al., 2022) could verify the sensitivity of the developed instruments to the first symptoms even before the bvFTD diagnosis, but also help to better understand the symptoms' progression starting from their onset. The early identification of the symptoms is important to apply precise symptoms' therapy as soon as possible.

Further studies including brain-images data are necessary to provide information whether BDQ domain-specific scores and scores of the BASIT-ToM are associated with the specific brain pathologies. As described above different behavioural domains are associated with different brain regions. So, it needs to be investigated whether a high score in e.g., apathy as reported by informants is associated with the atrophy of the corresponding brain area. Or whether BASIT-ToM scores on affective and cognitive ToM are able to discriminate between different brain pathologies associated with these skills.

The use of standardised and reality-close instruments for both behavioural symptoms and ToM can also contribute to investigation of the association between social cognition impairments and behavioural symptoms in bvFTD. It is often discussed in the literature that such behavioural symptoms of bvFTD as disinhibition, loss of empathy but also apathy could be caused (at least partly) by impairments of social cognition (Desmarais et al., 2018; Dilcher et al., 2023; Johnen & Bertoux, 2019). However, appropriate assessment tools were lacking to investigate these associations deeper. BDQ and BASIT-ToM can help to approach this topic. Insights in these associations can be especially interesting for the development of new therapies for bvFTD symptoms. So, Cotelli et al. (2018) demonstrated that stimulation of certain brain areas can influence ToM ability in bvFTD patients. Also, medications studies, for example including oxytocin, showed promising improvements in social cognition and behaviour in

bvFTD patients (Finger et al., 2015; Jesso et al., 2011). The effects of social cognition therapies on behavioural symptoms in bvFTD patients requires further research, especially in longitudinal studies. The BDQ and the BASIT-ToM enabling standardised assessment between different time points can be used in such studies to monitor therapies' effects.

In addition to the extensive research on bvFTD diagnostic methods or social cognition assessment, it is crucial to ensure the knowledge transfer to the clinical routine. So, study of Shinagawa et al. (2016) showed that misdiagnoses of bvFTD are very often among non-bvFTD specialists mainly due to lack of familiarity with core diagnostic symptoms. And the survey of Quesque et al. (2022) about social cognition assessment among French neuropsychologists revealed that only 35.7% of professionals who graduated more than five years ago received a training on social cognition assessment. Consequently, the reported confidence when assessing social cognition was on average 2.7 (out of 5), compared to 4.3 when assessing memory or executive functions (Quesque et al., 2022). This illustrates the need for more training on social cognition assessment and bvFTD diagnostic methods for not only clinicians in education, but also for professionals with many years of practice experiences. Only ensuring that research findings will be implemented in clinical practice can improve diagnosis accuracy of the bvFTD and other brain disorders.

Conclusion

Within this dissertation two novel diagnostic tools for bvFTD were developed, i.e., (1) BDQ assessing bvFTD typical behavioural symptoms reported by patients' close persons, and (2) BASIT-ToM, a performance-based test assessing ToM abilities close to real-life setting. These instruments on one side operationalise the current diagnostic criteria for bvFTD and on other side consider new recommendations on additional diagnostic domains to be evaluated in brain disorders, and particularly in bvFTD. The implementation of the developed tools in the clinical routine can increase diagnosis accuracy of the bvFTD and contribute to better understanding of the development and progression of the bvFTD symptoms. The novel knowledge can promote the revision of the current diagnostic criteria that can result in higher diagnostic accuracy, earlier symptoms' therapies and better well-being of the patients and their close persons.

References

- Adolphs, R. (2001). The neurobiology of social cognition. *Current opinion in neurobiology*, *11*(2), 231-239. [https://doi.org/10.1016/S0959-4388\(00\)00202-6](https://doi.org/10.1016/S0959-4388(00)00202-6)
- Ahmed, R. M., Halliday, G., & Hodges, J. R. (2021). Hypothalamic symptoms of frontotemporal dementia disorders. *Handbook of clinical neurology*, *182*, 269-280. <https://doi.org/10.1016/B978-0-12-819973-2.00019->
- Ahmed, R. M., Irish, M., Kam, J., Van Keizerswaard, J., Bartley, L., Samaras, K., . . . Piguet, O. (2014). Quantifying the eating abnormalities in frontotemporal dementia. *JAMA neurology*, *71*(12), 1540-1546. <https://doi:10.1001/jamaneurol.2014.1931>
- Ahmed, R. M., Latheef, S., Bartley, L., Irish, M., Halliday, G. M., Kiernan, M. C., . . . Piguet, O. (2015). Eating behavior in frontotemporal dementia: Peripheral hormones vs hypothalamic pathology. *Neurology*, *85*(15), 1310-1317. <https://doi.org/10.1212/WNL.0000000000002018>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (Vol. 5): American psychiatric association Washington, DC.
- 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>
- Arioli, M., Crespi, C., & Canessa, N. (2018). Social cognition through the lens of cognitive and clinical neuroscience. *BioMed research international*, *2018*. <https://doi.org/10.1155/2018/4283427>
- Baez, S., Pinasco, C., Roca, M., Ferrari, J., Couto, B., García-Cordero, I., . . . Matallana, D. (2019). Brain structural correlates of executive and social cognition profiles in behavioral variant frontotemporal dementia and elderly bipolar disorder. *Neuropsychologia*, *126*, 159-169. <https://doi.org/10.1016/j.neuropsychologia.2017.02.012>
- Barbosa, I. G., Leite, F. d. M. C., Bertoux, M., Guimarães, H. C., Mariano, L. I., Gambogi, L. B., . . . de Souza, L. C. (2023). Social cognition across bipolar disorder and behavioral-variant frontotemporal dementia: an exploratory study. *Brazilian Journal of Psychiatry*, *45*, 132-136. <https://doi.org/10.47626/1516-4446-2022-2935>
- Barker, M. S., Gottesman, R. T., Manoochchri, M., Chapman, S., Appleby, B. S., Brushaber, D., . . . Fields, J. A. (2022). Proposed research criteria for prodromal behavioural variant frontotemporal dementia. *Brain*, *145*(3), 1079-1097. <https://doi.org/10.1093/brain/awab365>
- Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger

- syndrome. *Journal of Child psychology and Psychiatry*, 38(7), 813-822. <https://doi.org/10.1111/j.1469-7610.1997.tb01599.x>
- Barth, J. T., Pliskin, N., Axelrod, B., Faust, D., Fisher, J., Harley, J. P., . . . Ricker, J. (2003). Introduction to the NAN 2001 definition of a clinical neuropsychologist NAN policy and planning committee. *Archives of Clinical Neuropsychology*, 18(5), 551-555. <https://doi.org/10.1093/arclin/18.5.551>
- Bertoux, M., de Souza, L. C., O'Callaghan, C., Greve, A., Sarazin, M., Dubois, B., & Hornberger, M. (2016). Social cognition deficits: The key to discriminate behavioral variant frontotemporal dementia from Alzheimer's disease regardless of amnesia? *Journal of Alzheimer's Disease*, 49(4), 1065-1074. <https://doi.org/10.3233/JAD-150686>
- Bertoux, M., Delavest, M., de Souza, L. C., Funkiewiez, A., Lépine, J.-P., Fossati, P., . . . Sarazin, M. (2012). Social cognition and emotional assessment differentiates frontotemporal dementia from depression. *Journal of Neurology, Neurosurgery & Psychiatry*, 83(4), 411-416. <https://doi.org/10.1136/jnnp-2011-301849>
- Bertoux, M., Flanagan, E. C., Hobbs, M., Ruiz-Tagle, A., Delgado, C., Miranda, M., . . . Hornberger, M. (2018). Structural anatomical investigation of long-term memory deficit in behavioral frontotemporal dementia. *Journal of Alzheimer's Disease*, 62(4), 1887-1900. <https://doi.org/10.3233/JAD-170771>
- Białocka-Pikul, M., Kołodziejczyk, A., & Bosacki, S. (2017). Advanced theory of mind in adolescence: Do age, gender and friendship style play a role? *Journal of adolescence*, 56, 145-156. <https://doi.org/10.1016/j.adolescence.2017.02.009>
- Bora, E., Walterfang, M., & Velakoulis, D. (2015). Theory of mind in behavioural-variant frontotemporal dementia and Alzheimer's disease: a meta-analysis. *Journal of Neurology, Neurosurgery & Psychiatry*, 86(7), 714-719. <https://doi.org/10.1136/jnnp-2014-309445>
- Bozeat, S., Gregory, C. A., Ralph, M. A. L., & Hodges, J. R. (2000). Which neuropsychiatric and behavioural features distinguish frontal and temporal variants of frontotemporal dementia from Alzheimer's disease? *Journal of Neurology, Neurosurgery & Psychiatry*, 69(2), 178-186. <https://doi.org/10.1136/jnnp.69.2.178>
- Broe, M., Hodges, J., Schofield, E., Shepherd, C., Kril, J., & Halliday, G. (2003). Staging disease severity in pathologically confirmed cases of frontotemporal dementia. *Neurology*, 60(6), 1005-1011. <https://doi.org/10.1212/01.wnl.0000052685.09194.39>
- Buhl, C., Stokholm, J., & Gade, A. (2013). Clinical utility of short social cognitive tests in early differentiation of behavioral variant frontotemporal dementia from Alzheimer's disease.

- Dementia and geriatric cognitive disorders extra*, 3(1), 376-385.
<https://doi.org/10.1159/000355123>
- Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A. R., Langner, R., & Eickhoff, S. B. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Structure and Function*, 217, 783-796.
<https://doi.org/10.1007/s00429-012-0380-y>
- Cotelli, M., Adenzato, M., Cantoni, V., Manenti, R., Alberici, A., Enrici, I., . . . Padovani, A. (2018). Enhancing theory of mind in behavioural variant frontotemporal dementia with transcranial direct current stimulation. *Cognitive, Affective, & Behavioral Neuroscience*, 18, 1065-1075. <https://doi.org/10.3758/s13415-018-0622-4>
- 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>
- Critchley, H. D. (2005). Neural mechanisms of autonomic, affective, and cognitive integration. *Journal of comparative neurology*, 493(1), 154-166. <https://doi.org/10.1002/cne.20749>
- Cummings, J. L., Mega, M., Gray, K., Rosenberg-Thompson, S., Carusi, D. A., & Gornbein, J. (1994). The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. *Neurology*, 44(12), 2308-2308.
<https://doi.org/10.1212/wnl.44.12.2308>
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *Journal of Personality and Social Psychology*, 10, 85. Retrieved from https://www.uv.es/~friasnav/Davis_1980.pdf
- Delbeuck, X., Pollet, M., Pasquier, F., Bombois, S., & Moroni, C. (2022). The clinical value of the faux pas test for diagnosing behavioral-variant frontotemporal dementia. *Journal of Geriatric Psychiatry and Neurology*, 35(1), 62-65.
<https://doi.org/10.1177/0891988720964253>
- Dermody, N., Wong, S., Ahmed, R., Piguet, O., Hodges, J. R., & Irish, M. (2016). Uncovering the neural bases of cognitive and affective empathy deficits in Alzheimer's disease and the behavioral-variant of frontotemporal dementia. *Journal of Alzheimer's Disease*, 53(3), 801-816. <https://doi.org/10.3233/JAD-160175>
- Desmarais, P., Lanctôt, K. L., Masellis, M., Black, S. E., & Herrmann, N. (2018). Social inappropriateness in neurodegenerative disorders. *International psychogeriatrics*, 30(2), 197-207. <https://doi.org/10.1017/S1041610217001260>

- Dilcher, R., Malpas, C. B., O'Brien, T. J., & Vivash, L. (2023). Social Cognition in Behavioral Variant Frontotemporal Dementia and Pathological Subtypes: A Narrative Review. *Journal of Alzheimer's Disease*(Preprint), 1-20. <https://doi.org/10.3233/JAD-221171>
- Dodich, A., Cerami, C., Cappa, S. F., Marcone, A., Golzi, V., Zamboni, M., . . . Iannaccone, S. (2018). Combined socio-behavioral evaluation improves the differential diagnosis between the behavioral variant of frontotemporal dementia and Alzheimer's disease: in search of neuropsychological markers. *Journal of Alzheimer's Disease*, *61*(2), 761-772. <https://doi.org/10.3233/JAD-170650>
- Dodich, A., Crespi, C., Santi, G. C., Cappa, S. F., & Cerami, C. (2021). Evaluation of discriminative detection abilities of social cognition measures for the diagnosis of the behavioral variant of frontotemporal dementia: a systematic review. *Neuropsychology Review*, *31*, 251-266. <https://doi.org/10.1007/s11065-020-09457-1>
- Ducharme, S., Dols, A., Laforce, R., Devenney, E., Kumfor, F., Van Den Stock, J., . . . Vijverberg, E. (2020). Recommendations to distinguish behavioural variant frontotemporal dementia from psychiatric disorders. *Brain*, *143*(6), 1632-1650. <https://doi.org/10.1093/brain/awaa018>
- 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*, 623-636. <https://doi.org/10.1007/s10803-006-0107-0>
- Eddy, C. M. (2019). What do you have in mind? Measures to assess mental state reasoning in neuropsychiatric populations. *Frontiers in Psychiatry*, *10*, 459572. <https://doi.org/10.3389/fpsy.2019.00425>
- Englund, B., Brun, A., Gustafson, L., Passant, U., Mann, D., Neary, D., & Snowden, J. (1994). Clinical and neuropathological criteria for frontotemporal dementia. *J Neurol Neurosurg Psychiatry*, *57*(4), 416-418. <https://doi.org/10.1136/jnnp.57.4.416>
- Eslinger, P., Dennis, K., Moore, P., Antani, S., Hauck, R., & Grossman, M. (2005). Metacognitive deficits in frontotemporal dementia. *Journal of Neurology, Neurosurgery & Psychiatry*, *76*(12), 1630-1635. <https://doi.org/10.1136/jnnp.2004.053157>
- Fernandez-Duque, D., Baird, J. A., & Black, S. E. (2009). False-belief understanding in frontotemporal dementia and Alzheimer's disease. *Journal of clinical and experimental neuropsychology*, *31*(4), 489-497. <https://doi.org/10.1080/13803390802282688>
- Fernández-Matarrubia, M., Matías-Guiu, J. A., Cabrera-Martín, M. N., Moreno-Ramos, T., Valles-Salgado, M., Carreras, J. L., & Matías-Guiu, J. (2017). Episodic memory

- dysfunction in behavioral variant frontotemporal dementia: a clinical and FDG-PET study. *Journal of Alzheimer's Disease*, 57(4), 1251-1264. <https://doi.org/10.3233/JAD-160874>
- Ferretti, V., & Papaleo, F. (2019). Understanding others: emotion recognition in humans and other animals. *Genes, Brain and Behavior*, 18(1), e12544. <https://doi.org/10.1111/gbb.12544>
- Finger, E. C., MacKinley, J., Blair, M., Oliver, L. D., Jesso, S., Tartaglia, M. C., . . . Pasternak, S. (2015). Oxytocin for frontotemporal dementia: a randomized dose-finding study of safety and tolerability. *Neurology*, 84(2), 174-181. <https://doi.org/10.1212/WNL.0000000000001133>
- Frith, C. D. (2004). Schizophrenia and theory of mind. *Psychological medicine*, 34(3), 385-389. <https://doi.org/10.1017/s0033291703001326>
- Frith, C. D. (2008). Social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1499), 2033-2039. <https://doi.org/10.1098/rstb.2008.0005>
- Giovagnoli, A. R., Bell, B., Erbetta, A., Paterlini, C., & Bugiani, O. (2019). Analyzing theory of mind impairment in patients with behavioral variant frontotemporal dementia. *Neurological Sciences*, 40, 1893-1900. <https://doi.org/10.1007/s10072-019-03911-6>
- Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., . . . Boeve, B. F. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, 76(11), 1006-1014. <https://doi.org/10.1212/WNL.0b013e31821103e6>
- Gregory, C., Lough, S., Stone, V., Erzinclioglu, S., Martin, L., Baron-Cohen, S., & Hodges, J. R. (2002). Theory of mind in patients with frontal variant frontotemporal dementia and Alzheimer's disease: theoretical and practical implications. *Brain*, 125(4), 752-764. <https://doi.org/10.1093/brain/awf079>
- Happé, F. G. (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>
- Harris, J. M., Gall, C., Thompson, J. C., Richardson, A. M., Neary, D., Du Plessis, D., . . . Jones, M. (2013). Sensitivity and specificity of FTDC criteria for behavioral variant frontotemporal dementia. *Neurology*, 80(20), 1881-1887. <https://doi.org/10.1212/WNL.0b013e318292a342>
- Healey, M. L., & Grossman, M. (2018). Cognitive and affective perspective-taking: evidence for shared and dissociable anatomical substrates. *Frontiers in Neurology*, 9, 372314. <https://doi.org/10.3389/fneur.2018.00491>

- Henry, J. D., Cowan, D. G., Lee, T., & Sachdev, P. S. (2015). Recent trends in testing social cognition. *Current opinion in psychiatry*, 28(2), 133-140. <https://doi.org/10.1097/YCO.0000000000000139>
- 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. <https://doi.org/10.1037/a0030677>
- Henry, J. D., Phillips, L. H., & Von Hippel, C. (2014). A meta-analytic review of theory of mind difficulties in behavioural-variant frontotemporal dementia. *Neuropsychologia*, 56, 53-62. <https://doi.org/10.1016/j.neuropsychologia.2013.12.024>
- Hogan, D. B., Jetté, N., Fiest, K. M., Roberts, J. I., Pearson, D., Smith, E. E., . . . Maxwell, C. J. (2016). The prevalence and incidence of frontotemporal dementia: a systematic review. *Canadian Journal of Neurological Sciences*, 43(S1), S96-S109. <https://doi.org/10.1017/cjn.2016.25>
- Hughes, C. P., Berg, L., Danziger, W., Coben, L. A., & Martin, R. L. (1982). A new clinical scale for the staging of dementia. *The British journal of psychiatry*, 140(6), 566-572. <https://doi.org/10.1192/bjp.140.6.566>
- Hynes, C. A., Baird, A. A., & Grafton, S. T. (2006). Differential role of the orbital frontal lobe in emotional versus cognitive perspective-taking. *Neuropsychologia*, 44(3), 374-383. <https://doi.org/10.1016/j.neuropsychologia.2005.06.011>
- Ibañez, A., & Manes, F. (2012). Contextual social cognition and the behavioral variant of frontotemporal dementia. *Neurology*, 78(17), 1354-1362. <https://doi.org/10.1212/WNL.0b013e3182518375>
- Jesso, S., Morlog, D., Ross, S., Pell, M. D., Pasternak, S. H., Mitchell, D. G., . . . Finger, E. C. (2011). The effects of oxytocin on social cognition and behaviour in frontotemporal dementia. *Brain*, 134(9), 2493-2501. <https://doi.org/10.1093/brain/awr171>
- Johnen, A., & Bertoux, M. (2019). Psychological and cognitive markers of behavioral variant frontotemporal dementia—A clinical neuropsychologist's view on diagnostic criteria and beyond. *Frontiers in Neurology*, 10, 594. <https://doi.org/10.3389/fneur.2019.00594>
- Johnson, J. K., Diehl, J., Mendez, M. F., Neuhaus, J., Shapira, J. S., Forman, M., . . . Neumann, M. (2005). Frontotemporal lobar degeneration: demographic characteristics of 353 patients. *Archives of neurology*, 62(6), 925-930. <https://doi.org/10.1001/archneur.62.6.925>
- Josephs, K. A., Whitwell, J. L., & Jack Jr, C. R. (2008). Anatomic correlates of stereotypes in frontotemporal lobar degeneration. *Neurobiology of aging*, 29(12), 1859-1863. <https://doi.org/10.1016/j.neurobiolaging.2007.04.027>

- Kalbe, E., Schlegel, M., Sack, A. T., Nowak, D. A., Dafotakis, M., Bangard, C., . . . 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>
- Kertesz, A., Davidson, W., & Fox, H. (1997). Frontal behavioral inventory: diagnostic criteria for frontal lobe dementia. *Canadian Journal of Neurological Sciences*, *24*(1), 29-36. <https://doi.org/10.1017/s0317167100021053>
- Khosdelazad, S., Jorna, L. S., McDonald, S., Rakers, S. E., Huitema, R. B., Buunk, A. M., & Spikman, J. M. (2020). Comparing static and dynamic emotion recognition tests: Performance of healthy participants. *PloS one*, *15*(10), e0241297. <https://doi.org/10.1371/journal.pone.0241297>
- Kipps, C. M., Nestor, P. J., Acosta-Cabronero, J., Arnold, R., & Hodges, J. R. (2009). Understanding social dysfunction in the behavioural variant of frontotemporal dementia: the role of emotion and sarcasm processing. *Brain*, *132*(3), 592-603. <https://doi.org/10.1093/brain/awn314>
- Koyama, A., Hashimoto, M., Fukuhara, R., Ichimi, N., Takasaki, A., Matsushita, M., . . . Ikeda, M. (2018). Caregiver burden in semantic dementia with right-and left-sided predominant cerebral atrophy and in behavioral-variant frontotemporal dementia. *Dementia and geriatric cognitive disorders extra*, *8*(1), 128-137. <https://doi.org/10.1159/000487851>
- Krendl, A. C., & Betzel, R. F. (2022). Social cognitive network neuroscience. *Social cognitive and affective neuroscience*, *17*(5), 510-529. <https://doi.org/10.1093/scan/nsac020>
- Krudop, W. A., Dols, A., Kerssens, C. J., Eikelenboom, P., Prins, N. D., Möller, C., . . . Van Der Flier, W. M. (2017). The pitfall of behavioral variant frontotemporal dementia mimics despite multidisciplinary application of the FTDC criteria. *Journal of Alzheimer's Disease*, *60*(3), 959-975. <https://doi.org/10.3233/JAD-170608>
- Kumfor, F., Honan, C., McDonald, S., Hazelton, J. L., Hodges, J. R., & Piguet, O. (2017). Assessing the “social brain” in dementia: applying TASIT-S. *Cortex*, *93*, 166-177. <https://doi.org/10.1016/j.cortex.2017.05.022>
- Kumfor, F., Irish, M., Leyton, C., Miller, L., Lah, S., Devenney, E., . . . Piguet, O. (2014). Tracking the progression of social cognition in neurodegenerative disorders. *Journal of Neurology, Neurosurgery & Psychiatry*, *85*(10), 1076-1083. <https://doi.org/10.1136/jnnp-2013-307098>
- Kumfor, F., Zhen, A., Hodges, J. R., Piguet, O., & Irish, M. (2018). Apathy in Alzheimer's disease and frontotemporal dementia: distinct clinical profiles and neural correlates. *Cortex*, *103*, 350-359. <https://doi.org/10.1016/j.cortex.2018.03.019>

- LaMarre, A. K., Rascovsky, K., Bostrom, A., Toofanian, P., Wilkins, S., Sha, S. J., . . . Laforce, R. J. (2013). Interrater reliability of the new criteria for behavioral variant frontotemporal dementia. *Neurology*, 80(21), 1973-1977. <https://doi.org/10.1212/WNL.0b013e318293e368>
- Li, X., Wang, K., Wang, F., Tao, Q., Xie, Y., & Cheng, Q. (2013). Aging of theory of mind: The influence of educational level and cognitive processing. *International Journal of Psychology*, 48(4), 715-727. <https://doi.org/10.1080/00207594.2012.673724>
- Lichtenstein, M. L., Stewart, P. V., Kirchner, H. L., Finney, G., & Feldman, H. H. (2023). Exploring social cognition tests to differentiate frontotemporal dementia from depression: a two-step pilot study. *Alzheimer Disease & Associated Disorders*, 37(2), 145-148. <https://doi.org/10.1097/WAD.0000000000000526>
- Lima-Silva, T. B., Bahia, V. S., Carvalho, V. A., Guimarães, H. C., Caramelli, P., Balthazar, M. L., . . . Nitrini, R. (2015). Neuropsychiatric symptoms, caregiver burden and distress in behavioral-variant frontotemporal dementia and Alzheimer's disease. *Dementia and geriatric cognitive disorders*, 40(5-6), 268-275. <https://doi.org/10.1159/000437351>
- Loi, S. M., Eratne, D., Goh, A. M., Wibawa, P., Farrand, S., Kelso, W., . . . Velakoulis, D. (2021). A 10 year retrospective cohort study of inpatients with younger-onset dementia. *International Journal of Geriatric Psychiatry*, 36(2), 294-301. <https://doi.org/10.1002/gps.5424>
- Magno, M. A., Canu, E., Agosta, F., & Filippi, M. (2022). Measuring social cognition in frontotemporal lobar degeneration: a clinical approach. *Journal of Neurology*, 1-18. <https://doi.org/10.1007/s00415-021-10889-9>
- Magno, M. A., Canu, E., Filippi, M., & Agosta, F. (2022). Social cognition in the FTL spectrum: evidence from MRI. *Journal of Neurology*, 1-14. <https://doi.org/10.1007/s00415-021-10892-0>
- Mariano, L. I., Caramelli, P., Guimarães, H. C., Gambogi, L. B., Moura, M. V. B., Yassuda, M. S., . . . de Souza, L. C. (2020). Can social cognition measurements differentiate behavioral variant frontotemporal dementia from Alzheimer's disease regardless of apathy? *Journal of Alzheimer's Disease*, 74(3), 817-827. <https://doi.org/10.3233/JAD-190861>
- Mariano, L. I., O'Callaghan, C., Guimarães, H. C., Gambogi, L. B., da Silva, T. B., Yassuda, M. S., . . . Teixeira, A. L. (2020). Disinhibition in frontotemporal dementia and Alzheimer's disease: a neuropsychological and behavioural investigation. *Journal of the International Neuropsychological Society*, 26(2), 163-171. <https://doi.org/10.1017/S1355617719000973>

- Massimo, L., Powers, C., Moore, P., Vesely, L., Avants, B., Gee, J., . . . Grossman, M. (2009). Neuroanatomy of apathy and disinhibition in frontotemporal lobar degeneration. *Dementia and geriatric cognitive disorders*, 27(1), 96-104. <https://doi.org/10.1159/000194658>
- Mathersul, D., McDonald, S., & Rushby, J. A. (2013). Understanding advanced theory of mind and empathy in high-functioning adults with autism spectrum disorder. *Journal of clinical and experimental neuropsychology*, 35(6), 655-668. <https://doi.org/10.1080/13803395.2013.809700>
- McDonald, S., Bornhofen, C., Shum, D., Long, E., Saunders, C., & Neulinger, K. (2006). Reliability and validity of The Awareness of Social Inference Test (TASIT): a clinical test of social perception. *Disability and rehabilitation*, 28(24), 1529-1542. <https://doi.org/10.1080/09638280600646185>
- McDonald, S., & Flanagan, S. (2004). Social perception deficits after traumatic brain injury: interaction between emotion recognition, mentalizing ability, and social communication. *Neuropsychology*, 18(3), 572. <https://doi.org/10.1037/0894-4105.18.3.572>
- McDonald, S., Flanagan, S., Martin, I., & Saunders, C. (2004). The ecological validity of TASIT: A test of social perception. *Neuropsychological rehabilitation*, 14(3), 285-302. <https://doi.org/10.1080/09602010343000237>
- McDonald, S., Flanagan, S., Rollins, J., & Kinch, J. (2003). TASIT: A new clinical tool for assessing social perception after traumatic brain injury. *The Journal of head trauma rehabilitation*, 18(3), 219-238. <https://doi.org/10.1097/00001199-200305000-00001>
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Jr, Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., Mohs, R. C., Morris, J. C., Rossor, M. N., Scheltens, P., Carrillo, M. C., Thies, B., Weintraub, S., & Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association*, 7(3), 263–269. <https://doi.org/10.1016/j.jalz.2011.03.005>
- Mendez, M. F. (2022). Behavioral variant frontotemporal dementia and social and criminal transgressions. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 34(4), 328-340. <https://doi.org/10.1176/appi.neuropsych.21080224>
- Mioshi, E., Hsieh, S., Savage, S., Hornberger, M., & Hodges, J. R. (2010). Clinical staging and disease progression in frontotemporal dementia. *Neurology*, 74(20), 1591-1597. <https://doi.org/10.1212/WNL.0b013e3181e04070>

- Molenberghs, P., Johnson, H., Henry, J. D., & Mattingley, J. B. (2016). Understanding the minds of others: A neuroimaging meta-analysis. *Neuroscience & Biobehavioral Reviews*, *65*, 276-291. <https://doi.org/10.1016/j.neubiorev.2016.03.020>
- Moosbrugger, H., & Kelava, A. (2020). Qualitätsanforderungen an Tests und Fragebogen („Gütekriterien“). *Testtheorie und Fragebogenkonstruktion*, 13-38. https://doi.org/10.1007/978-3-662-61532-4_2
- Musa, G., Slachevsky, A., Muñoz-Neira, C., Méndez-Orellana, C., Villagra, R., González-Billault, C., . . . Lillo, P. (2020). Alzheimer’s disease or behavioral variant frontotemporal dementia? Review of key points toward an accurate clinical and neuropsychological diagnosis. *Journal of Alzheimer's Disease*, *73*(3), 833-848. <https://doi.org/10.3233/JAD-190924>
- Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D., Black, S., . . . Albert, M. (1998). Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology*, *51*(6), 1546-1554. <https://doi.org/10.1212/wnl.51.6.1546>
- Osborne-Crowley, K. (2020). Social cognition in the real world: Reconnecting the study of social cognition with social reality. *Review of General Psychology*, *24*(2), 144-158. <https://doi.org/10.1177/1089268020906483>
- Ossenkoppele, R., Singleton, E. H., Groot, C., Dijkstra, A. A., Eikelboom, W. S., Seeley, W. W., . . . Rabinovici, G. D. (2022). Research criteria for the behavioral variant of Alzheimer disease: a systematic review and meta-analysis. *JAMA neurology*, *79*(1), 48-60. <https://doi.org/10.1001/jamaneurol.2021.4417>
- Pardini, M., Gialloreti, L. E., Mascolo, M., Benassi, F., Abate, L., Guida, S., . . . Krueger, F. (2013). Isolated theory of mind deficits and risk for frontotemporal dementia: a longitudinal pilot study. *Journal of Neurology, Neurosurgery & Psychiatry*, *84*(7), 818-821. <https://doi.org/10.1136/jnnp-2012-303684>
- Peet, B. T., Castro-Suarez, S., & Miller, B. L. (2021). The neuropsychiatric features of behavioral variant frontotemporal dementia. *Frontotemporal dementias: emerging milestones of the 21st Century*, 17-31. https://doi.org/10.1007/978-3-030-51140-1_2
- Peet, B. T., Spina, S., Mundada, N., & La Joie, R. (2021). Neuroimaging in frontotemporal dementia: heterogeneity and relationships with underlying neuropathology. *Neurotherapeutics*, *18*, 728-752. <https://doi.org/10.1007/s13311-021-01101-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](https://doi.org/10.1016/0022-0965(85)90051-7)

- Perry, D. C., Sturm, V. E., Seeley, W. W., Miller, B. L., Kramer, J. H., & Rosen, H. J. (2014). Anatomical correlates of reward-seeking behaviours in behavioural variant frontotemporal dementia. *Brain*, *137*(6), 1621-1626. <https://doi.org/10.1093/brain/awu075>
- Perry, D. C., Whitwell, J. L., Boeve, B. F., Pankratz, V. S., Knopman, D. S., Petersen, R. C., . . . Josephs, K. A. (2012). Voxel-based morphometry in patients with obsessive-compulsive behaviors in behavioral variant frontotemporal dementia. *European Journal of Neurology*, *19*(6), 911-917. <https://doi.org/10.1111/j.1468-1331.2011.03656.x>
- Pickering-Brown, S. M. (2007). The complex aetiology of frontotemporal lobar degeneration. *Experimental neurology*, *206*(1), 1-10. <https://doi.org/10.1016/j.expneurol.2007.03.017>
- Piguet, O., Hornberger, M., Mioshi, E., & Hodges, J. R. (2011). Behavioural-variant frontotemporal dementia: diagnosis, clinical staging, and management. *The Lancet Neurology*, *10*(2), 162-172. [https://doi.org/10.1016/S1474-4422\(10\)70299-4](https://doi.org/10.1016/S1474-4422(10)70299-4)
- Piguet, O., Kumfor, F., & Hodges, J. (2017). Diagnosing, monitoring and managing behavioural variant frontotemporal dementia. *Medical Journal of Australia*, *207*(7), 303-308. <https://doi.org/10.5694/mja16.01458>
- Poletti, M., Enrici, I., & Adenzato, M. (2012). Cognitive and affective Theory of Mind in neurodegenerative diseases: neuropsychological, neuroanatomical and neurochemical levels. *Neuroscience & Biobehavioral Reviews*, *36*(9), 2147-2164. <https://doi.org/10.1016/j.neubiorev.2012.07.004>
- Poos, J. M., Jiskoot, L. C., Papma, J. M., van Swieten, J. C., & van den Berg, E. (2018). Meta-analytic review of memory impairment in behavioral variant frontotemporal dementia. *Journal of the International Neuropsychological Society*, *24*(6), 593-605. <https://doi.org/10.1017/S1355617718000115>
- Pose, M., Cetkovich, M., Gleichgerricht, E., Ibáñez, A., Torralva, T., & Manes, F. (2013). The overlap of symptomatic dimensions between frontotemporal dementia and several psychiatric disorders that appear in late adulthood. *International review of psychiatry*, *25*(2), 159-167. <https://doi.org/10.3109/09540261.2013.769939>
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and brain sciences*, *1*(4), 515-526. <http://doi:10.1017/S0140525X00076512>
- Quesque, F., Nivet, M., Etchepare, A., Wauquiez, G., Prouteau, A., Desgranges, B., & Bertoux, M. (2022). Social cognition in neuropsychology: A nationwide survey revealing current representations and practices. *Applied Neuropsychology: Adult*, 1-14. <https://doi.org/10.1080/23279095.2022.2061859>

- 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>
- Ranasinghe, K. G., Rankin, K. P., Lobach, I. V., Kramer, J. H., Sturm, V. E., Bettcher, B. M., . . . Shany-Ur, T. (2016). Cognition and neuropsychiatry in behavioral variant frontotemporal dementia by disease stage. *Neurology*, 86(7), 600-610. <https://doi.org/10.1212/WNL.0000000000002373>
- Rankin, K. P. (2021). Measuring behavior and social cognition in FTLD. *Frontotemporal dementias: emerging milestones of the 21st Century*, 51-65. https://doi.org/10.1007/978-3-030-51140-1_4
- Rankin, K. P., Baldwin, E., Pace-Savitsky, C., Kramer, J., & Miller, B. (2005). Self awareness and personality change in dementia. *Journal of Neurology, Neurosurgery & Psychiatry*, 76(5), 632-639. <https://doi.org/10.1136/jnnp.2004.042879>
- Rankin, K. P., Salazar, A., Gorno-Tempini, M. L., Sollberger, M., Wilson, S. M., Pavlic, D., . . . Miller, B. L. (2009). Detecting sarcasm from paralinguistic cues: anatomic and cognitive correlates in neurodegenerative disease. *Neuroimage*, 47(4), 2005-2015. <https://doi.org/10.1016/j.neuroimage.2009.05.077>
- Rankin, K. P., Toller, G., Gavron, L., La Joie, R., Wu, T., Shany-Ur, T., . . . Miller, B. L. (2021). Social behavior observer checklist: patterns of spontaneous behaviors differentiate patients with neurodegenerative disease from healthy older adults. *Frontiers in Neurology*, 12, 683162. <https://doi.org/10.3389/fneur.2021.683162>
- Rascovsky, K., & Grossman, M. (2013). Clinical diagnostic criteria and classification controversies in frontotemporal lobar degeneration. *International review of psychiatry*, 25(2), 145-158. <https://doi.org/10.3109/09540261.2013.763341>
- Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., . . . Onyike, C. U. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*, 134(9), 2456-2477. <https://doi.org/10.1093/brain/awr179>
- Reber, J., & Tranel, D. (2019). Frontal lobe syndromes. *Handbook of clinical neurology*, 163, 147-164. <https://doi.org/10.1016/B978-0-12-804281-6.00008-2>
- Reus, L. M., Vijverberg, E. G., Tijms, B. M., Ten Kate, M., Gossink, F., Krudop, W. A., . . . van der Flier, W. M. (2018). Disease trajectories in behavioural variant frontotemporal dementia, primary psychiatric and other neurodegenerative disorders presenting with behavioural change. *Journal of psychiatric research*, 104, 183-191. <https://doi.org/10.1016/j.jpsychires.2018.07.014>

- Rosness, T. A., Engedal, K., & Chemali, Z. (2016). Frontotemporal dementia: an updated clinician's guide. *Journal of Geriatric Psychiatry and Neurology*, 29(5), 271-280. <https://doi.org/10.1177/0891988716654986>
- Russell, E. (2012). *The scientific foundation of neuropsychological assessment: With applications to forensic evaluation*: Elsevier.
- Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: evidence for rapid and involuntary computation of what other people see. *Journal of experimental psychology: human perception and performance*, 36(5), 1255. <https://doi.org/10.1037/a0018729>
- Samtani, S., Meka, A., & Siette, J. (2023). Beyond memory: exploring the value of social cognition for older adults with neurocognitive disorders. *Frontiers in Psychiatry*, 14, 1209745. <https://doi.org/10.3389/fpsy.2023.1209745>
- Schlaffke, L., Lissek, S., Lenz, M., Juckel, G., Schultz, T., Tegenthoff, M., . . . Brüne, M. (2015). Shared and nonshared neural networks of cognitive and affective theory-of-mind: A neuroimaging study using cartoon picture stories. *Human brain mapping*, 36(1), 29-39. <https://doi.org/10.1002/hbm.22610>
- Seeley, W. W., Crawford, R., Rascovsky, K., Kramer, J. H., Weiner, M., Miller, B. L., & Gorno-Tempini, M. L. (2008). Frontal paralimbic network atrophy in very mild behavioral variant frontotemporal dementia. *Archives of neurology*, 65(2), 249-255. <https://doi.org/10.1001/archneurol.2007.38>
- Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., . . . Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience*, 27(9), 2349-2356. <https://doi.org/10.1523/JNEUROSCI.5587-06.2007>
- Seeley, W. W., Zhou, J., & Kim, E.-J. (2012). Frontotemporal dementia: what can the behavioral variant teach us about human brain organization? *The Neuroscientist*, 18(4), 373-385. <https://doi.org/10.1177/1073858411410354>
- Setién-Suero, E., Murillo-García, N., Sevilla-Ramos, M., Abreu-Fernández, G., Pozueta, A., & Ayesa-Arriola, R. (2022). Exploring the relationship between deficits in social cognition and neurodegenerative dementia: a systematic review. *Frontiers in aging neuroscience*, 14, 778093. <https://doi.org/10.3389/fnagi.2022.778093>
- Shamay-Tsoory, S. G., & 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>

- Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain*, *132*(3), 617-627. <https://doi.org/10.1093/brain/awn279>
- Shany-Ur, T., Poorzand, P., Grossman, S. N., Growdon, M. E., Jang, J. Y., Ketelle, R. S., . . . Rankin, K. P. (2012). Comprehension of insincere communication in neurodegenerative disease: lies, sarcasm, and theory of mind. *Cortex*, *48*(10), 1329-1341. <https://doi.org/10.1016/j.cortex.2011.08.003>
- Shinagawa, S., Catindig, J. A., Block, N. R., Miller, B. L., & Rankin, K. P. (2016). When a little knowledge can be dangerous: false-positive diagnosis of behavioral variant frontotemporal dementia among community clinicians. *Dementia and geriatric cognitive disorders*, *41*(1-2), 99-108. <https://doi.org/10.1159/000438454>
- Snowden, J., Gibbons, Z., Blackshaw, A., Doubleday, E., Thompson, J., Craufurd, D., . . . Neary, D. (2003). Social cognition in frontotemporal dementia and Huntington's disease. *Neuropsychologia*, *41*(6), 688-701. [https://doi.org/10.1016/s0028-3932\(02\)00221-x](https://doi.org/10.1016/s0028-3932(02)00221-x)
- Sparks, A., McDonald, S., Lino, B., O'Donnell, M., & Green, M. J. (2010). Social cognition, empathy and functional outcome in schizophrenia. *Schizophrenia research*, *122*(1-3), 172-178. <https://doi.org/10.1016/j.schres.2010.06.011>
- 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>
- Strikwerda-Brown, C., Ahmed, R. M., Piguet, O., & Irish, M. (2022). Try to see it my way—Examining the relationship between visual perspective taking and theory of mind in frontotemporal dementia. *Brain and Cognition*, *157*, 105835. <https://doi.org/10.1016/j.bandc.2021.105835>
- Strikwerda-Brown, C., Ramanan, S., & Irish, M. (2019). Neurocognitive mechanisms of theory of mind impairment in neurodegeneration: a transdiagnostic approach. *Neuropsychiatric disease and treatment*, 557-573. <https://doi.org/10.2147/NDT.S158996>
- Tesar, B., Deckert, M., Schmoeger, M., & Willinger, U. (2020). Electrophysiological correlates of basic and higher order cognitive and affective theory of mind processing in emerging and early adulthood—An explorative event-related potentials study to investigate first-, second-, and third-order theory of mind processing based on visual cues. *Frontiers in Human Neuroscience*, *14*, 79. <https://doi.org/10.3389/fnhum.2020.00079>

- Toller, G., Brown, J., Sollberger, M., Shdo, S. M., Bouvet, L., Sukhanov, P., . . . Rankin, K. P. (2018). Individual differences in socioemotional sensitivity are an index of salience network function. *Cortex*, *103*, 211-223. <https://doi.org/10.1016/j.cortex.2018.02.012>
- Torralva, T., Gleichgerrcht, E., Ardila, M. J. T., Roca, M., & Manes, F. F. (2015). Differential cognitive and affective theory of mind abilities at mild and moderate stages of behavioral variant frontotemporal dementia. *Cognitive and Behavioral Neurology*, *28*(2), 63-70. <https://doi.org/10.1097/WNN.0000000000000053>
- Torralva, T., Roca, M., Gleichgerrcht, E., Bekinschtein, T., & Manes, F. (2009). A neuropsychological battery to detect specific executive and social cognitive impairments in early frontotemporal dementia. *Brain*, *132*(5), 1299-1309. <https://doi.org/10.1093/brain/awp041>
- Van den Stock, J., De Winter, F.-L., Stam, D., Van de Vliet, L., Huang, Y.-A., Dries, E., . . . Vandenberghe, M. (2019). Reduced tendency to attribute mental states to abstract shapes in behavioral variant frontotemporal dementia links with cerebellar structural integrity. *NeuroImage: Clinical*, *22*, 101770. <https://doi.org/10.1016/j.nicl.2019.101770>
- Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Human brain mapping*, *30*(3), 829-858. <https://doi.org/10.1002/hbm.20547>
- Vijverberg, E. G., Dols, A., Krudop, W. A., Peters, A., Kerssens, C. J., van Berckel, B. N., . . . Prins, N. D. (2016). Diagnostic accuracy of the frontotemporal dementia consensus criteria in the late-onset frontal lobe syndrome. *Dementia and geriatric cognitive disorders*, *41*(3-4), 210-219. <https://doi.org/10.1159/000444849>
- Woolley, J. D., Khan, B. K., Murthy, N. K., Miller, B. L., & Rankin, K. P. (2011). The diagnostic challenge of psychiatric symptoms in neurodegenerative disease: rates of and risk factors for prior psychiatric diagnosis in patients with early neurodegenerative disease. *Journal of Clinical Psychiatry*, *72*(2), 126. <https://doi.org/10.4088/JCP.10m06382oli>
- World Health Organization. (2020, December 9). The top 10 causes of death. Retrieved from <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>
- World Health Organization. (2022). International Statistical Classification of Diseases and Related Health Problems (11th edition). <https://icd.who.int/>.
- Younes, K., & Miller, B. L. (2020). Frontotemporal dementia: neuropathology, genetics, neuroimaging, and treatments. *Psychiatric Clinics*, *43*(2), 331-344. <https://doi.org/10.1016/j.psc.2020.02.006>

Appendix A. Diagnostic criteria for frontotemporal dementia (Englund et al., 1994)

Core diagnostic features

Behavioural disorders

- Insidious onset and slow progression
- Early loss of personal awareness (neglect of personal hygiene and grooming)
- Early loss of social awareness (lack of social tact, misdemeanours such as shoplifting)
- Early signs of disinhibition (such as unrestrained sexuality, violent behaviour, inappropriate jocularity, restless pacing)
- Mental rigidity and inflexibility
- Hyperorality (oral/dietary changes, overeating, food fads, excessive smoking and alcohol consumption, oral exploration of objects)
- Stereotyped and preservative behaviour (wandering, mannerisms such as clapping, singing, dancing, ritualistic preoccupation such as hoarding, toileting, and dressing)
- Utilisation behaviour (unrestrained exploration of objects in the environment)
- Distractibility, impulsivity, and impersistence
- Early loss of insight into the fact that the altered condition is due to a pathological change of own mental state

Affective Symptoms

- Depression, anxiety, excessive sentimentality, suicidal and fixed ideation, delusion (early and evanescent)
- Hypochondriasis, bizarre somatic preoccupation (early and evanescent)
- Emotional unconcern (emotional indifference and remoteness, lack of empathy and sympathy, apathy)
- Amimia (inertia, asponaneity)

Speech disorder

- Progressive reduction of speech (asponaneity and economy of utterance)
- Stereotypy of speech (repetition of limited repertoire of words, phrases, or themes)
- Echolalia and perseveration
- Late mutism

Spatial orientation and praxis preserved (intact abilities to negotiate the environment)

Physical signs

- Early primitive reflexes
- Early incontinence
- Late akinesia, rigidity, tremor
- Low and labile blood pressure

Investigations

- Normal EEG despite clinically evident dementia
 - Brain imaging (structural or functional, or both): predominant frontal or anterior temporal abnormality, or both
 - Neuropsychology (profound failure on “frontal lobe” tests in the absence of severe amnesia, aphasia, or perceptual spatial disorders)
-

Supportive diagnostic features

- Onset before 65
- Positive family history of similar disorders in a first degree relative
- Bulbar palsy, muscular weakness and wasting, fasciculations (motor neuron disease)

Diagnostic exclusion feature

- Abrupt onset with ictal events
- Head trauma related to onset
- Early severe amnesia
- Early spatial disorientation, lost in surroundings, defective localisation of objects
- Early severe apraxia
- Logoclonic speech with rapid loss of train of thought
- Myoclonus
- Cortical bulbar and spinal deficits
- Cerebellar ataxia
- Chorea-athetosis
- Early, severe, pathological EEG
- Brain imaging (predominant post-central structural or functional deficit. Multifocal cerebral lesions on CT or MRI)
- Laboratory tests indicating brain involvement or inflammatory disorder (such as multiple sclerosis, syphilis, AIDS and herper simplex encephalitis)

Relative diagnostic exclusion features

- Typical history of chronic alcoholism
 - Sustained hypertension
 - History of vascular disease (e.g., angina, claudication)
-

Appendix B. Diagnostic criteria for frontotemporal dementia (Neary et al., 1998)

I. Core diagnostic features

- A. Insidious onset and gradual progression
 - B. Early decline in social interpersonal conduct
 - C. Early impairment in regulation of personal conduct
 - D. Early emotional blunting
 - E. E. Early loss of insight
-

II. Supportive diagnostic features

- A. Behavioural disorder
 - 1. Decline in personal hygiene and grooming
 - 2. Mental rigidity and inflexibility
 - 3. Distractibility and impersistence
 - 4. Hyperorality and dietary changes
 - 5. Perseverative and stereotyped behaviour
 - 6. Utilization behaviour
 - B. Speech and language
 - 1. Altered speech output
 - a. Aspontaneity and economy of speech
 - b. Press of speech
 - 2. Stereotypy of speech
 - 3. Echolalia
 - 4. Perseveration
 - 5. Mutism
 - C. Physical signs
 - 1. Primitive reflexes
 - 2. Incontinence
 - 3. Akinesia, rigidity, and tremor
 - 4. Low and labile blood pressure
 - D. Investigations
 - 1. Neuropsychology: significant impairment on frontal lobe tests in the absence of severe amnesia, aphasia, or perceptuospatial disorder
 - 2. Electroencephalography: normal on conventional EEG despite clinically evident dementia
 - 3. Brain imaging (structural and/or functional): predominant frontal and/or anterior temporal abnormality
-

III. Supportive feature

- A. Onset before 65 years: positive family history of similar
 - B. Bulbar palsy, muscular weakness and wasting, disorder in first-degree relative fasciculations (associated motor neuron disease present in a minority of patients)
-

IV. Diagnostic exclusion features

A. Historical and clinical

1. Abrupt onset with ictal events
2. Head trauma related to onset
3. Early, severe amnesia
4. Spatial disorientation
5. Logoclonic, festinant speech with loss of train of thought
6. Myoclonus
7. Corticospinal weakness
8. Cerebellar ataxia
9. Choreoathetosis

B. Investigations

1. Brain imaging: predominant postcentral structural or functional deficit; multifocal lesions on CT or MRI
 2. Laboratory tests indicating brain involvement of metabolic or inflammatory disorder such as MS, syphilis, AIDS, and herpes simplex encephalitis
-

Relative diagnostic exclusion features

A. Typical history of chronic alcoholism

B. Sustained hypertension

C. History of vascular disease (e.g., angina, claudication)

Note. All core features must be present to fulfil the criteria for diagnosis. The diagnosis becomes more likely when more supportive features are present. The physical features should be regarded as “supportive” rather than as necessary conditions for diagnosis. All exclusion features must be absent.

**Appendix C. Diagnostic criteria for behavioural variant frontotemporal dementia
(Rascovsky et al., 2011)**

I. Neurodegenerative disease

The following symptom must be present to meet criteria for bvFTD

A. Shows progressive deterioration of behaviour and/or cognition by observation or history (as provided by a knowledgeable informant).

II. Possible bvFTD

Three of the following behavioural/cognitive symptoms (A–F) must be present to meet criteria. Ascertainment requires that symptoms be persistent or recurrent, rather than single or rare events.

A. Early* behavioural disinhibition [one of the following symptoms (A.1–A.3) must be present]:

- A.1. Socially inappropriate behaviour
- A.2. Loss of manners or decorum
- A.3. Impulsive, rash or careless actions

B. Early apathy or inertia [one of the following symptoms (B.1–B.2) must be present]:

- B.1. Apathy
- B.2. Inertia

C. Early loss of sympathy or empathy [one of the following symptoms (C.1–C.2) must be present]:

- C.1. Diminished response to other people's needs and feelings
- C.2. Diminished social interest, interrelatedness or personal warmth

D. Early perseverative, stereotyped or compulsive/ritualistic behaviour [one of the following symptoms (D.1–D.3) must be present]:

- D.1. Simple repetitive movements
- D.2. Complex, compulsive or ritualistic behaviours
- D.3. Stereotypy of speech

E. Hyperorality and dietary changes [one of the following symptoms (E.1–E.3) must be present]:

- E.1. Altered food preferences
- E.2. Binge eating, increased consumption of alcohol or cigarettes
- E.3. Oral exploration or consumption of inedible objects

F. Neuropsychological profile: executive/generation deficits with relative sparing of memory and visuospatial functions [all of the following symptoms (F.1–F.3) must be present]:

- F.1. Deficits in executive tasks
 - F.2. Relative sparing of episodic memory
 - F.3. Relative sparing of visuospatial skills
-

III. Probable bvFTD

All of the following symptoms (A–C) must be present to meet criteria.

- A. Meets criteria for possible bvFTD
 - B. Exhibits significant functional decline (by caregiver report or as evidenced by Clinical Dementia Rating Scale or Functional Activities Questionnaire scores)
 - C. Imaging results consistent with bvFTD [one of the following (C.1–C.2) must be present]:
 - C.1. Frontal and/or anterior temporal atrophy on MRI or CT
 - C.2. Frontal and/or anterior temporal hypoperfusion or hypometabolism on PET or SPECT
-

IV. Behavioural variant FTD with definite FTLN Pathology

Criterion A and either criterion B or C must be present to meet criteria.

- A. Meets criteria for possible or probable bvFTD
 - B. Histopathological evidence of FTLN on biopsy or at post-mortem
 - C. Presence of a known pathogenic mutation
-

V. Exclusionary criteria for bvFTD

Criteria A and B must be answered negatively for any bvFTD diagnosis. Criterion C can be positive for possible bvFTD but must be negative for probable bvFTD.

- A. Pattern of deficits is better accounted for by other non-degenerative nervous system or medical disorders
 - B. Behavioural disturbance is better accounted for by a psychiatric diagnosis
 - C. Biomarkers strongly indicative of Alzheimer's disease or other neurodegenerative process
-

Note. 'Early' refers to symptom presentation within the first 3 years.

Appendix D. Study I: Supplementary materials

Supplementary Material

The Behavioural Dysfunction Questionnaire discriminates behavioural variant frontotemporal dementia from Alzheimer's disease dementia and major depressive disorder

Journal of Neurology

Anna Semenkova^{1,2}, Olivier Piguet³, Andreas Johnen⁴, Matthias L. Schroeter^{5,6}, Jannis Godulla^{5,6}, Christoph Linnemann⁷, Markus Mühlhauser⁷, Thomas Sauer⁷, Markus Baumgartner⁸, Sarah Anderl-Straub⁹, Markus Otto^{9,10}, Ansgar Felbecker¹¹, Reto W. Kressig¹, Manfred Berres¹², and Marc Söllberger^{1,13}

1 Memory Clinic, University Department of Geriatric Medicine FELIX PLATTER, Basel, Switzerland

2 Faculty of Psychology, University of Basel, Switzerland

3 The University of Sydney, School of Psychology and Brain and Mind Centre, NSW, Australia

4 Clinic for Neurology, Münster University Hospital, Münster, Germany

5 Clinic for Cognitive Neurology, University Hospital Leipzig, Germany

6 Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

7 University Psychiatric Clinic, Basel, Switzerland

8 Memory Clinic Sonnweid, Wetzikon, Switzerland

9 Department of Neurology, University of Ulm, Ulm, Germany

10 Department of Neurology, University Hospital Halle, Germany

11 Clinic of Neurology und Neurophysiology, Canton Hospital St. Gallen, Switzerland

12 Faculty of Mathematics and Technology, University of Applied Sciences Koblenz, Germany

13 Department of Neurology, University Hospital Basel, Switzerland

Correspondence concerning this article should be addressed to Marc Sollberger, Memory Clinic, University Department of Geriatric Medicine FELIX PLATTER, Burgfelderstrasse 101, 4055 Basel, Switzerland. Email: marc.sollberger@felixplatter.ch

Table of content

Supplementary A. Development of the Behavioural Dysfunctional Questionnaire	4
Supplementary B. Inclusion and exclusion criteria of patients and healthy participants	16
Supplementary C. Item scores of the Behavioural Dysfunction Questionnaire of healthy participants	18
Supplementary D. Internal consistency of the Behavioural Dysfunctional Questionnaire.....	21
Supplementary E. Discriminatory power of the domain scores and the BDQ-Global Score between bvFTD and Non-bvFTD patients	22
References	23

Supplementary A. Development of the Behavioural Dysfunctional Questionnaire

The revised diagnostic criteria for behavioural variant frontotemporal dementia (bvFTD), involve six domains with two to three subdomains (Table A1) [7]. The first five domains (A–E) include behavioural disorders, whereas the sixth domain (F) includes a cognitive disorder (i.e., primary executive dysfunction).

Table A1. *Domains and subdomains of the diagnostic criteria for behavioural variant frontotemporal dementia*

A. Early* behavioural disinhibition

- A.1. Socially inappropriate behaviour
- A.2. Loss of manners or decorum
- A.3. Impulsive, rash or careless actions

B. Early* apathy or inertia

- B.1. Apathy
- B.2. Inertia

C. Early* loss of sympathy or empathy

- C.1. Diminished response to other people's needs and feelings
- C.2. Diminished social interest, interrelatedness or personal warmth

D. Early* perseverative, stereotyped or compulsive/ritualistic behaviour

- D.1. Simple repetitive movements
- D.2. Complex, compulsive or ritualistic behaviours
- D.3. Stereotypy of speech

E. Hyperorality and dietary changes

- E.1. Altered food preferences
- E.2. Binge eating, increased consumption of alcohol or cigarettes
- E.3. Oral exploration or consumption of inedible objects

F. Neuropsychological profile: executive/generation deficits with relative sparing of memory and visuospatial functions

- F.1. Deficits in executive tasks
 - F.2. Relative sparing of episodic memory
 - F.3. Relative sparing of visuospatial skills
-

* "early" refers to symptom presentation within the first three years

A domain is considered as affirmed if at least one symptom of this domain is persistent or recurrent [7]. Rascovsky, et al. [7] provided examples of symptoms of each subdomain. Please see for example the subdomain “socially inappropriate behaviour” (Fig. A1).

Figure A1. *Exemplary symptoms of the subdomain “socially inappropriate behaviour”*

A.1. Socially inappropriate behaviour

Examples of behaviours that violate social norms include inappropriately approaching, touching or kissing strangers, verbal or physical aggression, public nudity or urination, inappropriate sexual acts and criminal behaviour (such as theft or shoplifting).

Development of the questions

We aimed to operationalise the five behavioural domains of the diagnostic criteria for bvFTD [7] by developing an informant questionnaire. To do this, we kept the structure of the five domains and their subdomains and added the exemplary symptoms from the Appendix of the consensus paper [7] as items to the respective subdomain. First, we translated and retranslated the subdomains and symptoms by native English and German speakers, respectively. Please see, as an example, the subdomain “socially inappropriate behaviour” in English and German language in Table A2.

Table A2. *Subdomain “socially inappropriate behaviour” according to Rascovsky, et al. [7] in English and German*

<i>A.1. Socially inappropriate behaviour</i>	<i>A.1. Sozial unangebrachtes Verhalten</i>
Examples of behaviours that violate social norms include inappropriately approaching, touching or kissing strangers, verbal or	Beispiele von Verhalten, das soziale Normen verletzt, sind unpassende Annäherungen, Berühren oder Küssen von

physical aggression, public nudity or Fremden, verbale oder körperliche urination, inappropriate sexual acts and Aggression, öffentliche Nacktheit oder criminal behaviour (such as theft or öffentliches Urinieren, unpassende sexuelle shoplifting). Handlungen und kriminelles Verhalten (wie Diebstahl oder Ladendiebstahl).

Second, we generated questions based on each translated subdomain, e.g., “Zeigt sie/er sozial unangebrachtes Verhalten, wie zum Beispiel ...” / “Does she/he show socially inappropriate behaviour, such as ...” (Fig. A2, orange colour).

Third, we added the symptoms that belong to the corresponding subdomain and named them “items”. For example, “Unangemessene Annäherungen, wie fremde Personen anfassen oder körperlich ganz nahe kommen” / “Inappropriate approaches, such as touching strangers or getting very close physically ” was named item 1.1 (Fig. A2, blue colour).

Figure A2. Operationalisation of the subdomain “socially inappropriate behaviour”

<p>1 A.1. Sozial unangebrachtes Verhalten Beispiele von Verhalten, das soziale Normen verletzt, sind unpassende Annäherungen, Berühren oder Küssen von Fremden, verbale oder körperliche Aggression, öffentliche Nacktheit oder öffentliches Urinieren, unpassende sexuelle Handlungen und kriminelles Verhalten (wie Diebstahl oder Ladendiebstahl).</p>	<p>1. Zeigt sie/er <u>sozial unangebrachtes Verhalten</u>, wie zum Beispiel:</p> <p>1.1 Unangemessene Annäherungen, wie fremde Personen anfassen oder körperlich ganz nahe kommen</p> <p>1.2 Berühren oder Küssen von fremden Personen</p> <p>1.3 Verbale Aggressionen, wie beschuldigen, anschreien, usw.</p> <p>1.4 Körperliche Aggressionen, wie schlagen, stossen, kratzen, usw.</p> <p>1.5 Sich in der Öffentlichkeit entkleiden</p> <p>1.6 Urinieren in der Öffentlichkeit</p> <p>1.7 Unangemessene sexuelle Handlungen, wie sich in der Öffentlichkeit selbst befriedigen / masturbieren</p> <p>1.8 Kriminelles Verhalten (wie z.B. Diebstahl)</p>
--	---

Generation of item scores

We decided to use a six-point (0-5) Likert scale to score each item. Similarly to the Frontal Behavioral Inventory (FBI) [4], we decided to use both, severity and frequency, to measure the degree of behavioural disorders, as some disorders are represented best by severity such as apathy or empathy and others best by frequency such as compulsive behaviour. We described each point of the Likert scale as follows

0 – no;

1 – very mild / rare (less than once a month);

2 – mild / occasionally (approximately once a month but not weekly);

3 – moderate / sometimes (about once a week);

4 – severe / often (several times a week but not daily);

5 – very severe / very often (daily)

Four of the five behavioural domains (A-D) are marked with the time criterion “early”. “Early” refers to symptom presentation within the first three years [7]. We operationalized this criterion by asking the informants about both, the respective symptom onset and the clinical onset.

Examining the clarity of the BDQ and its feasibility in administration

To check the BDQ for its clarity and feasibility in administration, we asked several employees of the Memory Clinic Basel and 20 patients’ caregivers of different ages, sexes and education to read each question carefully for its clarity and to fill out the questionnaire by thinking of a relative’s behaviour. Afterwards, we interviewed them about the clarity of the questionnaire’s instruction, the questions and the answer options. Their comments and suggestions were analysed and applied, if judged as appropriate. In the following, you see the structure of the BDQ using the example of the subdomain “socially inappropriate behaviour” (Table A3).

Table A3 Structure of the subdomain “socially inappropriate behaviour”

1.	Zeigt sie/er sozial <u>unangebrachtes</u> Verhalten, wie zum Beispiel:	NEIN	Wenn JA, seit wann? (Anzahl Wochen, Monate oder Jahre)	Aktuelle Häufigkeit/Schweregrad				
				Sehr selten/ Sehr leicht (Weniger als 1x/Monat)	Selten/ Leicht (Ca. 1x/Monat, aber nicht wöchentlich)	Manchmal / Mittel (Ca. 1x/Woche)	Häufig/ Stark (Mehrere Male pro Woche, aber nicht täglich)	Sehr häufig/ Sehr stark (Mind. 1x/Tag)
1.1	Unangemessene Annäherungen, wie fremde Personen anfassen oder körperlich ganz nahe kommen	0		1	2	3	4	5
1.2	Berühren oder Küssen von fremden Personen	0		1	2	3	4	5
1.3	Verbale Aggressionen, wie beschuldigen, anschreien, usw.	0		1	2	3	4	5
1.4	Körperliche Aggressionen, wie schlagen, stossen, kratzen, usw.	0		1	2	3	4	5
1.5	Sich in der Öffentlichkeit entkleiden	0		1	2	3	4	5
1.6	Urinieren in der Öffentlichkeit	0		1	2	3	4	5
1.7	Unangemessene sexuelle Handlungen, wie sich in der Öffentlichkeit selbst befriedigen / masturbieren	0		1	2	3	4	5
1.8	Kriminelles Verhalten (wie z.B. Diebstahl)	0		1	2	3	4	5

The final version of BDQ included 56 items corresponding to 15 questions (Table A4). As the number of symptoms differs between subdomains in the diagnostic criteria [7], the number of items between questions differs too. In addition to answering the items, informants had the opportunity to note additional symptoms (question 16). At the end of the questionnaire, the informants were asked about the patient’s first symptom(s) and the clinical onset (question 17).

Table A4. *Structure of the Behavioural Dysfunction Questionnaire*

Behavioural domains	Subdomains	Questions	Items
A. Early* behavioural disinhibition	A1. Socially inappropriate behaviour	1	1.1 – 1.8
	A2. Loss of manners or decorum	2	2.1 – 2.9
	A3. Impulsive, rash or careless actions	3	3.1 – 3.5
B. Early* apathy or inertia	B1. Apathy	4	4.1 – 4.2
	B2. Inertia	5	5.1 – 5.2
C. Early* loss of sympathy or empathy	C1. Diminished response to other people's needs and feelings	6	6.1 – 6.2
	C2. Diminished social interest, interrelatedness or personal warmth	7, 8	7.1 – 7.3 8.1 – 8.3
D. Early* perseverative, stereotyped or compulsive/ritualistic behaviour	D1. Simple repetitive movements	9	9.1 – 9.9
	D2. Complex, compulsive or ritualistic behaviours	10	10.1 – 10.7
	D3. Stereotypy of speech	11	11.1
E. Hyperorality and dietary changes [one of the following symptoms]	E1. Altered food preferences	12	12.1 – 12.2
	E2. Binge eating, increased consumption of alcohol or cigarettes	13, 14	
	E3. Oral exploration or consumption of inedible objects	15	
Additional questions	Other behavioural disorders	16	
	First symptom(s) and time of their start	17	

* "early" refers to symptom presentation within the first three years

In the following, you find the list of questions and items of the BDQ in German (Table A5) and in English (Table A6).

Table A5. Behavioural Dysfunction Questionnaire (German version)

1. Zeigt sie/er sozial unangebrachtes Verhalten, wie zum Beispiel:

- 1.1. Unangemessene Annäherungen, wie fremde Personen anfassen oder körperlich ganz nahe kommen
- 1.2. Berühren oder Küssen von fremden Personen
- 1.3. Verbale Aggressionen, wie beschuldigen, anschreien, usw.
- 1.4. Körperliche Aggressionen, wie schlagen, stossen, kratzen, usw.
- 1.5. *Sich in der Öffentlichkeit entkleiden **
- 1.6. Urinieren in der Öffentlichkeit
- 1.7. *Unangemessene sexuelle Handlungen, wie sich in der Öffentlichkeit selbst befriedigen / masturbieren **
- 1.8. Kriminelles Verhalten (wie z.B. Diebstahl)

2. Ist Ihnen bei ihr/ihm der Verlust von Umgangsformen oder des Anstands aufgefallen, wie zum Beispiel:

- 2.1. Unpassendes Lachen
- 2.2. Fluchen oder Schreien
- 2.3. Beleidigungen
- 2.4. Unhöfliche oder sexuell anzügliche Bemerkungen
- 2.5. Mangel an Anstand (z.B. nicht in der Warteschlange anstehen können)
- 2.6. Fehlende Achtung der Privatsphäre
- 2.7. Keine angemessene Reaktion auf soziale Signale (z.B. die Person redet weiter, obwohl ihr/ihm signalisiert wird aufzuhören)
- 2.8. Mangelnde Körperpflege (z.B. sie/er trägt übelriechende, schmutzige, verschlissene oder unpassende Kleidung)
- 2.9. Unhöfliches Benehmen in der Öffentlichkeit, wie furzen, an den Geschlechtsteilen kratzen, in den Zähnen herumstochern, spucken oder rülpsen

3. Zeigt sie/er impulsive, unbedachte oder achtlose Handlungen, wie zum Beispiel:

- 3.1. Rücksichtsloses (Auto-)Fahren
- 3.2. *Neu begonnenes Glücksspiel **
- 3.3. *Stehlen von Nahrungsmitteln oder glänzenden Objekten **
- 3.4. Unbedachtes Kaufen oder Verkaufen von Objekten
- 3.5. Unbedachtes Preisgeben von persönlichen Daten wie z.B. die Nummer der Kreditkarte

4. Haben Sie bemerkt, dass ihr/ihm der Antrieb fehlt, wie zum Beispiel:

- 4.1. Mangel an Spontanität

4.2. Vermindertes oder fehlendes Interesse für Tätigkeiten, die ihr/ihm früher wichtig waren

5. Ist Ihnen bei ihr/ihm träges Verhalten aufgefallen, wie zum Beispiel:

5.1. Aufforderungen sind notwendig, damit alltägliche Verrichtungen (wie z.B. Zähneputzen) begonnen oder ausgeführt werden

5.2. Gespräche werden nicht begonnen oder aufrechterhalten

6. Ist Ihnen aufgefallen, dass sie/er weniger auf die Bedürfnisse und Gefühle anderer eingeht, wie zum Beispiel:

6.1. Gleichgültigkeit gegenüber Schmerz oder Leid anderer

6.2. Verletzende Bemerkungen zum Schmerz oder Leid anderer Personen

7. Ist Ihnen bei ihr/ihm ein vermindertes Interesse an sozialen Kontakten und Beziehungen aufgefallen, wie zum Beispiel:

7.1. Vermindertes Interesse an der Gesellschaft anderer

7.2. Vermeiden von Blickkontakt

7.3. Abnahme von sozialem Engagement

8. Ist Ihnen bei ihr/ihm eine Abnahme von Wärme im zwischenmenschlichen Umgang aufgefallen, wie zum Beispiel:

8.1. Vermeiden von körperlichem Kontakt wie z.B. Berührung oder Umarmung von Freunden und Verwandten

8.2. Emotionale Distanziertheit, d. h. nicht mehr auf äussere Einflüsse positiver oder negativer Art emotional reagieren

8.3. Gefühlskälte

9. Haben Sie beobachtet, dass sie/er wiederholt die gleichen Bewegungen ausführt, wie zum Beispiel:

9.1. Reiben der Hände

9.2. Klopfen mit Händen oder Füßen

9.3. Klatschen der Hände

9.4. Sich kratzen

9.5. Sich an Haut oder Kleidern zupfen

9.6. Summen

9.7. *Mit dem Stuhl schaukeln **

9.8. Räuspern

9.9. Mit den Lippen schmatzen

10. Ist Ihnen bei ihr/ihm ein zwanghaftes oder ritualisiertes Verhalten aufgefallen, wie zum Beispiel:

10.1. Zwanghaftes Zählen

- 10.2. Zwanghafte Reinigungsrituale
- 10.3. Zwanghaftes Sammeln oder Horten
- 10.4. Zwanghaftes Kontrollieren
- 10.5. Zwanghaftes Auf-die-Toilette-Gehen
- 10.6. Zwanghaftes Anordnen von Gegenständen
- 10.7. Zwanghaftes Gehen bestimmter Strecken

11. Haben Sie bei ihr/ihm wiederholt sprachliche Auffälligkeiten bemerkt, wie zum Beispiel:

- 11.1. Zwanghaftes Wiederholen von Wörtern, Sätzen oder Erzählungen

12. Sind Ihnen bei ihr/ihm veränderte Vorlieben für Nahrungsmittel aufgefallen, wie zum Beispiel:

- 12.1. Verstärktes Verlangen nach Süßigkeiten
- 12.2. Einschränkung auf den Konsum bestimmter Nahrungsmittel

13. Sind Ihnen bei ihr/ihm Essanfälle aufgefallen?

14. Hat sie/er neu mit dem Konsum von Zigaretten oder Alkohol begonnen, oder den Konsum von Zigaretten oder Alkohol erhöht?

15. Haben Sie bei ihr/ihm beobachtet, dass sie/er nicht essbare Gegenstände in den Mund nimmt oder isst? *

16. Falls Ihnen Verhaltensauffälligkeiten bei Ihrer/Ihrem Angehörigen aufgefallen sind, die wir nicht erfragt haben, bitten wir Sie, diese hier anzugeben

17. Seit wann besteht eine Veränderung – sei es in der Kraft/Motorik (z.B. Gehen oder Koordination), im Denken oder Sprechen, in der Stimmung, im Verhalten oder anderweitig – bei Ihrer/Ihrem Angehörigen?

Welche Form der Veränderung war dies?

* excluded items (affirmed by less than 5% of the informants of patients with behavioural variant frontotemporal dementia)

Table A6. Behavioural Dysfunction Questionnaire (English version)

1. Does she/he show socially inappropriate behaviour, such as:

- 1.1. Inappropriate approaches, such as touching strangers or getting very close physically
- 1.2. Touching or kissing strangers
- 1.3. Verbal aggression, such as blaming, yelling at, etc.
- 1.4. Physical aggression, such as hitting, pushing, scratching, etc.
- 1.5. *Undressing in public **

1.6. Urinating in public

1.7. *Inappropriate sexual acts, such as pleasuring oneself / masturbating in public **

1.8. Criminal behaviour (such as stealing)

2. Have you noticed in her/him a loss of manners or etiquette, such as:

2.1. Inappropriate laughter

2.2. Swearing or yelling

2.3. Offensive comments

2.4. Rude or sexually suggestive comments

2.5. Lack of etiquette (e.g., not being able to wait in line)

2.6. Lack of respect for privacy

2.7. Failure to respond appropriately to social cues (e.g., continuing to talk even though it was signalled to stop)

2.8. Lack of personal hygiene (e.g., she/he wears malodorous, stained, torn or inappropriate clothing)

2.9. Rude behaviour in public, such as farting, scratching private parts, picking teeth, spitting, or belching

3. Does she/he show impulsive, thoughtless, or careless actions, such as:

3.1. Reckless (car-)driving

3.2. *Newly started gambling **

3.3. *Stealing food or shiny objects **

3.4. Unwise buying or selling of objects

3.5. Careless disclosure of personal data such as credit card number

4. Have you noticed that she/he lacks drive, such as:

4.1. Lack of spontaneity

4.2. Decreased or lack of interest in activities that used to be important to her/him

5. Have you noticed any sluggish behaviour in her/him, such as:

5.1. Prompts are necessary for everyday tasks (such as brushing teeth) to be started or performed

5.2. Conversations are not initiated or maintained

6. Have you noticed that she/he is less responsive to the needs and feelings of others, such as:

6.1. Ignorance of pain or suffering of others

6.2. Making hurtful comments leading to other person's pain or suffering

7. Have you noticed in her/him a decreased interest in social contacts and relationships, such as:

7.1. Decreased interest in the company of others

7.2. Avoiding eye contact

7.3. Decrease in social engagement

8. Have you noticed in her/him a decreased warmth in interpersonal interactions, such as:

8.1. Avoiding physical contact, such as touching or hugging friends and relatives

8.2. Emotional detachment, i.e. no longer reacting emotionally to external stimuli of a positive or negative nature

8.3. Emotional coldness

9. Have you observed that she/he repeatedly performs the same movements, such as:

9.1. Rubbing hands

9.2. Tapping with hands or feet

9.3. Clapping hands

9.4. Scratching oneself

9.5. Tugging at skin or clothes

9.6. Humming

9.7. *Rocking with chair **

9.8. Clearing throat

9.9. Smacking of lips

10. Have you noticed any compulsive or ritualistic behaviours in her/him, such as:

10.1. Compulsive counting

10.2. Compulsive cleaning rituals

10.3. Compulsive collecting or hoarding

10.4. Compulsive controlling

10.5. Compulsive going to the toilet

10.6. Compulsive arrangement of objects

10.7. Compulsive walking of certain routes

11. Have you noticed any repeatedly language abnormalities in her/him, such as:

11.1. Compulsive repetition of words, phrases, or narratives

12. Have you noticed any changes in her/his food preferences, such as:

12.1. Increased craving for sweets

12.2. Restriction on consumption of certain foods

13. Have you noticed her/him having any binge eating episodes?

14. Has she/he newly started consuming cigarettes or alcohol, or increased the usual consumption of cigarettes or alcohol?

15. Have you observed her/him putting non-edible items in the mouth or eating them?*

16. If you have noticed behavioural abnormalities in your loved one that we have not asked about, please indicate them here

17. When have the first changes in your loved one started (whether in strength / motor function [e.g., walking or coordination], thinking or speaking, mood, behaviour, or otherwise)?

What kind of changes did you observe?

* excluded items (affirmed by less than 5% of the informants of patients with behavioural variant frontotemporal dementia)

Supplementary B. Inclusion and exclusion criteria of patients and healthy participants

Patients with behavioural variant frontotemporal dementia (bvFTD) and Alzheimer's disease dementia (ADD) were recruited from two Swiss (i.e., Memory Clinic, University Department of Geriatric Medicine FELIX PLATTER, Basel; Clinic of Neurology and Neurophysiology, Canton Hospital St. Gallen) and three German (i.e., Clinic for Neurology, Münster University Hospital; Clinic for Cognitive Neurology, University Hospital Leipzig; Department of Neurology, University of Ulm) memory clinics. The diagnosis was established by a multidisciplinary team consisting of neurologists, neuropsychologists, and psychiatrists, who performed comprehensive neuropsychological and neuroimaging assessments. Inclusion criteria were the diagnosis of at least probable bvFTD [7] or probable ADD [5], and availability of a reliable informant, who has regular contact with the patient. Exclusion criteria were major neurocognitive disorder at moderate or severe stage according to DSM-5, history of or current drug and/or alcohol abuse as well as drug- and/or alcohol-related disorder according to ICD-10, history of severe depressive episode or current depressive episode according to ICD-10, history of or current major psychiatric disorders according to ICD-10, traumatic brain injury, systemic disorders or brain diseases that could result in behavioural changes.

Patients with major depressive disorder (MDD) were recruited from three Swiss institutions (i.e., Memory Clinic, University Department of Geriatric Medicine FELIX PLATTER, Basel; Clenia Schlössli AG, Clinic of Psychiatry and Psychotherapy, Oetwil am See; University Psychiatric Clinic, Basel). The inclusion criterion was the diagnosis of an at least moderate depressive episode according to ICD-10. Exclusion criteria were a neurocognitive disorder according to DSM-5, history of or current drug or/and alcohol abuse as well as drug- and alcohol-related disorder according to ICD-10, any other major psychiatric disorders according to ICD-10, traumatic brain injury, systemic disorders or brain diseases that could result in behavioural changes.

Healthy participants were recruited from the participant pool of the Memory Clinic, University Department of Geriatric Medicine FELIX PLATTER Basel, Switzerland. Inclusion

criteria were at least seven years of education, German and/or Swiss German as mother tongue and self-report of good health. Exclusion criteria were conditions with potential negative influence on the behaviour, including cognitive deficits [i.e., Montreal Cognitive Assessment [6] score below the demographically-adjusted fifth percentile for cognitively healthy individuals [8]], systemic or brain diseases, psychiatric disorders according the ICD-10, traumatic brain injury, chronic pain, history of or regular intake of any psychoactive drugs (except benzodiazepines for sleep) and severe sensory and/or motor deficits. Additionally, participants were checked for signs of depressive mood by use of the Beck Depression Inventory (score \geq ten points) [1] or the Geriatric Depression Scale (score \geq five points) [9].

Supplementary C. Item scores of the Behavioural Dysfunction Questionnaire of healthy participants

We analysed the item scores of the 414 healthy subjects (Table C1). Of the 56 items, four items (1.5, 1.7, 3.2, and 3.3) were never affirmed by healthy participants' informants. Items were generally (95.2%) negated. Items were affirmed in 3% as "very mild" or "rare", in 0.9% as "mild" or "occasionally"; in 0.5% as "moderate" or "sometimes"; in 0.3% as "severe" or "often" and in 0.1% as "very severe" or "very often".

Table C1. Answer pattern (in percentages) of informants of healthy participants

	Item score						99-%*
	0	1	2	3	4	5	
<i>Domain A. Early behavioural disinhibition</i>							
Item 1.1	98.79	0.97	0	0	0.24	0	1
Item 1.2	98.06	1.94	0	0	0	0	1
Item 1.3	85.23	10.65	1.94	1.45	0.73	0	3
Item 1.4	99.27	0.48	0	0.24	0	0	0
Item 1.5	100	0	0	0	0	0	0
Item 1.6	99.76	0.24	0	0	0	0	0
Item 1.7	100	0	0	0	0	0	0
Item 1.8	98.79	1.21	0	0	0	0	1
Item 2.1	94.19	5.57	0.24	0	0	0	1
Item 2.2	88.35	7.28	2.91	0.49	0.73	0.24	3
Item 2.3	91.77	5.57	1.21	0.48	0.97	0	3
Item 2.4	97.82	1.21	0.73	0.24	0	0	1
Item 2.5	92.98	5.33	1.21	0.24	0.24	0	2
Item 2.6	96.13	3.63	0.24	0	0	0	1
Item 2.7	87.62	8.74	1.21	1.7	0.73	0	3
Item 2.8	97.82	1.7	0.24	0.24	0	0	1
Item 2.9	97.09	1.94	0.24	0.24	0	0.49	1
Item 3.1	94.4	4.14	0.73	0.49	0.24	0	2
Item 3.2	100	0	0	0	0	0	0
Item 3.3	100	0	0	0	0	0	0
Item 3.4	95.41	3.62	0.97	0	0	0	1
Item 3.5	99.03	0.97	0	0	0	0	0

Domain B. Early apathy or inertia

Item 4.1	81.55	9.95	5.58	2.91	0	0	3
Item 4.2	86.96	9.18	2.66	1.21	0	0	3
Item 5.1	97.58	1.69	0.24	0.48	0	0	1
Item 5.2	90.58	6.04	1.69	0.97	0.24	0.48	3
<i>Domain C. Early loss of sympathy or empathy</i>							
Item 6.1	94.19	3.87	1.69	0.24	0	0	2
Item 6.2	95.64	3.39	0.73	0.24	0	0	1
Item 7.1	87.68	7.25	3.86	0.48	0.72	0	3
Item 7.2	94.2	4.35	0.97	0.24	0	0.24	2
Item 7.3	92.27	5.56	1.69	0.48	0	0	2
Item 8.1	94.9	3.16	0.24	1.21	0	0.49	3
Item 8.2	94.43	3.39	1.45	0.48	0.24	0	2
Item 8.3	93.69	4.61	0.97	0.49	0.24	0	2
<i>Domain D. Early perseverative, stereotyped or compulsive/ritualistic behaviour</i>							
Item 9.1	96.36	1.21	0	1.21	0.73	0.49	4
Item 9.2	96.36	2.18	0	0.49	0.73	0.24	3
Item 9.3	98.54	0.73	0.24	0.49	0	0	1
Item 9.4	94.9	2.18	0.49	1.21	0.49	0.73	4
Item 9.5	97.57	1.21	0.24	0.49	0.24	0.24	2
Item 9.6	95.87	2.43	0.97	0.24	0.24	0.24	2
Item 9.7	98.79	0.97	0	0.24	0	0	1
Item 9.8	93.19	3.16	0.97	0.97	1.22	0.49	4
Item 9.9	99.27	0.24	0	0.24	0.24	0	0
Item 10.1	98.31	0.48	0.72	0.48	0	0	2
Item 10.2	97.34	1.21	0.48	0.24	0.72	0	2
Item 10.3	93.72	2.9	0.97	1.45	0.72	0.24	3
Item 10.4	93.24	3.14	1.69	1.45	0.48	0	3
Item 10.5	98.31	0.72	0.24	0.48	0	0.24	1
Item 10.6	96.38	1.93	0.72	0.72	0.24	0	2
Item 10.7	99.52	0.24	0	0.24	0	0	0
Item 11.1	95.41	1.69	1.21	0.72	0.72	0.24	3
<i>Domain E. Hyperorality and dietary changes</i>							
Item 12.1	87.44	3.62	2.42	2.42	2.42	1.69	5
Item 12.2	93.24	2.66	0.97	0.97	0.97	1.21	5
Item 13	93.95	4.6	1.45	0	0	0	2
Item 14	96.13	1.94	0.48	0.48	0.73	0.24	3
Item 15	99.27	0.48	0.24	0	0	0	0

* 99-% refers to the 99-percentile of the respective item score. 0 – no; 1 – very mild / rare (less than once a month); 2 – mild / occasionally (approximately once a month but not weekly); 3 – moderate / sometimes (about once a week); 4 – severe / often (several times a week but not daily); 5 – very severe / very often (daily).

We analysed the answer pattern of healthy participants' informants to calculate the BDQ-Global Domain Score (BDQ-GDS). BDQ-GDS represents the number of affirmed behavioural domains (0-5). According to Rascovsky, et al. [7], a domain can be considered as present if at least one of the behavioural symptoms (i.e., items) is "persistent or recurrent rather than single or rare event". In the BDQ, this corresponds to an item score ≥ 3 (moderate / sometimes). However, as several items were affirmed with the scores greater than three also for healthy participants, we decided to add "greater than the 99-percentile in healthy subjects" as an additional criterion for domain affirmation. This later criterion was applied in 19 of the 50 items (38%). For example, item 1.3 is only considered as pathological in case of a score ≥ 4 (Table C1). Notably, 99-percentile of two items (i.e., 12.1, 12.2), both from the domain "Hyperorality and dietary changes", reached a score of five.

Supplementary D. Internal consistency of the Behavioural Dysfunctional Questionnaire

We analysed the internal consistency of the BDQ in the patients' sample ($N = 131$) by use of the Kuder-Richardson Formula 20 (KR-20). We used KR-20, because items scores showed a strong positive skewness.

The five domains overall showed an excellent internal consistency of .92 (Table D1). Three of the five domains ("early behavioural disinhibition", "early loss of sympathy/empathy" and "early perseverative/stereotyped behaviour") showed good internal consistency ($\alpha = .76$ to $.86$). Domain "early apathy/inertia" had an acceptable ($\alpha = .67$) and domain "hyperorality and dietary changes" had a poor internal consistency ($\alpha = .54$). A major reason for the low internal consistencies of the last two domains is their small number of items (i.e., four items). In addition, the poor internal consistency of domain "hyperorality and dietary changes" reflects its different behavioural disorders [altered food preferences (2 items); binge eating (1 item); increased consumption of alcohol or cigarettes (1 item)].

Table D1. *Internal consistencies of the Behavioural Dysfunctional Questionnaire at domain and global level in the patients' sample*

Domains	KR-20	Number of items
Early behavioural disinhibition	.86	18
Early apathy / inertia	.67	4
Early loss of sympathy / empathy	.84	8
Early perseverative, stereotyped or compulsive behaviour	.76	16
Hyperorality and dietary changes	.54	4
Global (all domains)	.92	50

KR-20 = Kuder-Richardson Formula 20

Supplementary E. Discriminatory power of the domain scores and the BDQ-Global Score between bvFTD and Non-bvFTD patients

To examine whether behavioural domain scores differ in their discriminatory power between bvFTD and Non-bvFTD patients, we run univariate logistic regression with each domain mean score, followed by ROC analyses. All domain mean scores separated the two groups acceptable to excellent [3] with AUC ranging between 77.27% (early loss of sympathy/empathy) and 84.46% (early behavioural disinhibition) (Table E1). Using the Delong's method [2], we found similar discriminatory power across the domain scores. Likewise, the BDQ-Global Score separated the two groups similarly to three of the five domain scores (i.e., early behavioural disinhibition, early perseverative, stereotyped or compulsive behaviour, and hyperorality/dietary changes). In contrast, the BDQ-Global Score separated the two groups better than the domain score of "early apathy/inertia" and the domain score of "early loss of sympathy/empathy" ($p < .05$).

Table E1. Discriminatory power of each domain score and the BDQ-Global Score between bvFTD and Non-bvFTD patients

Domains	Area under the curve
Early behavioural disinhibition	84.46 (CI: 77.07–91.85)
Early apathy / inertia	79.61 (CI: 70.47–88.75)
Early loss of sympathy / empathy	77.27 (CI: 67.72–86.83)
Early perseverative, stereotyped or compulsive behaviour	80.91 (CI: 71.42–90.4)
Hyperorality and dietary changes	81.5 (CI: 72.78–90.23)
BDQ-Global Score	85.98 (CI: 78.73–93.22)

CI = 95% confidence interval

References

1. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J (1961) An inventory for measuring depression. *Archives of general psychiatry* 4:561-571
2. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44:837-845
3. Hosmer Jr. DW, Lemeshow S, Sturdivant RX (2013) *Applied Logistic Regression*, Third Edition. John Wiley & Sons, Inc.
4. Kertesz A, Davidson W, Fox H (1997) Frontal behavioral inventory: diagnostic criteria for frontal lobe dementia. *Canadian Journal of Neurological Sciences* 24:29-36
5. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Jr., Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R, Mohs RC, Morris JC, Rossor MN, Scheltens P, Carrillo MC, Thies B, Weintraub S, Phelps CH (2011) The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association* 7:263-269
6. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H (2005) The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society* 53:695-699
7. Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, Van Swieten JC, Seelaar H, Dopper EG, Onyike CU (2011) Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain* 134:2456-2477
8. Thomann AE, Goettel N, Monsch RJ, Berres M, Jahn T, Steiner LA, Monsch AU (2018) The Montreal Cognitive Assessment: normative data from a German-speaking cohort and comparison with international normative samples. *Journal of Alzheimer's Disease* 64:643-655
9. Yesavage JA, Sheikh JI (1986) 9/1 Geriatric depression scale (GDS) recent evidence and development of a shorter version. *Clinical gerontologist* 5:165-173

Appendix E. Study II: Supplementary materials

Elektronisches Supplement

ESM 1. Zusätzliche Informationen zu den Auswahl- und Teilnahmebedingungen der Studienteilnehmenden

Für Deutschland wurden die Kontaktdaten über die Webseite der *Deutschen Alzheimer Gesellschaft e. V.* (DAIzG; <https://www.deutsche-alzheimer.de/unser-service/gedaechtnissprechstunden.html>) ermittelt. Die Kontaktdaten der schweizerischen Memory-Kliniken (MK) wurden über die Webseite des Vereins *Swiss Memory Clinics* (<https://www.swissmemoryclinics.ch/de/home/mitglieder>) akquiriert; diejenigen der österreichischen MK über die Webseite des *Bundesministeriums für Soziales, Gesundheit, Pflege und Konsumentenschutz* (BMSGPK) (<https://www.gesundheit.gv.at/krankheiten/gehirn-nerven/demenz/gedaechtnisambulanzen>). Die gesammelten Adressen wurden anschließend um Institutionen, die in den Auflistungen mehrfach aufgeführt wurden, zusammengehörten oder keine, respektive fehlerhafte E-Mail-Kontaktdaten aufwiesen, bereinigt. Alle Teilnehmenden gaben im Vorfeld eine informierte Einwilligung zur freiwilligen Teilnahme an der Umfrage. Die Institutionen erhielten keine Aufwandsentschädigung.

Elektronisches Supplement

ESM 2. Gründe, keine Tests zur Emotionserkennung durchzuführen, die jeweils von nur einer Institution genannt wurden

Aus dem Optionsfeld:

- Die Testergebnisse haben keine Relevanz für die Therapie

Aus dem Freitextfeld:

- Fehlende Validierung
- Fehlende Normierung und Validierung
- Zu wenig diagnostische Informationen
- Zu wenige Tests und fehlende Bezugsadressen
- Standardisierte Testverfahren sind nicht notwendig; subjektive Beurteilung ist genauer
- Untersuchung der Emotionserkennung spielt in der Demenzdiagnostik eine untergeordnete Rolle
- Schwerpunkt der Institution liegt auf psychosozialen Interventionen

Appendix F. Study III: Supplementary materials

Supplemental Material

Development of the Basel Version of the Awareness of Social Inference Test – Theory of Mind (BASIT-ToM) in Healthy Adults

Table of Contents

Appendix A	6
<i>Table A1</i>	6
Evaluation Criteria of the TASIT-SIM Scenes for Scene Selection.....	6
Appendix B	7
<i>Table B1.1</i>	7
Evaluation and Selection of the Honesty Scenes from TASIT-SIM Form A and Form B.....	7
<i>Table B1.2</i>	9
Evaluation and Selection of the Simple Sarcasm Scenes from TASIT-SIM Form A and Form B..	9
<i>Table B1.3</i>	11
Evaluation and Selection of the Paradoxical Sarcasm Scenes from TASIT-SIM Form A and Form B.....	11
<i>Table B2</i>	13
Selected Scenes for the BASIT-ToM Film Project.....	13
Appendix C	14
<i>Document C1</i>	14
Screenplays of the BASIT-ToM Scenes.....	14
<i>Table C1</i>	28
Main Changes of the Script (TASIT-SIM → BASIT-ToM).....	28
Appendix D	29
<i>Table D1</i>	29
Distribution of the Actors to the BASIT-ToM and Corresponding TASIT-SIM Scenes.....	29
Appendix E	30
<i>Document E1</i>	30
Background Information of each BASIT-ToM Intensity Version Scene for the Actors.....	30

Appendix F	47
<i>Table F1: Scene H#1</i>	<i>47</i>
Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind	47
<i>Table F2: Scene H#2</i>	<i>49</i>
Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind	49
<i>Table F3: Scene H#3</i>	<i>51</i>
Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind	51
<i>Table F4: Scene H#4</i>	<i>53</i>
Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind	53
<i>Table F5: Scene sS#1</i>	<i>55</i>
Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind	55
<i>Table F6: Scene sS#2</i>	<i>57</i>
Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind	57
<i>Table F7: Scene sS#3</i>	<i>59</i>
Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind	59
<i>Table F8: Scene sS#4</i>	<i>61</i>
Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind	61
<i>Table F9: Scene pS#1</i>	<i>63</i>
Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind	63

<i>Table F10: Scene pS#2</i>	65
Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind	65
<i>Table F11: Scene pS#3</i>	67
Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind	67
<i>Table F12: Scene pS#4</i>	69
Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind	69
<i>Table F13: Scene Practice(pS)</i>	71
Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind	71
Appendix G	73
<i>Table G1</i>	73
Main Differences between the BASIT-ToM and TASIT-SIM	73
Appendix H	75
<i>Document H1</i>	75
Details of the Programming and Data Storage	75
Appendix I	79
<i>Figure I1</i>	79
Exemplary Representation of the BASIT-ToM Computer-Based Application Process	79
Appendix J	80
<i>Table J.1:</i>	80
Item Characteristic Curves of the Intensity Version Scenes for the Message Type Honesty	80
<i>Table J.2:</i>	86
Item Characteristic Curves of the Intensity Version Scenes for the Message Type Paradoxical Sarcasm	86

<i>Table J.3:</i>	93
Item Characteristic Curves of the Intensity Version Scenes for the Message Type Simple	
Sarcasm.....	93
Appendix K	102
<i>Table K.1:</i>	102

Appendix A

Table A1

Evaluation criteria of the TASIT-SIM scenes for scene selection

- 1) The topic of the film scene is suitable for the age groups over 60
 - 2) The topic of the film scene is suitable for people from the German-speaking cultural area
 - 3) The film scene appears realistic and comprehensible
 - 4) The content of the film scene is contemporary
 - 5) Set design appears realistic
 - 6) The message type portrayed appears appropriate for the story of the scene
 - 7) The message type is portrayed by different modalities (e.g. gestures, facial expressions, linguistic characteristics)
 - 8) Office or living room setting is possible
 - 9) A change of location is not necessary during the scene
 - 10) Gender ratio in the cast is unimportant
 - 11) Number of actors
 - 12) Duration of the scene is not too short
-

Note. TASIT - SIM = The Awareness of Social Inference Test - Social Inference Minimal.

Evaluation criteria	Scene ID									
	A1	A4	A7	A11	A14	B1	B4	B7	B11	B14
A change of location is not necessary during the scene	1	1	1	1	1	1	1	1	1	1
Gender ratio in the cast is unimportant [†]	1	1	1	1	1	1	1	1	1	1
Number of actors [†]	2	1	2	2	2	2	2	2	2	2
Duration of the scene is not too short [†]	1	1	1	1	1	1	1	1	1	1
Total score	8/9	8/9	8/9	9/9	7/9	8/9	8/9	9/9	8/9	7/9

Note. TASIT-SIM = The Awareness of Social Inference Test – Social Inference Minimal; A = TASIT-SIM form A, B = TASIT-SIM form B. Scoring: 0 = criterion was not fulfilled; 1 = criterion was fulfilled. Selected scenes are in bold. [†]Criterion is not part of the total score.

Evaluation criteria	Scene ID									
	A2	A6	A9	A10	A13	B2	B6	B9	B10	B13
A change of location is not necessary during the scene	1	1	1	1	1	1	1	1	1	1
Gender ratio in the cast is unimportant [†]	1	1	1	1	1	1	1	1	0	0
Number of actors [†]	2	2	2	2	2	2	1	2	2	2
Duration of the scene is not too short [†]	1	1	1	1	1	1	1	1	1	1
Total score	8/9	8/9	8/9	9/9	7/9	8/9	8/9	9/9	8/9	7/9

Note. TASIT-SIM = The Awareness of Social Inference Test – Social Inference Minimal; A = TASIT-SIM form A, B = TASIT-SIM form B. Scoring: 0 = criterion was not fulfilled; 1 = criterion was fulfilled. Selected scenes are in bold. [†]Criterion is not part of the total score.

Evaluation criteria	Scene ID									
	A3	A5	A8	A12	A15	B3	B5	B8	B12	B15
A change of location is not necessary during the scene	1	1	1	1	1	1	1	1	1	1
Gender ratio in the cast is unimportant [†]	1	1	1	1	1	0	1	0	1	1
Number of actors [†]	2	2	2	2	2	2	2	2	2	2
Duration of the scene is not too short [†]	1	0	1	0	1	0	1	1	1	1
Total score	8/9	8/9	8/9	9/9	7/9	8/9	8/9	9/9	8/9	7/9

Note. TASIT-SIM = The Awareness of Social Inference Test – Social Inference Minimal; A = TASIT-SIM form A, B = TASIT-SIM form B. Scoring: 0 = criterion was not fulfilled; 1 = criterion was fulfilled. Selected scenes are in bold. [†]Criterion is not part of the total score.

Table B2*Selected Scenes for the BASIT-ToM Film Project*

Message type	Scene 1	Scene 2	Scene 3	Scene 4	Scene 5
Honesty	B11	A11	B1	B7	
Simple Sarcasm	A6	B2	B10	B13	
Paradoxical Sarcasm	A3	A15	B5	B8	B15

Note. BASIT-ToM = Basel Version of the Awareness of Social Inference Test – Theory of Mind. The scene code consists of a letter indicating the form of The Awareness of Social Inference Test – Social Inference Minimal (i.e. form A or form B) and a number indicating the respective scene number.

Appendix C

Document C1

Screenplays of the BASIT-ToM Scenes

23.

23 ÜBUNGSSZENE 2 (NEUES HAUS, B2-PS-15) - WOHN-/ESSZIMMER 23

Beide Personen sitzen am Tisch. Die Frau sitzt an einem Kreuzworträtsel, der Mann liest Zeitung und schaut hin und wieder zu seiner Frau auf. Dabei besteht kein direkter Blickkontakt zwischen den beiden.

F1

Hast du schon Simones neues Haus
gesehen?

F1 (CONT'D)

Hat sie es nicht in einer
wunderschönen Farbe streichen
lassen?

M2 schaut unbeteiligt weiter in seine Zeitung.

M2

Also ich hab wirklich keine
Ahnung warum sie diese Farbe für
die Wände ausgesucht hat.

F1

Die Wände sind wirklich schön,
aber nicht annähernd so schön wie
das Grün der Haustüre.

M2

Vielleicht ist sie ja
farbenblind.

F1

Hmm.

24.

24 EINFACHER SAKRASMUS 1 (SEMINAR, A2-ES-6) - BÜRO 24

Die Schreibtische sind leicht auf die Seite geschoben, an der Rückwand des Zimmers ist ein Whiteboard angebracht. Die Frau steht am Whiteboard und ist dabei, etwas aufzuzeichnen, als der Mann den Raum betritt. Er spricht F1 vorsichtig an.

M3

Hey! Es tut mir leid, ich kann das Seminar am Freitag nicht übernehmen.

F1

Was?

M3 nähert sich vorsichtig.

M3

Du hast mich doch gefragt, ob ich dein Seminar am Freitag übernehmen kann. Es tut mir sehr leid, ich kann es doch nicht machen.

F1 arbeitet am Whiteboard weiter und schaut den Mann nur einaml kurz an.

F1

Ach, sei nicht albern. Du musst dich deshalb nicht schlecht fühlen. Ich habe dich ja erst letzte Woche gefragt.

M3

Ja, ich weiss, dass ich gesagt habe, dass ich es machen könnte. Aber als ich heute noch mal in meinen Kalender gesehen habe, habe ich gemerkt, dass ich es nicht einrichten kann.

F1 dreht sich vom Whiteboard weg und schaut den Mann direkt an.

F1

Oh, ich weiss. Du bist schwer beschäftigt. Es war nicht fair von mir zu erwarten, dass du das Seminar übernehmen kannst.

Sie dreht sich wieder weg und arbeitet weiter am Whiteboard.

M3

Ich hätte das wirklich gerne für dich gemacht.

(CONTINUED)

CONTINUED:

25.

F1 schreibt weiterhin an das Whiteboard, kurzer Blick zu M3.

F1

Schon gut. Es war ein bisschen zu viel verlangt.

M3

In einer anderen Woche mache ich das gerne. Nur diese Woche geht es leider nicht. Ich hoffe es macht dir nichts aus.

F1 schaut den Mann direkt an.

F1

Oh nein, das macht mir natürlich nichts aus. Ich habe ja reichlich Zeit jemand anderen zu finden. Ich kann es auch immer noch selbst machen. Ich habe jede Menge Zeit!

26.

25 EINFACHER SARKASMUS 2 (WOCHENENDAUSFLUG, B2-ES-2) - WOZI
25

F3 und M1 decken gemeinsam den Esstisch.

M1
Würde es dir etwas
ausmachen, wenn meine
Schwester dieses Wochenende
mit in die Berge fährt? Es
geht ihr wirklich nicht gut
zur Zeit.

F3 schaut den Mann kurz an.

F3
Ich verstehe nicht warum du
überhaupt fragst? Du weisst ja
wie ich zu deiner Schwester
stehe.

M1 schaut die Frau an.

M1
Sie muss wirklich mal weg.

F3 schaut ihn nicht an.

F3
Müssen wir das nicht alle...?
Warum laden wir nicht auch noch
ein paar andere Leute ein? Was
macht denn deine Mutter dieses
Wochenende?

M1
Nein. Ich will nur, dass meine
Schwester kommt.

F3
Komm schon! Wir könnten ein
richtiges Familienwochenende
draus machen! Warum laden wir
nicht auch noch deinen Bruder und
seine Freundin ein?

M1
Jetzt übertreibst du aber!

27.

26 EINFACHER SARKASMUNS 3 (AUSGEHEN, B2-ES-10) - WOZI 26

F4 zieht sich gerade eine Jacke an und nimmt ihre Tasche vom Tisch. M2 betritt das Zimmer.

 M2
Wohin gehst du?

 F4
Zum Essen.

 M2
Frag mich nicht, ob ich mitkommen will!

 F4
Mach dir keine Sorgen, das werde ich nicht! Wie sehe ich aus?

 M2
Schrecklich.

 F4
Vielen Dank.

 F4 (CONT'D)
Ich komme nicht all zu spät zurück.

 M2
Du kommst also zurück?

 F4
Natürlich.

 M2
Wenn du nach 22 Uhr heimkommst, stehst du vor einer geschlossenen Haustür.

 F4
Ok, vielleicht sollte ich dann gar nicht nach Hause kommen.

28.

27 EINFACHER SARKASMUS 4 (HEMD, B2-ES-13) - WOZI 27

M4 hat ein neues Hemd an und bindet sich während des Gesprächs eine Krawatte um. F2 macht sich ausgehbereit.

M4
Was denkst du, soll ich das Hemd
anziehen, das meine Mutter mir
geschenkt hat?

F2 hält kurz inne.

F2
Natürlich. Wenn du möchtest.

M4 schaut zu F2.

M4
Nein, komm schon. Was denkst du?

F2 unterbricht ihre Aktion nicht.

F2
Trag es ruhig, wenn Du es tragen
willst.

M4
Ich weiss nicht, warum dir das
Hemd nicht gefällt.

F2 atmet hörbar aus.

F2
Ich habe nicht gesagt, dass es
mir nicht gefällt.

M4
Aber du willst nicht, dass ich es
anziehe. In Ordnung, dann werde
ich es nicht anziehen.

F2
Nein, nein. Es ist doch ein
wunderschönes Hemd und du siehst
darin sehr attraktiv aus.

29.

28 PARADOXER SARKASMUS 1 (BERICHT, A2-PS-3) - BÜRO 28

F3 sitzt am Schreibtisch im Büro und arbeitet am Computer.
Der Mann betritt das Büro und bringt ihr einige
Unterlagen.

M2

Das war ja ein wahnsinnig langer
Bericht, den du geschrieben hast.
Der war tonnenschwer.

F3 blickt nicht von ihrer Arbeit auf.

F3

Ja, nur ein paar kurze Nächte.

M2

Das muss dich das ganze
Wochenende gekostet haben.

F3 unterbricht ihre Arbeit, schaut M2 an.

F3

Nein, nein! Ich hatte ein total
entspanntes Wochenende!

30.

29 PARADOXER SARKASMUS 2 (OHRRINGE, A2-PS-15) - WOZI 29

M3 steht am Tisch und bügelt sein Hemd auf. F2 macht sich ausgehfertig.

F2

Hast du die Ohrringe gesehen, die mir meine Schwester geschenkt hat?

M3

Ich hätte wetten können, dass Sie dafür ein Vermögen ausgegeben hat!

F2 richtet sich weiterhin.

F2

Ja, mindestens!

M3

Die kannst du nur zu einem ganz besonderen Anlass tragen.

31.

30 PARADOXER SARKASMUS 3 (ESSEN, B2-PS-5) - WOHN-/ESSZIMMER
30

F1 und M4 sitzen am Esstisch. Die Teller sind leer
gegessen. M4 legt das Besteck auf den Teller und lehnt
sich auf dem Stuhl zurück.

F1

Wie hat es dir geschmeckt?

M4

Furchtbar! Ich glaube, ich wurde
gerade vergiftet.

F1

Dir muss es wirklich geschmeckt
haben. Normalerweise nimmst du
dir keine zweite Portion.

M4

Ich glaube, das war das
Schlimmste, was ich je gegessen
habe.

32.

31 PARADOXER SARKASMUS 4 (NACHTSCHICHT, B2-PS-8) - WOZI 31

F4 deckt den Esstisch. M1 kommt ins Zimmer und hilft seiner Frau den Tisch zu decken.

M1

Ach übrigens, ich habe schlechte Neuigkeiten. Mein Chef hat mich gebeten, am Samstag die Nachtschicht zu übernehmen.

F4 unterbricht das Tischdecken kurz, schaut M1 an.

F4

Oh ok, das ist in Ordnung. Gerade noch rechtzeitig. Ich rufe einfach im Restaurant an und sage die Reservierung ab.

M1 hilft nicht weiter den Tisch zu decken.

M1

Ich hätte nicht gedacht, dass dich das so ärgert.

F4

Ich bin nicht verärgert. Ich bin froh darüber, einen Abend lang dem gesellschaftlichen Trubel entgehen zu können.

33.

32 EHRlichkeit 1 (THEATERSTÜCK, B2-E-11) - WOHN-/ESSZIMMER
32

F1 und M1 sitzen auf dem Sofa. Sie unterhalten sich über ein Theaterstück, das sie gesehen haben.

F1
War das nicht ein tolles
Theaterstück?

M1
Ja, es war grossartig.

F1
Ich fand, es war hervorragend. Es
hat mich total gefesselt.

M1
Ja, mich auch.

F1
Hat Dich das Ende auch
überrascht?

M1
Oh ja!

F1
Ich fand die Schauspieler sehr
gut. Vor allem die
Hauptdarstellerin.

M1
Ja, sie war unglaublich. Und ihr
Partner... Was für eine Leistung!

F1
Hmmm, schade, dass das die letzte
Vorstellung war. Ich hätte das
Stück gerne noch mal gesehen.

M1
(bestätigend)
Ja, wirklich schade.

34.

33 EHRlichkeit 2 (EINTRITTSKARTEN, A2-E-11) - BÜRO 33

F2 steht vor dem Schreibtisch und reicht M2 Formulare. Der Mann lehnt am Schreibtisch und unterschreibt sie.

F2
Ach, übrigens: ich habe euch die
Eintrittskarten für Samstagabend
organisiert.

M2
(erstaunt)
Ah danke.

F2
(neutral)
Deine Frau hat mir gesagt, dass
ihr hingehen möchtet.

M2
Absolut! Vielen Dank!

F2
Es war ganz schön schwierig die
Eintrittskarten zu kriegen.

M2 schaut von den Formularen auf.

M2
Ich weiß das zu schätzen.

35.

34 EHRlichkeit 3 (ANSTRENGENDER TAG, B2-E-1) - BÜRO 34

Beide Protagonisten stehen erschöpft im Büro. F3 hält Unterlagen in der Hand.

M3

Puh, bin ich erschöpft.

F3

Oh, das wundert mich nicht. Du hast so hart gearbeitet.

M3

Wir beide sind ja seit heute Morgen um 8 Uhr am Schuft.

F3

Ja ich weiss. Ich weiss auch all das zu schätzen, was du geleistet hast.. und zwischendurch hast du auch noch Anrufe beantwortet und Kaffee gekocht.

M3

Ich bin froh, dass du das zu schätzen weisst.

F3

Absolut! Du warst eine grosse Hilfe.

M3

Willst du, dass ich morgen wieder komme?

F3

Oh, das wäre grossartig.

M3

Ok. Dann bis morgen.

F3

Danke.

36.

35 EHRlichkeit 4 (BEFÖRDERUNG, B2-E-7) - BÜRO

35

M4 und F4 teilen sich ein gemeinsames Büro. Ihre Schreibtische stehen so, dass sie während der Arbeit ab und zu ganz gut miteinander reden können.

M4

Ich wüsste zu gerne ob der Chef dieses Jahr mit meiner Arbeit zufrieden war.

F4

Keine Bange! Jeder weiss, dass du viel geleistet hast!

M4

Meinst du? Ich hätte diese Beförderung wirklich gerne.

F4

Oh, daran zweifle ich nicht. Ich wäre nicht überrascht, wenn du schon seit langem auf der Liste stehen würdest.

M4

Sicher?

F4

Warum nicht? Du hast dich so engagiert, hast so viel gearbeitet und so viele gute Entscheidungen getroffen.

M4

Nun ja, ich habe mir Mühe gegeben.

F4

Ohne dich wäre die Firma sicher den Bach runter gegangen. Die Firma braucht dich!

Table C1*Main Changes of the Script (TASIT-SIM → BASIT-ToM)*

Scene	Text removed	Text added
H#1 (B11)	Actor repeats what another actor says several times in TASIT-SIM. These repetitions have been deleted. M: " I feel as if I can see it another thousand times"	The deleted repetitions were replaced by "oh yeah".
H#2 (A11)		F: "I need another signature here."
H#3 (B1)	F: "...in between making coffee and answering telephone calls"	
H#4 (B7)	F: "...you're indispensable"	
Practice scene (B15)		F: "...and these window sills!"

Note. TASIT-SIM = The Awareness of Social Inference Test – Social Inference Minimal; BASIT-ToM = Basel Version of the Awareness of Social Inference Test – Theory of Mind. Scene code consists of a BASIT-ToM scene number (scene) and a letter (message type, H = Honesty). Scene code in the brackets denotes a TASIT-SIM scene consisting of form A or form B and a scene number. M = male actor, F = female actor.

Appendix D

Table D1

Distribution of the Actors to the BASIT-ToM and Corresponding TASIT-SIM Scenes

Scene	BASIT-ToM	TASIT-SIM
Honesty		
B11	M1+F1	M+F
A11	M2+F2	M+M
B1	F3+M3	M+F
B7	F4+M4	M+F
Simple Sarcasm		
A6	F1+M3	M+F
B2	M1+F3	M+F
B10	M2+F4	M+F
B13	F2+M4	M+F
Paradoxical Sarcasm		
A3	F3+M2	M+F
A15	M3+F2	M+F
B5	M4+F1	M+F
B8	F4+M1	M+F
B15 (practice scene)	F1+M2	M+F

Note. The four female actors are labelled F1 to F4 and the four male actors are labelled M1 to M4. The actors who communicate the respective message type (i.e., Honesty, Simple Sarcasm, or Paradoxical Sarcasm) are printed in bold. BASIT-ToM = Basel Version of the Awareness of Social Inference Test – Theory of Mind; TASIT-SIM = The Awareness of Social Inference Test – Social Inference Minimal.

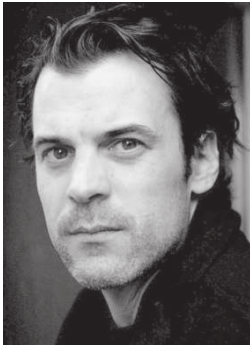
Appendix E

Document E1

Background Information of each BASIT-ToM Intensity Version Scene for the Actors

Honesty 1 (B11) – Theatre

Main Actor: **M1**



Communication Partner: **F1**



Background Information

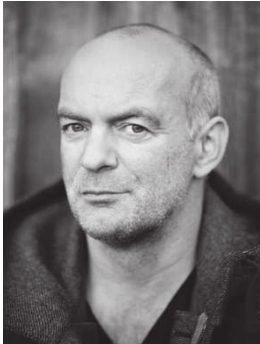
- M1 and F1 are married. In their free time, they do a lot together and like to chat about what they have experienced together.
- M1 and F1 are teachers

Intensities

- **Low intensity:** M1 and F1 have seen a play the week before. M1 and F1 are sitting on the sofa. M1 is reading a book and F1 is reading a magazine that has something written about the play. They casually talk about the play. M1 concentrates on his book during the conversation.
- **Medium intensity:** M1 and F1 have seen a play the night before. They are sitting on the sofa, drinking wine and talking. They rarely go to the theatre, but found the play good.
- **High intensity:** M1 and F1 have just returned home from a visit to the theatre, are sitting on the sofa and drinking wine. Both are theatre lovers and overwhelmed by the play. They reminisce and talk animatedly.

Honesty 2 (A11) – Tickets

Main Actor: M2



Communication Partner: F2



Background Information

- M2 is the boss of a company.
- F2 is M2's secretary. She supports his work in the company and takes on private organizational tasks, such as booking tickets.

Intensities

- **Low intensity:** F2 has organized tickets for M2 and his wife for an almost sold-out play. F2 gives him the tickets and casually clarifies to him that it was very difficult to get the tickets. He also casually thanks her while signing forms that F2 hands him. M2 likes to watch a play from time to time, but is not a fan of it. However, he is glad to be able to make his wife happy by going to the theatre.
- **Medium intensity:** F2 has organized tickets for a gala dinner for M2 and his wife. F2 is pleased and relieved that she was able to organize seats for M2 and his wife. M2 is very much looking forward to the evening and thanks F2 for her efforts.
- **High intensity:** M2 and his wife are big Rolling Stones fans. F2 was able to organize tickets for the last Rolling Stones concert. Now, she proudly shows them to her

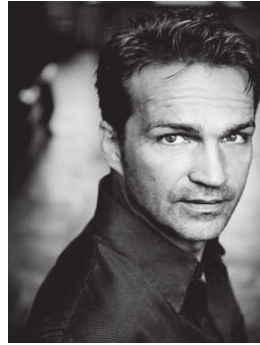
husband. M2 feels excited anticipation and is full of gratitude that F2 organized the tickets for him.

Honesty 3 (B1) – Hard Day

Main Actor: **F3**



Communication Partner: **M3**



Background Information

- F3 is the head of a small law firm with few employees.
- M3 is a lawyer.

Intensities

- **Low intensity:** In the law firm, old files have to be sorted out and destroyed once a year. This work has to be done in addition to the daily work tasks. M3 works as a freelancer in the law office. F3 has asked him to support her in this annual cleanup. While she concentrates on sorting out the last files, M3 says goodbye. It is actually a matter of course for both of them that M3 will help again the next day. Hence, the conversation is held casually.
- **Medium intensity:** The law firm is preparing for a difficult trial that will take place in three days. M3, as a part-time worker, only works 3 days a week. F3 has asked him to come to work on extra days so they can manage the workload. F3 is very happy that M3 helps. For M3 it is a matter of course that he supports his colleagues. Nevertheless, he is happy that his efforts are appreciated.

- **High intensity:** The law firm has to move this week. The employee who should have helped organize the move has called in sick at short notice. M3 has come back from his holiday to help. F3 is very relieved that M3 is willing to do this and thanks him profusely for his help. For M3 it is a matter of course that he supports his colleagues. Nevertheless, he is pleased that his efforts are appreciated.

Honesty 4 (B7) – Promotion

Main Actor: F4



Communication Partner: M4



Background information

- M4 is a team leader and wants to be promoted to head of department.
- F4 is a member of M4's team.
- They work in the district office.

Intensities

- **Low intensity:** The conversation takes place while they are working together. Meanwhile, M4 wants to talk to F4 for a bit and therefore brings up the subject of the promotion. However, he is not very worried that it might not work out. In the previous years, the boss was always satisfied with his work. F4 encourages M4 while they both continue to concentrate on their work. The whole conversation has the character of small talk.

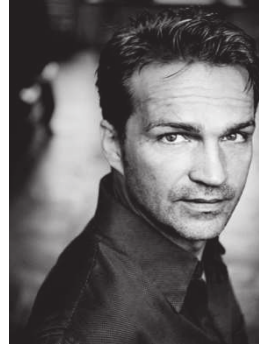
- **Medium intensity:** M4 and F4 are friends. M4 is worried that his boss was not satisfied with his work and someone else will be promoted. F4 wants to reassure and encourage him. They focus on the conversation and not on their work.
- **High intensity:** F4 is in love with M4. She thinks he is a great employee and a promotion is long overdue. When M4 expresses concerns about the possible promotion, she seizes the opportunity and wants to endear herself to him with her compliments.

Simple Sarcasm 1 (A6) – “Seminar”

Main Actor: F1



Communication Partner: M3



Background Information

- F1 and M3 both work as research fellows at a university of applied sciences for business administration. In addition to their work in research, they give seminars and lectures to students and therefore have a lot to deal with professionally.
- F1 is a single mother with two children. F1 can only reconcile work and family life with a great deal of organization and good planning. Her children often have to take a back seat and are frequently looked after by other people. As a result, F1 has a guilty conscience. She often feels stressed.
- M3 is single, has no children and many hobbies. He is able to balance work and leisure time well. He feels neither stressed nor overwhelmed. M3 is ambitious and wants to pursue a career at the university.
- Situation: Two weeks before the start of a seminar, F1 asked M3 to hold the seminar for her. M3 has agreed to her request. Two days before the seminar, M3 cancels.

Intensities

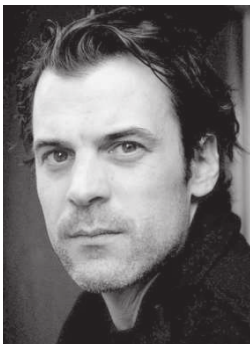
- **Low intensity:** There was no negative history between F1 and M3. They appreciate each other as colleagues, but are not friends in their free time. M3's cancellation annoys her because she relied on his commitment. If she does not find a replacement, she will have to hold the seminar herself.

- **Medium intensity:** F1 has the feeling that she has to perform better overall than M3. She feels that this is unfair and is therefore annoyed by M3 in advance. His cancellation makes her angry.
- **High intensity:** M3 is known at university for not doing anything for his colleagues; he does not keep to agreements and is only concerned about his own advantage. That is why many colleagues and F1 are angry with him. F1 has promised her child to go to his school play. That is why she asked M3 to take over the seminar for her. She has explained the urgency and told him how important it is that she can rely on him. His last-minute cancellation leaves her stunned and very angry.

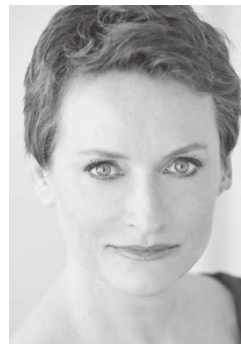
M3 is sincere in his apology. However, F1's problems do not affect him much. Since he wants to keep another appointment at short notice, he cancels F1 without a guilty conscience. This remains the same across all intensities.

Simple Sarcasm 2 (B2) – Weekend away

Main Actor: **M1**



Communication Partner: **F3**



Background Information

- F3 and M1 have been in a relationship for 10 years. They are not married and have no children. Both are very busy at work and have little free time together.
- M1 is a café owner. He spends a lot of time in his café and mostly has to work on weekends. He is a family man. His family tries to solve all problems together. Being there for the other siblings is a high priority for him. His sister very often has minor

problems for which she asks him for advice and therefore comes over. The little free time F3 and he have together is therefore often spent with his sister.

- F3 is an estate agent. She can arrange her working hours freely, but often has to work on weekends as well. She is not a family person. She likes her sister-in-law but thinks she exaggerates her problems and comes over too often. She would like to spend more of her limited free time alone with her partner.

Intensities

- **Low intensity:** It is not a special weekend outing. M1 and F3 had a small discussion beforehand about his sister's frequent appearances. However, F3 already expected M3 to ask her about his sister. Nevertheless, she is now a bit annoyed when she hears from M1 that his sister should also come along to their weekend trip.
- **Medium intensity:** F3 and M1 had already had a small argument the day before because F3 feels that M1's sister takes up too much space in their relationship. Therefore, she is angry that his sister should now also come along to the weekend trip.
- **High intensity:** Beforehand, there were almost daily arguments between F3 and M1 about the fact that they spend too little time together. In addition, F3 thinks that M1's sister is more important than her, which annoys her more and more and puts a lot of strain on their relationship. M1 has often promised weekend trips together and, contrary to his promises, has always invited his sister. Because F3 was very angry about this, M1 promised her that they would celebrate their ten-year anniversary as a couple on a weekend trip. Now F3 is stunned and very angry that M3 is already thinking about inviting his sister again.

M3 acts as if there has never been a discussion about his sister's frequent appearances. He is worried about his sister and therefore seriously considers taking

her with them. He does not understand why F3 has a problem with this. His behaviour remains the same across all intensities.

Simple Sarcasm 3 (B10) – Going out

Main Actor: **M2**



Communication Partner: **F4**



Background Information

- M2 is an artist. He is confident, articulate and good with people.
- F4 is a gallery owner. She is a confident, content person.

Intensities

- **Low intensity:** F4 and M2 have been a happy couple for many years. They harmonize very well. Their conversations are usually full of small but affectionate banter. They like to tease each other. They do not have to give their partner any linguistic or visual clues that what they say is meant ironically. They each understand their partner's irony. Irony has become such a part of their shared conversation that they hardly notice that they are being ironic with each other.
- **Medium intensity:** F4 and M2 have only been a couple and living together for a short time. They have a lot of fun together and flirt a lot. They often do this through an ironic communication style. They give their partner slight linguistic and visual cues (e.g. smiling) that what is said is meant ironically.

- **High intensity:** F4 and M2 have only been a couple for a short time and have been living together for a week. They have a lot of fun together and flirt a lot. F4 came home from work late last night, when M2 was already asleep. Today she is going out alone to meet a friend for dinner; she has been doing this for a while. M2 now teases her about it, although he knows why she goes out and whom she meets. Additionally, he is also teasing her because she had already gone out yesterday and he was home alone. Quite tough in the first week of living together! Accordingly, he gives strong linguistic and visual indications that what is said is meant ironically. F4 understands M4's irony and reacts with irony.

Simple Sarcasm 4 (B13) – Shirt

Main Actor: **F2**



Communication Partner: **M4**



Background Information

- F2 and M4 have been married for many years. M4 gets a shirt from his mother every year for his birthday. F2 thinks these shirts are ugly.
- F2 is a team leader in a company.
- M4 is a cook.
- Situation: F2 and M4 get dressed to go to the Christmas party of F2's company.

Intensities

- **Low intensity:** F2 and M4 have no quarrel beforehand and are looking forward to the upcoming party. He tries to choose a nice shirt and wants her to like it. She has given

up explaining to him, which shirts can and cannot be worn. Therefore, she "emotionlessly" to pityingly smiles at his shirt choice.

- **Medium intensity:** F2 and M4 want to go to the Christmas party of F2's company. At previous events, F2 has always been ashamed of her husband's poor attire. F2 has clearly told her husband several times not to wear the shirts he got from his mother. She has bought him new, fancy shirts and is annoyed when he wants to wear one of his mother's shirts again. M4 cannot understand why his wife does not like the shirts his mother regularly gives him. He wants to please his wife (F2) and therefore seriously asks her opinion.

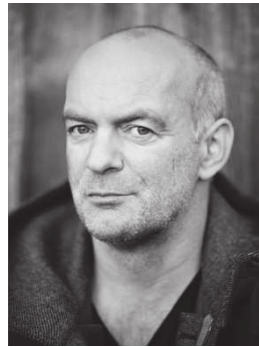
- **High intensity:** Immediately before the scene, F2 and M4 had a major argument. She is still angry and therefore immediately goes off the deep end when M4 also wants to wear one of his mother's ugly shirts to her Christmas party. M4 wants to appease his wife after the argument by approaching her and asking for her opinion. In doing so, he also wants to signal to her that he is happy to go to the party.

Paradoxical Sarcasm 1 (A3) – Report

Main Actor: F3



Communication Partner: M2



Background Information

- F3 is a management consultant; she works efficiently and reliably and is hardworking.
- M2 is the boss of the management consultancy. As the boss, he sets clear guidelines and has high expectations of his employees.

Intensities

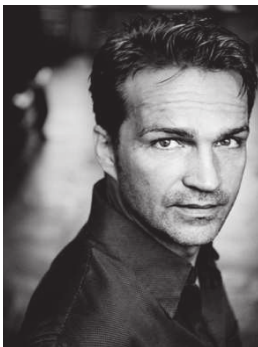
- Situation: M2 is the boss of a management consultancy and is F3's superior. On Friday, M2 asked F3 to finish a report by Monday. On Monday morning, there is a very long report on his desk. He actually appreciates short reports, but sees the amount of work F3 put into the report over the weekend.
- **Low intensity:** M2 and F3 appreciate each other and have a collegial relationship. F3 is used to writing long reports for her boss at short notice. She has no problem with this, yet she would have liked to spend the weekend doing something else. M2 wants to convey appreciation and thanks to her by asking.
- **Medium intensity:** There is a clear hierarchical separation between M2 and F3. M2 knew that F3 hardly had time to write a report on the weekend. Nevertheless, he gave her this task because the report was urgent. Accordingly, F3 is annoyed. M2 realizes

that the work for the report, given its length, was probably more time-consuming than he had imagined and now wants to convey appreciation to F3 by asking.

- **High intensity:** M2 and F3 do not have a good relationship. M2 knew that F3 did not have time to write a report at the weekend because F3 had planned to go to her mother's birthday party over the weekend. F3's mother was turning 70 and F3 had been looking forward to the party for weeks. F3 then had to write the report at night. Accordingly, she is angry with her boss.

Paradoxical Sarcasm 2 (A15) – Earrings

Main Actor: **M3**



Communication Partner: F2



Background Information

- F2 and M3 are married and wealthy.
- M3 is an architect.
- F2 is a fashion designer.

Intensities

- **Low intensity:** F2 and M3 have a good relationship with F2's sister. F2 was happy to receive a gift from her. However, neither of them likes the earrings. F2's sister likes to shop at the flea market. F2 and M3 amuse themselves a little about the earrings.

- **Medium intensity:** F2 and M3 have a good relationship with F2's sister. However, she is very thrifty and has a tendency to give cheap trinkets. These earrings, which were obviously cheap, are a typical gift. Accordingly, F2 and M3 have already expected such a gift and are very amused.
- **High intensity:** For F2 and M3, expensive gifts are important. They consider cheap gifts as worthless and embarrassing. They despise F2's sister and laugh at her and her embarrassing gift.

Paradoxical Sarcasm 3 (B5) – Meal

Main Actor: **M4**



Communication Partner: **F1**



Background Information

- F1 is a journalist
- M4 is a computer scientist

Intensities

- **Low intensity:** F1 and M4 have been a happy couple for many years. They harmonize very well and understand each other blindly. Their conversations are usually full of small but affectionate asides. They like to tease each other. They do not have to give their partner any linguistic or visual clues that what they say is meant ironically. They each understand their partner's irony. Irony has become such a part of their shared conversation that they hardly notice that they are being ironic with each other.

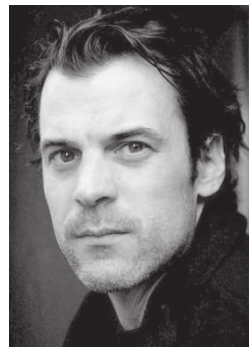
- **Medium intensity:** F1 and M4 have only been a couple and living together for a short time. They have a lot of fun together and flirt a lot. They often do this through an ironic communication style. They give their partner slight linguistic and visual cues (e.g. smiling) that what is said is meant ironically.
- **High intensity:** F1 and M4 have only been a couple for a short time and have been living together for a week. They have a lot of fun together and flirt a lot. They often do this through an ironic style of communication. Now F1, who is a passionate cook, has cooked something new and is of course curious how M4 liked it. M4 knows about F3's curiosity and wants to tease her a little in the form of strong irony. He gives correspondingly strong linguistic and visual indications that what is said is meant ironically. F1 also understands M4's irony and plays along.

Paradoxical Sarcasm 4 (B8) – Night Shift

Main Actor: **F4**



Communication Partner: **M1**



Background Information

- F4 and M1 are married.
- M1 is an air traffic controller and works in shifts. He works a lot of overtime and very often fills in for other colleagues. He has very little free time.

- F4 works in a travel agency. She works normal office hours and does not have to work overtime.

Intensities

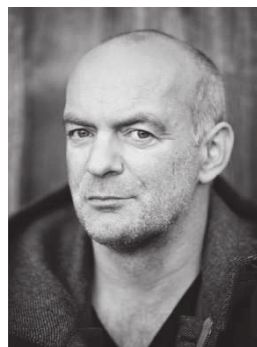
- **Low intensity:** It is not a special restaurant visit. Nevertheless, F4 would have liked to go out with her husband. She is a bit disappointed with M1 for not managing to turn down his boss.
- **Medium intensity:** F4 and M1 have already had a small argument the day before about M1 always agreeing to short-term night shifts and working overtime. Therefore, she is angry that he is cancelling the long-planned visit to the restaurant again.
- **High intensity:** Beforehand, there was an almost daily argument between F4 and M1 about M1 working too much and not resisting overtime and night shifts scheduled at short notice. The restaurant visit was planned by M1 as a reconciliation and was supposed to take place on their wedding day. Now F3 is bewildered and very angry that M1 is cancelling their evening together again.

Practice Scene: Paradoxical Sarcasm (B15) – New House

Main Actor: F1



Communication Partner: M2



Background Information

- F1 and M2 have been married for 20 years and live together in an apartment.
- Simone is M2's cousin and lives in a neighbouring village. She is currently building a single-family house with her husband.

Intensities

- **Low intensity:** F1 generally likes Simone. She has no problem with her and is not jealous that Simone is able to build a new house. However, F1 thinks that Simone has questionable taste and no sense of aesthetics. Therefore, she makes fun of Simone's choice of colour.
- **Medium intensity:** F1 does not like Simone very much, so she ridicules Simone's new house.
- **High intensity:** F1 cannot stand Simone. She is jealous that Simone can afford to build her own home. F1 thinks that Simone does not deserve to own a home. In order to depreciate Simone, she viciously gossips about her at every opportunity.
- M2 also thinks that Simone did not choose a good colour for her house. However, M2 finds his wife is gossiping excessively. He is not interested in his cousin and her house. For this reason, he is not interested in the "gossip" with his wife and wants to end it as quickly as possible with the sentence "maybe she is colour-blind" so that he can continue reading the newspaper in peace. M2's reactions hardly change across all intensities.

Appendix F

Table F1: Scene H#1

Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Möchte er ihr vermitteln, dass sie unterschiedlicher Meinung sind?	Does he want to convey to her that they disagree?	cogn	2 nd	No	Is Michael agreeing with Ruth about the movie?	cogn	2 nd	yes	a
War er von dem Theaterstück gelangweilt?	Was he bored by the play?	aff	1 st	No	Is he openly pleased that he saw the movie?	aff	1 st	yes	b
Versucht er ihr zu sagen, dass er die Schauspieler schlecht findet?	Is he trying to say he thought the actors were bad?	aff	1 st	No	Is he trying to say he thought the actors were good?	aff	1 st	yes	c

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Denkt er, dass ihr das Theaterstück gefallen hat?	Does he think she enjoyed the play?	aff	2 nd	yes	Does he think the movie was bad?	aff	1 st	no	d
Ist er in sie verliebt?	Is he in love with her?	hyper		no					e

Note. BASIT-ToM = Basel Version of the Awareness of Social Inference Test – Theory of Mind; TASIT-SIM = The Awareness of Social Inference Test – Social Inference Minimal; cogn = cognitive Theory of Mind; aff = affective Theory of Mind, hyper = hypermentalization.

Table F2: Scene H#2

Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Möchte er ihr vermitteln, dass er ihr für die Organisation der Eintrittskarten dankbar ist?	Is Michael trying to show he appreciates Gary getting the tickets?	aff	1 st	yes	Is Michael trying to show he appreciates Gary getting the tickets?	aff	1 st	yes	a
Freut er sich, dass sie ihm die Eintrittskarten organisiert hat?	Is he happy that she organised the tickets for him	aff	1 st	yes	Is Michael annoyed Gary got him the tickets?	aff	1 st	no	b
Versucht er ihr zu sagen, dass er sich über die Eintrittskarten freut?	Is he trying to say he's pleased about the tickets?	aff	1 st	yes	Is he trying to say he's pleased about the tickets?	aff	1 st	yes	c

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Denkt sie, dass er kein Interesse an den Eintrittskarten hat?	Does she think that he is not interested in the tickets?	aff	2 nd	no	By the end of the scene, does Gary think Michael wants to go?	cogn	2 nd	yes	d
Möchte sie sich seine Freundschaft erkaufen?	Does he want to bribe him to become friends with him?	hyper		no					e

Note. BASIT-ToM = Basel Version of the Awareness of Social Inference Test – Theory of Mind; TASIT-SIM = The Awareness of Social Inference Test – Social Inference Minimal; cogn = cognitive Theory of Mind; aff = affective Theory of Mind, hyper = hypermentalization.

Table F3: Scene H#3

Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Möchte sie, dass er sich geschätzt fühlt?	Is Ruth trying to make Michael feel appreciated?	aff	2 nd	yes	Is Ruth trying to make Michael feel appreciated?	aff	2 nd	yes	a
Ärgert sie sich über ihn?	Is she annoyed with him?	aff	1 st	no	Is she annoyed with him?	aff	1 st	no	b
Versucht sie ihm zu sagen, dass er ihr keine Hilfe war?	Is she trying to tell him that he has not been a big help?	cogn	1 st	no	Is he trying to say that he has been a big help?	cogn	1 st	yes	c
Denkt er, dass sie möchte, dass er am nächsten Tag wiederkommt?	Does he think she wants him to come back the next day?	cogn	2 nd	yes	Is she annoyed with him?	aff	1 st	no	d

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Bedankt sie sich bei ihm, damit er ihr am nächsten Tag erneut hilft?	Does she thank him so that he will help her again the next day?	hyper		no					e

Note. BASIT-ToM = Basel Version of the Awareness of Social Inference Test – Theory of Mind; TASIT-SIM = The Awareness of Social Inference Test – Social Inference Minimal; cogn = cognitive Theory of Mind; aff = affective Theory of Mind, hyper = hypermentalization.

Table F4: Scene H#4

Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Möchte sie ihm vermitteln, dass er kaum Chancen auf eine Beförderung hat?	Is she trying to convey to him that he has little chance of being promoted?	cogn	1 st	no	Is Ruth sending Michael up about his chances of a promotion?	cogn	1 st	yes	a
Würde sie ihm die Beförderung missgönnen?	Would she begrudge him the promotion?	aff	1 st	no	Would she like him to get the promotion?	aff	1 st	yes	b
Versucht sie ihm zu sagen, dass er gut gearbeitet hat?	Is she trying to say he has worked really well?	cogn	1 st	yes	Is she trying to say he has worked really well?	cogn	1 st	yes	c

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Denkt er, dass sie ihre Komplimente ernst meint?	Does he think she is serious about her compliments?	cogn	2 nd	yes	Does she think he deserves a promotion?	aff	1 st	yes	d
Lobt sie ihn, um sich berufliche Vorteile zu verschaffen?	Is she praising him to gain a professional advantage?	hyper		no					e

Note. BASIT-ToM = Basel Version of the Awareness of Social Inference Test – Theory of Mind; TASIT-SIM = The Awareness of Social Inference Test – Social Inference Minimal; cogn = cognitive Theory of Mind; aff = affective Theory of Mind, hyper = hypermentalization.

Table F5: Scene sS#1

Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Möchte sie, dass er sich trotz seiner Absage gut fühlt?	Is he trying to make him feel OK despite his cancellation?	aff	2 nd	no	Is Ruth trying to make Michael feel OK?	aff	2 nd	no	a
Ärgert sie sich über seine Absage?	Is she annoyed by his cancellation?	aff	1 st	yes	Is she annoyed with him?	aff	1 st	yes	b
Versucht sie ihm zu sagen, dass seine Absage ihr Probleme verursacht?	Is she trying to tell him that his cancellation is causing her problems?	cogn	1 st	yes	Is she trying to say that he is causing a big problem?	cogn	1 st	yes	c

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Denkt er, dass seine Absage für sie kein Problem darstellt?	Does he think that his cancellation is not a problem for her?	cogn	2 nd	no	Does she believe he is too busy to take the class?	cogn	2 nd	no	d
Glaubt sie, dass er ihr absagt, weil er ihr schaden möchte?	Does she think he is cancelling on her because he wants to harm her?	hyper		no					e

Note. BASIT-ToM = Basel Version of the Awareness of Social Inference Test – Theory of Mind; TASIT-SIM = The Awareness of Social Inference Test – Social Inference Minimal;

cogn = cognitive Theory of Mind; aff = affective Theory of Mind, hyper = hypermentalization.

Table F6: Scene sS#2

Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Möchte er weitere Verwandte einladen?	Does he want to invite other relatives?	cogn	1 st	no	Is he seriously suggesting they invite other family members?	cogn	1 st	no	a
Ärgert er sich über sie?	Is he annoyed with her?	aff	1 st	yes	Is he annoyed with her?	aff	1 st	yes	b
Versucht er ihr zu sagen, dass ihre Verwandten mitkommen sollen?	Is he trying to say he want her relatives to come?	cogn	1 st	no	Is he trying to say he doesn't want her relatives to come?	cogn	1 st	yes	c
Denkt sie, er möchte, dass ihre Verwandten mitkommen?	Does she think he wants her relatives to come?	cogn	2 nd	no	Does she think he wants her relatives to come?	cogn	2 nd	no	d

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Denkt er, dass sie den Besuch ihrer Schwester anspricht, um ihn zu provozieren?	Does he think she mentioning her sister's visit to provoke him?	hyper		no					e

Note. BASIT-ToM = Basel Version of the Awareness of Social Inference Test – Theory of Mind; TASIT-SIM = The Awareness of Social Inference Test – Social Inference Minimal;
cogn = cognitive Theory of Mind; aff = affective Theory of Mind, hyper = hypermentalization.

Table F7: Scene sS#3

Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Möchte er, dass sie sich schlecht fühlt?	Is he trying to make her feel bad?	aff	2 nd	no	Is Gary trying to make Ruth feel bad?	aff	2 nd	no	a
Ist er wütend auf sie?	Is he angry with her?	aff	1 st	no	Is he angry with her?	aff	1 st	no	b
Versucht er ihr zu sagen, dass sie gut aussieht?	Is he trying to say that she looks good?	cogn	1 st	yes	Is he trying to say that she looks ok?	cogn	1 st	yes	c
Denkt sie, dass er scherzt?	Does she think he is joking with her?	cogn	2 nd	yes	Does she think he is joking with her?	cogn	2 nd	yes	d

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Ist er in sie verliebt?	Is he in love with her?	hyper		no					e

Note. BASIT-ToM = Basel Version of the Awareness of Social Inference Test – Theory of Mind; TASIT-SIM = The Awareness of Social Inference Test – Social Inference Minimal; cogn = cognitive Theory of Mind; aff = affective Theory of Mind, hyper = hypermentalization.

Table F8: Scene sS#4

Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Möchte sie ihm vermitteln, dass sein Hemd schrecklich aussieht?	Is she trying to make him understand that his shirt looks awful?	cogn	1 st	yes	Is Ruth reassuring Gary that the shirt is nice?	cogn	1 st	no	a
Ärgert sie sich darüber, dass er das Hemd trägt?	Is she annoyed that he is wearing the shirt?	aff	1 st	yes	Is she happy for him to wear the shirt?	aff	1 st	no	b
Versucht sie ihm zu sagen, dass sie das Hemd schön findet?	Is she trying to say the shirt is beautiful?	cogn	1 st	no	Is she trying to say the shirt is awful?	cogn	1 st	yes	c

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Denkt er, dass sie das Hemd schrecklich findet?	Does he think she finds the shirt awful?	cogn	2 nd	yes	Does she think the shirt is OK?	cogn	1 st	no	d
Denkt sie, dass er das Hemd anziehen möchte, um sie zu provozieren?	Is she thinking he wants to wear the shirt to provoke her?	hyper		no					e

Note. BASIT-ToM = Basel Version of the Awareness of Social Inference Test – Theory of Mind; TASIT-SIM = The Awareness of Social Inference Test – Social Inference Minimal;
cogn = cognitive Theory of Mind; aff = affective Theory of Mind, hyper = hypermentalization.

Table F9: Scene pS#1

Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Möchte sie ihm vermitteln, dass das Schreiben des Berichts aufwendig war?	Is she trying to convey to him that writing the report took a lot of work?	cogn	1 st	yes	Is Ruth denying the report took a lot of work?	cogn	1 st	no	a
Ist sie unzufrieden, weil sie am Wochenende arbeiten musste?	Does Ruth seem unhappy about working all weekend?	aff	1 st	yes	Does Ruth seem happy about working all weekend?	aff	1 st	no	b
Versucht sie ihm zu sagen, dass sie ein anstrengendes Wochenende hatte?	Is she trying to tell him that she had an exhausting weekend?	cogn	1 st	yes	Is she trying to say she had a lazy, relaxing weekend?	cogn	1 st	no	c

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Denkt er, dass es für sie in Ordnung war, am Wochenende zu arbeiten?	Does he think it was OK for her working all weekend?	aff	2 nd	yes / no*	Does Michael think she took it easy on the weekend?	cogn	1 st	no	d
Spricht er sie auf den Bericht an, um sie zu provozieren?	Is he talking to her about the report to provoke her?	hyper		no					e

Note. BASIT-ToM = Basel Version of the Awareness of Social Inference Test – Theory of Mind; TASIT-SIM = The Awareness of Social Inference Test – Social Inference Minimal;

cogn = cognitive Theory of Mind; aff = affective Theory of Mind, hyper = hypermentalization.

*correct response = yes (low intensity), correct response = no (medium and high intensities)

Table F10: Scene pS#2

Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Möchte er sich negativ über die Ohringe äussern?	Is he commenting negatively on the earrings?	cogn	1 st	yes	Is Michael being complimentary about the dress?	cogn	1 st	no	a
Amüsiert er sich über die Qualität der Ohringe?	Is she making fun of the quality of the earrings?	aff	1 st	yes	Does he like the dress?	aff	1 st	no	b
Versucht er ihr zu sagen, dass die Ohringe hochwertig aussehen?	Is he trying to tell her that the earrings look high quality?	cogn	1 st	no	Is he trying to say the dress looks cheap?	cogn	1 st	yes	c
Denkt er, dass ihr die Ohringe gefallen?	Does he think she likes the earrings?	aff	2 nd	no	Does he think Ruth's sister paid a lot for the dress?	cogn	1 st	no	d

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Können sie die Schwester nicht ausstehen?	Can they not stand the sister?	hyper		no					e

Note. BASIT-ToM = Basel Version of the Awareness of Social Inference Test – Theory of Mind; TASIT-SIM = The Awareness of Social Inference Test – Social Inference Minimal; cogn = cognitive Theory of Mind; aff = affective Theory of Mind, hyper = hypermentalization.

Table F11: Scene pS#3

Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Möchte er sich negativ über das Essen äussern?	Is he commenting negatively on the meal?	cogn	1 st	no	Is Gary trying to compliment Ruth on the meal?	cogn	1 st	yes	a
Hat er das Essen genossen?	Did he enjoy the meal?	aff	1 st	yes	Did he enjoy the meal?	aff	1 st	yes	b
Versucht er ihr zu sagen, dass das Essen schlecht geschmeckt hat?	Is he trying to tell her that the meal tasted awful?	cogn	1 st	no	Is he trying to say that the meal tasted awful?	aff	1 st	no	c
Denkt sie, dass ihm das Essen schlecht geschmeckt hat?	Does she think he did not like the meal?	cogn	2 st	no	Does he think she is a good cook?	cogn	1 st	yes	d

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Ist er in sie verliebt?	Is he in love with her?	hyper		no					e

Note. BASIT-ToM = Basel Version of the Awareness of Social Inference Test – Theory of Mind; TASIT-SIM = The Awareness of Social Inference Test – Social Inference Minimal; cogn = cognitive Theory of Mind; aff = affective Theory of Mind, hyper = hypermentalization.

Table F12: Scene pS#4

Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Möchte sie, dass er sich aufgrund seiner Absage schlecht fühlt?	Is she trying to make him feel bad about his cancellation?	aff	2 nd	yes	Is Ruth trying to make Gary feel OK about cancelling?	aff	2 nd	no	a
Ist sie von seiner Absage enttäuscht?	Is she disappointed by his cancellation?	aff	1 st	yes	Is she happy to cancel?	aff	1 st	no	b
Versucht sie ihm zu sagen, dass sie froh ist, an diesem Abend zu Hause bleiben zu können?	Is she trying to tell him that she is happy to stay at home that night?	aff	1 st	no	Is she trying to say she wanted to go out that night?	cogn	1 st	yes	c

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Denkt er, dass sie froh ist, an diesem Abend zu Hause bleiben zu können?	Does he think she is happy to stay at home that night?	aff	2 nd	no	<i>By the end of the scene:</i> Does he think she's upset about cancelling dinner?	aff	2 nd	yes	d
Glaubt sie, dass er eine Affäre hat und nur vorgibt nachts zu arbeiten?	Does she think he is cheating on her and just pretending to work nights?	hyper		no					e

Note. BASIT-ToM = Basel Version of the Awareness of Social Inference Test – Theory of Mind; TASIT-SIM = The Awareness of Social Inference Test – Social Inference Minimal; cogn = cognitive Theory of Mind; aff = affective Theory of Mind, hyper = hypermentalization.

Table F13: Scene Practice(pS)

Test Questions of the BASIT-ToM and TASIT-SIM with their Corresponding Correct Answers, Types and Orders of Theory of Mind

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Möchte sie sich positiv über Simones Haus äussern?	Is she being complementary about Simone's house?	cogn	1 st	no	Is Ruth being complementary about her friend's house?	cogn	1 st	no	a
Gefällt ihr die Farbgebung des Hauses?	Does she like the color scheme of the house?	aff	1 st	no	Does she disapprove of the colour scheme?	aff	1 st	yes	b
Versucht sie zu sagen, dass sie die Farbgebung des Hauses schrecklich findet?	Is she trying to say the color scheme of the house is dreadful?	cogn	1 st	yes	Is she trying to say the colour scheme is dreadful?	cogn	1 st	yes	c
Denkt er, dass ihr Simones neues Haus gefällt?	Does he think she likes Simone's new house?	aff	2 nd	no	Does she think her friend has good taste?	cogn	1 st	nein	d

BASIT-ToM					TASIT-SIM				
Question (German)	Question (English)	ToM type	ToM order	Correct answer	Question	ToM type	ToM order	Correct answer	Question ID
Können beide Simone nicht ausstehen?	Can they both not stand Simone?	hyper		nein					e

Note. BASIT-ToM = Basel Version of the Awareness of Social Inference Test – Theory of Mind; TASIT-SIM = The Awareness of Social Inference Test – Social Inference Minimal; cogn = cognitive Theory of Mind; aff = affective Theory of Mind, hyper = hypermentalization.

Appendix G

Table G1

Main Differences between the BASIT-ToM and TASIT-SIM

	BASIT-ToM	TASIT-SIM
Test length		
Test versions	One test form (scenes from the TASIT-SIM forms A and B are included)	Two parallel test forms, named A and B (form A has been used in the majority of published studies)
Number of scenes	9 (3x3) scenes and one practice scene	Form A: 15 (5x3) scenes Form B: 15 (5x3) scenes In total 30 scenes
Scenes		
Distribution of actors	Each actor portrays a message type only once.	Unevenly distributed across message types
Gender distribution	Evenly across message types. Either a male actor or a female actor portrays the message types.	Almost evenly across message types of the test scenes but across all scenes more male actors (61%)
Age of the actors	Middle-aged actors	Quite wide range of age, but rather actors in their twenties/thirties
Interaction scenes	Communication partners react in a realistic manner	Communication partners act in a neutral emotional state
Script	The names of the actors are omitted and replaced as "she" / "he"	Each actor bears a name which he/she keeps throughout all scenes

	BASIT-ToM	TASIT-SIM
Production design	Consistent, realistic settings: at home or at the office; camera perspective and setting dimensions convey realism	Inconsistent and partly unrealistic settings (partly black background, office or living room setting); camera perspective and setting dimensions do not convey realism (e.g., mostly long shots)
Test questions	5 test questions per each scene assigned to ToM types and orders (i.e., 1 st order affective ToM: $n=17$, 1 st order cognitive ToM: $n=17$, 2 nd order affective ToM: $n=10$, 2 nd order cognitive ToM: $n=8$, hypermentalization: $n=13$) Questions refer to scene's main actor and his/her communication partner	4 test questions per each scene divided based on content (i.e., 1 st order affective ToM: $n=18$, 1 st order cognitive ToM: $n=24$, 2 nd order affective ToM: $n=5$, 2 nd order cognitive ToM: $n=5$) Questions refer to the scene's main actor

Note. BASIT-ToM = Basel Version of the Awareness of Social Inference Test–Theory of Mind; TASIT-SIM = The Awareness of Social Inference Test–Social Inference Minimal; ToM = Theory of Mind

Appendix H

Document H1

Details of the Programming and Data Storage

BASIT-ToM intensity versions scenes were implemented as part of a suit of tests. The order was as follows:

1. BASIT-ER intensity version scenes (masked)
2. BASIT-ToM intensity version scenes
3. Gender discrimination test (by use of facial stimuli)
4. Facial Emotional Intensity Recognition Test – Congruent (FEIRT-C) (Chiu et al., 2016)
5. Facial Emotion Intensity Recognition Test – Congruent and Incongruent (FEIRT-CIC) (Chiu et al., 2018)

These five tests were included in an application written in Python 2.7 using version 1.84.2 of the PsychoPy package (Pierce, 2007, 2009). The application was installed on a workstation running Ubuntu 16.04 LTS.

The pseudo-random selection and order of scenes shown to the participants was created beforehand using R: First, a random sample containing message type (“Honesty”, “Simple Sarcasm”, “Paradoxical Sarcasm”) exactly four times was created, with the constraint that no successive message type was equal. For each participant ID the order of the four scenes per message type was randomised. Then, the intensities were randomly distributed: The 240 first scenes of each message type were randomly splitted in three parts of 80 scenes with low, medium, and high intensity, respectively. For each intensity of the first scene (e.g., Honesty#1_low) the second scene of the given message type was then randomly splitted into two parts of 40 scenes with the two other intensities (Honesty#2_medium and Honesty#2_high). The intensity of the third scene was then fixed to the intensity that was not yet assigned to the first two scenes. For the three fourth scenes (one per message type), a similar mechanism was used: The 240 scenes of the first message types were randomly splitted in three parts of 80 scenes with low, medium, and high intensities. For each intensity of the first message type, the second message type was randomly splitted into two parts of 40

scenes with the two other intensities. The intensity of the last message type was fixed by the intensity not yet assigned to the first two message types. Lastly, the intensity of the practice scene shown to each participant at the beginning of the test was randomly assigned, such that each intensity of the practice scene was shown to 80 participants.

This resulted in:

- Each participant viewed the practice scene and 12 scenes (4 per message type),
- Consecutive scenes with different message types,
- Each participant viewed at least once a low intensity, a medium intensity, and a high intensity version scene per message type
- Each participant viewed four low intensity, four medium intensity, and four high intensity version scenes in total.
- Each intensity version scene was shown to 80 participants.
- The first scene was the practice scene, shown in each intensity to 80 participants, respectively.

Both, the sequence of the five questions [i.e., four types of questions, originating from The Awareness of Social Inference Test – Social Inference Minimal, (i.e., “doing”, “saying”, “thinking”, “feeling”) and the hypermentalization question] phrased for each scene, and the “Yes”-“No”-answers phrased for each scene, were also pseudo-randomly fixed using R. For each message type (“Honesty”, “Simple Sarcasm”, “Paradoxical Sarcasm”), the correct answers for the “doing”-, “saying”-, “thinking”- and “feeling”-questions were randomly fixed to 50% “Yes”- and 50% “No”-answers over all four scenes and five types of questions per message type. Notably, in the “Paradoxical Sarcasm” scene “Bericht” (PS#1), the correct answer on the “thinking” question depends on the intensity of the scene (“Yes” for low intensity, otherwise “No”). The correct answer for all “hypermentalization” questions was set to “No”. The order of the five questions was randomly determined for each scene (independent of the intensity).

The results were stored in to 240 separate files containing the different sequences of scenes shown to the individual participants, together with the five questions and correct answers for each scene.

The sequence of the 13 scenes (1 practice scene, four scenes of three-interaction type) for a specific participant was loaded into the application at start. The respective sequence of the scenes was allocated to the participant's ID.

The test started with a screen summarizing the instructions to the participant. By clicking on the button "next", the practice scene started. Afterwards, the participant was asked whether anything was still unclear regarding the test procedure. If not, the participant started the test by clicking on the button "next". In the following, the 12 scenes were shown to the participant in the prepared pseudo-randomized order. After each scene, the participants were asked to answer the five Yes/No questions. The participants were shown a message that the test was completed, before the application started the next test.

Test results were saved as .csv file and transferred to a secure cloud server provided by the Clinical Trial Unit of the University Hospital Basel (UHBS), backed up by the IT Department of the UHBS.

References:

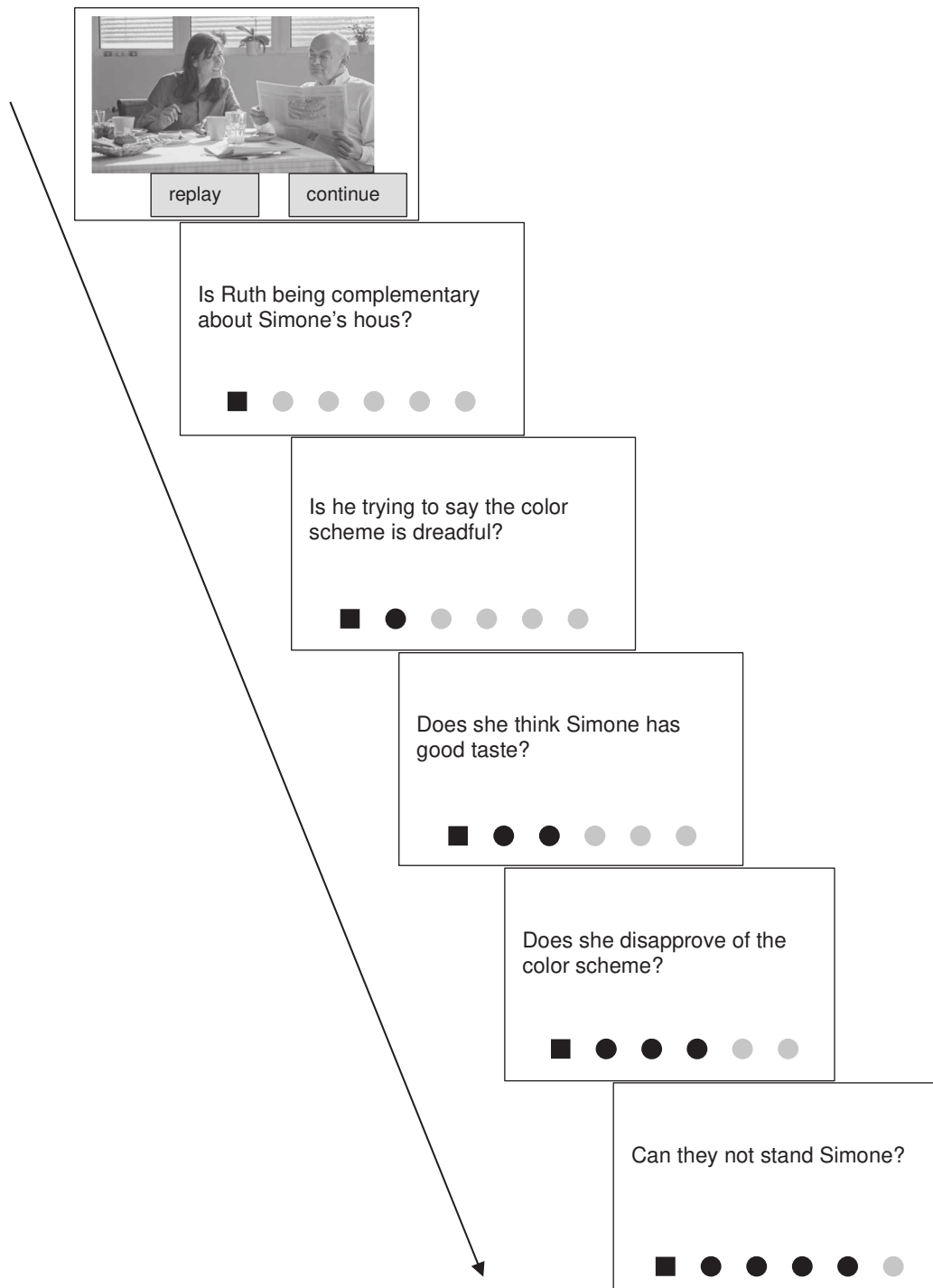
- masked. The Basel Version of the Awareness of Social Inference Test – Emotion Recognition (BASIT-ER): Preliminary Validation Analyses in Healthy Adults.
- Chiu, I., Piguet, O., Diehl-Schmid, J., Riedl, L., Beck, J., Leyhe, T., . . . Sollberger, M. (2016). Dissociation in Rating Negative Facial Emotions between Behavioral Variant Frontotemporal Dementia and Major Depressive Disorder. *Am J Geriatr Psychiatry*, 24(11), 1017-1027. doi:10.1016/j.jagp.2016.06.011
- Chiu, I., Piguet, O., Diehl-Schmid, J., Riedl, L., Beck, J., Leyhe, T., . . . Sollberger, M. (2018). Facial Emotion Recognition Performance Differentiates Between Behavioral Variant Frontotemporal Dementia and Major Depressive Disorder. *J Clin Psychiatry*, 79(1). doi:10.4088/JCP.16m11342

- Peirce J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, 2 (10), 1-8. doi:10.3389/neuro.11.010.2008
- Peirce, J. W. (2007). PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods*, 162 (1-2):8-13 doi:10.1016/j.jneumeth.2006.11.017

Appendix I

Figure I1

Exemplary Representation of the BASIT-ToM Computer-Based Application Process



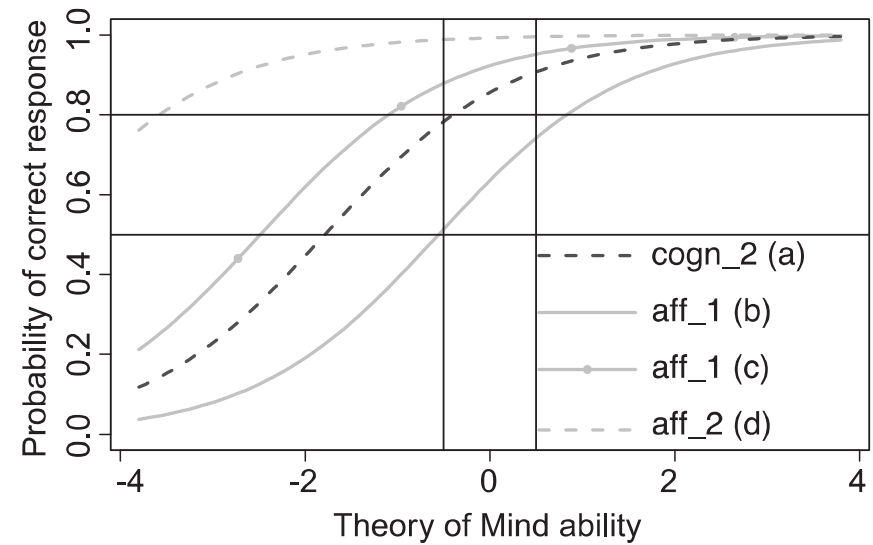
Note. The square represents the scene and the five circles the five questions. By clicking the black square, participants could rewatch the scene. By clicking a black circle, participants could reread the respective question and, if wanted, modify their answer.

Appendix J

Table J.1: Item Characteristic Curves of the Intensity Version Scenes for the Message Type Honesty

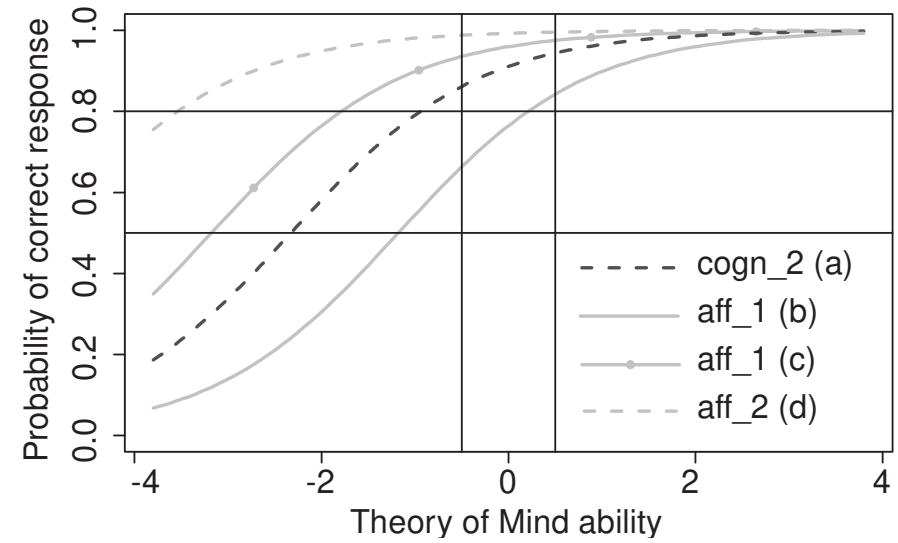
Intensity versions fitting the Rasch Model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

H#1_low **2**



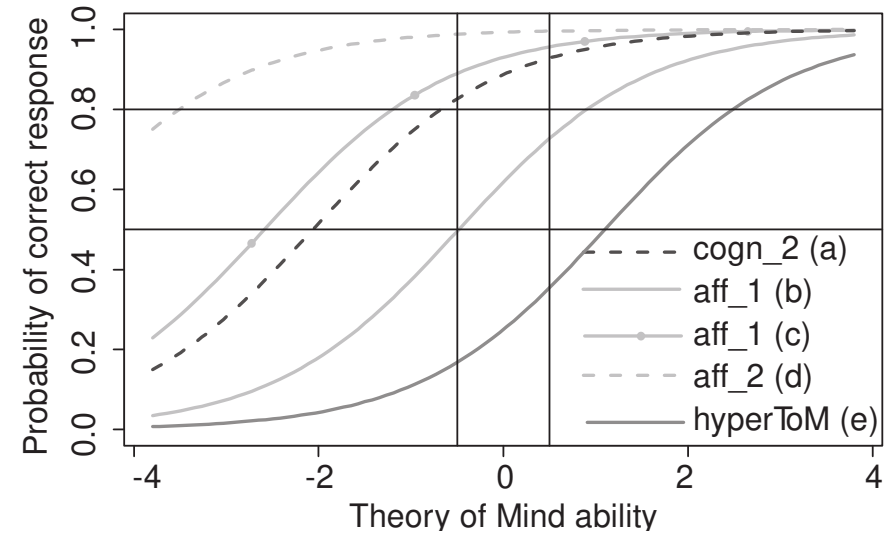
Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

H#1_medium 1

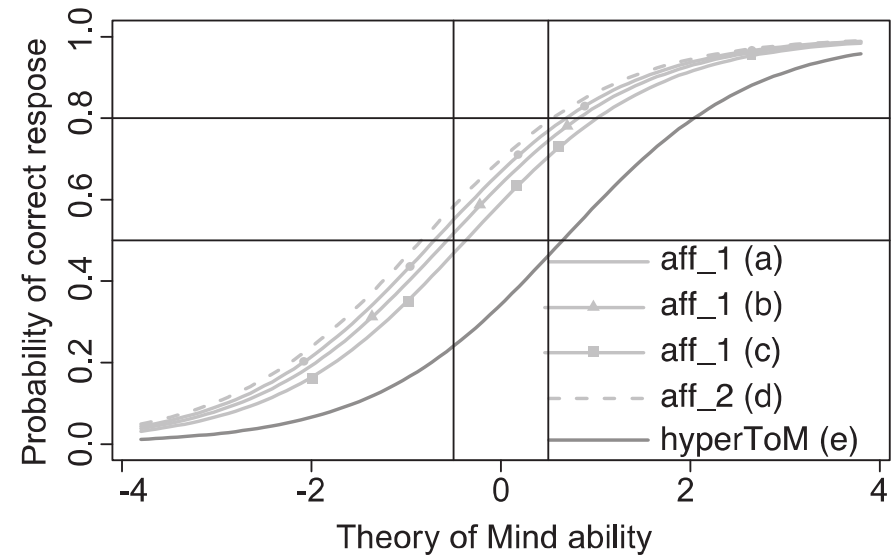


Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

H#1_high 1

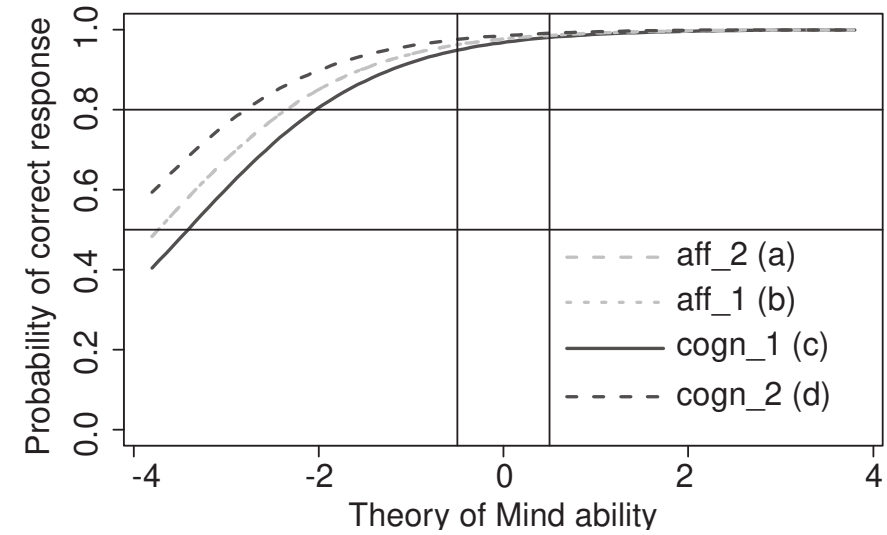


Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

H#2_low**4**

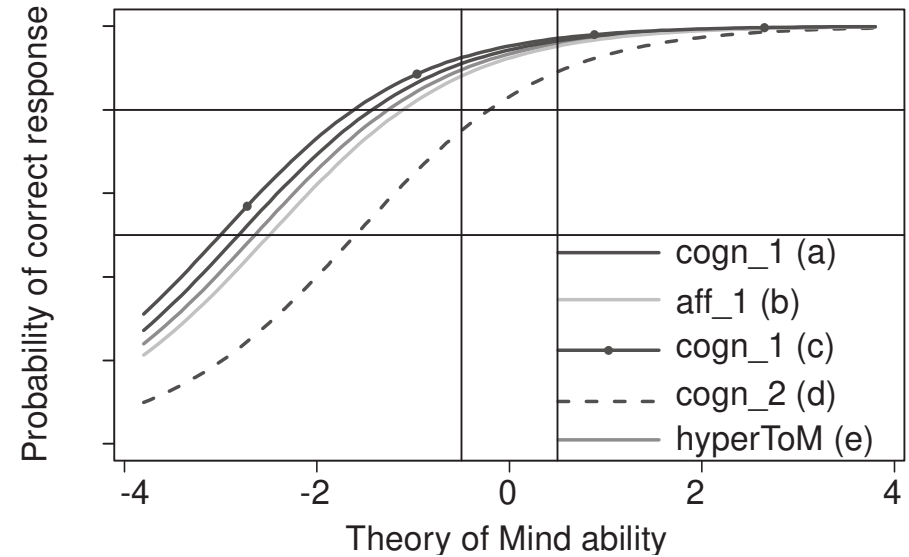
Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

H#3_low 0



Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

H#4_medium **1**

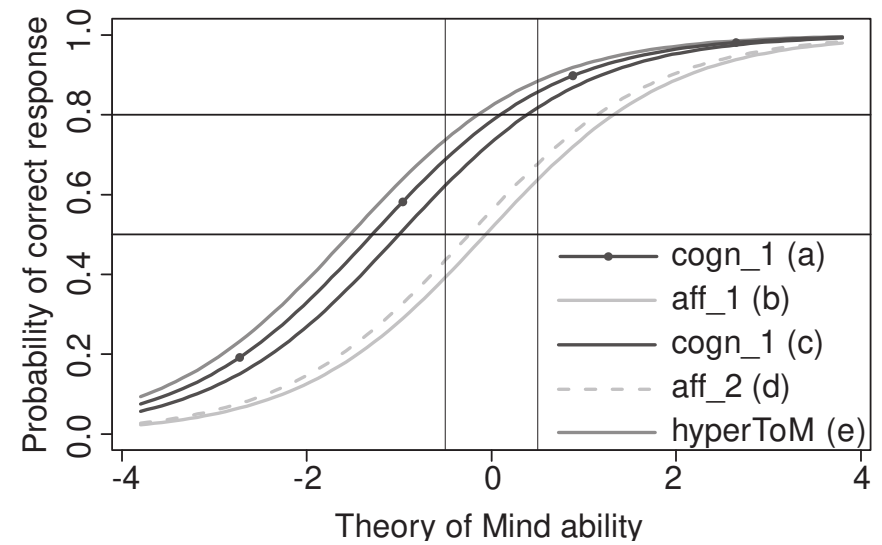


Note. ToM = Theory of Mind. Scene code consists of a letter, representing the message type (i.e., H=Honesty), the scene number, and the portrayed intensity (i.e., low, medium, high). Item characteristic curves of the hyperToM question is only depicted in presence of a model fit. Question code consists of abbreviations of type (i.e., cogn = cognitive ToM, aff = affective ToM, hyperToM = hypermentalization) and order (i.e., 1 = first-order ToM, 2 = second-order ToM) of ToM, and question's identification letter (a-d = ToM questions, e = hypermentalization question). Intensities versions selected for the BASIT-ToM are printed in bold.

Table J.2: Item Characteristic Curves of the Intensity Version Scenes for Message Type Paradoxical Sarcasm

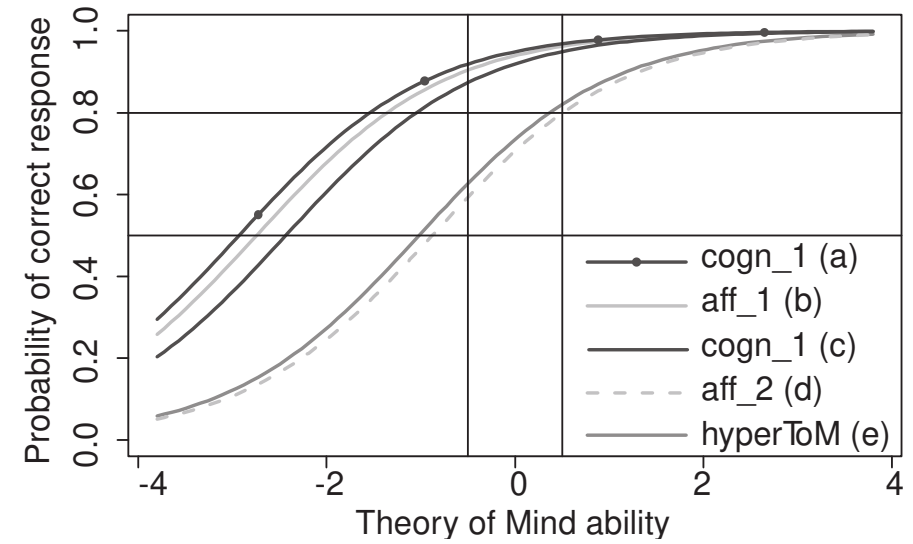
Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

pS#1_low **4**



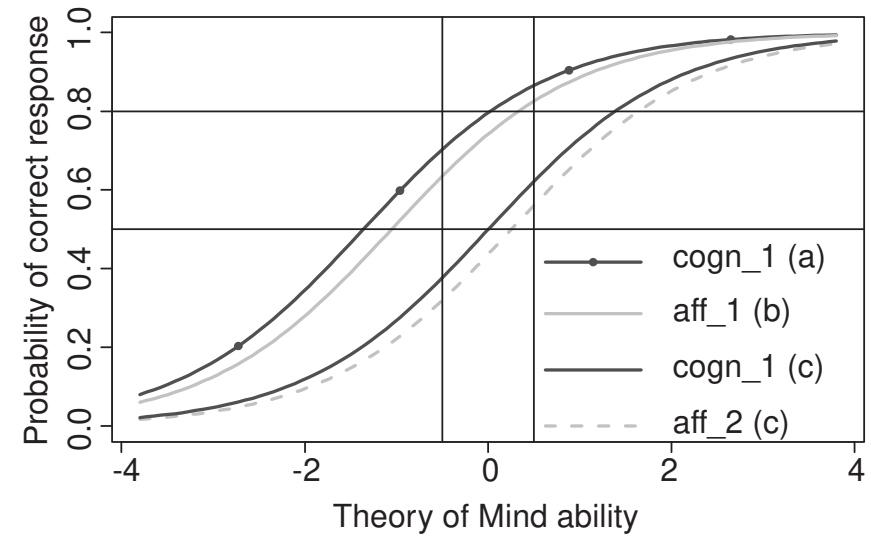
Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

pS#1_medium 1



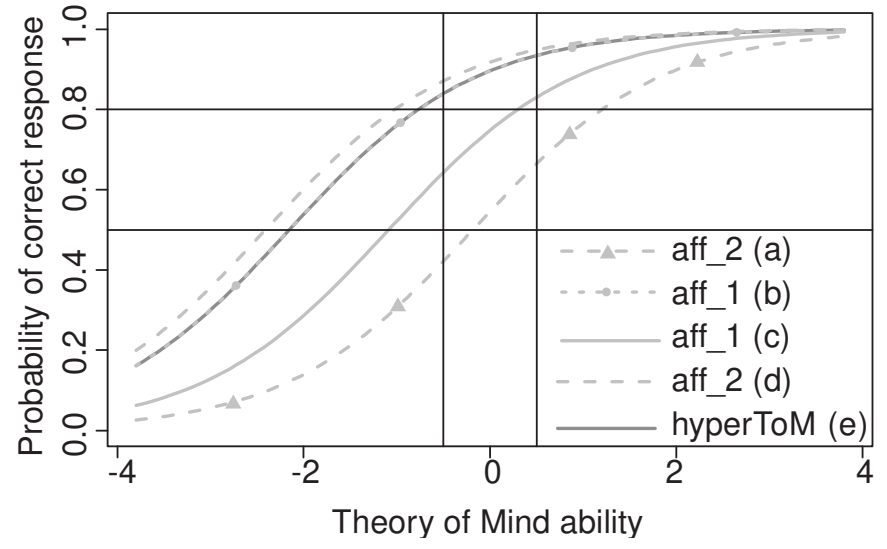
Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

pS#2_low 4



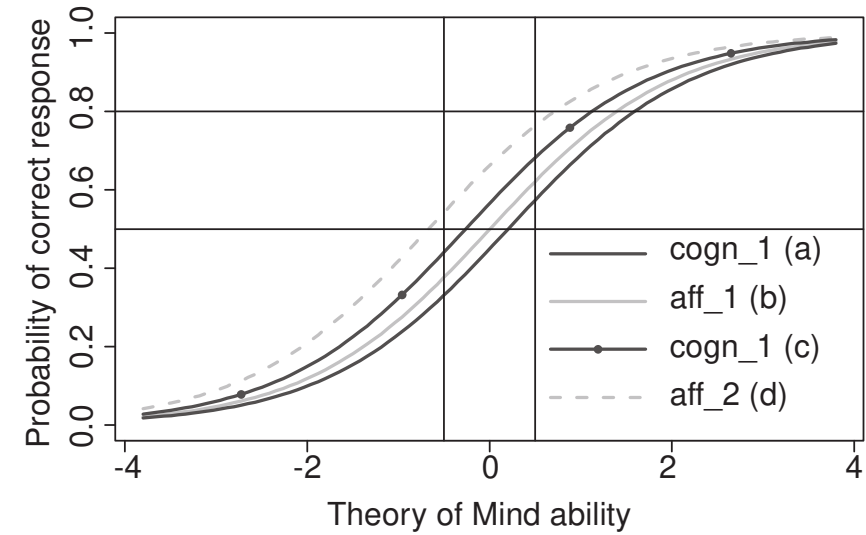
Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

pS#4_low† 2



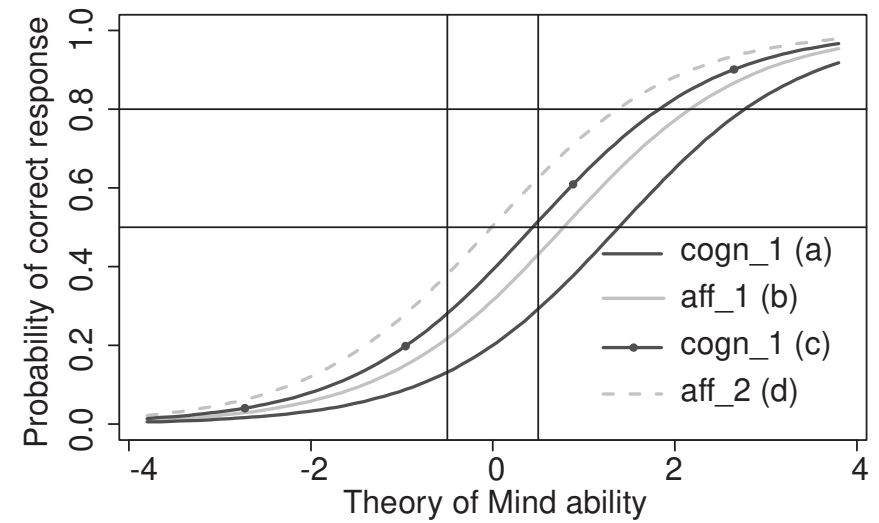
Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

Practice(pS)_low[‡] 4



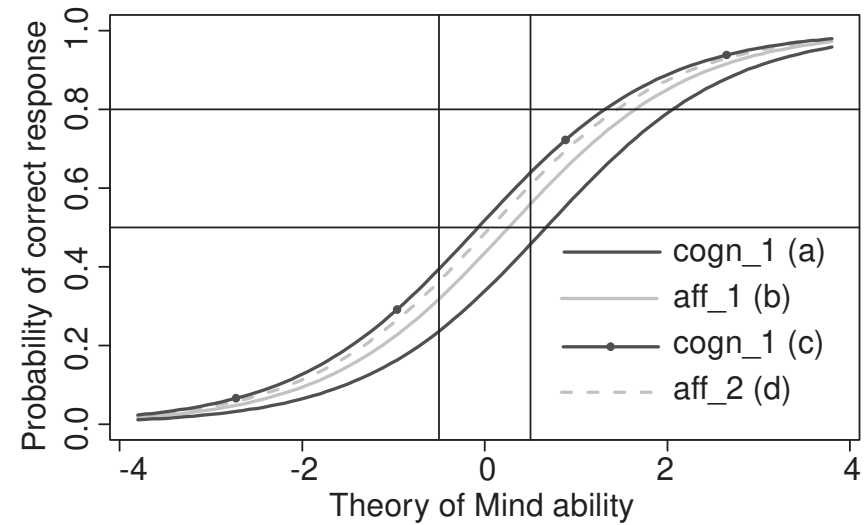
Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

Practice(pS)_medium 2



Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

Practice(pS)_high 3



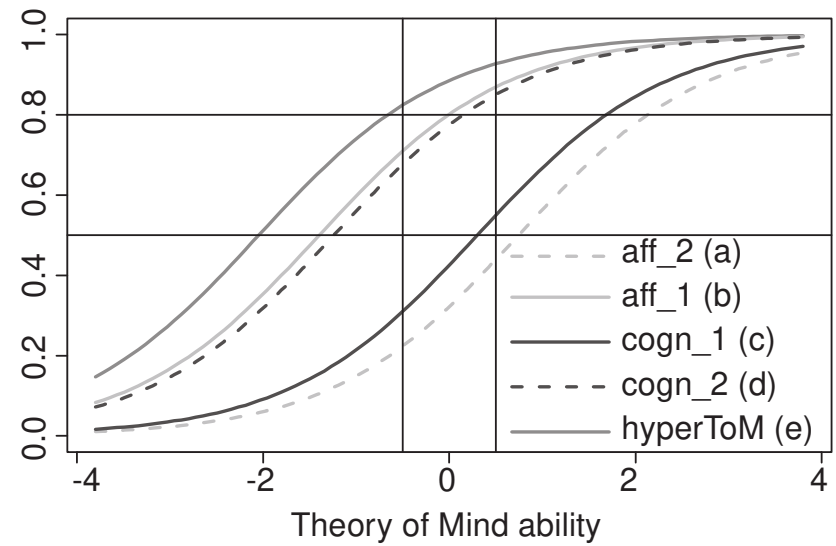
Note. ToM = Theory of Mind. Scene code consists of a letter, representing the message type (i.e., pS = Paradoxical Sarcasm), the scene number, and the portrayed intensity (i.e., low, medium, high). Item characteristic curves of the hyperToM question is only depicted in presence of a model fit. Question code consists of abbreviations of type (i.e., cogn = cognitive ToM, aff = affective ToM, hyperToM = hypermentalization) and order (i.e., 1 = first-order ToM, 2 = second-order ToM) of ToM, and question's identification letter (a-d = ToM questions, e = hypermentalization question). Intensities versions selected for the BASIT-ToM are printed in bold.

†pS#4_low will be used as practice scene in the Basel Version of The Awareness of Social Inference Test - Theory of Mind (BASIT-ToM).

*Practice(pS)_low will be used as a test scene for Paradoxical Sarcasm of the BASIT-ToM

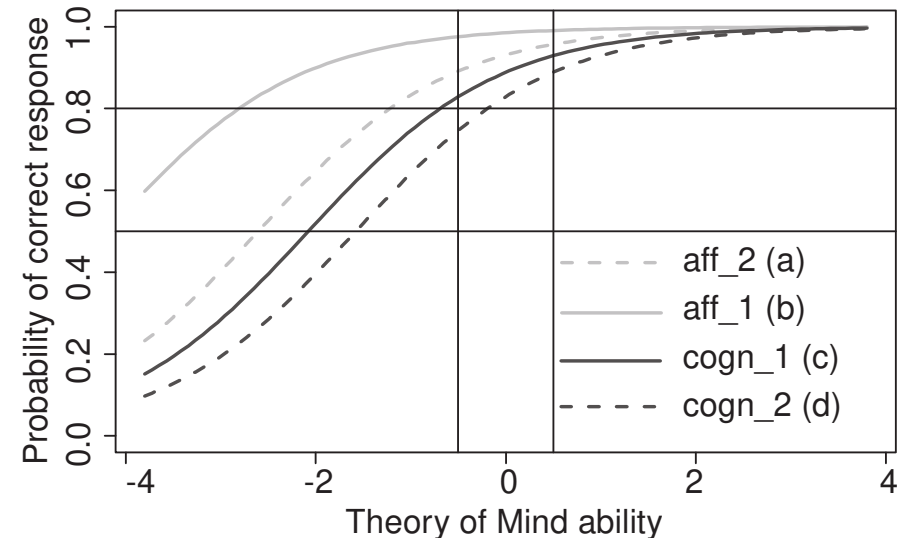
Table J.3: Item Characteristic Curves of the Intensity Version Scenes for Message Type Simple Sarcasm

Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

sS#1_low**3**

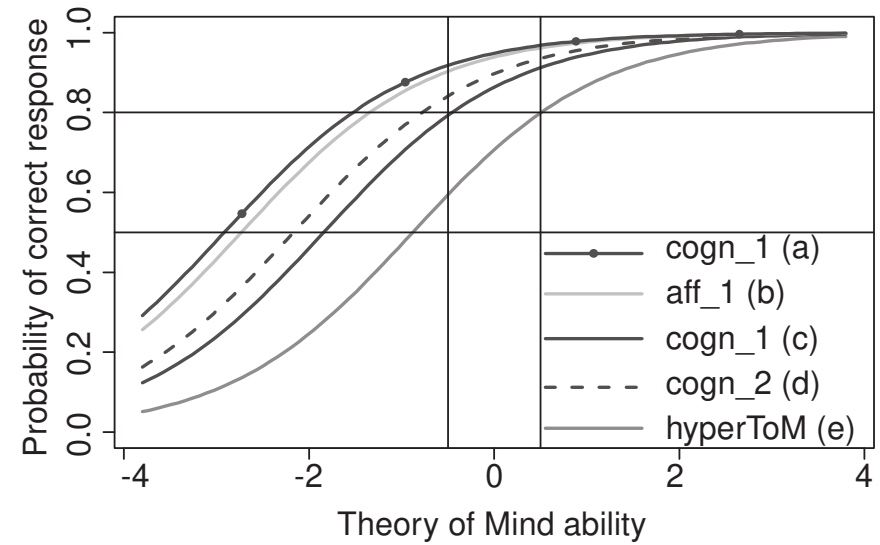
Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

sS#1_high 1



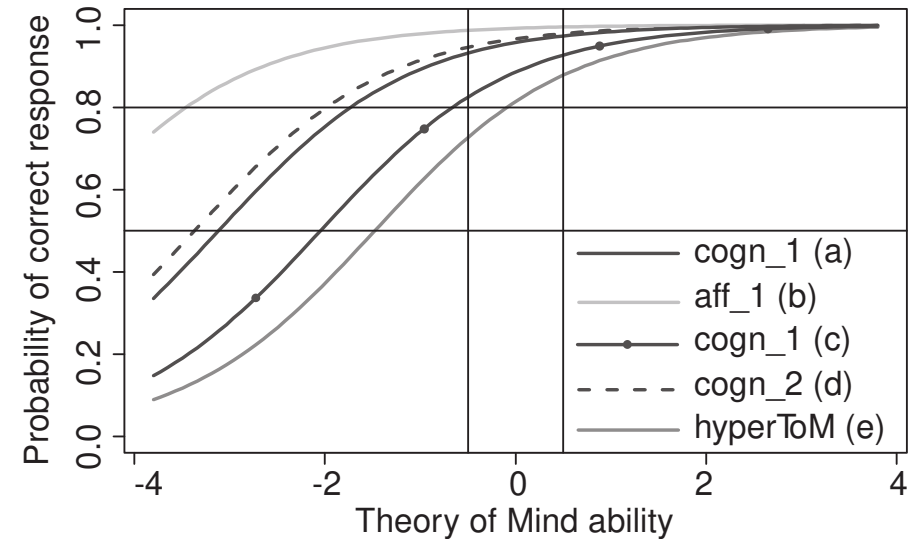
Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

sS#2_low 1



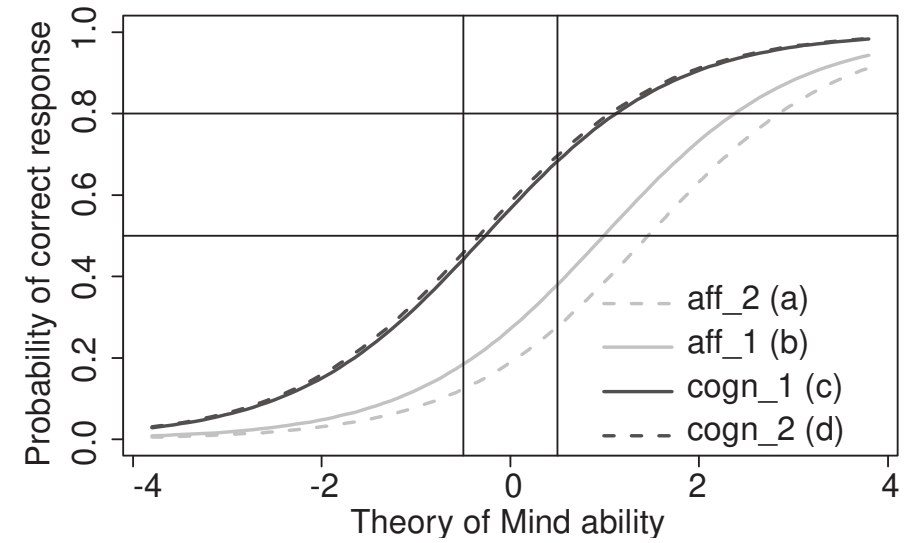
Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

sS#2_high 0



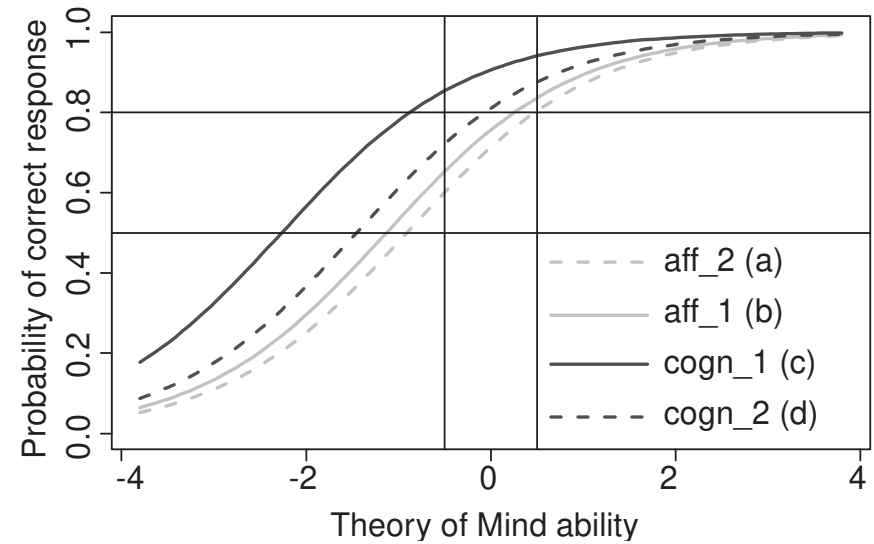
Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

sS#3_low 2



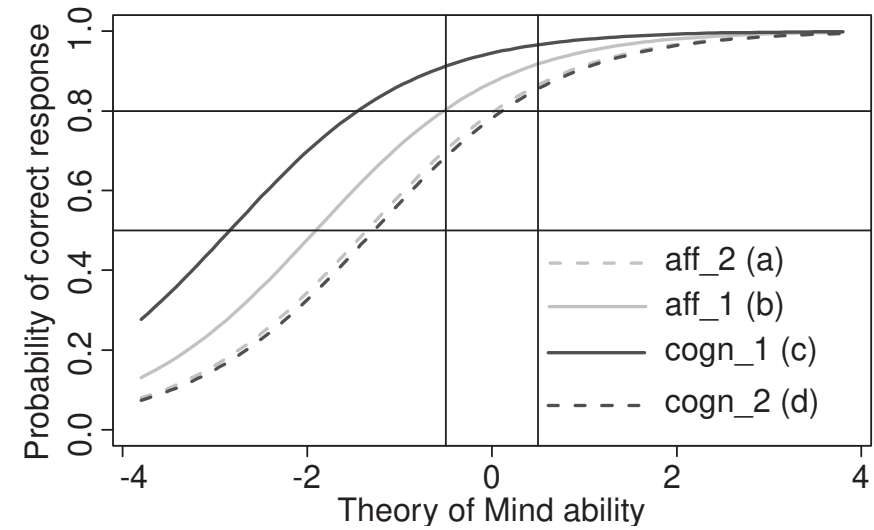
Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

sS#3_medium 3



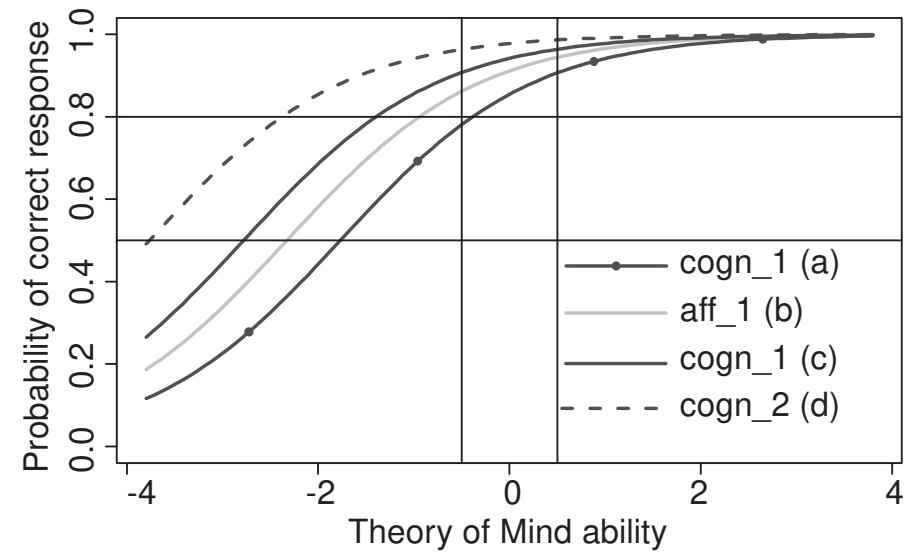
Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

sS#3_high 2



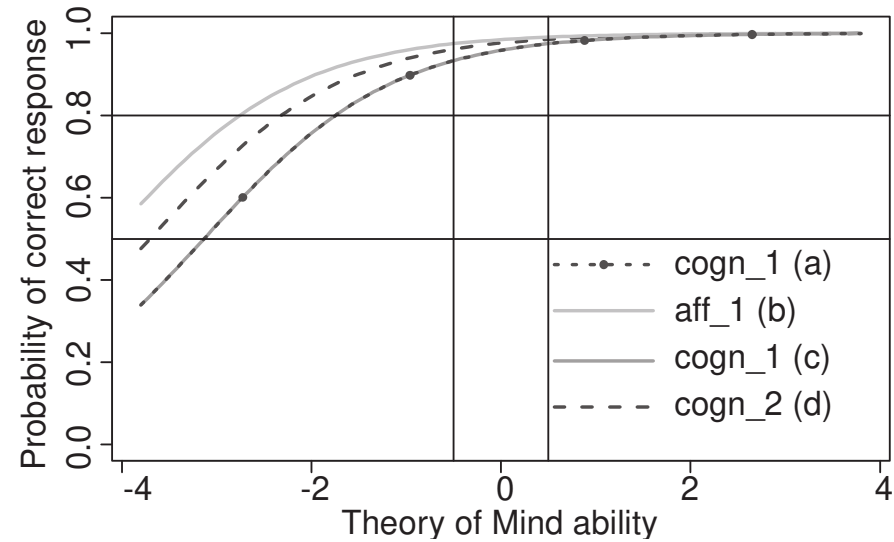
Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

sS#4_medium 1



Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves of the ToM and hyperToM questions of the intensity version fitting the Rasch model
--	---	---

sS#4_high 0



Note. ToM = Theory of Mind. Scene code consists of a letter, representing the message type (i.e., sS = Simple Sarcasm), the scene number, and the portrayed intensity (i.e., low, medium, high). Item characteristic curves of the hyperToM question is only depicted in presence of a model fit. Question code consists of abbreviations of type (i.e., cogn = cognitive ToM, aff = affective ToM, hyperToM = hypermentalization) and order (i.e., 1 = first-order ToM, 2 = second-order ToM) of ToM, and question's identification letter (a-d = ToM questions, e = hypermentalization question). Intensities versions selected for the BASIT-ToM are printed in bold.

Appendix K

Table K.1: Number of times each selected intensity version scene was watched by participants, depicted by the percentages of the respective 80 participants

Scene	Number of times the scene was watched in percentages [†]			
	1	2	3	4
H#1_low	91%	9%		
H#2_low	89%	11%		
H#4_medium	89%	11%		
sS#1_low	91%	9%		
sS#2_low	88%	13%		
sS#3_medium	91%	9%		
pS#1_low	88%	11%		1%
pS#2_low	58%	36%	6%	
pS#4_low [‡]	78%	23%		
Practice (pS_low) [§]	71%	25%	4%	

Note. Scene code consists of letters (i.e., H = Honesty, pS = Paradoxical Sarcasm, sS = Simple Sarcasm), the respective scene numbers and the intensity level.

[†]Rounded percentages may not necessarily add up to 100%.

[‡]pS#4_low (scene #4 containing Paradoxical Sarcasm at low intensity) will be used as practice scene.

[§]Practice(pS)_low will be used as a test scene for Paradoxical Sarcasm.