Universität Basel

Fakultät für Psychologie

Die Entwicklung der Basel Version of The Awareness of Social Inference Test (BASIT): Ein Innovatives Instrument zur Erfassung Sozial-Kognitiver Fähigkeiten

Band I von II

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

Marianne Somaya Jarsch

aus Friedrichshafen, Deutschland

Basel, 2021

Universität Basel

Fakultät für Psychologie

Genehmigt von der Fakultät für Psychologie auf Antrag von

Prof. Dr. phil. Andreas U. Monsch

Prof. Dr. phil. Klaus Opwis

Datum des Doktoratsexamen: 27.01.2022

Prof. Dr. phil. Jens Gaab

Dekan der Fakultät für Psychologie

Universität Basel

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:

- Jarsch, M., Monsch, A. U., Kressig, R. W., & Sollberger, M. (2021). Eine Lücke, die es zu schließen gilt: Die Untersuchung sozial-kognitiver Fähigkeiten an deutschsprachigen Memory Clinics und Gedächtnissprechstunden. *Zeitschrift für Neuropsychologie*. Manuscript submitted.

- Jarsch, M., Bartsch, K., Berres, M., Ryf, I., Jurisic, K., Sluka, C., Kressig, R. W., Monsch, A. U., McDonald, S., Piguet, O., & Sollberger, M. (2021). The Basel Version of the Awareness of Social Inference Test – Emotion Recognition (BASIT-ER): Preliminary Validation Analyses in Healthy Adults. *Neuropsychology*. https://doi.org/10.1037/neu0000786. Manuscript in press.

- Jarsch, M., Piguet, O., Berres, M., Sluka, C., Kressig, R. W., Monsch, A. U., McDonald, S., & Sollberger, M. (2021). The Basel Version of the Awareness of Social Inference Test – Theory of Mind (BASIT-ToM): Preliminary Validation Analyses in Healthy Adults. *Neuropsychology*. Manuscript submitted.

Basel, den 30.11.2021

M. Jarsch

Marianne Somaya Jarsch

Universität
Basel

Fakultät für
Psychologie

Für alle Frauen in der Wissenschaft

Danksagung

Mein ganz besonderer Dank gilt PD Dr. med. Marc Sollberger, der mir die Chance gegeben hat, aktiver Teil eines ganz besonderen und spannenden Projektes zu sein. Ich bedanke mich bei ihm für die gute Betreuung und die bemerkenswerte zeitliche Flexibilität, die er mir in den vergangenen Jahren ermöglicht hat. Während unserer Zusammenarbeit konnte ich aufgrund unseres konstruktiven Austauschs und der wertvollen fachlichen Diskussionen meine Expertise in den verschiedenen Phasen einer Testentwicklung kontinuierlich erweitern.

Ich bedanke mich herzlich bei meinem Erstbetreuer Prof. Dr. phil. Andreas U. Monsch, der mir schon während meines Masterstudiums eine wertvolle Sichtweise auf die Neuropsychologie vermittelt und meine Begeisterung für diesen wichtigen Fachbereich der Psychologie geweckt hat. Insbesondere bedanke ich mich bei ihm für seine zielgerichtete Unterstützung, damit ich meinen „straffen Zeitplan“ zur erfolgreichen Beendigung meines Doktorats umzusetzen kann.

Mein Dank gilt auch Prof. Dr. phil. Klaus Opwis, der äusserst unterstützend und freundlich die Zweitbetreuung meines Doktorats übernommen und sich ebenfalls bereit erklärt hat mich bei meiner zeitlichen Planung zu unterstützen.

Meinem Ehemann Christian möchte ich von Herzen dafür danken, dass wir im Leben jede Entscheidung gemeinsam treffen und wir beide auf diese Art unsere beruflichen und privaten Ziele in Einklang bringen können.

Ich freue mich sehr, dass ich mich an dieser Stelle auch bei meiner Mama Eva Albrecht und meinen Grosseltern Johanna und Franz Glock bedanken darf, die mich seit meiner Kindheit bei allen Entscheidungen ermutigen, bestärken und auf allen Ebenen unterstützen.

Abschliessend bedanke ich mich auch bei meinen wunderbaren Freundinnen Anna-Sophie, Melanie, Sonja, Antje und Heike für ihr Interesse, ihre wertvollen Tipps und ihre kontinuierliche Unterstützung.

Inhaltsverzeichnis

1 Zusammenfassung 1

2 Abstract 2

Teil 1 3

3 Einleitung 3

3.1 Soziale Kognition – Begriffsklärung und Diagnostische Relevanz bei Demenziellen Syndromen 4

3.1.1 Emotionserkennung (EE) 8

3.1.2 Theroy of Mind (ToM) 13

3.2 Bedeutung der Sozialen Kognition für die bvFTD und AD Diagnostik 17

3.2.1 Sozial-Kognitive Dysfunktion im Rahmen einer bvFTD 17

3.2.2 Sozial-Kognitive Dysfunktion im Rahmen einer AD 18

3.2.3 Manuskript 1 19

Eine Lücke, die es zu schließen gilt: Die Untersuchung sozial-kognitiver Fähigkeiten an deutschsprachigen Memory Clinics und Gedächtnissprechstunden 19

Teil 2 50

4 Das Schliessen der Lücke Zwischen Theorie und Praxis – Die Entwicklung eines Tests zur Untersuchung der EE- und ToM-Fähigkeiten, der den Anforderungen der MC/GS Entsprechen soll 50

4.1 Manuskript 2 52

The Basel Version of the Awareness of Social Inference Test – Emotion Recognition (BASIT-ER): Preliminary Validation Analyses in Healthy Adults 52

4.2 Manuskript 3 84

The Basel Version of the Awareness of Social Inference Test – Theory of Mind (BASIT-ToM): Preliminary Validation Analyses in Healthy Adults 84

Teil 3 125

5 Zusammenfassung der Erkenntnisse und Allgemeine Diskussion 125

5.1 Anforderung (1): Untersuchung der EE-Fähigkeit von Basis- und Komplexen Emotionen 126

5.1.1 Empfehlungen zum Weiteren Vorgehen in Zukünftigen Validierungsstudien des BASIT-ER 127

5.2 Anforderung (2): Untersuchung Verschiedener Konzepte der ToM-Fähigkeiten 130

5.2.1 Empfehlungen zum Weiteren Vorgehen in Zukünftigen Validierungsstudien des BASIT-ToM 131

5.3 Anforderung (3): Hohe Sensitivität zur Detektion Subtiler Defizite in den EE- und ToM-Fähigkeiten..... 132

5.3.1 Empfehlungen zum Weiteren Vorgehen in Zukünftigen Validierungsstudien des BASIT..... 133

5.4 Anforderung (4): Testleistung soll Prädiktiv für die EE- und ToM-Leistungen in Realen Sozialen Interaktionen sein 133

5.4.1 Empfehlungen zum Weiteren Vorgehen in Zukünftigen Validierungsstudien des BASIT..... 134

5.5 Anforderung (5): Durchführbarkeit mit Personen, die Beeinträchtigungen in Mehreren Kognitiven Domänen Aufweisen 135

5.5.1 Empfehlungen zum Weiteren Vorgehen in Zukünftigen Validierungsstudien des BASIT..... 136

5.6 Anforderung (6): Praktikable Testdauer..... 137

5.6.1 Empfehlungen zum Weiteren Vorgehen in Zukünftigen Validierungsstudien des BASIT..... 137

5.7 Normierungsstudie..... 138

5.8 Mögliche Weiterentwicklungen des BASIT 138

5.8.1 Erweiterung des BASIT um die Untersuchung Sozialen Verhaltens zu Ermöglichen 138

5.8.2 Einsatz für Verschieden Fähigkeitsniveaus 140

5.8.3 Differenzierte Untersuchung Verschiedener Sinneskanäle 141

5.8.4 Einsatz bei Anderen Erkrankungen 141

6 Fazit 142

8 Literaturverzeichnis 143

Anhang 155

Curriculum Vitae 155

Abbildungsverzeichnis

Abbildung 1 *Beispiele der Spezifischen Gesichtsausdrücke der Sechs Basisemotionen von Ekman und Friesen (1976)*.....**9**

Abbildung 2 *Rad der Emotionen in Anlehnung an Plutchik (1991)*.....**10**

Abbildung 3 *Darstellung der Vier Entwicklungsstufen der Präemotion Comfort (Wohlbefinden) nach Zinck & Newen (2008)*.....**12**

Abbildung 4 *Übersicht über die Konzeptuelle Einteilung der Theory of Mind-Fähigkeiten und Assoziierte Hirnregionen*.....**16**

Manuskript 2

The Basel Version of the Awareness of Social Inference Test - Emotion Recognition (BASIT - ER): Preliminary Validation Analyses in Healthy Adults

Figure 1 *Item Characteristic Curves of the Selected BASIT-ER Scenes*……**69**

Figure 2 *95% Confidence Intervals of the Estimated Difficulties of the Selected BASIT-ER Scenes*……...**73**

Manuskript 3

The Basel Version of the Awareness of Social Inference Test - Theory of Mind (BASIT - ToM): Preliminary Validation Analyses in Healthy Adults

Figure 1 *Item Characteristic Curves of the Selected BASIT-ToM Scenes for Message Type Honesty*……**105**

Figure 2 *Item Characteristic Curves of the Selected BASIT-ToM Scenes for Message Type Paradoxical Sarcasm*……**106**

Figure 3 *Item Characteristic Curves of the Selected BASIT-ToM Scenes for Message Type Simple Sarcasm*……..**107**

Figure 4 *95% Confidence Intervals of the Estimated Difficulties of the Selected BASIT-ToM Scenes for Message Type Honesty*……..**110**

Figure 5 *95% Confidence Intervals of the Estimated Difficulties of the Selected BASIT-ToM Scenes for Message Type Paradoxical Sarcasm*……...**112**

Figure 6 *95% Confidence Intervals of the Estimated Difficulties of the Selected BASIT-ToM Scenes for Message Type Simple Sarcasm*……**114**

Tabellenverzeichnis

Tabelle 1 *Auflistung der Kognitiven Domänen Gemäss des DSM-5 (American Psychiatric Association, 2013) und ICD-11 (World Health Organization, 2021)* **6**

Tabelle 2 *Exemplarische Auflistung Sozial-Kognitiver Fähigkeiten mit Ihren Entsprechenden Arbeitsdefinitionen* **7**

Tabelle 3 *Darstellung der Drei Cluster der Emotion Love (Liebe) nach Shaver et al. (1987)* **11**

Manuskript 2

The Basel Version of the Awareness of Social Inference Test - Emotion Recognition (BASIT - ER): Preliminary Validation Analyses in Healthy Adults

Table 1 *Selected Intensity Version Scenes for the BASIT-ER for further Use in Clinical Populations* **71**

Manuskript 3

The Basel Version of the Awareness of Social Inference Test - Theory of Mind (BASIT - ToM): Preliminary Validation Analyses in Healthy Adults

Table 1 *Example of an adaption of a TASIT-SIM question into a BASIT-ToM question as well as an example of a BASIT-ToM hypermentalization question* **98**

Table 2 *Selected Intensity Version Scenes for the BASIT-ToM for Further Use in Clinical Populations* **108**

Table 3 *Types and Orders of Theory of Mind for Each Question of the Ten Selected Scenes* **109**

Abkürzungsverzeichnis

AD	Alzheimer Erkrankung (Alzheimer's Disease)
aff	affective Theory of Mind
BASIT	Basel Version of The Awareness of Social Inference Test
BASIT-ER	Basel Version of The Awareness of Social Inference Test-Emotion Recognition
BASIT-ToM	Basel Version of The Awareness of Social Inference Test-Theory of Mind
bvFTD	Verhaltensvariante der frontotemporalen Demenz
CERAD-Plus	Consortium to Establish a Registry for Alzheimer's Disease-Plus
cogn	cognitive Theory of Mind
DSM-5	Diagnostic and Statistical Manual of Mental Disorders-Fifth Edition
EE	Emotionserkennung
EKNZ	Ethikkommission Nordwest- und Zentralschweiz
ER	emotion recognition
GALC	Geneva Affect Label Coder
H	Honesty
h	high intensity
hyperToM	hypermentalization question
ICD-11	International Classification of Diseases 11th Revision
JARS	Journal Article Reporting Standards – Quantitative for non experimental designs
l	low intensity
m	medium intensity

MASC	Movie for the Assessment of Social Cognition
MC/GS	Memory Clinics und Gedächtnissprechstunden
MCI	Mild Cognitive Impairment
MDD	Major Depression Disorder
MoCA	Montreal Cognitive Assessment
pS	Paradoxical sarcasm
SEA	Social Cognition and Emotional Assessment
sS	simple sarcasm
TASIT	The Awareness of Social Inference Test
TASIT-EET	The Awareness of Social Inference Test-Emotion Evaluation Test
TASIT-SIM	The Awareness of Social Inference Test-Social Inference Minimal
ToM	Theory of Mind
TOP	Transparency and Openness Promotion

1 Zusammenfassung

Forschungsarbeiten der letzten zwei Jahrzehnte haben gezeigt, dass die Untersuchung der sozial-kognitiven Fähigkeiten Emotionserkennung (EE) und Theory of Mind (ToM) die Differenzialdiagnostik bei neurodegenerativ bedingten demenziellen Syndromen entscheidend verbessern kann. Jedoch scheinen sich sozial-kognitive Testverfahren, die sich im Forschungskontext als passend erwiesen haben, nur schwer in die klinische Praxis integrieren zu lassen. Aktuell finden daher die wichtigen Erkenntnisse der sozial-kognitiven Demenzforschung keine klinische Anwendung in den deutschsprachigen Memory Clinics und Gedächtnissprechstunden (MC/GS). In Teil 1 dieser Dissertationsthesis wird eine explorative Onlineumfrage vorgestellt, die von 87 MC/GS der Länder Deutschland, Österreich und Schweiz beantwortet wurde. Die Ergebnisse zeigen eindrücklich, dass die MC/GS die Notwendigkeit die EE- und ToM-Fähigkeiten im Rahmen der Demenzdiagnostik zu untersuchen, erkannt haben und sich praktikabel anwendbare und normierte Tests mit guten psychometrischen Eigenschaften wünschen. Darüber hinaus werden aus den gewonnenen Erkenntnissen der Onlineumfrage neun Anforderungen abgeleitet, welche die MC/GS an sozial-kognitive Testverfahren stellen. Anschliessend wird in Teil 2 dieser Dissertationsthesis die Entwicklung und präliminäre Validierung des sozial-kognitiven Tests *Basel Version of The Awareness of Social Inference Test* (BASIT) ausführlich beschrieben, der die Anforderungen der MC/GS erfüllen soll, um zukünftig in die Demenzdiagnostik integriert werden zu können. Der BASIT besteht aus einem EE-Subtest (BASIT-Emotion Recognition) und einem ToM-Subtest (BASIT-Theory of Mind), welche die sozial-kognitiven Fähigkeiten der Testteilnehmer mittels kurzer Filmszenen mit sozial-relevantem Inhalt untersuchen. Abschliessend wird in Teil 3 evaluiert, ob der BASIT die Anforderungen der MC/GS erfüllt und welche weiteren Entwicklungsschritte für zukünftige Validierungsstudien geplant werden könnten.

2 Abstract

Over the last two decades, research has shown that the examination of the social-cognitive skills Emotion Recognition (EE) and Theory of Mind (ToM) can decisively improve differenzial diagnosis in neurodegenerative dementia-related syndromes. Although applicable in the research context, social cognitive tests appear difficult to integrate into clinical practice. The important findings of social-cognitive dementia research are not currently clinically implemented in German-speaking memory clinics and memory consultations (MC/GS). Part 1 of this dissertation thesis presents an explorative online survey answered by 87 MC/GS in Germany, Austria and Switzerland. The results indicate that MC/GS have recognized the need to examine EE and ToM skills in the context of dementia diagnosis and desire applicable test and normed tests with good psychometric properties. In addition, nine requirements highlighted by the MC/GS for social cognitive tests are derived from the findings obtained from the online survey. Part 2 of this dissertation thesis describes the development and preliminary validation of the social-cognitive test *Basel Version of The Awareness of Social Inference Test* (BASIT) in detail, which is intended to meet the requirements of the MC/GS in order to be integrated into dementia diagnosis in the future. The BASIT consists of an EE subtest (BASIT-Emotion Recognition) and a ToM subtest (BASIT-Theory of Mind), which examine participants' social-cognitive abilities using short film scenes with socially relevant content. Finally, Part 3 evaluates whether the BASIT meets the requirements of the MC/GS and which further development steps may be planned for future validation studies.

Teil 1

3 Einleitung

Soziale Beziehungen stehen im Mittelpunkt des menschlichen Handelns (Fortier et al., 2018) und nehmen Einfluss auf alle Lebensbereiche. Die erfolgreiche Gestaltung sozialer Beziehungen über einen längeren Zeitraum hinweg hängt dabei massgeblich von der Verhaltensregulation und -steuerung durch sozial-kognitive Prozesse ab (Brüne & Juckel, 2010; Fortier et al., 2018; Keifer et al., 2020; Kennedy & Adolphs, 2012). Sind diese Prozesse jedoch gestört, können die auftretenden sozial-kognitiven Defizite zu Problemen in den Bereichen Verständnis und Kommunikation, Umgang mit anderen Menschen, Teilhabe am gesellschaftlichen Leben, Mobilität, Selbstversorgung und Tätigkeiten des alltäglichen Lebens führen (American Psychiatric Association, 2013; Nelis et al., 2011; Timmer et al., 2020). Die Untersuchung sozial-kognitiver Fähigkeiten ist daher angesichts ihrer Bedeutung für das individuelle psychosoziale Funktionsniveau unerlässlich, um die Entstehung und Veränderung von angemessenem und inadäquatem Sozialverhalten in verschiedenen sozialen Kontexten zu verstehen. Vor allem bei Patient:innen mit demenziellen Syndromen wie zum Beispiel der Verhaltensvariante der Frontotemporalen Demenz (bvFTD) oder der Alzheimer Erkrankung (AD) treten diese problematischen Kognitions- und Verhaltensänderungen häufig auf (Cotter et al., 2018; Desmarais et al., 2018; Kessels et al., 2021) und werden somit für die Diagnostik, Behandlung und die Beratung von Betroffenen und deren Betreuungspersonen relevant.

Die vorliegende Dissertationsthesis ist in drei Teile gegliedert. Teil 1 definiert zunächst den Begriff der sozialen Kognition. Anschliessend wird anhand des ersten Manuskripts die Notwendigkeit aufgezeigt, Untersuchungen sozial-kognitiver Fähigkeiten in den Diagnostikprozess demenzieller Syndrome zu integrieren. Zudem wird eine explorative Onlineumfrage mit deutschsprachigen Memory Clinics und Gedächtnissprechstunden (MC/GS) vorgestellt, die folgende drei Fragen beantworten sollte: (1) Werden an den deutschsprachigen Kompetenzzentren für Demenzdiagnostik

sozial-kognitive Untersuchungen durchgeführt? (2) Falls ja, sind die durchgeführten Tests zufriedenstellend? (3) Welche Anforderungen stellen die Diagnostiker:innen der befragten MC/GS an sozial-kognitive Tests?

Aus den Ergebnissen dieser Online-Umfrage liess sich die Notwendigkeit ableiten, einen deutschsprachigen Test zur Untersuchung sozial-kognitiver Fähigkeiten zu entwickeln, der den Anforderungen der Diagnostiker:innen standhält, damit dieser zukünftig Teil des Diagnostikprozesses demenzieller Syndrome werden kann. Anschliessend beschreibt Teil 2 dieser Dissertation ausführlich die Testentwicklung der *Basel Version of The Awareness of Social Inference Test* (BASIT) unter Berücksichtigung der zuvor ermittelten Anforderungen. Der letzte Teil schliesst mit einer allgemeinen Diskussion und einem Ausblick ab, in denen die Ergebnisse der drei vorliegenden Manuskripte in Zusammenhang gestellt und diskutiert werden. Darüber hinaus werden Ideen bezüglich zukünftiger Möglichkeiten und Weiterentwicklungen der BASIT-Testbatterie aufgezeigt.

3.1 Soziale Kognition – Begriffsklärung und Diagnostische Relevanz bei Demenziellen Syndromen

Soziale Kognition wird sowohl im DSM-5 (Diagnostic And Statistical Manual Of Mental Disorders, Fitfth Edition; American Psychiatric Association, 2013) als auch im neuen ICD-11 (International Classification of Diseases, 11. Überarbeitung; World Health Organization, 2021) als eine der sechs bzw. sieben kognitiven Domänen aufgeführt, die zur Erstellung eines differenzierten und ausführlichen neurokognitiven Profils untersucht werden sollten (siehe Tabelle 1). Derzeit existiert jedoch weder in der klinischen Demenzforschung noch in der Praxis eine einheitliche Definition dieser kognitiven Domäne. Üblicherweise werden unter sozialer Kognition jedoch alle informationsverarbeitenden Prozesse bzw. sozial-kognitiven Fähigkeiten zusammengefasst, die darauf abzielen, den mentalen Zustand und das Verhalten anderer wahrzunehmen und zu interpretieren, um das eigene Verhalten an die soziale Situation

anpassen zu können (Frith & Frith, 2007). Eine abschliessende Auflistung und allgemeingültige Definitionen aller zugehöriger sozial-kognitiver Fähigkeiten existiert ebenfalls nicht. Daher kann eine exemplarische Auflistung verschiedener sozial-kognitiver Fähigkeiten, die in den Bereichen der neuropsychologischen und psychiatrischen Klinik und Forschung relevant sind,
in Tabelle 2 eingesehen werden. Im Bereich der Demenzforschung scheinen vor allem Veränderungen der sozial-kognitiven Fähigkeiten der Theory of Mind (ToM) und der Emotionserkennung (EE) bei anderen Personen eine Rolle zu spielen (z. B. Cotter et al., 2018; Dodich et al., 2020), weshalb auf diese in den folgenden Abschnitten näher eingegangen wird und sie zudem Gegenstand der Testentwicklung dieser Dissertation waren.

Tabelle 1

Auflistung der Kognitiven Domänen, die Gemäss des DSM-5 (American Psychiatric Association, 2013) *und des ICD-11* (World Health Organization, 2021) *zur Umfassenden und Differenzierten Beurteilung Kognitiver Fähigkeiten Untersucht Werden Sollten.*

Neurokognitive Domänen	
DSM-5	ICD-11
Exekutivfunktionen	Exekutivfunktionen
Komplexe Aufmerksamkeit (inklusive psychomotorischer Verarbeitungsgeschwindigkeit)	Aufmerksamkeit
Lernen und Gedächtnis	Gedächtnis
Perzeptuell-motorische Funktionen	Visuoperzeptive oder visuospatiale Fähigkeiten
Soziale Kognition	Soziale Kognition und Urteilsvermögen
Sprache	Sprache
	Psychomotorische Verarbeitungsgeschwindigkeit

Anmerkung. DSM-5 = Diagnostic And Statistical Manual Of Mental Disorders, Fitfth Edition; ICD-11 = International Classification of Diseases, 11. Überarbeitung.

Tabelle 2

Exemplarische Auflistung Sozial-Kognitiver Fähigkeiten mit Ihren Entsprechenden Arbeitsdefinitionen

Sozial-kognitive Fähigkeiten	Arbeitsdefinition
Emotionsverarbeitung	Die Emotionsverarbeitung umfasst die Wahrnehmung und den Umgang mit Emotionen. Sie vereint drei Subdomänen: (1) Emotionserkennung, beinhaltet das Wahrnehmen und Identifizieren von Emotionen anhand paralinguistischer Hinweisreize, wie z. B. Mimik, Gestik, Prosodie; (2) Emotionen verstehen und (3) Umgang mit den wahrgenommenen Emotionen (z. B. Pinkham et al., 2014)
Theory of Mind	Das Konzept der Theory of Mind (ToM) beschreibt die Fähigkeit den mentalen Zustand anderer Lebewesen zu erkennen und diesen gleichzeitig von den eigenen Gefühlen, Gedanken und Absichten abgrenzen zu können (z. B. Henry et al., 2016)
Soziale Wahrnehmung	Die soziale Wahrnehmung bezieht sich auf das Entschlüsseln und Verarbeiten sozialer Hinweisreize, wie den sozialen Kontext oder soziales Wissen (z. B. Regel, Rollen) und wie diese das Verhalten anderer beeinflussen können (z. B. Pinkham et al., 2014)
Empathie	Unter Empathie können zwei Fähigkeiten subsumiert werden. (1) Die Fähigkeit zu verstehen was andere fühlen ohne diese Gefühle selbst zu erleben (kognitive Empathie) und (2) Die Fähigkeit tatsächlich nachzuempfinden was andere Individuen fühlen (affektive Empathie). Das Konzept der Empathie überlappt sich mit dem Konzept der ToM (z. B. Fortier et al., 2018)

3.1.1 Emotionserkennung (EE)

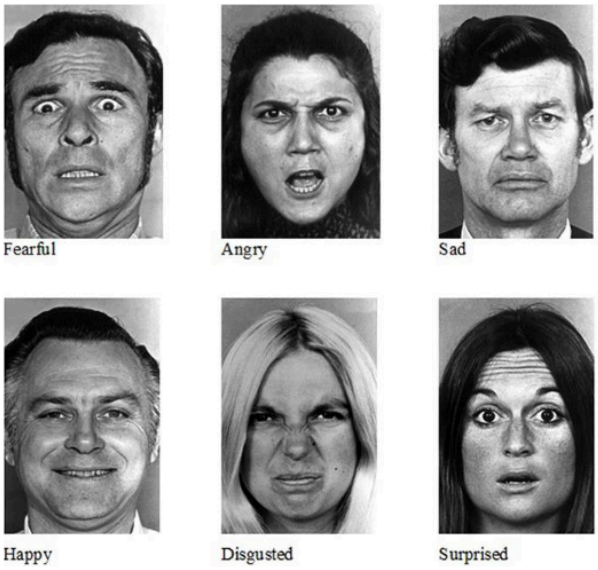
Die EE kann als eine Subdomäne der Fähigkeit Emotionen zu verarbeiten angesehen werden (z. B. Pinkham et al., 2014). Sie beinhaltet das Wahrnehmen und Identifizieren von Emotionen bei anderen Personen anhand linguistischer und paralinguistischer Merkmale, wie z. B. Mimik, Gestik, Prosodie oder Sprechgeschwindigkeit in sozialen Interaktionen (z. B. Ferretti & Papaleo, 2018; Pinkham et al., 2014). Zur Untersuchung der EE-Fähigkeit kann es hilfreich sein, Emotionen in benennbare Kategorien einzuteilen. Vertreter kategorialer Ansätze zur strukturellen Beschreibung von Emotionen benennen jeweils Basis- bzw. primäre Emotionen, sowie komplexe Emotionen, die sich aus den Kombinationen oder Weiterentwicklungen der jeweiligen Basis-/Primäremotionen zusammensetzen. Im Folgenden werden vier einflussreiche kategoriale Ansätze beschrieben, deren Emotionskategorisierung bei der Entwicklung des EE-Tests dieser Dissertation berücksichtigt wurde.

3.1.1.1 Klassifikation der Emotionen nach Paul Ekman

Ekman (1992) postuliert die sechs Basisemotionen Angst, Ekel, Freude, Trauer, Überraschung und Wut. Er geht davon aus, dass den Basisemotionen distinkte universelle Signale zugrunde liegen (z. B. universelle Emotionsentwicklung bei allen Individuen, universelle emotionsspezifische Gesichtsausdrücke, sowie distinkte physiologische Reaktionsmuster). Abbildung 1 zeigt Beispiele der spezifischen Gesichtsausdrücke für jede Basisemotion.

Abbildung 1

Beispiele der Spezifischen Gesichtsausdrücke der Sechs Basisemotionen von Ekman und Friesen (1976)



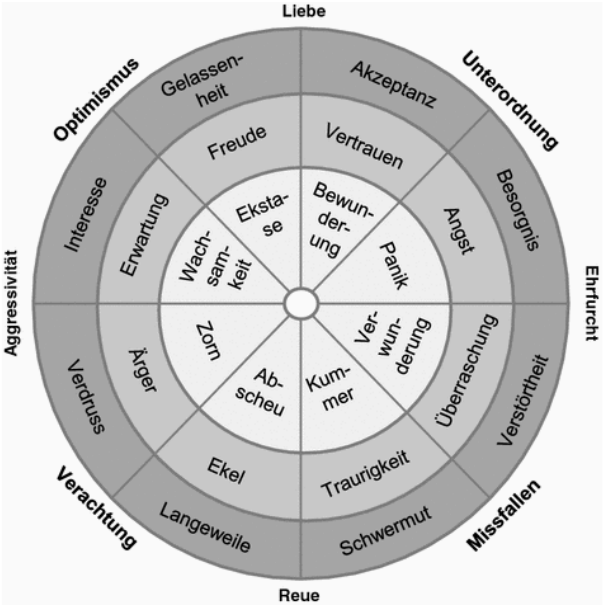
Anmerkung. Abbildung aus Lawrence, K., Campbell, R., & Skuse, D. (2015). Age, gender, and puberty influence the development of facial emotion recognition. *Frontiers in Psychology*, *6*, 761 (https://doi.org/10.3389/fpsyg.2015.00761). (Lawrence et al., 2015)

3.1.1.2 Klassifikation der Emotionen nach Robert Plutchik

Plutchik (1962) entwickelte das kategoriale Emotionsmodell „Rad der Emotionen“ (siehe Abbildung 2). Das Modell veranschaulicht die acht Primäremotionen Ärger, Angst, Ekel, Erwartung, Freude, Traurigkeit, Überraschung und Vertrauen, die in unterschiedlichen Intensitäten mit jeweils variierenden Emotionsbegriffen vorliegen können (z. B. kann Gelassenheit als schwächere und Ekstase als höhere Intensität von Freude angesehen werden). Je ähnlicher sich zwei Emotionen sind, desto näher liegen sie im Rad der Emotionen beieinander. Aus der Mischung der Basisemotionen ergeben sich gemischte bzw. sekundäre Emotionen (z. B. Freude + Akzeptanz = Liebe / Freundschaft).

Abbildung 2

Rad der Emotion in Anlehnung an Plutchik (1991)



Beachte. Das Rad der Emotionen ist in drei Ringe unterteilt. Im mittleren Ring befinden sich die acht Basisemotionen Vertrauen, Angst, Überraschung, Traurigkeit, Ekel, Ärger, Erwartung und Freude. Im inneren Ring werden parallel zur jeweiligen Basisemotion, Emotionsbegriffe benannt, die eine niedrigere Intensität der jeweiligen Basisemotion darstellen. Im äusseren Ring werden entsprechend Emotionsbegriffe der höheren Intensität der jeweiligen Basisemmotion genannt. Begriffe die ausserhalb des Ringes stehen sind Beispiele für gemischte/sekundäre Emotionen, die aus den Mischungen verschiedener Basisemotionen und ihrer Intensitäten entstehen (z. B. Freude + Akzeptanz = Liebe / Freundschaft). Abbildung aus Hoffmann, S., & Akbar, P. (2019). Emotion. In Konsumentenverhalten (pp. 53–69). Springer Fachmedien Wiesbaden (https://doi.org/10.1007/978-3-658-23567-3_4). (Hoffmann & Akbar, 2019)

3.1.1.3 Klassifikation der Emotionen nach Phillip Shaver und Kollegen

Shaver et al. (1987) konnten mithilfe einer hierarchischen Clusteranalyse von 213 Emotionsbegriffe die sechs Primäremotionen Angst, Freude, Liebe, Traurigkeit, Überraschung und Wut identifizieren, die sie als Basisemotionen postulierten. Jedes Cluster kann in zwei weitere Subcluster unterteilt werden, die als sekundäre bzw. tertiäre

Emotionen bezeichnet werden. Die Emotionsbegriffe innerhalb eines Clusters haben ähnliche Bedeutungen, können jedoch nicht als Synonyme angesehen werden. Eine Darstellung der sekundären und tertiären Emotionen der Emotion love (Liebe) wird in Tabelle 3 dargestellt.

Tabelle 3

Darstellung der drei Cluster der Emotion Love (Liebe) nach Shaver et al. (1987)

Cluster 1 Primäre Emotion	Cluster 2 Sekundäre Emotion	Cluster 3 Tertiäre Emotion
love	affection	adoration, fondness, liking, attraction, caring, tenderness, compassion, sentimentality
	lust	arousal, desire, passion, infatuation
	longing	longing

Anmerkung. Beispielhafter Auszug aus den Ergebnissen der hierarchischen Clusteranalyse der 213 Emotionsbegriffe von Shaver und Kollegen 1987. Die Tabelle wurde in Anlehnung an Shaver und Kollegen (2001) erstellt.

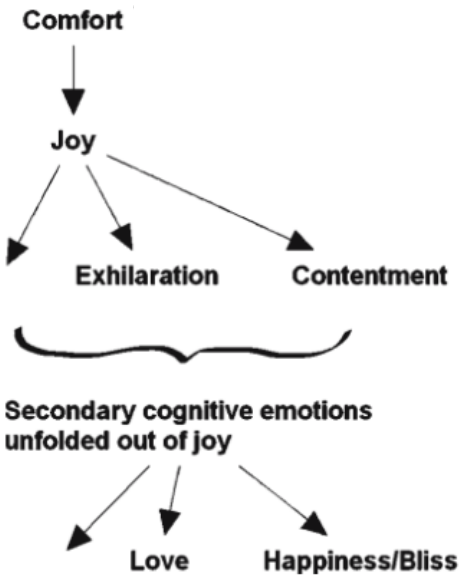
3.1.1.4 Klassifikation der Emotionen nach Alexandra Zinck und Albert Newen

Zinck & Newen (2008) ordnen Emotionen vier Entwicklungsstufen zu. Emotionen einer höheren Entwicklungsstufe gehen auch mit höherer kognitiver Komplexität einher, wobei jede Entwicklungsstufe als Weiterentwicklung der vorherigen angesehen werden kann. Auf der ersten Entwicklungsstufe werden Vor- bzw. Präemotionen verortet, die als wenig komplexe, unspezifische emotionale Reaktionen eine positive oder negative Bewertung einer Situation ermöglichen (z. B. Unbehagen oder Wohlbefinden). Die zweite Entwicklungsstufe beinhaltet die vier Basisemotionen Angst, Freude, Wut und Traurigkeit. Während die Basisemotionen mit universellen, distinkten Gesichtsausdrücken

einhergehen und ihre Verarbeitung automatisch und unbewusst stattfindet, erfolgt im Anschluss an ihre Wahrnehmung eine bewusste Verhaltensreaktion (z. B. Angst führt zu einer Fluchtreaktion). Die Emotionen der dritten Entwicklungsebene werden von Zinck und Newen (2008) als primäre kognitive Emotionen bezeichnet. Sie kommen durch einfache kognitive Bewertungen, wie z. B. dem Wahrnehmen einer Nichterfüllung von Bedürfnissen zustande. Auf der vierten Entwicklungsebene werden sekundäre kognitive Emotionen angesiedelt, deren Erkennen komplexe kognitive Prozesse voraussetzt, wie z. B. das Kennen und Verstehen sozialer Normen und Erwartungen oder das Vorhandensein eines Selbstkonzepts. Abbildung 3 stellt beispielhaft die vier Entwicklungsebenen der Präemotion Wohlbefinden (Comfort) dar.

Abbildung 3

Darstellung der Vier Entwicklungsstufen der Präemotion Comfort (Wohlbefinden)



Anmerkung. Die Abbildung ist ein Ausschnitt der Figure 1 aus Zinck, A., & Newen, A. (2008). Classifying emotion: a developmental account. *Synthese*, *161*(1), 1–25 (https://doi.org/10.1007/s11229-006-9149-2).

Die Unterscheidung zwischen verschiedenen Basisemotionen wird durch neuroanatomische Befunde gestützt, da ihre Wahrnehmung mit unterschiedlichen neuroanatomischen Substraten in Zusammenhang gebracht werden kann. Während zum Beispiel das Erkennen von Wut und Ekel mit der Integrität des orbitofrontalen Cortex und des Cortex insularis assoziiert wird, wird die Wahrnehmung von Trauer eher durch limbische und temporale Regionen unterstützt (siehe Übersicht Garcia-Cordero et al., 2021).

Die vier Theorien bzw. Klassifikationsmodelle verdeutlichen jedoch, dass derzeit weder ein Konsens über die Anzahl und die Benennung der Basisemotionen besteht, noch eine erschöpfende Auflistung aller wahrnehmbarer Emotionen existiert. Darüber hinaus zeigen sie, dass die menschliche Fähigkeit, Emotionen bei anderen Personen zu erkennen, nicht nur auf Basisemotionen beschränkt ist, sondern auch komplexere Emotionen erkannt werden können. Idealerweise sollte deshalb ein EE-Test über die Untersuchung der Fähigkeit Basisemotionen zu erkennen hinausgehen und auch Aufgaben zur Untersuchung der Erkennungsleistung komplexer Emotionen wie gemischter, primärer, sekundärer und tertiärer Emotionen bereitstellen, um die menschliche EE-Fähigkeit realistisch und umfassend abzubilden. Derzeit werden jedoch im Bereich der Demenzforschung hauptsächlich EE-Tests durchgeführt, die sich auf die Untersuchung der Basisemotionen fokussieren [z. B. der Ekman 60 Faces Test (Ekman & Friesen, 1976) oder der *The Awareness of Social Inference Test-Emotion Evaluation Test* (TASIT-EET; McDonald et al., 2003)].

3.1.2 Theroy of Mind (ToM)

Entsprechend der kognitiven Domäne Soziale Kognition existiert für die sozial-kognitive Fähigkeit ToM ebenfalls keine allgemeingültige Definition. In der Neuropsychologie hat sich jedoch eine gängige Arbeitsdefinition formiert, welche die ToM als menschliche Fähigkeit beschreibt, den mentalen Zustand anderer zu erkennen und diesen gleichzeitig von den eigenen Gefühlen, Gedanken und Absichten abgrenzen zu

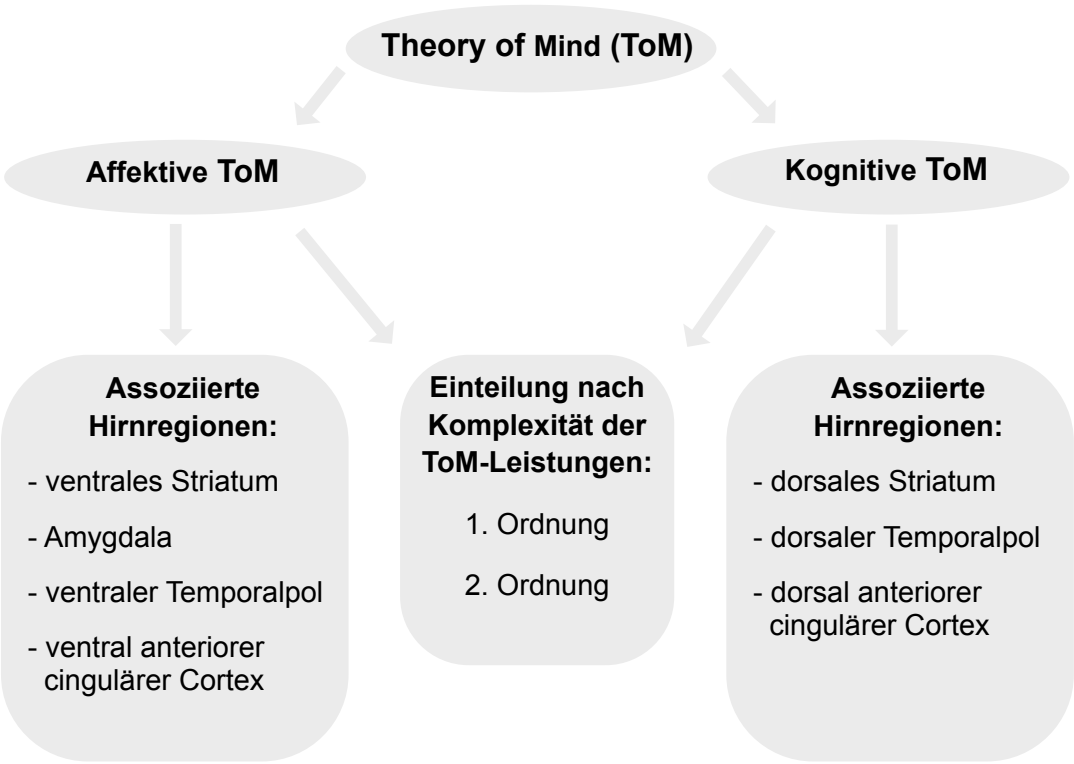
können (z. B. Healey & Grossman, 2018; Henry et al., 2016). Entsprechend der EE basieren auch ToM-Leistungen auf der Wahrnehmung und Interpretation linguistischer und paralinguistischer Hinweisreize (z. B. Gestik, Mimik, Prosodie, Sprechtempo) und lassen die Anpassung des eigenen Verhaltens an soziale Situationen zu (Frith & Frith, 2007). Individuen können unterschiedliche ToM-Formen ausbilden: Während affektive ToM-Leistungen Personen dazu befähigen die Gefühle und Stimmungslagen anderer zu erkennen, lässt die Ausbildung einer kognitiven ToM Schlussfolgerungen über die Gedanken, Überzeugungen und Absichten anderer Individuen zu (z. B. Healey & Grossman, 2018; Sebastian et al., 2012; Shamay-Tsoory et al., 2010). Befunde funktional bildgebender Verfahren unterstützen die Unterteilung in affektive und kognitive ToM. Während die affektive ToM eher ventromedialen und orbitofrontalen Regionen zugeordnet werden kann (ventrales Striatum, Amygdala, ventraler Temporalpol, ventraler anteriorer cingulärer Cortex), können die kognitiven Anteile der ToM eher dem dorsomedialen und dorsolateralem präfrontalen Cortex zugeordnet werden (dorsales Striatum, dorsaler Temporalpol, dorsaler anteriorer cingulärer Cortex) (zur Übersicht siehe Fortier et al., 2018; Healey & Grossman, 2018). Tatsächlich kann jedoch nicht nur die Art der ToM-Leistung, sondern auch ihre Komplexität variieren. Während bei einfachen ToM-Leistungen der mentale Zustand einer einzigen weiteren Person repräsentiert wird, beinhalten komplexe ToM-Leistungen Schlussfolgerungen über die ToM-Leistungen anderer Personen. Aus diesem Grund wird die Komplexität der ToM-Leistungen üblicherweise über verschiedene Ordnungen beschrieben, wobei derzeit vor allem die ersten zwei Ordnungen in klinischen Studien untersucht werden (z. B. Amodio & Frith, 2006; Canty, Neumann, Fleming, & Shum, 2015; Castelli et al., 2011; Wang et al., 2018). Eine ToM-Leistung erster Ordnung befähigt einen Beobachter A den affektiven oder kognitiven Zustand einer Person B zu erkennen (e.g., Meijering, Van Rijn, Taatgen, & Verbrugge, 2011). ToM-Leistungen zweiter Ordnung lassen zudem Mentalisierungen eines Beobachters A über die sozialen Schlussfolgerungen zu, die eine Person B über eine Person C trifft (z. B., Meijering et al., 2011). Abbildung 4 stellt eine Übersicht über die

konzeptuelle Einteilung der ToM-Fähigkeiten dar. Über diese Kategorisierungen hinaus, können ToM-Leistungen auch qualitativ bewertet werden, indem sie auf einem Kontinuum verortet werden, das von keiner Mentalisierung über reduzierte Mentalisierung (Hypomentalisierung) und korrekter Mentalisierung, bis hin zur Generierung zusätzlicher Annahmen über Intentionen und Emotionen, die nicht unmittelbar aus der sozialen Situation abgeleitet werden können (Hypermentalisierung) reicht (Abu-Akel & Bailey, 2000; Bertsch & Herpertz, 2021).

Umfassende Untersuchungen der ToM-Fähigkeiten einer Person sollten konsequenterweise neben den verschiedenen ToM-Formen (d.h. affektive und kognitive ToM) auch die Komplexität und die Qualität der ToM-Leistungen beinhalten. Meines Wissens nach existieren derzeit jedoch keine ToM-Tests, anhand derer sich simultan alle Komponenten der ToM-Fähigkeiten beurteilen lassen. Eine sehr differenzierte Beurteilung mehrerer ToM-Komponenten ermöglichen jedoch beispielsweise zwei filmszenebasierte Testverfahren. Das *Movie for the Assessment of Social Cognition* (MASC; Dziobek et al., 2006) lässt neben der Beurteilung der affektiven und kognitiven ToM-Fähigkeiten eine differenzierte Beurteilung der Qualität der gezeigten ToM-Mentalisierungen zu. Der *The Awareness of Social Inference Test* (TASIT; McDonald et al., 2003) ermöglicht durch die Durchführung der Subtests Social Inference Minimal und Social Inference Enriched zwar keine Beurteilung der Qualität der ToM-Leistungen, lässt jedoch Untersuchungen der affektiven und kognitiven ToM-Leistungen in jeweils erster und zweiter Ordnung zu.

Abbildung 4

Übersicht über die Konzeptuelle Einteilung der Theory of Mind-Fähigkeiten und Assoziierte Hirnregionen



Beachte. ToM=Theory of Mind. Die Aufzählungen der Hirnregionen, welche mit ToM-Fähigkeiten assoziiert werden entstammt 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 (https://doi.org/10.1016/J.NEUROL.2017.07.013).

3.2 Bedeutung der Sozialen Kognition für die bvFTD und AD Diagnostik

EE und ToM-Defizite, sowie ihre diagnostische Relevanz sind sowohl bei Patient:innen mit bvFTD als auch mit AD gut dokumentiert (siehe z. B. Übersichtsarbeiten von Cotter et al., 2018; Desmarais et al., 2018; Dodich et al., 2020; Henry et al., 2016).

3.2.1 Sozial-Kognitive Dysfunktion im Rahmen einer bvFTD

BvFTD ist eine neurodegenerative Erkrankung, die durch frühe progressive Änderungen des Verhaltens, des Sozialverhaltens, der Emotionsverarbeitung sowie durch spezifische kognitive Einschränkungen charakterisiert ist (Johnen & Bertoux, 2019; Rascovsky et al., 2011). Sozial-kognitive Defizite können schon in einer frühen Phase der Erkrankung auftreten und werden mit einer Degeneration des ventromedialen präfrontalen Kortex in Zusammenhang gebracht (siehe Übersicht Poletti et al., 2012).

3.2.1.1 ToM

Das Vorliegen von ToM-Defiziten bei bvFTD Patient:innen wurde anhand verschiedener Testparadigmen, wie zum Beispiel Faux-Pas Aufgaben, False-Belief Aufgaben oder filmszenenbasierte Untersuchen ausführlich dokumentiert und beschrieben (Cotter et al., 2018; Johnen & Bertoux, 2019; Poletti et al., 2012). Die Befunde deuten auf eine Dysfunktion aller ToM Komponenten mit unterschiedlicher Komplexität hin. So zeigen bvFTD Patient:innen sowohl verminderte Leistungen bei affektiven als auch bei kognitiven ToM-Aufgaben, in jeweils erster und zweiter Ordnung (siehe Übersichtsarbeiten von Fortier et al., 2018; Poletti et al., 2012). Darüber hinaus scheinen die Beeinträchtigungen der ToM-Fähigkeiten von bvFTD Patient:innen nicht oder nur in geringem Masse mit Dysfunktionen anderer kognitiver Bereiche assoziiert zu sein, weshalb von einem primären ToM Defizit ausgegangen werden kann (z. B. Henry et al., 2014; Johnen & Bertoux, 2019).

3.2.1.2 EE

Entsprechend der Befunde zu ToM-Defiziten, konnten bei bvFTD Patient:innen auch umfassende Beeinträchtigungen beim Erkennen verschiedener Basisemotionen wie Angst, Ekel, Freude, Trauer, Überraschung und Wut von gezeigt werden (Cotter et al., 2018; Fortier et al., 2018) .

3.2.2 Sozial-Kognitive Dysfunktion im Rahmen einer AD

Im Gegensatz zur bvFTD stehen zu Beginn einer AD nicht Verhaltensänderungen im Vordergrund, sondern ein früher und progressiver Verlust des episodischen Gedächtnisses, der aus neuroanatomischer Sicht mit einer initialen Degeneration des entorhinalen Kortex und des Hippocampus einhergeht (z. B. Christidi et al., 2018). Im Verlauf der AD breitet sich die Neurodegeneration auf den gesamten zerebralen Kortex aus (Braak & Braak, 1991), weshalb nachfolgend auch sozial-kognitive Fähigkeiten beeinträchtigt werden können (z. B. Christidi et al., 2018).

3.2.2.1 ToM

Zahlreiche Studien dokumentieren ToM-Defizite bei AD Patient:innen, jedoch werden die Befunde kontrovers diskutiert, sodass derzeit keine abschliessende Schlussfolgerung aus den Befunden abgeleitet werden kann (Fortier et al., 2018; Poletti et al., 2012). Während einige Studien bei AD Patient:innen einen hohen Einfluss exekutiver Funktionen auf ToM-Leistungen zeigen, gibt es ebenso Erklärungsmodelle und Befunde die ToM-Defizite ohne den Einfluss anderer kognitiver Domänen bei AD Patient:innen darstellen (siehe Übersichtsarbeit von Fortier et al., 2018). Insgesamt gibt es Hinweise darauf, dass AD Patient:innen eher ein Defizit bei komplexeren kognitiven ToM Aufgaben zweiter Ordnung zeigen, während ihre ToM-Leistungen bei kognitiven ToM Aufgaben erster Ordnung und bei affektiven ToM Aufgaben erster und zweiter Ordnung intakt zu sein scheinen (siehe Metaanalyse De Lucena et al., 2020).

3.2.2.2 EE

Die Befunde zur EE bei AD sind ebenfalls inkonsistent. Während einige Studien berichten, dass die EE-Fähigkeiten bei AD Patient:innen intakt sind, zeigt eine Metaanalyse von Klein-Koerkamp et al. (2012), dass eine Beeinträchtigung der EE auch nach Kontrolle anderer kognitiver Funktionen vorliegt.

3.2.3 Manuskript 1

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

Manuskript 1 zeigt zunächst in Form einer Übersichtsarbeit die Relevanz auf Untersuchungen sozial-kognitiver Fähigkeiten in die Differenzialdiagnostik neurodegenerativ bedingter demenzieller Syndrome zu integrieren. Es veranschaulicht, dass Erkenntnisse über individuelle sozial-kognitive Fähigkeiten einen inkrementellen Beitrag zur Frühdiagnostik, zur Differenzial- und Verlaufsdiagnostik im Erkrankungsstadium sowie zur Beratung und Therapie von Betroffenen mit demenziellen Syndromen und ihren Angehörigen liefern können. Aus diesen Erkenntnissen lässt sich die Notwendigkeit ableiten, Untersuchungen sozial-kognitiver Fähigkeiten in den Diagnostikprozess demenzieller Erkrankungen zu integrieren. Anschliessend wurde mittels einer explorativen Onlineumfrage evaluiert, ob 87 deutschsprachige MC/GS der Länder Deutschland, Österreich und Schweiz] die Relevanz sozial-kognitiver Untersuchungen erkannt haben und sozial-kognitive Testverfahren an ihrer Einrichtung durchführen. Die Ergebnisse zeigten, dass die Mehrheit der MC/GS (ca. 85%), welche den Fragebogen beantworteten, die Untersuchung der EE- und ToM- Fähigkeiten durchaus als wichtig erachteten. Jedoch erfolgte eine Testung dieser sozial-kognitiven Fähigkeiten bisher nur vereinzelt. Hierfür führten sie folgende Hauptgründe an: (1) Zu lange Testdauer, (2) Vorhandene Tests sind für Patient:innen mit Störungen in mehreren kognitiven Domänen ungeeignet und (3) Niedrige Sensitivität der sozial-kognitiven Tests. Die wenigen MC/GS, die sozial-kognitive Tests durchführten, zeigten sich mit diesen

weitgehend unzufrieden und begründeten dies anhand folgender Kritikpunkte: (1) Fehlende Normierung, (2) Fehlende Validierung, (3) Zu lange Testdauer und (4) Realitätsferne Stimuli. Basierend auf den Umfrageergebnissen werden abschliessend Empfehlungen für die Entwicklung sozial-kognitiver Tests ausgesprochen, damit zukünftig Neuentwicklungen und Testadaptionen den Anforderungen der Praxis standhalten und in die Diagnosealgorithmen integriert werden können.

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

A gap that needs to be closed: The assessment of social-cognitive abilities in German-speaking memory clinics

Marianne Jarsch[1,2], Andreas Monsch[1,2], Reto W. Kressig[1], Marc Sollberger[1,3]

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

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

[3]Department of Neurology, University Hospital Basel, Switzerland.

Korrespondierende Autorin:

Marianne Jarsch

Universitäre Altersmedizin FELIX PLATTER, Memory Clinic

Burgfelderstrasse 101, 4055 Basel, Schweiz

E-Mail: marianne.jarsch@felixplatter.ch

Zusammenfassung

Die Untersuchung sozial-kognitiver Fähigkeiten kann die Differenzialdiagnostik bei neurodegenerativ bedingten demenziellen Syndromen entscheidend verbessern. Sozial-kognitive Testverfahren, die sich im Forschungskontext als passend erwiesen haben, sind jedoch nur schwer in die klinische Praxis integrierbar. In der vorliegenden Studie führten wir eine Befragung zum aktuellen Stand der sozial-kognitiven Untersuchung an den Memory Clinics und Gedächtnissprechstunden (MC/GS) in Deutschland, Österreich und der Schweiz durch. Die Ergebnisse zeigen, dass die deutschsprachigen MC/GS die Bedeutung sozial-kognitiver Untersuchungen zwar erkannt haben, sozial-kognitive Tests jedoch nur von wenigen Institutionen durchgeführt und diese meist als nicht zufriedenstellend bewertet werden. Basierend auf unseren Studienergebnissen sprechen wir Empfehlungen für die Entwicklung sozial-kognitiver Tests aus, damit in Zukunft Neuentwicklungen den Anforderungen der Praxis standhalten und in die Diagnosealgorithmen integriert werden können.

Schlüsselwörter: Emotionserkennung, Theory of Mind, sozial-kognitive Fähigkeiten, Demenzdiagnostik, neuropsychologische Untersuchung

Abstract

The assessment of social-cognitive abilities can clearly improve the differential diagnosis of dementia syndromes caused by neurodegeneration. However, social-cognitive tests that have proven appropriate in research settings are difficult to integrate into clinical practice. In the present study, we conducted a survey on the status quo of social-cognitive assessments in memory clinics (MC/GS) in Germany, Austria and Switzerland. Our results show that although the German-speaking MC/GS acknowledge the importance of social-cognitive testing, social-cognitive tests are only carried out by a few institutions and are mainly considered as unsatisfactory. Based on our study results, we make recommendations for the development of social-cognitive tests, that can fulfill the demands of clinical practice and be integrated into diagnostic algorithms in future.

Keywords: Emotion recognition, Theory of Mind, social-cognitive abilities, diagnosis of dementia, neuropsychological assessment

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

Memory Clinics und Gedächtnissprechstunden (MC/GS) führen als multiprofessionelle Kompetenzzentren im deutschsprachigen Raum die Diagnostik, Behandlung und Beratung von Personen mit demenziellen Erkrankungen durch (Diehl et al., 2005; Lautenschlager & Kurz, 2009; Swiss Memory Clinics, 2021). Teil der interdisziplinären diagnostischen Abklärung sind ausführliche neuropsychologische Untersuchungen, die das Erstellen eines kognitiven Leistungsprofils ermöglichen und zur Frühdiagnostik sowie zur Beantwortung differenzialdiagnostischer Fragestellungen beitragen (Diehl et al., 2005). Die fünfte Auflage des *Diagnostischen und Statistischen Manuals Psychischer Störungen* (DSM-5; American Psychiatric Association, 2013) empfiehlt zur umfassenden Beschreibung des klinischen Bildes und der Beurteilung des Schweregrades einer vorliegenden neurokognitiven Störung die Untersuchung sechs kognitiver Domänen: (1) Sprache, (2) Lernen und Gedächtnis, (3) komplexe Aufmerksamkeit, (4) Exekutivfunktionen, (5) perzeptuell-motorische Fähigkeiten und (6) soziale Kognition (Emotionserkennung und Theory of Mind). Im deutschsprachigen Raum stehen für die Untersuchung der kognitiven Leistungen in den Bereichen (1) bis (5) etablierte validierte und normierte Testverfahren zur Verfügung, wie z.B. die CERAD-Plus (Monsch, 2015). Die Beurteilung der sozialen Kognition scheint sich jedoch deutlich schwieriger zu gestalten, da derzeit weder im DSM-5 (American Psychiatric Association, 2013) noch in der aktuellen Literatur normierte und validierte deutschsprachige Tests empfohlen werden und es keine einheitliche Definition dieser kognitiven Domäne gibt [eine Übersicht sozial-kognitiver Tests, die in der Forschung verwendet werden, kann bei Henry und Kollegen (2015) eingesehen werden]. Üblicherweise werden unter sozialer Kognition alle informationsverarbeitenden Prozesse bzw. sozial-kognitiven Fähigkeiten zusammengefasst, die darauf abzielen den mentalen Zustand und das Verhalten anderer wahrzunehmen und zu interpretieren, um das eigene Verhalten an die soziale Situation anpassen zu können (Frith & Frith, 2007). Im klinischen Kontext werden vor allem die

sozial-kognitiven Fähigkeiten der Emotionserkennung (EE) und Theory of Mind (ToM) erforscht (Cotter et al., 2018). Während EE die menschliche Fähigkeit beschreibt die Emotionen anderer wahrnehmen und benennen zu können, werden ToM-Leistungen als komplexe Fähigkeit zur Perspektivübernahme angesehen (Pinkham et al., 2014). Die Ausbildung einer ToM ermöglicht uns die Gedanken, Absichten und Emotionen anderer zu verstehen und diese von unserem eigenen mentalen Zustand abzugrenzen (Henry et al., 2016).

Im Folgenden wird die Relevanz der Untersuchung von EE- und ToM-Fähigkeiten für die Demenzdiagnostik und sozial-kognitive Beratung und Therapie näher beleuchtet. Anschließend stellen wir anhand unserer Expert:innen-Umfrage den aktuellen Stand der Erfassung sozial-kognitiver Fähigkeiten an deutschsprachigen MC/GS vor und leiten aus den Ergebnissen Empfehlungen zur Neuentwicklung sozial-kognitiver Tests für die klinische Praxis ab.

Die Untersuchung sozial-kognitiver Fähigkeiten im Kontext der Demenzdiagnostik

Etablierte neuropsychologische Testverfahren sind derzeit nur begrenzt in der Lage zwischen verschiedenen neurodegenerativ-bedingten demenziellen Syndromen zu diskriminieren, da die verwendeten Tests nicht alle betroffenen Hirnregionen und neuronale Netzwerke abdecken, die kognitive Prozesse repräsentieren. Insbesondere das „Salience Network“ (Seeley et al., 2007) und das „Semantic Appraisal Network“ (Yeo et al., 2011), die mit sozial-kognitiven Fähigkeiten assoziiert werden (z.B. Christidi et al., 2018; Toller et al., 2018) und bei bestimmten neurodegenerativen Erkrankungen (z.B. der frontotemporalen Lobärdegeneration) betroffen sind (Menon, 2011; Seeley et al., 2009), können nicht erfasst werden. Entsprechend ist der Einbezug der Testung sozial-kognitiver Fähigkeiten zur verbesserten Demenzdiagnostik von zentraler Bedeutung. In den letzten Jahren sind ausführliche Übersichtsarbeiten und Metaanalysen entstanden, die die Relevanz der Erfassung sozial-kognitiver Fähigkeiten in der Demenzdiagnostik herausarbeiten und betonen (z.B. Arioli et al., 2018; Cotter et al., 2018; Dodich et al.,

2020; Elamin et al., 2012; Henry et al., 2015, 2016; Johnen & Bertoux, 2019; ZhongQuan et al., 2020).

Frühdiagnostik

Dank einer guten Sensibilisierung der Hausärzt:innen für das Auftreten demenzieller Erkrankungen und der Entwicklung neuer effizienter Methoden zum case-finding (z.B. BrainCheck, Ehrensperger et al., 2014) und Screening (z.B. MoCA, Nasreddine et al., 2005) nehmen vermehrt Patient:innen mit leichten kognitiven Störungen (Mild Cognitive Impairment, MCI) die diagnostische Abklärung an einer MC/GS in Anspruch. Für diese Patientengruppe ist es bedeutsam frühzeitig den MCI-Subtyp und die verursachende Pathologie zu identifizieren, um mit passenden Therapie- und Beratungsmaßnahmen den Verlauf der Symptome positiv zu beeinflussen. Derzeit ist die Differenzialdiagnostik jedoch äußerst herausfordernd, da die standardmäßig durchgeführten neuropsychologischen Untersuchungen häufig nicht zwischen verschiedenen MCI-Subtypen unterscheiden (Dodich et al., 2021) und sich zudem die Symptome möglicher zugrunde liegender Pathologien überschneiden können (Johnen & Bertoux, 2019). Die Untersuchung sozial-kognitiver Fähigkeiten könnte hier eine sinnvolle Ergänzung darstellen. Aktuelle Studien zeigen, dass sowohl reduzierte ToM- als auch EE-Fähigkeiten als kognitive Marker Hinweise auf einen MCI-Subtyp und eine mögliche Progression hin zur Alzheimer Erkrankung (AD) liefern könnten (Dodich et al., 2021; McCade et al., 2013; Michaelian et al., 2021; ZhongQuan et al., 2020). Zudem könnte die Erfassung sozial-kognitiver Fähigkeiten dazu beitragen ein frühes Stadium der Verhaltensvariante der fronto-temporalen Demenz (bvFTD) von einer AD oder psychiatrischen Erkrankung abzugrenzen (Chiu et al., 2018; Cotter et al., 2018; Desmarais et al., 2018; Johnen & Bertoux, 2019).

Diagnostik im fortgeschrittenen demenziellen Erkrankungsstadium

Patient:innen mit neurodegenerativ-bedingten demenziellen Erkrankungen zeigen häufig reduzierte Fähigkeiten Emotionen zu erkennen und ToM-Leistungen zu vollziehen (z.B. Arioli et al., 2018; Cotter et al., 2018; Elamin et al., 2012; Henry et al., 2016). Die

Schwere dieser sozial-kognitiven Defizite ist vergleichbar mit den Störungen in anderen kognitiven Bereichen (Cotter et al., 2018). Bei verschiedenen Demenzformen und Schweregraden der Erkrankungen zeigen sich jedoch unterschiedliche EE- und ToM-Störungsmuster, weshalb EE- und ToM-Defizite als kognitive Marker in der Differenzial- und Verlaufsdiagnostik eingesetzt werden könnten (z.B. Castelli et al., 2011; Dodich et al., 2020; Poletti et al., 2012). Vor allem im Bereich der bvFTD-Diagnostik könnte die Untersuchung sozial-kognitiver Fähigkeiten einen zusätzlichen Vorteil bieten, da zumeist molekulare Biomarker zur Diagnosestellung fehlen (Ducharme et al., 2020; Johnen & Bertoux, 2019) und das klinische Erscheinungsbild einer bvFTD häufig psychiatrischen Erkrankungen (z.B. affektive Störungen, Schizophrenie) ähnelt (Ducharme et al., 2020; Woolley et al., 2011). Fehldiagnosen sind entsprechend häufig und treten bei ca. 50% der Patient:innen auf (Krudop et al., 2017; Shinagawa et al., 2016). Dodich und Kolleg:innen (2020) zeigten, dass sowohl EE- als auch ToM- Tests mit einer hohen Akkuratheit (71%-93.3%) zwischen einer bvFTD, AD und psychiatrischen Störungen diskriminieren können.

Relevanz sozial-kognitiver Defizite für Beratung und Therapie

Reduzierte sozial-kognitive Fähigkeiten werden mit einem veränderten Sozialverhalten und vermindertem psychosozialen Funktionsniveau in Zusammenhang gebracht (Brüne & Juckel, 2010; Hasson-Ohayon et al., 2017; Kennedy & Adolphs, 2012; Shimokawa et al., 2001). Bei Patient:innen mit demenziellen Syndromen treten diese problematischen Kognitions- und Verhaltensänderungen häufig auf (Cotter et al., 2018; Desmarais et al., 2018; Kessels et al., 2021) und werden somit auch für die Behandlung und die Beratung von Betroffenen und deren Betreuungspersonen relevant. Ein unangepasstes Sozialverhalten und vermindertes psychosoziales Funktionsniveau können die zwischenmenschlichen Beziehungen der Betroffenen, deren Lebensgestaltung (Nelis et al., 2011; Timmer et al., 2020) und die psychische Belastung der Betreuungspersonen (McCade et al., 2011, 2013; Nelis et al., 2011; Shimokawa et al., 2001) negativ beeinflussen. Die Folgen hiervon sind eine Reduktion der Lebensqualität und des Wohlbefindens der Betreuungspersonen (siehe Metaanlyse Martyr et al., 2018).

Psychoedukation und ein individuell angepasstes Training sozial-kognitiver Fähigkeiten könnten diese negativen Effekte möglicherweise abmildern.

Der differenzialdiagnostische Nutzen sozial-kognitiver Untersuchungen und die negativen Effekte sozialer Verhaltensänderungen auf das Wohlbefinden und die Lebensqualität der Betroffenen und ihrer Bezugspersonen, haben dazu geführt, dass internationale Demenz-Fachinitiativen (z.B. Boccardi et al., 2021; Ducharme et al., 2020) die Integration sozial-kognitiver Untersuchungen in die Diagnosealgorithmen ausdrücklich empfehlen. Bis dato konnten wir jedoch keine deutschsprachigen Testverfahren ermitteln, die einen routinemäßigen Einsatz in der Einzelfalldiagnostik einer MC/GS rechtfertigen würden, indem sie folgende fünf Kriterien erfüllen: (1) Angemessene Konstruktvalidität mit Schwerpunkt auf prädiktiver Validität für Sozialverhalten/psychosoziales Funktionsniveau und hohe Reliabilität, (2) Normierung unter Kontrolle der Einflüsse von demografischen Variablen wie Alter, Geschlecht und Ausbildung, (3) Validierung für spezifische Störungsbilder (z.B. bvFTD, AD), (4) Durchführbarkeit mit Personen, die kognitive Störungen in mehreren kognitiven Bereichen aufweisen (z.B. kurze Testdauer, geringe Belastung anderer kognitiver Kapazitäten) und (5) keine Decken- bzw. Bodeneffekte bei Gesunden, sodass auch leichte Defizite in den sozial-kognitiven Fähigkeiten identifiziert werden können. Wir bildeten daher eine Arbeitsgruppe, mit dem Ziel eine deutschsprachige Adaption des *The Awareness of Social Inference Test* (TASIT, McDonald et al., 2003) vorzunehmen, dessen Subtests für EE und ToM gute psychometrische Eigenschaften aufweisen. Die Basler Versionen des Emotion Evaluation Test (BASIT-Emotion Recognition; Jarsch et al., 2021) und des Sozialen Inferenz Tests des TASIT (BASIT-ToM; Jarsch, Piguet, et al., 2021) sollen in Zukunft als normierte, valide, reliable und sensitive Testverfahren zur Untersuchung sozial-kognitiver Fähigkeiten in der Demenzdiagnostik eingesetzt werden. Angesichts der zahlreich existierenden EE- und ToM Tests in der Forschung, scheint es jedoch problematisch zu sein, Tests zu entwickeln, die in die Diagnosealgorithmen einer MC/GS integriert werden können. Daher möchten wir die Anforderungen an einen sozial-kognitiven Test, die von

den deutschsprachigen Expert:innen für neurokognitive Diagnostik formuliert werden, bei unserer Testentwicklung berücksichtigen. Zu diesem Zweck entwickelten wir einen explorativen Online-Fragebogen, der zum einen den aktuellen Stand der sozial-kognitiven Diagnostik an deutschen, österreichischen und schweizerischen MC/GS erhebt und zum anderen spezifische Antworten zu den Anforderungen der Praxis an Sensitivität, Teststimuli, Antwortformat und Testdurchführung liefern kann. Basierend auf den Ergebnissen der Expert:innen-Umfrage werden wir Empfehlungen für Entwicklungen sozial-kognitiver Tests ausformulieren und diese in die Entwicklung der BASIT-Testbatterie miteinfließen lassen.

Methoden

Studienteilnehmer

Es wurden 198 deutschsprachige MC/GS (Deutschland: n=155, Schweiz: n=28, Österreich: n=15) zur Studienteilnahme angefragt, die im Januar 2021 in Deutschland, der Schweiz oder Österreich offiziell als deutschsprachige MC/GS ausgewiesen waren. Die Leitungspersonen der jeweiligen MC/GS erhielten per E-Mail einen Weblink, durch den sie auf den Online-Fragebogen weitergeleitet wurden. Falls die Leitungspersonen selbst keine Aussage zur Diagnostik sozial-kognitiver Fähigkeiten an ihren Institutionen treffen konnten, bestand die Möglichkeit den Fragebogen im Team auszufüllen bzw. ihn intern an die entsprechende Fachpersonen weiterzuleiten. Wir erhoben die Berufsgruppenzugehörigkeit der Fragebogenteilnehmenden, um zu verifizieren, dass die Fragen von Fachpersonen beantwortet wurden. Weitere Informationen über die Auswahl- und Teilnahmebedingungen der Studienteilnehmer können im elektronischen Zusatzmaterial (ESM) 1 eingesehen werden. In Absprache mit der Ethikkommission der Nordwest- und Zentralschweiz bedurfte es keiner Ethikbewilligung.

Expert:innen-Fragebogen

Der explorative Online-Fragebogen wurde mit der webbasierten Umfrage-Software LimeSurvey Version 3.25.8+210118 (LimeSurvey Project Team, 2021) erstellt. Die Druckversion des Fragebogens kann im ESM 2 eingesehen werden.

Anzahl und Format der Fragen

Der Fragebogen beinhaltete in Abhängigkeit des Antwortverhaltens maximal 22 Fragen mit den Antwortformaten 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 sozial-kognitiven Fähigkeiten. Die Institutionen wurden gebeten, die Relevanz von fünf sozial-kognitiven Fähigkeiten (EE, ToM, gezielter Umgang mit wahrgenommenen Emotionen, Soziale Wahrnehmung, Attributionsstil) auf einer vierstufigen Skala (nicht relevant bis relevant) zu beurteilen. Diese fünf sozial-kognitiven Fähigkeiten wurden durch einen Expert:innen-Konsens im Bereich der Schizophrenieforschung (Pinkham et al., 2014) als die wichtigsten Bereiche sozialer Kognition identifiziert. Zusätzlich bestand die Möglichkeit in einem Freitextfeld weitere relevante sozial-kognitive Fähigkeiten zu benennen. Die Antworten sollten Aufschluss darüber geben, ob die MC/GS die Untersuchung sozial-kognitiver Fähigkeiten grundsätzlich als wichtig erachten und welche Bereiche sozialer Kognition im Fokus stehen sollten.

Relevanz sozial-kognitiver Untersuchungen für Patientengruppen mit unterschiedlichen Erkrankungen. Die Institutionen wurden gebeten Patientengruppen mit verschiedenen Erkrankungen auszuwählen, für welche sie sozial-kognitive Diagnostik als relevant erachten.

Untersuchung der EE- und ToM-Fähigkeiten an deutschsprachigen MC/GS. Die Institutionen wurden gebeten anzugeben, ob sie Tests zur Untersuchung der EE- und ToM-Fähigkeiten durchführen und falls ja, diese in einem Freitextfeld zu benennen. Anschließend sollten sie ihre Zufriedenheit mit diesen Tests angeben. Die Antworten sollten uns Aufschluss darüber geben, ob und wie sozial-kognitive Untersuchungen in den MC/GS derzeit stattfinden. Zusätzlich sollten die Institutionen die Wahrscheinlichkeit einschätzen, dass sie einen neu entwickelten Test einsetzen, der sozial-kognitive Fähigkeiten praktikabel, valide und reliabel erfasst.

Gründe, die gegen die Durchführung von EE- und ToM-Tests angeführt werden. Die Institutionen konnten in einem Freitextfeld Gründe formulieren, die gegen die Durchführung sozialer-kognitiver Tests sprechen und/oder aus den folgenden fünf Gründen auswählen: (1) Tests sind für Patient:innen mit Störungen in mehreren kognitiven Domänen ungeeignet, (2) Die Durchführung dauert zu lange, (3) zu niedrige Test-Sensitivität, (4) keine Therapierelevanz der Testergebnisse, (5) Eine Durchführung wurde bis dato nicht in Erwägung gezogen. In die Analyse werden diejenigen Institutionen miteinbezogen, die aktuell keine sozial-kognitiven Tests durchführten, eine Durchführung jedoch in Erwägung gezogen hatten.

***Alternative Erfassung sozial-kognitiver Fähigkeiten*.** Institutionen, die aktuell keine sozial-kognitiven Tests im Rahmen der neurokognitiven Diagnostik durchführten, wurden gebeten anzugeben, ob sie diese dennoch im Rahmen von Patient:innen- und/oder Angehörigenbefragungen erfassen. Zusätzlich stand ihnen ein Freitextfeld zur Verfügung, um weitere alternative Erfassungsmethoden zu benennen. Die Angaben erlaubten es zu erfassen, ob die sozial-kognitiven Fähigkeiten von Patient:innen trotz fehlender Verwendung sozial-kognitiver Tests im Diagnoseprozess berücksichtigt werden.

Anforderungen an Testsensitivität, Teststimuli, Antwortformat und Testdurchführung. Die Anwendung eines Tests in der Einzelfalldiagnostik der klinischen Routine ist nicht ausschließlich von Qualitätskriterien wie vorhandener Normierung und Validierung abhängig, sondern auch von praktischen Kriterien, wie z.B. einer

angemessenen Durchführungsdauer oder Zielgruppenpassung. Aus diesem Grund interessierten wir uns für folgende Aspekte: (1) Anforderungen der MC/GS an die Durchführungsdauer von sozial-kognitiven Screening- bzw. ausführlichen Tests, (2) Wichtigkeit einer assistierten Testdurchführung (z.B. Hilfe bei der Navigation durch eine computerbasierte Applikation oder Vorlesen der Testfragen), (3) Wichtigkeit realistisch wirkender Teststimuli, (4) Wichtigkeit einer hohen Testsensitivität. In einem Freitextfeld am Ende des Fragebogens hatten die Institutionen zusätzlich die Möglichkeit uns weitere Punkte mitzuteilen. Die Institutionen konnten die Wichtigkeit der jeweiligen Anforderungen auf einer vierstufigen Skala von unwichtig bis wichtig einschätzen.

Analysen

Die Datenanalyse wurde mit IBM SPSS Statistics for Macintosh, Version 27.0.1.0 (IBM Corp., 2007) durchgeführt. Die Antworten der vollständig und teilweise ausgefüllten Fragebögen wurden mittels Häufigkeitsanalysen ausgewertet.

Ergebnisse

Anzahl teilnehmender MC/GS und Rücklaufquote

Von insgesamt 198 angeschriebenen deutschen (n=155), schweizerischen (n=28) und österreichischen (n=15) MC/GS nahmen 43.9% (n=87) an der Befragung teil. Die Rücklaufquote lag für deutsche MC/GS bei 41.9%, für schweizerische MC/GS bei 67.9% und für österreichische MC/GS bei 20%. Die teilnehmenden Institutionen teilten sich nach Ländern wie folgt auf: Deutschland (74.7%; n=65), Schweiz (21.8%, n=19) und Österreich (3.4%; n=3). 94.3% (n=82) der MC/GS füllten den Fragebogen vollständig aus, 5.7% (n=5) beantworten ihn teilweise.

Diagnostische Relevanz verschiedener sozial-kognitiver Fähigkeiten

Die Mehrheit der Institutionen bewertete die diagnostische Relevanz der fünf zur Auswahl gestellten sozial-kognitiven Fähigkeiten als „eher relevant“ bzw. „relevant“ [Emotionen erkennen 86.2% (n=75), ToM: 83.9% (n=73), soziale Wahrnehmung: 95.4%

(n=83), gezielter Umgang mit wahrgenommenen Emotionen: 77% (n=67), Attributionsstil: 73.6% (n=64)]. Die Auswertung des Freitextfeldes wird im ESM 3.1 beschrieben.

Relevanz sozial-kognitiver Untersuchungen für Patientengruppen mit unterschiedlichen Erkrankungen

92.7% (n=76) der MC/GS befürworteten sozial-kognitive Untersuchungen bei Patienten:innen mit neurodegenerativen Erkrankungen. Die Ergebnisse zu weiteren Patientengruppen werden im ESM 3.2 beschrieben.

Untersuchung der EE-Fähigkeit an deutschsprachigen MC/GS

86.0% (n=74) der MC/GS gaben an, dass sie keine Tests zur Untersuchung der EE-Fähigkeit durchführten. Von diesen hatten 58.1% (n=43) die Implementierung eines solchen Tests bis dato nicht in Erwägung gezogen. 33.8% (n=25) der Institutionen hatten sich trotz der Überlegung EE-Tests durchzuführen, aufgrund folgender Gründe dagegen entschieden: (1) Die Durchführung dauert zu lange [52.0% (n=13)], (2) Die vorhandenen EE-Tests, sind für Patienten:innen mit Störungen in mehreren kognitiven Domänen ungeeignet [44.0% (n=11)], (3) Die vorhandenen EE-Tests weisen eine niedrige Sensitivität auf [44.0% (n=11)]. 8.1% (n=6) der Institutionen machte keine Angaben, warum sie aktuell keine EE-Tests implementiert hatten. Weitere Gründe, die jeweils von einer Institution genannt wurden, werden im ESM 3.3 aufgelistet.

62.2 % (n=46) der MC/GS, die keine EE Tests durchführten, nahmen eine subjektive Einschätzung der EE-Fähigkeit ihrer Patient:innen vor, 28.4% (n=21) führten keine subjektive Einschätzung der EE-Fähigkeit durch und 27.0% (n=20) der MC/GS machten diesbezüglich keine Angaben. 80.4% (n=37) der Institutionen, die eine subjektive Beurteilung der EE-Fähigkeiten vornahmen, führten hierfür eine Kombination aus Angehörigen- und Patient:innenbefragung durch, 10.9% (n=5) befragten ausschließlich die Angehörigen der Betroffenen, 4.3% (n=2) führten ausschließlich Patient:innen-Befragungen durch und eine Institution (2.2%) beurteilte die EE-Fähigkeit ausschließlich auf der Basis einer klinischen Beobachtung. 2.2% (n=1) der MC/GS machten keine Aussage zur subjektiven Beobachtungsmethode der EE-Fähigkeit ihrer Patient:innen.

Von den 12 Institutionen, die EE-Tests durchführten, gaben 25% (n=3) an, dass EE-Tests ausschließlich im Rahmen von Forschungsprojekten zum Einsatz kamen. 50% (n=6) der MC/GS verwendeten statisches Bildmaterial (siehe ESM 3.4), die weiteren 50% machten bezüglich der verwendeten Tests keine Angaben. Die Häufigkeitsverteilung der Zufriedenheitsbewertung der Institutionen mit den verwendeten Tests war wie folgt: „eher zufrieden“/„zufrieden“ 58.3% (n=7), „neutral“ 8.3% (n=1), „eher unzufrieden“/„unzufrieden“ 33.3% (n=4).

Zehn Institutionen benannten für die durchgeführten EE-Tests folgenden Nachteile: (1) Es existieren keine Normierungen [41.7% (n=5)], (2) Die Durchführung ist mit einem hohen Zeitaufwand verbunden [25% (n=3)], (3) Die Tests sind zu wenig realitätsnah [25% (n=3)]. Die Vorteile und weitere Nachteile, die jeweils von weniger als 20% der Institutionen genannt wurden, werden im ESM 3.5 aufgelistet.

Untersuchung der ToM-Fähigkeiten an deutschsprachigen MC/GS

84.5% (n=71) der Institutionen gaben an, keine ToM-Tests im Rahmen der neurokognitiven Diagnostik durchzuführen. Von diesen zogen 52.1% (n=37) eine Durchführung nicht in Erwägung. 46.5% (n=33) der Institutionen hatten sich trotz der Überlegung ToM-Tests durchzuführen, aufgrund folgender Gründe dagegen entschieden: (1) Durchführung dauert zu lange [60.6% (n=20)], (2) Die Tests sind für Patient:innen mit mehreren kognitiven Störungen ungeeignet [39.4% (n=13)], (3) Vorhandene ToM-Tests weisen eine zu niedrige Sensitivität auf [27.3% (n=9)]. 1.4% (n=1) der MC/GS machten keine Angaben, warum sie aktuell keine ToM-Tests implementiert hatten. Weitere Gründe, die von weniger als 20% der Institutionen benannt wurden, werden im ESM 3.6 aufgelistet.

50.7% der Institutionen (n=36), die keine ToM Tests durchführten, nahmen eine subjektive Einschätzung der ToM-Fähigkeiten ihrer Patient:innen vor, 47.9% (n=34) machten keine subjektive Einschätzung und 1.4% (n=1) der Institutionen machte diesbezüglich keine Angaben. 77.8% (n=28) der MC/GS, die subjektive Beurteilungen vornahmen, führten eine Kombination aus Angehörigen- und Patient:innenbefragung

durch, 16.7% (n=6) befragten ausschließlich die Angehörigen der Betroffenen, 2.8% (n=1) führten ausschließlich Patient:innen-Befragungen durch und eine Institution (2.8%) beurteilte die ToM-Fähigkeit ausschließlich auf der Basis einer klinischen Beobachtung.

Von den 13 Institutionen, die ToM-Tests durchführten, gaben 15.4% (n=2) der Institutionen an diese ausschließlich für Forschungszwecke zu verwenden. 92.3% (n=12) der Institutionen führten Tests mit statischem Stimulusmaterial durch (siehe ESM 3.7). Eine Institution machte bezüglich des verwendeten ToM-Tests keine Angaben. Die Häufigkeitsverteilung der Zufriedenheitsbewertung der Institutionen mit den verwendeten Tests war wie folgt: „eher unzufrieden“/„unzufrieden“ 46.2% (n=6), „neutral“ 23.1% (n=3), „eher zufrieden“/„zufrieden“ 30.8% (n=4). Die Institutionen nannten für die verwendeten ToM-Tests folgende Kritikpunkte: (1) Fehlende Validierung [46.2% (n=6)], (2) Die Durchführung dauert zu lange [30.8% (n=4)]. Weitere Vor- und Nachteile, die jeweils von weniger als 20% der Institutionen genannt wurden, werden im ESM 3.8 aufgelistet.

Anforderungen an die Testentwicklung

Sensitivität

72.4% (n=63) der Institutionen beurteilten eine hohe Test-Sensitivität als „eher wichtig“/„wichtig“. Hingegen bewerteten 21.8% (n=19) der Institutionen eine hohe Sensitivität als „eher unwichtig“ bzw. „unwichtig“. 5.7% (n=5) der Institutionen machten diesbezüglich keine Angaben.

Durchführungsdauer

Die Häufigkeitsverteilung der bevorzugten maximalen Durchführungsdauer für einen Screeningtest zur Untersuchung sozial-kognitiver Fähigkeiten war wie folgt: Max. 10 Minuten 82.6% (n=72), max. 20 Minuten 11.5% (n=10), max. 30 Minuten 1.1% (n=1), keine Angabe 4.6% (n=4).

Für eine ausführliche sozial-kognitive Diagnostik bevorzugten 10.3% (n=9) max. 10 Minuten, 36.8% (n=32) max. 20 Minuten, 28.7% (n=25) max. 30 Minuten, 6.9% (n=6) max. 40 Minuten, 1.1% (n=1) max. 50 Minuten und 1.1% (n=1) über 1 Stunde Durchführungszeit. 4.6% (n=4) der Institutionen machten diesbezüglich keine Angaben.

Art der Durchführung

88.5% (n=77) der MC/GS hielt es für „eher wichtig“ bzw. „wichtig“, dass die Testleitung Patient:innen bei der Durchführung eines Tests assistieren kann (z.B. Vorlesen der Fragen, Hilfe bei der Navigation durch eine computerbasierte Applikation). 6.9% (n=6) der MC/GS beurteilten die Möglichkeit der assistierten Durchführung als „eher unwichtig“. 4.6% (n=4) der Institutionen machten diesbezüglich keine Angaben.

Test-Stimuli

88.5% (n=77) der MC/GS bewerteten realitätsnahe Test-Stimuli als „wichtig“ bzw. „eher wichtig“. 6.9% (n=6) der Institutionen bewerteten realitätsnahe Stimuli als „unwichtig“ bzw. „eher unwichtig“. 4.6% (n=4) der Institutionen machten diesbezüglich keine Angaben.

Antwortformat

Etwa jeweils ein Drittel der Institutionen bevorzugte entweder eine Kombination aus offenem und geschlossenem Antwortformat [37.9% (n=33)] oder ein geschlossenes Antwortformat mit Mehrfachauswahl [35.6%(n=31)]. 17.2% (n=15) der Institutionen bevorzugten ein geschlossenes Antwortformat mit Einfachauswahl. 5.7% (n=5) machten diesbezüglich keine Angaben.

Bedarf der MC/GS an einem praktikablen sozial-kognitiven Test mit guten psychometrischen Eigenschaften

89% (n=73) der MC/GS erachteten es als „eher wahrscheinlich“/„wahrscheinlich“ zukünftig einen praktikablen, validen und reliablen sozial-kognitiven Test mit realitätsnahen Stimuli an ihrer Institution einzusetzen. 11% der Institutionen (n=9) schätzten demgegenüber die Wahrscheinlichkeit als „eher unwahrscheinlich“/„unwahrscheinlich“ ein.

Diskussion

Die vorliegende Studie hat ein deutliches Auseinanderdriften zwischen Forschung und Praxis im Bereich der sozial-kognitiven Diagnostik aufgezeigt. Obwohl eindeutige Evidenz vorliegt, dass die Testung sozial-kognitiver Fähigkeiten zu einer Verbesserung der Demenzdiagnostik führen kann, findet derzeit keine flächendeckende Integration sozial-kognitiver Untersuchungen an deutschsprachigen MC/GS statt. Die Ergebnisse unserer Umfrage zeigen deutlich, dass die überwiegende Mehrheit der deutschsprachigen MC/GS die Wichtigkeit sozial-kognitiver Untersuchungen in der Demenzdiagnostik erkannt haben und praktikable, valide und normierte Testverfahren in die klinische Routine einführen möchten. Dementsprechend schätzen die Institutionen die von der SCOPE-Studie (Pinkham et al., 2014) vorgeschlagenen fünf sozial-kognitiven Hauptdomänen EE, ToM, soziale Wahrnehmung, Attributionsstil und gezielter Umgang mit wahrgenommenen Emotionen mehrheitlich als relevant/eher relevant ein. Diese Beurteilung deutet darauf hin, dass Veränderungen dieser fünf sozial-kognitiven Fähigkeiten eine wichtige Rolle bei neurokognitiven Demenzerkrankungen spielen und im Diagnoseprozess Beachtung finden sollten. Derzeit konzentriert sich die Demenz-Forschung jedoch fast ausschließlich auf Untersuchungen der EE- und ToM-Fähigkeiten. Vorläufige Evidenz (Johnen & Bertoux, 2019; Panchal et al., 2016; Possin et al., 2013) zeigt dennoch, dass auch Tests zur sozialen Wahrnehmung (z.B. Kennen und verstehen sozialer Normen und Regeln) in Zukunft eine größere Rolle in der Demenz-Differenzialdiagnostik spielen könnten. Die überwiegende Mehrheit der MC/GS (ca. 85%) führte aktuell keine EE- oder ToM-Tests durch. Ungefähr die Hälfte dieser Institutionen gab sogar an, eine Durchführung bis dato nicht in Erwägung gezogen zu haben. Die übrigen begründeten ihre Entscheidung gegen den Einsatz von EE- bzw. ToM-Tests hauptsächlich mit den folgenden drei Kritikpunkten: (1) Zu lange Testdauer, (2) Vorhandene Tests sind für Patient:innen mit Störungen in mehreren kognitiven Domänen ungeeignet und (3) Niedrige Sensitivität der vorhandenen Tests. Diese Befunde legen nahe, dass die sozial-kognitive Diagnostik in den institutsinternen Diagnosealgorithmen zu

wenig Beachtung findet. Verstärkter Informationsaustausch zwischen den deutschsprachigen MC/GS und Fortbildungen könnten die praktische Relevanz der sozial-kognitiven Diagnostik in den Fokus rücken und Forschungsprojekte zur Neu- oder Weiterentwicklung geeigneter Testverfahren anstoßen. Schon jetzt existieren vielversprechende Verfahren zur Untersuchung sozial-kognitiver Fähigkeiten bei Patient:innen mit neurodegenerativen Erkrankungen wie z.B. das *Mini-Social Cognition and Emotional Assessment* (Mini-SEA, Bertoux et al., 2012) und die *Congruent and Incongruent Facial Emotion Intensity Rating Task* (Chiu et al., 2018), die mit einer angemessenen Testdauer und Sensitivität auch bei Patient:innen mit leichten Störungen in mehreren kognitiven Domänen durchgeführt werden können. Bevor die genannten Testverfahren jedoch im deutschsprachigen Raum in die Diagnosealgorithmen integriert werden können, bedarf es weiterer Entwicklungsschritte, wie z.B. Anpassungen an die deutsche Sprache (Mini-SEA, Bertoux et al., 2012) oder Normierung (*Congruent and Incongruent Facial Emotion Intensity Rating Task,* Chiu et al., 2018). Nur wenige Institutionen (ca. 15%) führten EE- bzw. ToM-Tests durch. Von diesen gab ein relevanter Anteil an mit den verwendeten Tests eher unzufrieden/unzufrieden zu sein. Diese Einschätzung wird zusätzlich durch den Befund unterstützt, dass sowohl für die verwendeten EE- als auch für die ToM-Tests mehr Nachteile als Vorteile aufgelistet wurden. Zusammengefasst wurden vier Hauptkritikpunkte genannt: (1) Fehlende Normierung, (2) Fehlende Validierung, (3) Zu lange Testdauer und (4) Realitätsferne Stimuli. Diese Rückmeldungen zeigen auf, dass die aktuell in der klinischen Routine eingesetzten sozial-kognitiven Tests nicht die geforderten Qualitätsstandards neuropsychologischer Tests erfüllen. Daher greifen mehr als die Hälfte der MC/GS auf subjektive Erfassungsmethoden wie Patient:innen- und Angehörigenbefragungen zurück und die Wahrscheinlichkeit einen praktikablen, validierten und reliablen Test zukünftig einzusetzen, wird von den MC/GS mehrheitlich als „eher wahrscheinlich“/„wahrscheinlich“ beurteilt.

Die entstandene Lücke in der neurokognitiven Diagnostik der deutschsprachigen MC/GS kann nur geschlossen werden, indem zukünftige Testentwicklungen nicht nur die Anforderungen eines Forschungskontexts, sondern auch die der klinischen Praxis berücksichtigen. Basierend auf unseren Umfrageergebnissen können die folgenden Empfehlungen zur Neuentwicklung oder Adaption sozial-kognitiver Tests für die klinische Routine formuliert werden:

Praktikable Testdauer

Der Zeitaufwand für ein Screeningtool sozial-kognitiver Fähigkeiten sollte maximal 10 Minuten betragen. Eine ausführliche Diagnostik sollte möglichst eine Dauer von 20 Minuten nicht überschreiten.

Realistische Teststimuli

Die verwendeten Teststimuli sollten möglichst realitätsnah sein, damit das Antwortverhalten der Patient:innen prädiktiv für die spontane Verarbeitung sozialer Reize in realen sozialen Interaktionen ist. Es ist wahrscheinlich, dass realistische dynamische Teststimuli hierfür besser geeignet sind als statische Stimuli (Westerhof-Evers et al., 2014), da dynamische Stimuli eine erhöhte Aktivität in Hirnregionen auslösen können, die mit Gesichtserkennung, EE und emotionaler Verarbeitung assoziiert sind (Arsalidou et al., 2011; Risko et al., 2012). Wir empfehlen daher die Verwendung von dynamischen Stimuli (z.B. Filmsequenzen, Virtual Reality Applikationen), die Bereiche der sozialen Kognition mittels realistisch wirkender sozial-relevanter Situationen untersuchen.

Antwortformat mit Mehrfachauswahl

Das Antwortverhalten der Patient:innen kann nur dann ein realistisches Abbild der Verarbeitung sozialer Reize sein, wenn das Antwortformat des Tests dies zulässt. Da in einer realistischen Situation mehrere Emotionen, Intentionen und Gedanken gleichzeitig wahrgenommen werden können (Jarsch, Bartsch, et al., 2021; Jarsch, Piguet, et al., 2021), sollte das Antwortformat entweder eine Kombination aus geschlossenen Antwortoptionen und Freitextfeld oder geschlossene Antworten mit Mehrfachauswahl darstellen.

Hohe Sensitivität

Wir empfehlen Tests mit einer möglichst hohen Sensitivität zu entwickeln, die leichte Defizite und Veränderungen sozial-kognitiver Fähigkeiten detektieren können. Sensitive Tests könnten somit einen prognostischen Nutzen sowohl in den frühen als auch in späteren Demenzstadien haben (z.B. Dodich et al., 2021; Johnen & Bertoux, 2019).

Assistierte Testdurchführung

Im Wissen um das meist höhere Lebensalter der zu untersuchenden Patient:innen und dem Vorliegen kognitiver und teils körperlicher Beeinträchtigungen, sollte die Testdurchführung assistiert erfolgen, z.B. in Form von Unterstützung bei der Navigation durch computerbasierten Applikationen oder mittels Vorlesen der Aufgabenstellungen, Fragen und Antwortoptionen.

Normierung und Validierung

Sozial-kognitive Tests sollten unter Berücksichtigung der Einflüsse von demografischen Variablen wie Alter, Geschlecht und Ausbildung für die deutschsprachige Population normiert werden, um die individuelle Leistung der Patient:innen angemessen interpretieren zu können. Wir empfehlen zudem, die vorhandene Evidenz über demenzspezifische sozial-kognitive Störungsmuster bei der Testentwicklung zu berücksichtigen, um möglichst gute psychometrische Eigenschaften zu erzielen. Anschließend sollte der entwickelte Test für das entsprechende demenzielle Syndrom (z.B. bvFTD) validiert werden.

Die genannten Empfehlungen zur Testentwicklung werden wir bei der Finalisierung unserer deutschsprachigen Adaption des TASIT (Jarsch, Bartsch, et al., 2021; Jarsch, Piguet, et al., 2021) bestmöglich berücksichtigen, um den MC/GS des deutschsprachigen Raumes einen praktikablen, normierten, validen, reliablen und sensitiven Test zur realitätsnahen Untersuchung sozial-kognitiver Fähigkeiten zur Verfügung stellen zu können.

Wir bewerten sowohl die Gesamt-Rücklaufquote unseres Fragebogens, als auch die länderspezifischen Rücklaufquoten für die Länder Schweiz und Deutschland als zufriedenstellend. In Österreich fiel die Rücklaufquote jedoch gering aus, wofür nur schwer Gründe benannt werden können. Die Ergebnisse unserer Fragebogenstudie liefern somit einen umfassenden Einblick in den Status quo der sozial-kognitiven Untersuchungen in der Schweiz und in Deutschland. Aussagen über die sozial-kognitive Diagnostik der österreichischen MC/GS können jedoch nur mit Einschränkung getroffen werden.

Fazit

Die vorliegende Arbeit zeigt anhand aktueller Forschungsergebnisse und der Wahrnehmung der MC/GS die Relevanz sozial-kognitiver Untersuchungen im Rahmen der neuropsychologischen Demenzdiagnostik auf. Die Antworten auf unseren Expert:innen-Fragebogen machten jedoch deutlich, dass an deutschsprachigen MC/GS derzeit keine bzw. nur wenig zufriedenstellende sozial-kognitive Untersuchungen durchgeführt werden. Es besteht die Notwendigkeit normierte, validierte, reliable und sensitive Tests zu entwickeln, die den Anforderungen der klinischen Praxis standhalten. Nur so können die theoretischen Errungenschaften der sozial-kognitiven Demenzforschung erfolgreich in die Klinik transferiert und der Diagnoseprozess verbessert werden.

Interessenskonflikte und finanzielle Unterstützung

Die Autor:innen geben an, dass keine Interessenskonflikte bestehen. MJ 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 and Axovant Sciences finanziell unterstützt.

Literaturverzeichnis

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*. American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

Arioli, M., Crespi, C., & Canessa, N. (2018). Social Cognition through the Lens of Cognitive and Clinical Neuroscience. *BioMed Research International*, *2018*, 1–18. https://doi.org/10.1155/2018/4283427

Arsalidou, M., Morris, D., & Taylor, M. J. (2011). Converging Evidence for the Advantage of Dynamic Facial Expressions. *Brain Topography*, *24*(2), 149–163. https://doi.org/10.1007/s10548-011-0171-4

Bertoux, M., Delavest, M., de Souza, L. C., Funkiewiez, A., Lépine, J.-P., Fossati, P., Dubois, B., & Sarazin, M. (2012). Social Cognition and Emotional Assessment differentiates frontotemporal dementia from depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, *83*(4), 411–416. https://doi.org/10.1136/jnnp-2011-301849

Boccardi, M., Monsch, A. U., Ferrari, C., Altomare, D., Berres, M., Bos, I., Buchmann, A., Cerami, C., Didic, M., Festari, C., Nicolosi, V., Sacco, L., Aerts, L., Albanese, E., Annoni, J., Ballhausen, N., Chicherio, C., Démonet, J., Descloux, V., … Kliegel, M. (2021). Harmonizing neuropsychological assessment for mild neurocognitive disorders in Europe. *Alzheimer's & Dementia*, alz.12365. https://doi.org/10.1002/alz.12365

Brüne, M., & Juckel, G. (2010). Soziale Kognition bei Schizophrenien Mentalisierung und psychosoziales Funktionsniveau. *Nervenarzt*, *81*, 339–346. https://doi.org/10.1007/s00115-009-2841-y

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. https://doi.org/10.1080/13607863.2010.513038

Chiu, I., Piguet, O., Diehl-Schmid, J., Riedl, L., Beck, J., Leyhe, T., Holsboer-Trachsler, E., Kressig, R. W., Berres, M., Monsch, A. U., & Sollberger, M. (2018). Facial Emotion Recognition Performance Differentiates Between Behavioral Variant Frontotemporal Dementia and Major Depressive Disorder. *The Journal of Clinical Psychiatry*, *79*(1), 16m11342. https://doi.org/10.4088/JCP.16m11342

Christidi, F., Migliaccio, R., Santamaría-García, 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., 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

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

Diehl, J., Monsch, A. U., Marksteiner, J., Romero, B., Wolter, D., Calabrese, P., Mau-Erer, C., Hampel, H., Staehelin, H. B., & Kurz, A. (2005). *Diagnostik und Therapie dementieller Erkrankungen-Erfahrungen aus deutschsprachigen Memory-Kliniken*. www.memoryclinic.ch

Dodich, A., Crespi, C., Santi, G. C., Cappa, S. F., & Cerami, C. (2020). Evaluation of Discriminative Detection Abilities of Social Cognition Measures for the Diagnosis of the Behavioral Variant of Frontotemporal Dementia: a Systematic Review. *Neuropsychology Review*, 1–16. https://doi.org/10.1007/s11065-020-09457-1

Dodich, A., Crespi, C., Santi, G. C., Marcone, A., Iannaccone, S., Perani, D., Cappa, S. F., & Cerami, C. (2021). Emotion Recognition Deficits in the Differential Diagnosis of Amnestic Mild Cognitive Impairment: A Cognitive Marker for the Limbic-Predominant Phenotype. *Journal of the International Neuropsychological Society*, 1–7.

https://doi.org/10.1017/S1355617721000254

Ducharme, S., Dols, A., Laforce, R., Devenney, E., Kumfor, F., van den Stock, J., Dallaire-Théroux, C., Seelaar, H., Gossink, F., Vijverberg, E., Huey, E., Vandenbulcke, M., Masellis, M., Trieu, C., Onyike, C., Caramelli, P., de Souza, L. C., Santillo, A., Waldö, M. L., … Pijnenburg, Y. (2020). Recommendations to distinguish behavioural variant frontotemporal dementia from psychiatric disorders. *Brain*, *143*(6), 1632–1650. https://doi.org/10.1093/brain/awaa018

Ehrensperger, M. M., Taylor, K. I., Berres, M., Foldi, N. S., Dellenbach, M., Bopp, I., Gold, G., Von Gunten, A., Inglin, D., Müri, R., Rüegger, B., Kressig, R. W., & Monsch, A. U. (2014). *BrainCheck - a very brief tool to detect incipient cognitive decline: optimized case-finding combining patient- and informant-based data*. 1–12. https://doi.org/10.1186/s13195-014-0069-y

Elamin, M., Pender, N., Hardiman, O., & Abrahams, S. (2012). Social cognition in neurodegenerative disorders: a systematic review. *Journal of Neurology, Neurosurgery, and Psychiatry*, *83*(11), 1071–1079. https://doi.org/10.1136/jnnp-2012-302817

Frith, C. D., & Frith, U. (2007). Social Cognition in Humans. *Current Biology*, *17*(16), R724–R732. https://doi.org/10.1016/J.CUB.2007.05.068

Hasson-Ohayon, I., Mashiach-Eizenberg, M., Arnon-Ribenfeld, N., Kravetz, S., & Roe, D. (2017). Neuro-cognition and social cognition elements of social functioning and social quality of life. *Psychiatry Research*, *258*, 538–543. https://doi.org/10.1016/J.PSYCHRES.2017.09.004

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., 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. https://doi.org/10.1038/nrneurol.2015.229

IBM SPSS Statistics for Macintosh (27.0.1.0). (2007). IBM Corp.

Jarsch, M., Bartsch, K., Berres, M., Ryf, I., Jurisic, K., Sluka, C., Kressig, R. W., Monsch, A. U., McDonald, S., Piguet, O., & Sollberger, M. (in press). The Basel Version of the Awareness of Social Inference Test – Emotion Recognition (BASIT-ER): Preliminary Validation Analyses in Healthy Adults. *Neuropsychology*. https://doi.org/10.1037/neu0000786

Jarsch, M., Piguet, O., Sluka, C., Huber, N., Kressig, R. W., Monsch, A. U., Berres, M., & Sollberger, M. (2021). The Basel Version of the Awareness of Social Inference Test – Theory of Mind (BASIT-ToM): Preliminary Validation Analyses in Healthy Subjects. *Neuropsychology*. Manuscript submitted.

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

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

Kessels, R. P. C., Waanders-Oude Elferink, M., & Tilborg, I. (2021). Social cognition and social functioning in patients with amnestic mild cognitive impairment or Alzheimer's dementia. *Journal of Neuropsychology*, *15*(2), 186–203. https://doi.org/10.1111/jnp.12223

Krudop, W. A., Dols, A., Kerssens, C. J., Eikelenboom, P., Prins, N. D., Möller, C., Schouws, S., Rhebergen, D., van Exel, E., van der Flier, W. M., Sikkes, S., Scheltens, P., Stek, M. L., & Pijnenburg, Y. A. L. (2017). The Pitfall of Behavioral Variant Frontotemporal Dementia Mimics Despite Multidisciplinary Application of the FTDC Criteria. *Journal of Alzheimer's Disease : JAD*, *60*(3), 959–975. https://doi.org/10.3233/JAD-170608

Lautenschlager, N. T., & Kurz, A. (2009). Gedächtnissprechstunden (Memory-Kliniken). In

Demenzen in Theorie und Praxis (pp. 311–325). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-35486-4_21

LimeSurvey Project Team. (2021). *LimeSurvey: An Open Source survey tool*. LimeSurvey GmbH. https://www.limesurvey.org

Martyr, A., Nelis, S. M., Quinn, C., Wu, Y.-T., Lamont, R. A., Henderson, C., Clarke, R., Hindle, J. V., Thom, J. M., Jones, I. R., Morris, R. G., Rusted, J. M., Victor, C. R., & Clare, L. (2018). Living well with dementia: a systematic review and correlational meta-analysis of factors associated with quality of life, well-being and life satisfaction in people with dementia. *Psychological Medicine*, *48*(13), 2130–2139. https://doi.org/10.1017/S0033291718000405

McCade, D., Savage, G., Guastella, A., Hickie, I. B., Lewis, S. J. G., & Naismith, S. L. (2013). Emotion Recognition in Mild Cognitive Impairment: Relationship to Psychosocial Disability and Caregiver Burden. *Journal of Geriatric Psychiatry and Neurology*, *26*(3), 165–173. https://doi.org/10.1177/0891988713491832

McCade, D., Savage, G., & Naismith, S. L. (2011). Review of Emotion Recognition in Mild Cognitive Impairment. *Dementia and Geriatric Cognitive Disorders*, *32*(4), 257–266. https://doi.org/10.1159/000335009

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.

Menon, V. (2011). Large-scale brain networks and psychopathology: a unifying triple network model. *Trends in Cognitive Sciences*, *15*(10), 483–506. https://doi.org/10.1016/J.TICS.2011.08.003

Michaelian, J. C., Duffy, S. L., Mowszowski, L., Guastella, A. J., McCade, D., McKinnon, A. C., & Naismith, S. L. (2021). Poorer Theory of Mind in Amnestic Mild Cognitive Impairment Is Associated with Decreased Functional Connectivity in the Default Mode Network. *Journal of Alzheimer's Disease*, *81*(3), 1079–1091. https://doi.org/10.3233/JAD-201284

Monsch, A. U. (2015). Case-finding in der hausärztlichen Praxis und neuropsychologische Diagnostik an einer Memory Clinic. *Therapeutische Umschau*, *72*(4), 213–217. https://doi.org/10.1024/0040-5930/a000667

Nasreddine, Z. S., Phillips, N. A., BÃ©dirian, 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. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Nelis, S. M., Clare, L., Martyr, A., Markova, I., Roth, I., Woods, R. T., Whitaker, C. J., & Morris, R. G. (2011). Awareness of social and emotional functioning in people with early-stage dementia and implications for carers. *Aging & Mental Health*, *15*(8), 961–969. https://doi.org/10.1080/13607863.2011.575350

Panchal, H., Paholpak, P., Lee, G., Carr, A., Barsuglia, J. P., Mather, M., Jimenez, E., & Mendez, M. F. (2016). Neuropsychological and Neuroanatomical Correlates of the Social Norms Questionnaire in Frontotemporal Dementia Versus Alzheimer's Disease. *American Journal of Alzheimer's Disease and Other Dementias*, *31*(4), 326–332. https://doi.org/10.1177/1533317515617722

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*(4), 813–823. https://doi.org/10.1093/schbul/sbt081

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

Possin, K. L., Feigenbaum, D., Rankin, K. P., Smith, G. E., Boxer, A. L., Wood, K., Hanna, S. M., Miller, B. L., & Kramer, J. H. (2013). Dissociable executive functions in behavioral variant frontotemporal and Alzheimer dementias. *Neurology*, *80*(24),

2180–2185. https://doi.org/10.1212/WNL.0b013e318296e940

Risko, E. F., Laidlaw, K., Freeth, M., Foulsham, T., & Kingstone, A. (2012). Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. *Frontiers in Human Neuroscience*, *6*, 143. https://doi.org/10.3389/fnhum.2012.00143

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*(1), 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., Reiss, A. L., & Greicius, M. D. (2007). *Behavioral/Systems/Cognitive Dissociable Intrinsic Connectivity Networks for Salience Processing and Executive Control*. https://doi.org/10.1523/JNEUROSCI.5587-06.2007

Shimokawa, A., Yatomi, N., Anamizu, S., Torii, S., Isono, H., Sugai, Y., & Kohno, M. (2001). Influence of Deteriorating Ability of Emotional Comprehension on Interpersonal Behavior in Alzheimer-Type Dementia. *Brain and Cognition*, *47*(3), 423–433. https://doi.org/10.1006/BRCG.2001.1318

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

Swiss Memory Clinics. (2021). https://www.swissmemoryclinics.ch/

Timmer, M. L., Jacobs, B., Schonherr, M. C., Spikman, J. M., & van der Naalt, J. (2020). The Spectrum of Long-Term Behavioral Disturbances and Provided Care After Traumatic Brain Injury. *Frontiers in Neurology*, *11*, 246. https://doi.org/10.3389/fneur.2020.00246

Toller, G., Brown, J., Sollberger, M., Shdo, S. M., Bouvet, L., Sukhanov, P., Seeley, W. W., Miller, B. L., & 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

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. https://doi.org/10.1080/13803395.2014.966661

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*. https://www.psychiatrist.com/jcp/depression/diagnostic-challenge-psychiatric-symptoms-neurodegenerative/

Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fisch, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*(3), 1125–1165. https://doi.org/10.1152/jn.00338.2011

ZhongQuan, Y., PanWen, Z., Hui, Z., YuanYuan, S., HaiChun, S., JianGuo, Z., & PingLei, P. (2020). Theory of mind in Alzheimer's disease and amnestic mild cognitive impairment: a meta-analysis. *Neurological Sciences*, *41*(5), 1027–1039. https://doi.org/10.1007/s10072-019-04215-5

Teil 2

4 Das Schliessen der Lücke Zwischen Theorie und Praxis – Die Entwicklung eines Tests zur Untersuchung der EE- und ToM-Fähigkeiten, welcher den Anforderungen der MC/GS Entsprechen soll

Die gewonnen Erkenntnisse aus Manuskript 1, verdeutlichen, dass die Notwendigkeit besteht, ein objektives sozial-kognitives Testverfahren zu entwickeln, das idealerweise folgende Anforderungen erfüllen sollte[1]:

(1) Untersuchung der EE-Fähigkeit von Basis- und komplexen Emotionen
(2) Untersuchung verschiedener Konzepte der ToM-Fähigkeiten (affektive und kognitive ToM in jeweils erster und zweiter Ordnung, Qualität der ToM-Leistungen)
(3) Hohe Sensitivität zur Detektion subtiler Defizite in den EE- und ToM-Fähigkeiten
(4) Testleistung soll prädiktiv für die EE- und ToM-Leistungen in realen sozialen Interaktionen sein
(5) Durchführbarkeit mit Personen, die Beeinträchtigungen in mehreren kognitiven Domänen aufweisen
(6) Praktikable Testdauer (ca. 20 Minuten)
(7) Validierung für spezifische Störungsbilder (z. B. bvFTD, AD)
(8) Hohe Reliabilität
(9) Normierung für die deutschsprachige Population unter Kontrolle der Einflüsse demografischer Variablen wie Alter, Geschlecht und Ausbildung

[1] Die Reihenfolge der Anforderungen ist frei gewählt und entspricht keiner Hierarchie.

Im Folgenden wird die Entwicklung und präliminäre Validierung des sozial-kognitiven Tests BASIT ausführlich beschrieben, der die neun Anforderungen der MC/GS erfüllen soll. In die aktuelle Entwicklungsphase wurden die Anforderungen (1) – (6) einbezogen. Die Anforderungen (7) – (9) werden erst in den nachfolgenden Validierungsstudien berücksichtigt werden. Der BASIT besteht aus einem EE-Subtest (BASIT-Emotion Recognition) und einem ToM-Subtest (BASIT-Theory of Mind), welche die sozial-kognitiven Fähigkeiten der Testteilnehmer mittels kurzer realistischer Filmszenen mit sozial-relevantem Inhalt untersuchen. Als Basis der Testentwicklung wurde der TASIT (McDonald et al., 2003) verwendet, der im englischsprachigen Raum ein etabliertes und validiertes Testverfahren mit guten psychometrischen Eigenschaften zur Untersuchung der EE- und ToM-Fähigkeiten darstellt. Der TASIT weist neben Stärken, wie z. B. seinem dynamischen realistischn Stimulusmaterial mit geringer Komplexität oder seiner Fähigkeit zwischen verschiedenen neurodegenerativen Erkrankungen diskriminieren zu können (z.B Kipps et al., 2009; Kumfor et al., 2017), auch einige Schwächen auf [z. B. eine lange Durchführungsdauer und Deckeneffekte bei gesunden Probanden (z. B. Henry et al., 2014; Westerhof-Evers et al., 2014)]. Durch eine aufwendige und sorgfältige Testentwicklung soll der BASIT auf den positiven Testeigenschaften des TASIT aufbauen und seine Limitationen überwinden.

4.1 Manuskript 2

The Basel Version of the Awareness of Social Inference Test – Emotion Recognition (BASIT-ER): Preliminary Validation Analyses in Healthy Adults

Manuskript 2 beschreibt detailliert die Entwicklung und präliminäre Validierung des deutschsprachigen EE - Subtest des BASIT (BASIT-ER) an 240 gesunden Erwachsenen.

The Basel Version of the Awareness of Social Inference Test – Emotion Recognition (BASIT-ER): Preliminary Validation Analyses in Healthy Adults

Marianne Jarsch[1,2], Kaja Bartsch[1,2], Manfred Berres[3], Isabelle Ryf[1,2], Kristina Jurisic[1,2], Constantin Sluka[4], Reto W. Kressig[1], Andreas U. Monsch[1,2], Skye McDonald[5], Olivier Piguet[6], and Marc Sollberger[1,7]

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

[2]Faculty of Psychology, University of Basel, Switzerland;

[3]Faculty of Mathematics and Technology, University of Applied Sciences Koblenz, Germany

[4]Department of Clinical Research, University of Basel and University Hospital Basel, Switzerland

[5]University of New South Wales Psychology, Sydney, Australia

[6]School of Psychology and Brain and Mind Centre, The University of Sydney, Australia

[7]Department of Neurology, University Hospital Basel, Switzerland.

Author Note

Marianne Jarsch https://orcid.org/0000-0003-1871-8435

Kaja Bartsch https://orcid.org/0000-0002-7785-9417

Manfred Berres https://orcid.org/0000-0003-0409-8554

Isabelle Ryf https://orcid.org/0000-0003-1796-5493

Kristina Jurisic https://orcid.org/0000-0002-6804-0610

Constantin Sluka https://orcid.org/0000-0001-7672-4986

Reto W. Kressig https://orcid.org/0000-0002-3215-7681

Andreas U. Monsch https://orcid.org/0000-0003-1692-1560

Skye McDonald https://orcid.org/0000-0003-0723-6094

Olivier Piguet https://orcid.org/0000-0002-6696-1440

Marc Sollberger https://orcid.org/0000-0001-8863-3731

Two exemplary test scenes of the BASIT-ER at three different emotional intensities can be found at https://doi.org/10.6084/m9.figshare.13507245.

MJ 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 the National Health and Medical Research Council of Australia Senior Research Fellowship grant 1103258. All authors report no conflict of interest.

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 Auchter, Anna-Lisa Krieg, Valentina Arnold, Julia Desax, Nerissa Dettling, Riane Baur, Vera Nina Looser, Anne-Catherine Ruegg, Teresa Müggler, Nathalie Müller, Vaisnavi Varatharajan, Sophie von Achenbach and Andreas Gassilloud for help in participant recruitment and data collection.

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

Abstract

Objective: Emotion recognition (ER) is commonly impaired in many brain disorders. Therefore, its reliable and ecologically valid examination is a crucial part of any comprehensive neuropsychological assessment. In this regard, an established English-language test identifying deficits in ER is “The Awareness of Social Inference Test - Emotion Evaluation Test” (TASIT-EET). However, no comparable German-language ER test currently exists. In this study, we aimed to develop and preliminarily validate the first German-language adaption of TASIT-EET in healthy adults.

Method: We selected 22 scenes [three scenes per emotion (i.e., anger, disgust, fear, happiness, sadness, surprise, and neutral) and one practice scene] out of the 56 TASIT-EET scenes. In collaboration with a film institute and professional actors, we filmed each scene (except neutral) at three different emotional intensities. Then, we administered these intensity version scenes to 240 cognitively and mentally healthy participants, equally distributed in sex and age, ranging from 35 to 92 years.

Results: By applying Rasch analysis, the one intensity version per scene was selected that showed neither floor nor ceiling effects in participants whose ability in ER was in the medium range.

Conclusions: We have conducted a preliminary validation of the first German-language adaption of TASIT-EET, i.e., the “Basel Version of the Awareness of Social Inference Test - Emotion Recognition” (BASIT-ER). The BASIT-ER comprises the strengths of the TASIT-EET, while it overcomes its limitations such as ceiling effects in healthy participants and the simple response format. Next, the BASIT-ER needs to be validated in clinical populations.

Keywords: Social cognition, emotion recognition, healthy subjects, dynamic stimuli, test development

Key Points

Objective

The aim of this study was to develop and provide preliminary validation of the Basel Version of the Awareness of Social Inference Test - Emotion Recognition (BASIT-ER) in a sample of healthy adults.

Findings

We succeeded in generating a dynamic, German-language emotion recognition test, which is neither too easy nor too difficult for healthy adults.

Importance

The preliminary validation of the BASIT-ER is an important step towards the development of a German-language test that allows the assessment of subtle emotion recognition deficits in a realistic way and within a practicable timeframe in clinical settings.

Next Steps

Next, the BASIT-ER needs to be validated in clinical populations with neuropsychological disorders.

The Basel Version of the Awareness of Social Inference Test – Emotion Recognition (BASIT-ER): Preliminary Validation Analyses in Healthy Adults

Social cognition refers to a wide range of cognitive processes, ranging from perception of socio-emotional stimuli to social behavior (Henry et al., 2016). Consequently, social cognitive abilities such as emotion recognition (ER) play a fundamental role in human life. They enable individuals to interact successfully with others in the long term (Kennedy & Adolphs, 2012) and are crucial for their perceived quality of life (e.g., Hasson-Ohayon et al., 2017; Maat et al., 2012; Yogarajah & Mula, 2019). Accordingly, social cognitive abilities need to be assessed in clinical settings. Thus, social cognition has been included as one of the six clinical domains of neurocognitive function in the Diagnostic and Statistical Manual for Mental Disorders (DSM-5; American Psychiatric Association, 2013). According to DSM-5, clinicians need to assess ER, next to cognitive and affective theory of mind, for any comprehensive diagnosis of a neurocognitive disorder.

Indeed, ER is impaired in the majority of neurocognitive disorders either due to traumatic brain injuries (e.g., Babbage et al., 2011) or due to brain diseases such as neurodegenerative (e.g., Bora, 2017; Bora et al., 2016; Torres Mendonça De Melo Fádel et al., 2019), neurodevelopmental (e.g., Fonagy & Goodyer, 2008), vascular (e.g., Adams et al., 2019), neuroinflammatory (e.g., Cotter et al., 2016) or psychiatric disorders (e.g., Loi et al., 2013; Savla et al., 2013). In addition, existing evidence shows that the assessment of ER plays a crucial role in the diagnosis of specific neurocognitive disorders such as behavioral variant frontotemporal dementia (e.g., Dodich et al., 2021, Johnen & Bertoux, 2019). Sensitive and reliable tests with high construct validity available for clinical populations that assess ER within a practicable timeframe are therefore essential. Ideally, individual performance in an ER test predicts the ability to recognize emotions in real life. In this regard, dynamic, realistic and socially relevant stimuli (e.g. film scenes) meet this demand more than static stimuli do (Westerhof-Evers, 2014), possibly as they elicit higher activity in brain regions associated with face recognition, ER and emotional processing

than static stimuli (e.g., Arsalidou et al., 2011, Risko et al., 2012). Regarding validated tests for German-speaking patients that contain dynamic, multimodal emotional stimuli, only the Geneva Emotion Recognition Test (Schlegel et al., 2014) is available, to the best of our knowledge. In that test, actors, who express meaningless syllables, portray emotions in short video clips by looking directly into the camera. These actors, however, do not act in a scene (e.g., answering a phone call) or interact with another person as in a real-world situation. Therefore, a German-language test that examines ER based on dynamic, multimodal stimuli in social interactions is needed.

We decided to develop a German-language adaption of the validated and well-known English-language, dynamic Emotion Evaluation Test (EET) of The Awareness of Social Inference Test (TASIT) (McDonald et al., 2003). TASIT-EET comprises two sets of short realistic scenes in which actors portray six basic emotions (i.e., anger, disgust, fear, happiness, sadness, and surprise) or a neutral emotional state in real-world situations with little cognitive complexity. It was designed as a criterion-based test that can identify clear deficits in ER in adults. Indeed, several studies have shown that TASIT-EET is able to discriminate between patients with different brain disorders including neurodegenerative diseases (i.e., behavioral and language variants of frontotemporal dementia (Kipps et al., 2009, Kumfor et al., 2014, Yang et al., 2021), progressive supranuclear palsy (Yang et al., 2021), Alzheimer's disease dementia (Kumfor et al., 2014, Yang et al., 2021)), traumatic brain injuries (McDonald et al., 2003, 2006, 2018) and schizophrenia (Sparks et al., 2010), and healthy controls. It has adequate reliability measures and evidence for construct and ecological validity has been shown (for review McDonald, 2012). However, there are also some limitations. First, because of its duration, the test does not lead itself to be easily used in clinical settings (Henry et al., 2015; Westerhof-Evers et al., 2014). Second, the protagonists in the videos portray the emotional stimuli in high intensity, resulting in ceiling effects in healthy adults (McDonald et al., 2006), which prevents the detection of subtle ER deficits. Moreover, one might speculate that emotional stimuli of high intensity lead to high arousal (e.g., Deckert et al., 2020), which, in turn, might negatively influence ER.

Third, TASIT-EET scenes are inconsistent in scene design and camera perspective and appear somewhat outdated. Indeed, the emotional portrayals at times do not seem authentic (e.g., close-ups of the actors talking to the camera or portrayal of emotions in high intensity). Moreover, actors were cast in varying role relationships to each other. For example, in one scene, a female actor plays a secretary with the male actor playing her supervisor, whereas in another scene, the same actors are portrayed as spouses. It is possible that these aspects may have an impact on the task performance. Additionally, actors were not assigned to the different emotions equally, which may lead participants to associate an actor with a particular emotion. Finally, TASIT-EET is a simple, forced-choice labelling task that prevents participants being able to provide information about other potentially identified emotions (e.g., mixed (Plutchik, 1962, 1980) or secondary emotions (Shaver et al., 2001) or the perceived intensity of the emotions portrayed.

Thus, the current study aimed to overcome these limitations by creating an adapted version of TASIT-EET for German speakers, followed by a preliminary validation in healthy participants. We named the adapted German-language version the “Basel Version of the Awareness of Social Inference Test (BASIT)-ER”.

Methods

The Awareness of Social Inference Test – Emotion Evaluation Test (TASIT-EET) (McDonald et al., 2003)

TASIT-EET examines the ability to recognize six basic emotions (i.e., anger, disgust, fear, happiness, sadness and surprise) and neutral emotional states in social situations (i.e., dialogue interactions or non-dialogue interactions such as phone calls or direct talks to the camera), based on linguistic and paralinguistic cues (e.g., gestures, facial expressions, prosody). Two parallel test forms exist, form A and form B. Each form contains 28 spontaneous or scripted scenes [four scenes per emotion (4x7)] of 15 to 60 seconds duration in quasi-randomized order. Using a forced choice paradigm, participants are asked to decide which emotion is portrayed in the scene.

TASIT-EET was developed to clearly differentiate between cognitively healthy people with “average social skills” and individuals with clear deficits in ER (McDonald et al., 2003). Accordingly, actors were asked to portray emotional intensities at such high levels that healthy people should be able to recognize the portrayed emotions (McDonald et al., 2003).

Development of the BASIT-ER Scenes

Selection of Scenes from the TASIT-EET

We decided to develop one, rather than two test forms, given the negligible practice effects on TASIT-EET scores (McDonald et al., 2006). First, we selected 22 scenes (3x7 emotions for test scenes and one practice scene) out of the 56 TASIT-EET scenes (28 form A scenes and 28 form B scenes) (McDonald et al., 2003) based on eight evaluation criteria that focused on different aspects of cinematographic quality, target group fit and cinematic feasibility (see Supplementary Material A). Three raters (MJ, IR and MS) rated first each scene independently and finally selected the 22 scenes in consensus (see Supplementary Material B1). The selected TASIT-EET scenes are depicted in the Supplementary Material B2.

Conceptualization of the BASIT-ER 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-EET as follows:

Screenplay. First, the transcription of TASIT-EET scenes (Westerhof-Evers et al., 2014) was adapted to the German language without change of content. However, we reduced sarcastic content from scenes B23 and B28 and reduced emotion incongruent content from scene A9 to avoid incorrect answers either due to cognitive deficits other than ER (when thinking of later use of the test in clinical populations) or due to potential distractors that could unsystematically increase the difficulty of ER. Second, in contrast to TASIT-EET, we omitted naming the actors to exclude any potential bias associated with naming. Accordingly, we cast all dialogue-scenes with one woman and one man. BASIT-ER

screenplays and description of the main changes in the text are presented in the Supplementary Material C.

Cast and Acting of Actors. We cast the scenes with four female and four male, middle-aged, professional German-speaking actors. Each actor portrayed a type of emotion only once to avoid any actor-specific associations to an emotion. Apart from two pairs of actors who played two scenes together, pairs of actors only played one scene together. Equality of actors' sex across the different emotions was not possible as we shot three scenes per emotion.

To make the dialogue-scenes as realistic as possible, not only the main actors portraying the emotion, but also their conversation partners acted in a realistic way. This type of acting differs from that in TASIT-EET, in which conversation partners all acted in a neutral way. In real life, however, one would not expect another person to react in a neutral way to someone who expresses strong emotions such as anger or fear. A conversation partner's neutral behavior may even confuse a viewer with preserved ER capacity. The distribution of the actors to each of the 22 scenes is presented in the Supplementary Material D.

Realism of Scenes and Camera Work. We placed emphasis on a high degree of realism in the film scenes through realistic set designs, representing either 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 Emotions. Apart from the neutral emotion scenes, actors were asked to portray emotions at three 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 emotions. After the shooting, MJ, IR, MS and five master's degree psychology students selected the best scenes for each intensity by consensus. Two examples of these scenes (Fear#1 and Sadness#1), each at low, medium and high level of intensity, can be found at https://doi.org/10.6084/m9.figshare.13507245

(Jarsch et al., 2021). Background information of each BASIT-ER scene and main differences between TASIT-EET and BASIT-ER scenes are described in the Supplementary Material E.

Selection of the BASIT-ER Intensity Version Scenes for Use in Clinical Populations

Participants. Next, we administered the BASIT-ER intensity version scenes to 240 cognitively and mentally healthy Central European participants (50% women) ranging from 35 to 92 years of age to select scenes for use in clinical populations, i.e., 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 both, age within the sample and sex by age. Each group consisted of 48 participants (50% women; at maximum three participants were at the same age). Inclusion criteria were age equal or greater than 35 years, a minimum total education of seven years, German and/or Swiss German as mother tongue 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., Beck Depression Inventory (Beck et al., 1961) score ≥ ten points, if aged <65 years, or Geriatric Depression Scale (Yesavage & Sheikh, 1986) score ≥ five points, if aged ≥ 65 years), cognitive deficits [i.e., Montreal Cognitive Assessment (Nasreddine 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 anesthesia within the last three months and severe sensory and motor deficits. Participants were recruited via the Basel Memory Clinic's Registry of Healthy Individuals Interested to Participate in Research, announcements, flyers and local social network groups. The study was approved by the Ethics Committee of Northwestern and Central Switzerland (EKNZ) and all participants provided written informed consent.

Application of the BASIT-ER Scenes. BASIT-ER 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 Supplementary Material F.

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 21 test scenes were then administered in a pseudo-randomized order (not the same emotion twice in a row). Each basic emotion was shown in three scenes; in each scene at a different emotional intensity. Accordingly, each participant watched each basic emotion once at low, once at medium and once at high intensity with different scenes representing each intensity. For example, participant#1 watched scene Anger#1 at low intensity, scene Anger#2 at medium intensity and scene Anger#3 at high intensity, whereas participant#2 watched scene Anger#1 at medium intensity, scene Anger#2 at high intensity and scene Anger#3 at low intensity. Consequently, each intensity version (i.e., three intensities per scene) was watched by 80 participants.

Following each scene, participants were asked to state which emotion(s) they had identified. They could click on one or several buttons, each button labelled with the name of one of seven (i.e., six basic emotions and neutral) emotions, and they could also click on a button called “other emotion” (see Supplementary Material F). This paradigm differs from TASIT-EET, in which participants are asked to choose only one of the six basic emotions or neutral (McDonald et al., 2003). Our paradigm, comprising a structured and unstructured (i.e., free text) format, allowed for the examination of the range of perceived emotions in each scene. Even though, the actors were asked to portray a specific basic emotion in a scene, we wanted to know in detail, how participants perceived each scene to judge the quality of the scenes in terms of emotion portrayal. In other words, this type of

response format allowed us to check whether the portrayal of the intended basic emotion had worked and to assess the range of emotions that each scene represents. After a participant had pressed the button for “other emotion”, he/she said the “other emotion” term aloud. Then, the examiner decided according to an algorithm (see Supplementary Material G for more information) whether to classify the term as an emotion or not. If the examiner classified the term as an emotion, he/she typed the name in a free text field on a second screen. When the examiner classified the term as no emotion, he/she notified the participant and did not enter the term.

If the participant had labelled or named only one emotion, he/she was asked on the next screen to rate the intensity of this emotion by clicking on one of the seven circles, increasing in size from left to right, ranging from “no *emotion*” (smallest circle on the left) to “maximum *emotion*” (largest circle on the right) with equal distances between circumferences of adjacent circles (see Supplementary Material F). If the participant had labelled and/or named more than one emotion, he/she was asked to state whether one of these emotions was the strongest. If he/she answered with ”Yes”, he/she was asked to select the strongest emotion and to rate its intensity. If he/she answered with “No”, no rating took place, and the next scene started.

There was no time limit for answering the questions. In contrast to the paradigm used in TASIT-EET, participants were able to watch 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’ answer, about potential difficulties in understanding the respective scene’s content. Difficulties in assessing a scene may not be exclusively due to difficulty in ER, but potentially also due to difficulties in other cognitive domains such as psychomotor speed or working memory. This type of information is valuable given that the aim is to administer the BASIT-ER to individuals with cognitive deficits. Similarly, participants could change their answers within a question block.

Scoring of the Participants' Answers. We defined a participant's answer as correct if the identified emotion(s) was/were categorized within the spectrum of the portrayed emotion.

In terms of "other emotion", we needed first to define whether an "other emotion" term could be assigned to the spectrum of the portrayed emotion. We proceeded as follows: The "other emotion" term was entered into the Geneva Affect Label Coder (GALC; Scherer, 2005). If the emotion term could not be assigned to the spectrum of the portrayed emotion by the GALC, we searched influential emotion theories (Plutchik, 1962, 1980, 2003; Shaver et al., 2001; Zinck & Newen, 2008). If the search turned out negative, we searched the German-language dictionary "Duden" (https://www.duden.de) for a synonym of the named "other emotion" term. If we found a synonym, we checked whether it fitted to the spectrum of the portrayed emotion by searching again the GALC and, if necessary, the emotion theories. The algorithm in identifying 'other emotion' is described in the Supplementary Material H.

We classified a participant's answer as correct if one of the three following criteria was fulfilled:

1. The participant only labelled the portrayed emotion (e.g., fear) or rated the portrayed emotion as the strongest out of several labelled and/or named emotions.
2. The participant named only one "other emotion" term that belonged to the spectrum of the portrayed emotion (e.g., panic, which belongs to the spectrum of fear) or rated an "other emotion" term that belonged to the spectrum of the portrayed emotion as the strongest out of several labelled and/or named emotions.
3. The participant labelled and/or named several emotions without rating one emotion as the strongest, but more than half of these emotions could be assigned to the spectrum of the portrayed emotion. For example, the portrayed emotion is fear and the participant names and labels the following

three emotions: anger, fear, and panic. The answer would be classified as correct as two (fear and panic) of the three emotions are within the spectrum of the portrayed emotion.

Detailed description of the GALC (Scherer, 2005) and the algorithm for emotion ascertainment is provided in the Supplementary Material I.

Data Analysis. The aim of the data analysis of the 57 intensity version scenes [6(basic emotions) x 3(scenes) x 3(intensities) = 54 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 ability in ER was in the medium range. Regarding the neutral scenes, we aimed, if possible (no ceiling or floor effects), to include all three scenes to achieve a balance in the number of selected scenes per emotion (i.e., three scenes per emotion). To achieve this type of scene selection, we took the following three-steps approach:

Step 1. We conducted Rasch analysis (Rasch, 1960) by applying the R ltm package, version 1.1-1 (Rizopoulos, 2006) with RStudio (R Studio Team, 2016) in R version 3.5.1 (R Core Team, 2020) and the package ggplot2*,* version 3.3.0 (Wickham et al., 2020) to analyze the relations between the estimated participants' ability parameters and the probabilities of correct answering the respective intensity version. In each of the six basic emotion analyses, we included the respective nine intensity versions (3 scenes per emotion x 3 intensities). In the practice scene analysis, we included the respective three intensity versions and in the neutral emotion analysis, we included the three neutral scenes. Within each of these analyses, separate analyses were carried out for each scene. In order to check the model fit of the data, we conducted the ltm package parametric Bootstrap goodness-of-fit test using Pearson's $\chi 2$, based on 201 data sets (the original data set plus 200 simulated datasets).

Step 2. Based on the graphical assessment of the item characteristic curve (ICC) plots, we aimed to choose one intensity per scene that predicted a correct answering probability between 0.5 and 0.8 in participants with a medium ability around 0 [-0.5,0.5] in ER.

Step 3. Next, we examined the 95% confidence intervals (CI) of the estimated difficulties of the selected intensity version scenes in relation to the required difficulty interval. 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 ER 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.

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-ER stimulus material. The original data sets, the R syntax and the BASIT-ER scenes can be provided by the corresponding author on request. The study's design and its analysis were not pre-registered.

Results

Selection of Scenes

Step 1

The parametric Bootstrap goodness-of-fit test showed an acceptable fit of the Rasch model for the practice scene intensity versions and all intensity versions of the six basic emotions by showing a non-significant *p*-value ($p>.05$), but not for the three neutral scenes. ***Step 2***

We succeeded in selecting one intensity version of the practice scene and one intensity version of each basic emotion scene that predicted correct response probabilities between 0.5 and 0.8 in participants with ER ability around 0 apart from Anger#1_high (> 0.8) (see Figure 1A-G). We could not determine the relations between the estimated participants' ability parameters and the probabilities of correctly answering the Neutral scenes because of the missing model fit. Importantly, however, most participants

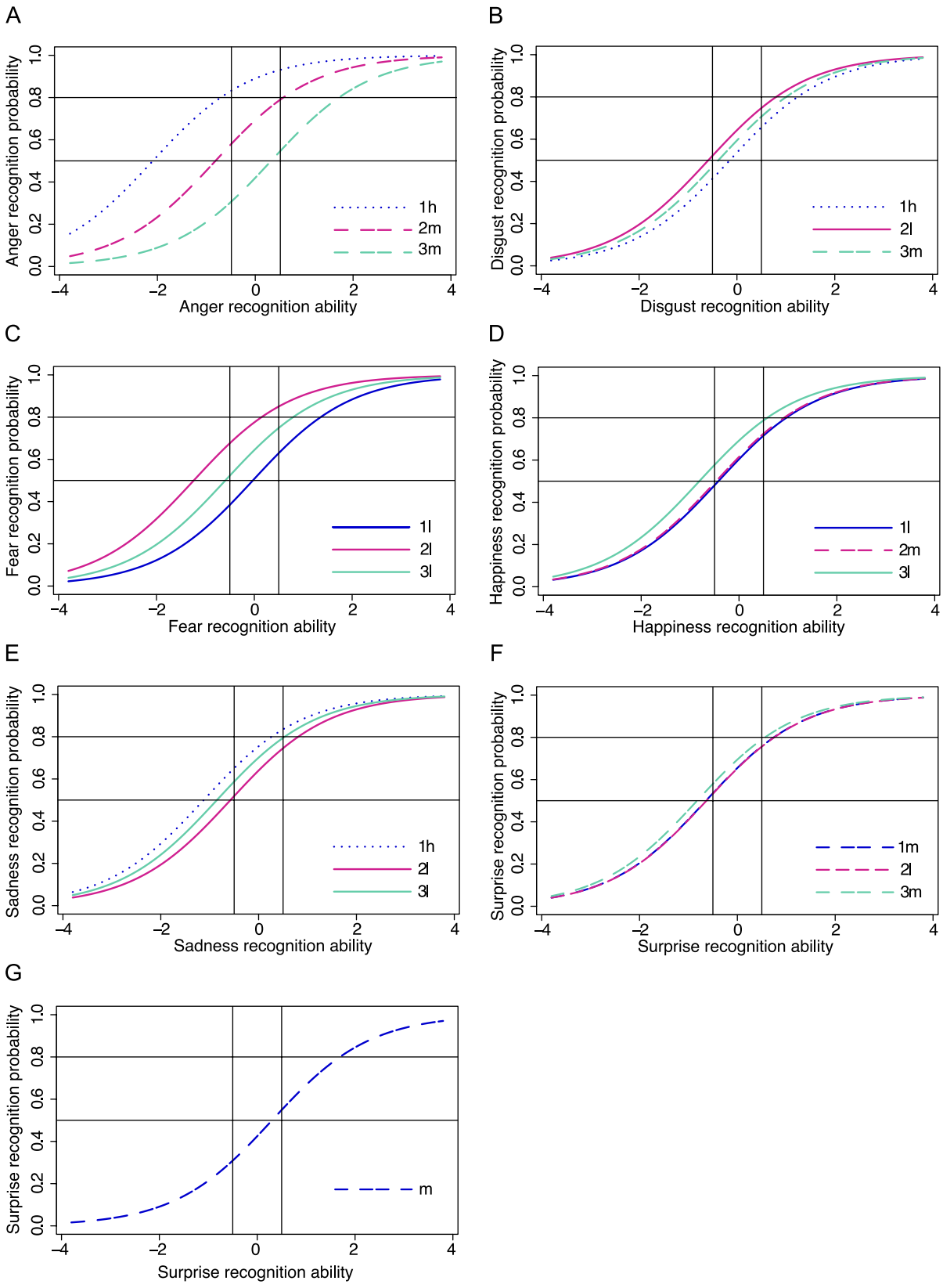
recognized the emotion "neutral" in each of these scenes [scene Neutral#1: 95% correct, scene Neutral#2: 81% correct, scene Neutral#3: 91% correct)]. These findings demonstrate that the scenes are useful to test the ability to recognize "neutral", which is why we kept them. See tables L3.19 – L3.21 in the Supplementary Material L for more information. All selected scenes are shown in Table 1. The item characteristic curves of all intensity versions of all scenes are shown in Supplementary Material K.

Step 3

Analysis of the 95% confidence intervals of the three intensities per scene generally revealed variations in ER difficulties according to their intensity levels (see Supplementary Material J). For three scenes (i.e., Anger#1, Anger#3, Disgust#3), we observed three difficulty levels, whereby ER was most difficult at the low intensity level, easiest at the high intensity level, with ER difficulty being in between for the medium intensity scenes. For 13 scenes, only two levels of difficulty could be distinguished: In five of these (i.e., Disgust#1, Fear#3, Happiness#2, Sadness#3, Surprise#3), significant differences in difficulty were present between low and high intensity levels, but not between medium and low or between medium and high intensity levels. In five scenes (i.e., Anger#2, Fear#1, Happiness#1, Sadness#2, Surprise#2), versions at the medium and high intensity levels did not differ in difficulty, but both significantly differed from the versions at low intensity level. For two scenes (i.e., Disgust#2, Happiness#3), the low intensity level versions differed in difficulty from the medium intensities, whereas high intensities did not differ from low and medium intensities. For one scene (i.e., Sadness#1), low and medium intensities did not differ, whereas both differed significantly in difficulty from high intensity. Only in two scenes (i.e., Fear#2, Surprise#1), the three intensities did not differ in their level of difficulty.

Figure 1

Item Characteristic Curves of the Selected BASIT-ER Scenes



Note. Item characteristic curves of the selected BASIT-ER scenes for the six basic emotions [i.e., A: anger, B: disgust, C: fear, D: happiness, E: sadness, and F: surprise], and G: the practice (surprise) scene. All scenes show neither floor nor ceiling effects for participants with emotion recognition ability in the medium range. Scene code consists of a number (scene) and a letter (emotional intensity). l=low intensity; m=medium intensity; h=high intensity. BASIT-ER = Basel Version of the Awareness of Social Inference Test - Emotion Recognition.

Table 1

Selected Intensity Version Scenes for the BASIT-ER for further Use in Clinical Populations

Basic emotion	Intensity version		
	Scene 1	Scene 2	Scene 3
Practice (Surprise)	Medium		
Anger	High*	Medium	Medium
Disgust	High	Low	Medium
Fear	Low	Low	Low
Happiness	Low	Medium	Low
Sadness	High	Low	Low
Surprise	Medium	Low	Medium
Neutral	Included	Included	Included

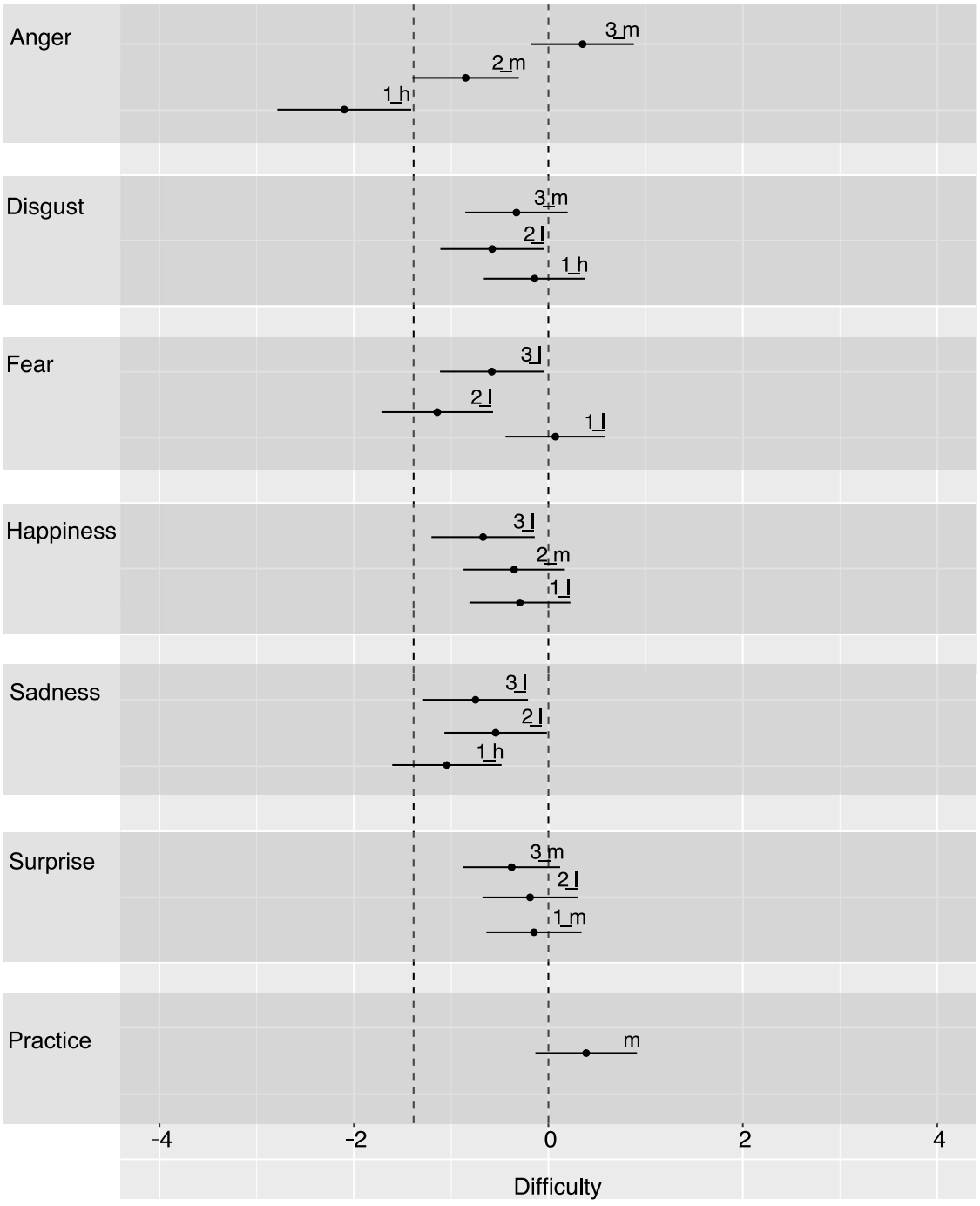
Note. BASIT-ER = Basel Version of the Awareness of Social Inference Test - Emotion Recognition.

*scene is outside the area of correct answering probability between 0.5 and 0.8 in participants with medium emotion recognition ability (-0.5 to 0.5).

Next, we calculated the range of difficulty that scenes were required to have (i.e., [-1.386,0]) for response probabilities in the range of 0.5 to 0.8 for participants with a medium ability in ER. Then we examined, how well the 95% confidence intervals (CI) of the estimated difficulties of the selected scenes covered the required range of difficulty (Figure 2). The estimated difficulties were within the required range in all selected Disgust, Happiness, Sadness and Surprise scenes and the Anger#2_medium, Fear#2_low, and Fear#3_low scenes [i.e., in 15 (83%) of the 18 basic emotion scenes]. We found higher than required difficulty estimations in the Fear#1_low scene (coverage by the lower half of the CI) and in the Anger#3_medium scene (marginal coverage by the lower CI). A lower than required difficulty estimation was found in the Anger#1_high scene (no coverage by the CI). The 95% CI of the estimated difficulties of all intensity version scenes are shown in Supplementary Material J.

Figure 2

95% Confidence Intervals of the Estimated Difficulties of the Selected BASIT-ER Scenes



Note. 95% confidence intervals of the estimated difficulties in emotion recognition of the 18 selected basic emotion scenes and the practice scene. Scene codes consist of a number (scene) and a letter (emotional intensity). l=low intensity; m=medium intensity; h=high intensity. The dashed lines illustrate the required range of difficulty [-1.386,0].

BASIT-ER = Basel Version of the Awareness of Social Inference Test - Emotion Recognition.

Number of Times the Selected Scenes were Watched

The majority of the selected test scenes (83%) were watched only once. 17% were watched at least twice, i.e., 13% were watched twice, 3% three times and less than 1% four times.

Emotion Profiles of the Selected Scenes

We analyzed the “other emotion” terms provided by the participants whose responses were classified as correct according to our scoring algorithm in the selected 21 test scenes. Of these participants, between 2% (in Neutral#1) and 52% (in Sadness#2_low) (*Median=10%, IQR=6)* named an “other emotion” term. The percentages of participants whose answers were classified as correct based on their proffered “other emotion” term ranged from 0% (Neutral#1, Neutral#2) to 50% (Sadness#2_low) (*Median=6%, IQR=9)*.

At each of the selected test scenes, at least three emotions were identified by participants whose responses were classified as correct. The number of identified emotions across scenes was distributed as follows: Three emotions in 24% (n=5), four emotions in 14% (n=3), five emotions in 24% (n=5), six emotions in 28% (n=6), eight emotions in 5% (n=1) and eleven emotions in 5% (n=1). Surprise was the only basic emotion, which was identified more than 50% of the time as an additional emotion beyond the portrayed emotion and this occurred in four scenes (i.e., Anger#2_medium, Anger#3_medium, Fear#1_low and Happiness#2_medium). More information on emotion profiles and percentages of labelled and/or named emotions of every selected scene are shown in Supplementary Material L.

Discussion

In this study, we developed and provided preliminary validation of the BASIT-ER, the first German-language TASIT-EET adaption in a sample of 240 healthy adults with a large age range. To the best of our knowledge, this is the first German-language dynamic, multimodal test on ER in realistic situations without ceiling or floor effects in healthy adults. We adapted TASIT-EET to overcome some methodological and practicable limitations such as inconsistencies in cinematic realization, ceiling effects in healthy participants, the simple response format and the long-time of administration.

We identified 22 scenes (one practice item and three scenes apiece for the six basic emotions and neutral) that will comprise the BASIT-ER. We found that all selected scenes were suitable for assessing the recognition of the respective emotion based on a thorough emotion response assessment, which included provision for selection of several emotions and additional free text format. As validation of the quality of the scenes, we found that gradual changes in portrayed emotional intensities related to difficulties in ER as measured by the correct answering probabilities.

In terms of emotion response patterns, we found that at least three emotions were identified by participants whose answers were classified as correct at each of the selected test scenes. Notably, surprise was the only basic emotion, which was identified - in addition to the portrayed emotion - in over 50% of the time in four of the selected scenes. This finding is in line with other studies showing that surprise - an emotional experience based on appraisal of novelty and unexpectedness - often accompanies other emotions (e.g., Fontaine et al., 2007).

In contrast to TASIT-EET, participants had the option to watch the scenes repetitively in case they did not have understood the content of the scene. Notably, participants used this opportunity sometimes; 17% of the selected scenes were watched more than once. It may be argued that in real life, situations are only experienced once and the option of multiple watching may make the task less realistic. However, real life situations generally last longer and include more context than our scenes, which likely

facilitates ER. More importantly, by giving participants' the opportunity to watch each scene until they have understood the scene, we minimize the risk that low ER test scores are due to cognitive dysfunction other than ER. This point is particularly relevant in regard to patients with neuropsychological disorders, who make up the target population for whom the BASIT-ER was designed.

Finally, our study should not be limited to the development of the BASIT-ER. We have generated a well-characterized set of dynamic emotional stimuli of varying difficulty from a sample of 240 healthy people with a large age range, evenly balanced in age and sex. Depending on the research question, type of participants and time available for testing, researchers can choose the stimuli and paradigms they favor.

Our study contains some limitations. First, no model fit was obtained for the Neutral scenes, which limits the conclusions about the relationship between the probability of correct recognition of neutral emotional states in these scenes and participants' ability in neutral ER. Thus, we could not select neutral scenes that were neither too difficult nor too easy for healthy participants based on Rasch analyses. Most participants, however, recognized the emotion "neutral" in these scenes. Therefore, these scenes appear suitable for assessing the ability to recognize "neutral". Second, the age range for the sample was very large, which may be seen as a limitation, given the association between age and ER (e.g., Hayes et al., 2020). However, the balanced distribution of ages within the sample, the large sample size and the pseudorandomized scene assignment to each participant resulted in an almost balanced distribution of age years within each intensity version scene (range of medians: 55.4-64.2; range of IQR: 21.4-30.9). In case of a completely balanced distribution of age years within intensity version scenes, the median age would have been 59.6 with an IQR of 25.1. Therefore, the possible influence of age on ER should not been considered a confounding factor in scene selection.

As a next step, the BASIT-ER needs to be validated in a clinical sample. In doing so, the free response format will be removed from the task, as this option is neither practical nor objectively feasible in a clinical context. The validation study will examine the

reliability, construct and ecological validity of the BASIT-ER. Moreover, it will guide us in the final selection of BASIT-ER scenes for later use in clinical settings.

In conclusion, we have developed and preliminarily validated the first German-language adaption of TASIT-EET, called the BASIT-ER, in healthy adults. The ultimate aim is for the BASIT-ER, ideally in a shortened version, to become a core part of any comprehensive German-language neuropsychological exam in clinical routine.

References

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). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*. https://doi.org/10.1176/appi.books.9780890425596

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. https://doi.org/10.1037/amp0000191

Arsalidou, M., Morris, D., & Taylor, M. J. (2011). Converging Evidence for the Advantage of Dynamic Facial Expressions. *Brain Topography*, *24*(2), 149–163. https://doi.org/10.1007/s10548-011-0171-4

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*(3), 277–285. https://doi.org/10.1037/a0021908

Beck, A., Ward, C., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An Inventory for Measuring Depression. *Archives of General Psychiatry*, *4*, 561-571. https://doi.org/10.1001/ARCHPSYC.1961.01710120031004

Bora, E. (2017). Meta-analysis of social cognition in amyotrophic lateral sclerosis. *Cortex*, *88*, 1–7. https://doi.org/10.1016/j.cortex.2016.11.012

Bora, E., Velakoulis, D., & Walterfang, M. (2016). Meta-Analysis of Facial Emotion Recognition in Behavioral Variant Frontotemporal Dementia: Comparison With Alzheimer Disease and Healthy Controls. *Journal of Geriatric Psychiatry and Neurology*, *29*(4), 205–211. https://doi.org/10.1177/0891988716640375

Cotter, J., Firth, J., Enzinger, C., Kontopantelis, E., Yung, A. R., Elliott, R., & Drake, R. J. (2016). Social cognition in multiple sclerosis: A systematic review and meta-analysis. *Neurology*, *87*(16), 1727-1736. https://doi.org/10.1212/WNL.0000000000003236

Deckert, M., Schmoeger, M., Auff, E., & Willinger, U. (2020). Subjective emotional arousal: an explorative study on the role of gender, age, intensity, emotion regulation difficulties, depression an anxiety symptoms, and meta-emotion. *Psychological Research*, *84*(7), 1857-1876. https://doi.org/10.1007/s00426-019-01197-z

Dodich, A., Crespi, C., Santi, G. C., Marcone, A., Iannaccone, S., Perani, D., Cappa, S. F., & Cerami, C. (2021). Emotion Recognition Deficits in the Differential Diagnosis of Amnestic Mild Cognitive Impairment: A Cognitive Marker for the Limbic-Predominant Phenotype. *Journal of the International Neuropsychological Society*, 1–7. https://doi.org/10.1017/S1355617721000254

Fonagy, P., & Goodyer, I. (Eds.). (2008). Social Cognition and Developmental Psychopathology. https://doi.org/10.1093/med/9780198569183.001.0001

Fontaine, J. R., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological science*, *18*(12), 1050-1057. https://doi.org/10.1111/j.1467-9280.2007.02024.x

Hasson-Ohayon, I., Mashiach-Eizenberg, M., Arnon-Ribenfeld, N., Kravetz, S., & Roe, D. (2017). Neuro-cognition and social cognition elements of social functioning and social quality of life. *Psychiatry Research*, *258*, 538–543. https://doi.org/10.1016/J.PSYCHRES.2017.09.004

Hayes, G. S., McLennan, S. N., Henry, J. D., Phillips, L. H., Terrett, G., Rendell, P. G., ... & Labuschagne, I. (2020). Task characteristics influence facial emotion recognition age-effects: A meta-analytic review. *Psychology and aging*, *35*(2), 295.

https://doi.org/10.1037/pag0000441

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., 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. https://doi.org/10.1038/nrneurol.2015.229

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

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

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

Kumfor, F., Irish, M., Leyton, C., Miller, L., Lah, S., Devenney, E., Hodges, J. R. & Piguet, O. (2014). Tracking the progression of social cognition in neurodegenerative disorders. Journal of Neurology, Neurosurgery & Psychiatry, 85(10), 1076-1083. http://dx.doi.org/10.1136/jnnp-2013-307098

Loi, F., Vaidya, J. G., & Paradiso, S. (2013). Recognition of emotion from body language among patients with unipolar depression. *Psychiatry Research*, *209*(1), 40–49. https://doi.org/10.1016/J.PSYCHRES.2013.03.001

Maat, A., Fett, A.-K., & Derks, E. (2012). Social cognition and quality of life in schizophrenia. *Schizophrenia Research*, *137*(1–3), 212–218. https://doi.org/10.1016/J.SCHRES.2012.02.017

McDonald, S. (2012). New Frontiers in Neuropsychological Assessment: Assessing Social Perception Using a Standardized Instrument, The Awareness of Social Inference Test. *Australian Psychologist*, *47*(1), 39–48. https://doi.org/10.1111/j.1742-9544.2011.00054.x

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., 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

McDonald, S., Rushby, J. A., Dalton, K. I., Allen, S. K., & Parks, N. (2018). The role of abnormalities in the corpus callosum in social cognition deficits after Traumatic Brain Injury. *Social Neuroscience*, *13*(4), 471–479. https://doi.org/10.1080/17470919.2017.1356370

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. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C.D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., Ishiyama, J., …, & Yarkoni, T. (2015). SCIENTIFIC STANDARDS. Promoting an open research culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

Peirce, J. W. (2007). PsychoPy--Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13.

https://doi.org/10.1016/j.jneumeth.2006.11.017

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*, 10. https://doi.org/10.3389/neuro.11.010.2008

Plutchik, R. (1962). *The emotions: Facts, theories and a new model*. New York: Crown Publishing Group/Random House.

Plutchik, R. (1980). *Emotion, a psychoevolutionary synthesis*. New York: Harper & Row.

Plutchik, R. (2003). *Emotions and life : perspectives from psychology, biology, and evolution*. Washington, DC: American Psychological Association.

Jarsch, M., Bartsch, K., Berres, M., Ryf, I., Jurisic, K., Sluka, C., Kressig, R. W., Monsch, A. U., Piguet, O., & Sollberger, M. (2021). *Exemplary Scenes of the Basel Version of The Awareness of Social Inference Test-Emotion Evaluation Test (BASIT-EET) at three emotion intensities*. https://doi.org/10.6084/m9.figshare.13507245

R Core Team. (2020). R: A Language and Environment for Statistical Computing. Retrieved from https://www.r-project.org/

R Studio Team. (2016). *RStudio: Integrated Development Environment for R*. Boston, MA: RStudio, Inc.

Rasch, G. (1960). *Probabilistic Models for some Intelligence and Attainment Tests*. Copenhagen: Paedagogike Institut.

Risko, E. F., Laidlaw, K., Freeth, M., Foulsham, T., & Kingstone, A. (2012). Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. *Frontiers in Human Neuroscience*, *6*, 143. https://doi.org/10.3389/fnhum.2012.00143

Rizopoulos, D. (2006). ltm : An R Package for Latent Variable Modeling and Item Response Theory Analyzes. Journal of Statistical Software, 17(5), 1–25. https://doi.org/10.18637/jss.v017.i05

Savla, G. N., Vella, L., Armstrong, C. C., Penn, D. L., & Twamley, E. W. (2013). Deficits in Domains of Social Cognition in Schizophrenia: A Meta-Analysis of the Empirical Evidence. *Schizophrenia Bulletin*, *39*(5), 979–992.

https://doi.org/10.1093/schbul/sbs080

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, *44*(4), 695–729. https://doi.org/10.1177/0539018405058216

Schlegel, K., Grandjean, D., & Scherer, K. R. (2014). Introducing the Geneva emotion recognition test: an example of Rasch-based test development. *Psychological assessment*, *26*(2), 666. https://doi.org/10.1037/a0035246

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (2001). Emotional Knowledge: Further Exploration of a Prototype Approach. In W. G. Parrott (Ed.), *Emotions in Social Psychology: Essential Readings* (pp. 26–56). Philadelphia: Taylor & Francis.

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

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. https://doi.org/10.3233/JAD-180080

Torres Mendonça De Melo Fádel, B., Santos De Carvalho, R. L., Belfort Almeida Dos Santos, T. T., & Dourado, M. C. N. (2019). Facial expression recognition in Alzheimer's disease: A systematic review. *Journal of Clinical and Experimental Neuropsychology*, *41*(2), 192–203. https://doi.org/10.1080/13803395.2018.1501001

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. https://doi.org/10.1080/13803395.2014.966661

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., & Dunnington, D. (2020). *ggplot2: create elegant data visualisations using the grammar of graphics*. https://CRAN.R-project.org/package=ggplot2

Yang, W. F., Toller, G., Shdo, S., Kotz, S. A., Brown, J., Seeley, W. W., ... & Rankin, K. P. (2021). Resting functional connectivity in the semantic appraisal network predicts accuracy of emotion identification. *NeuroImage: Clinical*, *31*, 102755. https://doi.org/10.1016/j.nicl.2021.102755

Yesavage, J. A., & Sheikh, J. I. (1986). Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. *Clinical Gerontologist*, *5*(1–2), 165–173. https://doi.org/10.1300/J018v05n01_09

Yogarajah, M., & Mula, M. (2019). Social cognition, psychiatric comorbidities, and quality of life in adults with epilepsy. *Epilepsy & Behavior*, *100*, 106321. https://doi.org/10.1016/j.yebeh.2019.05.017

Zinck, A., & Newen, A. (2008). Classifying emotion: a developmental account. *Synthese*, *161*(1), 1–25. https://doi.org/10.1007/s11229-006-9149-2

4.2 Manuskript 3

The Basel Version of the Awareness of Social Inference Test – Theory of Mind (BASIT-ToM): Preliminary Validation Analyses in Healthy Adults

Manuskript 3 beschreibt detailliert die Entwicklung und präliminäre Validierung des deutschsprachigen ToM - Subtest des BASIT (BASIT-ToM) an 240 gesunden Erwachsenen.

The Basel Version of the Awareness of Social Inference Test – Theory of Mind (BASIT-ToM): Preliminary Validation Analyzes in Healthy Adults

Marianne Jarsch[1,2], Olivier Piguet[3], Manfred Berres[4], Constantin Sluka[5], Reto W. Kressig[1], Andreas U. Monsch[1,2], Skye McDonald[6], and Marc Sollberger[1,7]

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

[2]Faculty of Psychology, University of Basel, Switzerland;

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

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

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

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

[7]Department of Neurology, University Hospital Basel, Switzerland.

Author Note

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

Reto W. Kressig https://orcid.org/0000-0002-3215-7681

Andreas U. Monsch https://orcid.org/0000-0003-1692-1560

Skye McDonald https://orcid.org/0000-0003-0723-6094

Marc Sollberger https://orcid.org/0000-0001-8863-3731

One exemplary test scene of the BASIT-ER at three intensity levels of the portrayed paralinguistic cues can be found at https://doi.org/10.6084/m9.figshare.17074997

MJ 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 the National Health and Medical Research Council of Australia Senior Research Leadership Fellowship grant 2008020. All authors report no conflict of interest.

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 Auchter, 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.

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

Abstract

Objective: 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 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 and provide preliminary validation of the first German-language adaption of TASIT-SIM in healthy adults.

Method: 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.

Results: By applying Rasch analysis, we selected intensity versions that showed neither floor nor ceiling effects in the majority of the ToM questions in participants whose ToM abilities were in the medium range.

Conclusions: We have conducted a preliminary validation of the first German-language adaption of TASIT-SIM, i.e., the “Basel Version of the Awareness of Social Inference Test – Theory of Mind” (BASIT-ToM). 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 clinical populations.

Keywords: Social cognition, Theory of Mind, healthy subjects, dynamic stimuli, test development

Key Points

Question

The aim of this study was to develop and provide preliminary validation of the Basel Version of the Awareness of Social Inference Test – Theory of Mind (BASIT-ToM) in a sample of healthy adults.

Findings

We succeeded in generating a dynamic, German-language test of theory of mind, which is neither too easy nor too difficult for healthy adults.

Importance

The preliminary validation of the BASIT-ToM is an important step towards the development of a German-language test that allows the assessment of subtle theory of mind deficits in a realistic way and within a practicable timeframe in clinical settings.

Next Steps

Next, the BASIT-ToM needs to be validated in clinical populations with neuropsychological disorders.

The Basel Version of The Awareness of Social Inference Test – Theory of Mind (BASIT-ToM): Preliminary Validation Analyzes in Healthy Adults

Social cognition refers to a wide range of cognitive processes, ranging from perception of socio-emotional stimuli to social behavior (Henry et al., 2016). Consequently, social cognitive abilities such as theory of mind (ToM) or emotion recognition play a fundamental role in human life. They enable individuals to interact successfully with others in the long term (Kennedy & Adolphs, 2012) and are crucial for their perceived quality of life (e.g., Hasson-Ohayon et al., 2017; Maat et al., 2012; Yogarajah & Mula, 2019).

ToM enables the adoption of cognitive (cognitive ToM) and affective (affective ToM) perspectives of others (e.g., Sebastian et al., 2012; Shamay-Tsoory & Aharon-Peretz, 2007; Shamay-Tsoory, Harari, Aharon-Peretz, & Levkovitz, 2010) by identifying others' cognitive mental and affective mental states, respectively. 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 (e.g., Henry, von Hippel, Molenberghs, Lee, & Sachdev, 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 from another person (e.g., Henry, von Hippel, Molenberghs, Lee, & Sachdev, 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 (e.g., Amodio & Frith, 2006; Canty, Neumann, Fleming, & Shum, 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) (e.g., Meijering, Van Rijn, Taatgen, & Verbrugge, 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) (e.g., Meijering et al., 2011).

Lastly, apart from content and degree of cognitive complexity within the concept of ToM, ToM can be conceptualised 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).

Deficits in ToM occur in a wide range of brain disorders, including traumatic brain injuries, neurodevelopmental and neurodegenerative brain disorders 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 perspective 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, multimodal emotional 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 46 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 for use in patients with neurodevelopemental and psychiatric disorders (Dziobek et al., 2006; Fossati et al., 2018). However, given its high cognitive complexity, the test is unsuitable 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, the test takes 30-45 minutes in patients with mild to moderate traumatic brain injuries [J. Quinting, personal communication, October 29, 2021; Quinting et al. (2020)], which precludes its use in the 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 validated and well-known English-language, dynamic Social Inference Test (SIM) of The Awareness of Social Inference Test (TASIT) (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., behavioral variant frontotemporal dementia (e.g., Kipps et al., 2009; Kumfor et al., 2017), semantic variant primary progressive aphasia (Rankin et al., 2009), supranuclear palsy (Ghosh et al., 2012), neuroinflammatory diseases [e.g., multiple sclerosis (e.g., Genova & McDonald, 2020; Sofologi et al., 2019)], traumatic brain injuries (e.g., McDonald, 2012; McDonald et al., 2017) and psychiatric diseases such as schizophrenia and bipolar disorder (e.g., Patel et al., 2020; Quidé et al.,

2020). It has adequate reliability and there is evidence for construct and ecological validity (for review see McDonald, 2012; McDonald, Flanagan, Martin, & Saunders, 2004).

TASIT-SIM, however, contains limitations that need to be addressed (see also (Jarsch et al., in press). 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 may result in ceiling effects in the assessment of healthy participants and prevent 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.

Thus, the current study aimed to overcome these limitations by creating an adapted version of TASIT-SIM for German speakers, followed by a preliminary validation in healthy participants. We named the adapted German-language version the "Basel Version of the Awareness of Social Inference Test [BASIT-Theory of Mind (BASIT-ToM)]".

Methods

The Awareness of Social Inference Test – Social Inference (TASIT-SI)

TASIT – Social Inference (SI) comprises two parts [i.e., SI-Minimal (SIM) and SI-Enriched (SIE)] (McDonald et al., 2003). The 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 incongruency cannot be perceived, statements might be perceived as honest, meaningless, bizarre or as a lie (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).

The 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 the 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., honest, simple sarcasm, paradoxical sarcasm)]. The SI-E comprises 16 scenes [eight scenes per two different message types (i.e., sarcasm, lie)]. The scenes last 15 to 60 seconds and are administered in a quasi-randomised order.

In both SI parts, 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 sarcasticly, or who lies. Participants can answer each question with “yes”, “no”, or “don’t 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 parts 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, McDonald, Sufani, Hine, & Kumfor, 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. We renamed the adapted version 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., in press).

Selection of the Scenes from TASIT-SIM

One test form was developed, rather than two, given the negligible practice effects on TASIT-SIM scores (McDonald et al., 2003). First, we selected 13 film scenes (4x3 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 Supplemental Material A1). Three raters (MJ, MS, and a master's degree psychology student) rated each scene independently, then agreed on a joint-rating and finally selected the 13 scenes in consensus (see Supplemental Material B, Tables B1.1-B1.3). The selected TASIT-SIM scenes are depicted in the Supplemental Material B, Table B2.

Conceptualisation of the BASIT-ToM Scenes

We developed the conceptualisation of the scenes and the cinematographic realisation 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., in press), we omitted naming the actors to exclude any potential bias associated with naming. All dialogue-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 Supplemental Material 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 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 (e.g., 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, a conversation partner's neutral behavior may confuse a viewer with preserved ToM capacity. The distribution of the actors to each of the 13 scenes is presented in the Supplemental Material 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 exchange types. After the shooting, MJ, MS and six master's degree psychology students selected the best scenes for each intensity by consensus. One example of these scenes (pS#4) at low, medium and high levels of intensity, can be found at (https://doi.org/10.6084/m9.figshare.17074997)

Background information for each BASIT-ToM scene is described in the Supplemental Material 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 about beliefs. Other question types also tapped ToM to varying degrees (S. McDonald, personal communication, November 15, 2021; McDonald et al., 2003).

We decided to categorize 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)]. Consequently, separate analyses of types and complexity should be feasible in the BASIT-ToM. This is critical in regard to the future use of the BASIT-ToM in clinical

samples. Scores reflecting different kinds of ToM ability, which partly reflect different neuroanatomic substrates (e.g., 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 (e.g., Lancaster et al., 2019; Poletti et al., 2012; Rossetto et al., 2018; Yildirim et al., 2020) than an overall ToM score.

Moreover, we created a hypermentalization question (hyperToM) for each scene to assess whether an individual overattributes another person's cognitive and/or affective mental states. 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; Henry et al., 2014).

In TASIT-SIM, participants could answer each question with "yes", "no", or "don't 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 "don't 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 to towards "yes" or "no" answers, respectively. An 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 newly developed hypermentalization question related to the same scene as the ToM question, are depicted in Table 1. All questions of BASIT-ToM and TASIT-SIM with their corresponding answers, types and orders of ToM as well as the newly developed hypermentalization questions are found in the Supplemental Material F. The main

differences between TASIT-SIM and BASIT-ToM scenes are described in the Supplemental Material G

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 BASIT-ToM hypermentalization question related to the same scene as the ToM 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 talking to her about the report to provoke her?	hyperToM		no

Note. ToM=Theory of Mind; 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; cogn=cognitive ToM; aff=affective ToM; hyperToM= hypermentalization.

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 for use in clinical populations, i.e., 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). Inclusion and exclusion criteria have been presented elsewhere (Jarsch et al., in press). 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 Supplemental Material 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 familiarise the participant with the procedure and to clarify any potential questions. The 12 test scenes were then administered in a pseudo-randomised 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 x 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.

Following each scene, participants were required to answer the four ToM questions and the hyperToM question with “yes” or “no”. There was no time limit for answering the questions. Participants were allowed to watch 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’s content. Difficulties in understanding a scene’s content may be not exclusively due to difficulties in ToM, but potentially also due to difficulties in more basic cognitive domains such as psychomotor speed or working memory. This type of information is valuable given that the aim is to administer the BASIT-ToM to individuals with cognitive deficits. Similarly, participants could change their answers within a question block. An exemplary representation of the BASIT-ToM computer based application process is depicted in Supplemental Material H.

Data Analysis. The aim of the data analysis of the 39 intensity version scenes [3(message types) x 4(scenes) x 3(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 prioritised 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’ ability parameters and the probabilities of correct answering the four ToM questions of the respective intensity version.

In each of the three message type analyzes, we included the respective twelve intensity versions (4 scenes per message type x 3 intensities). In the practice scene analysis, we included the respective three intensity versions. Separate analyzes were carried out for each message type intensity version scene and each practice intensity version scene. In order to check the model fit of the data, we conducted the ltm package parametric Bootstrap goodness-of-fit test using Pearson's $\chi 2$, 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 (CI) of the estimated difficulties of the selected intensity versions in relation to the required difficulty interval. 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.

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 original data sets, the R syntax and the BASIT-ToM scenes can be provided by the corresponding author on request. The study's design and its analysis were not pre-registered.

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 analyzes 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 Supplemental Material 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 x 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] (Fig. 1 A-C for the three H scenes, Fig. 2 A-D for the three pS scenes and Fig. 3 A-C for the sS scenes). All selected intensity versions are shown in Table 2 and their respective ToM types and orders are shown in Table 3.

The number of the ToM questions that met the required response probabilities varied between the selected scenes. In 8 (80%) of the 10 selected intensity version scenes at least half of the questions met the required response probalities, 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 scenes 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); Fig. 4 for H scenes, Fig. 5 for pS scenes and Fig. 6 for sS scenes]. The estimated difficulties were considered within the required range of difficulty when at least half of the CI covered the required range. In 2 of the 10 selected scenes (20%; i.e., H#2_low, pS#1_low), all four questions were within the required range. In another 2 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 3 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 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 2 scenes (i.e., H#4_medium, sS#2_low), all questions were below the required difficulty (too easy).

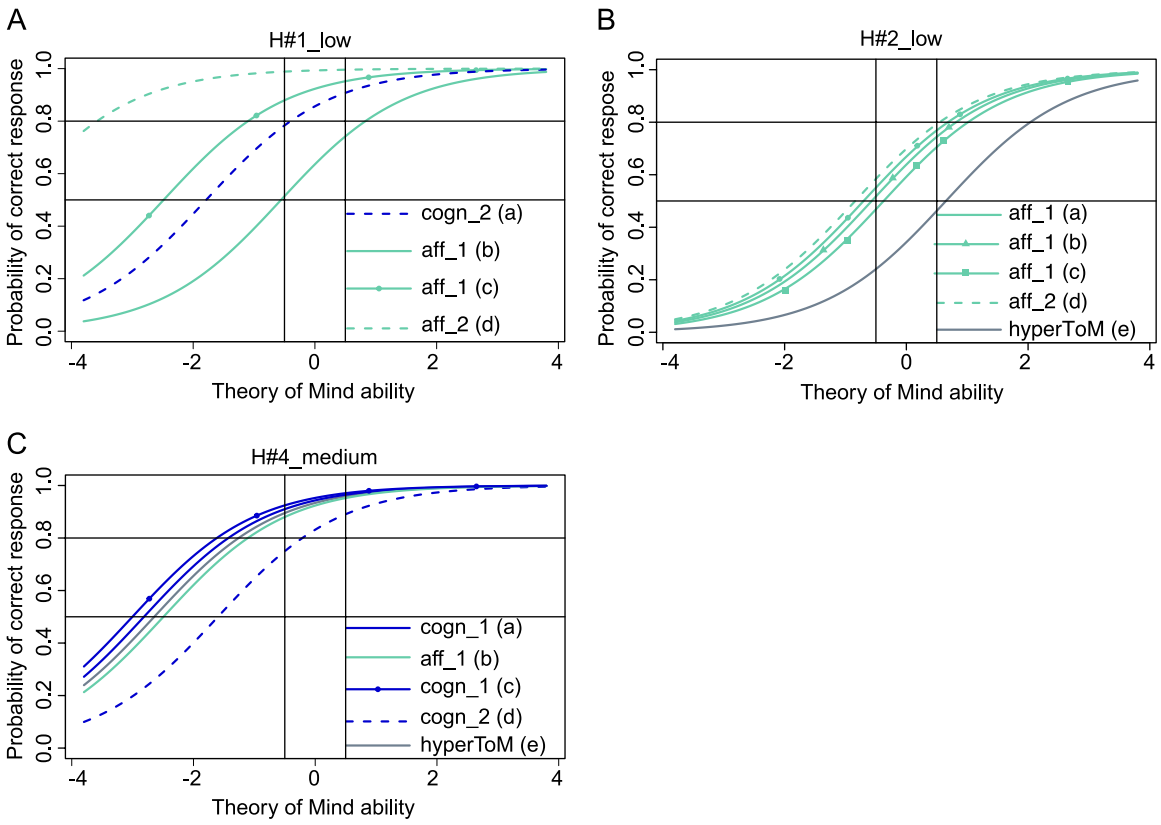
Notably, we found no systematic differences in terms of difficulty between first- and second-order ToM questions and between cognitive and affective ToM questions, respectively (Fig. 4-6).

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 6 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 2 [i.e., pS#1_low (Fig. 2A), sS#2_low (Fig. 3B)] of these scenes, the response probability of the hyperToM question was within the required range of 0.5 to 0.8 (Fig. 5-6); in 3 scenes [i.e., H#4_medium (Fig. 1C), pS#4_low (Fig. 2C), sS#1_low (Fig. 3A)], it was above 0.8 [too easy, Fig. 4-6] and in 1 scene [i.e., H#2_low (Fig. 1B)], it was below 0.5 [too difficult, Fig. 4]. A model fit was no longer present in 4 scenes (i.e., H#1_low, pS#2_low, practice(pS)_low, sS#3_medium).

Figure 1

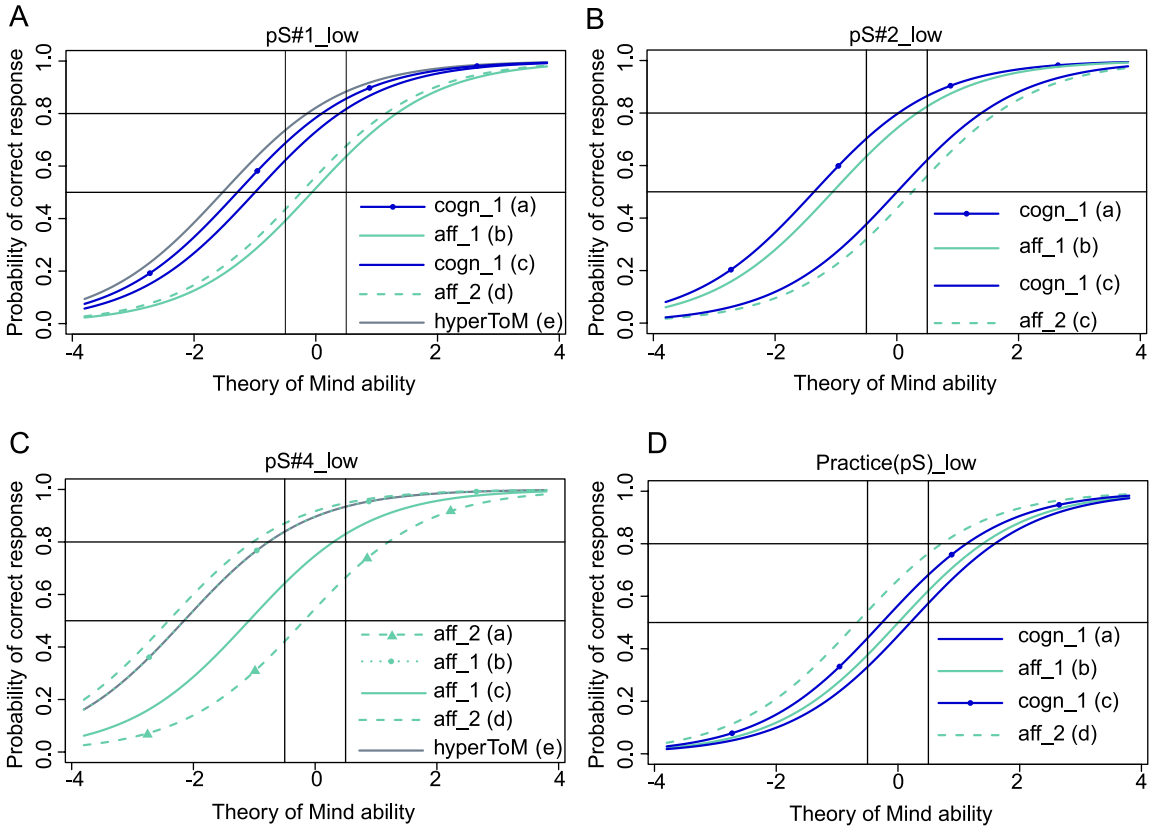
Item Characteristic Curves of the Selected BASIT-ToM Scenes for Message Type Honesty



Note. 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: H#1_low, B: H#2_low, C: 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) 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). BASIT-ToM=Basel Version of The Awareness of Social Inference Test - Theory of Mind.

Figure 2

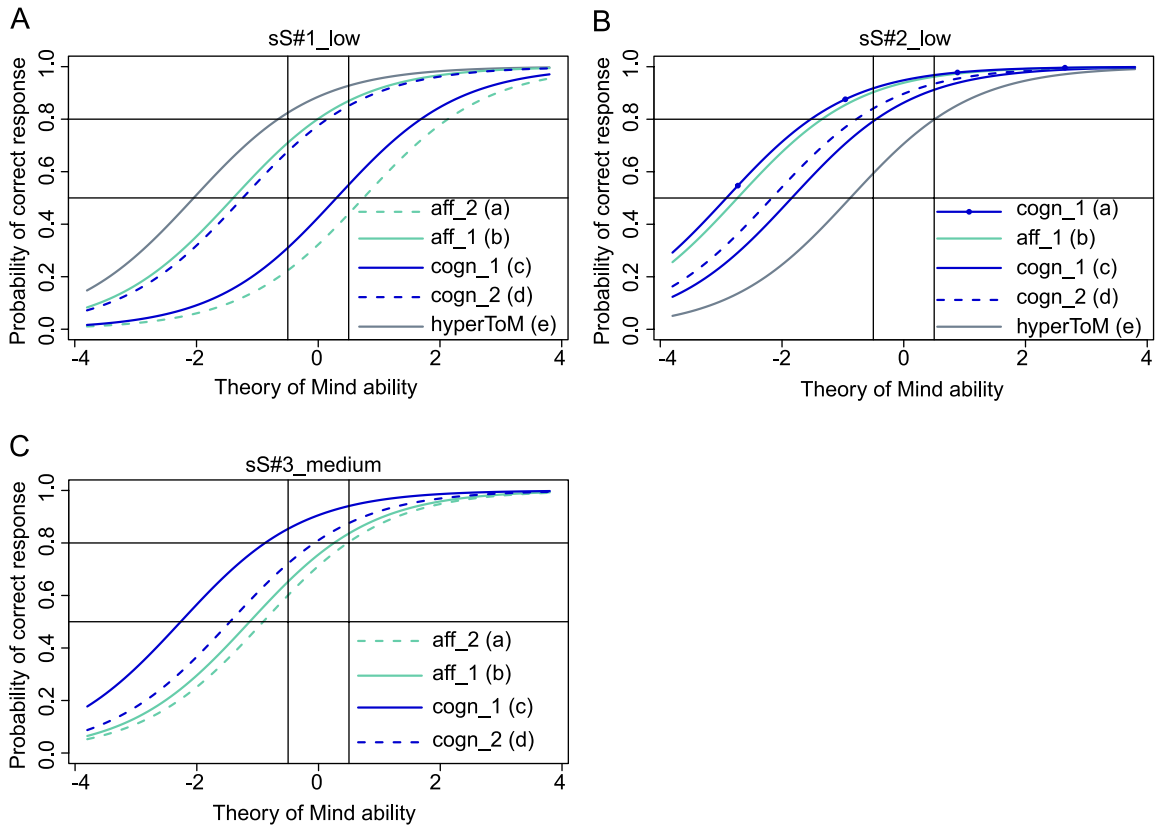
Item Characteristic Curves of the Selected BASIT-ToM Scenes for Message Type Paradoxical Sarcasm



Note. 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); 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 3

Item Characteristic Curves of the Selected BASIT-ToM Scenes for Message Type Simple Sarcasm



Note. 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). 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*	Practice (pS)_low†

Note. BASIT-ToM=Basel Version of The Awareness of Social Inference Test-Theory of Mind; pS=paradoxical sarcasm; n/a=not available.

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

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

Table 3

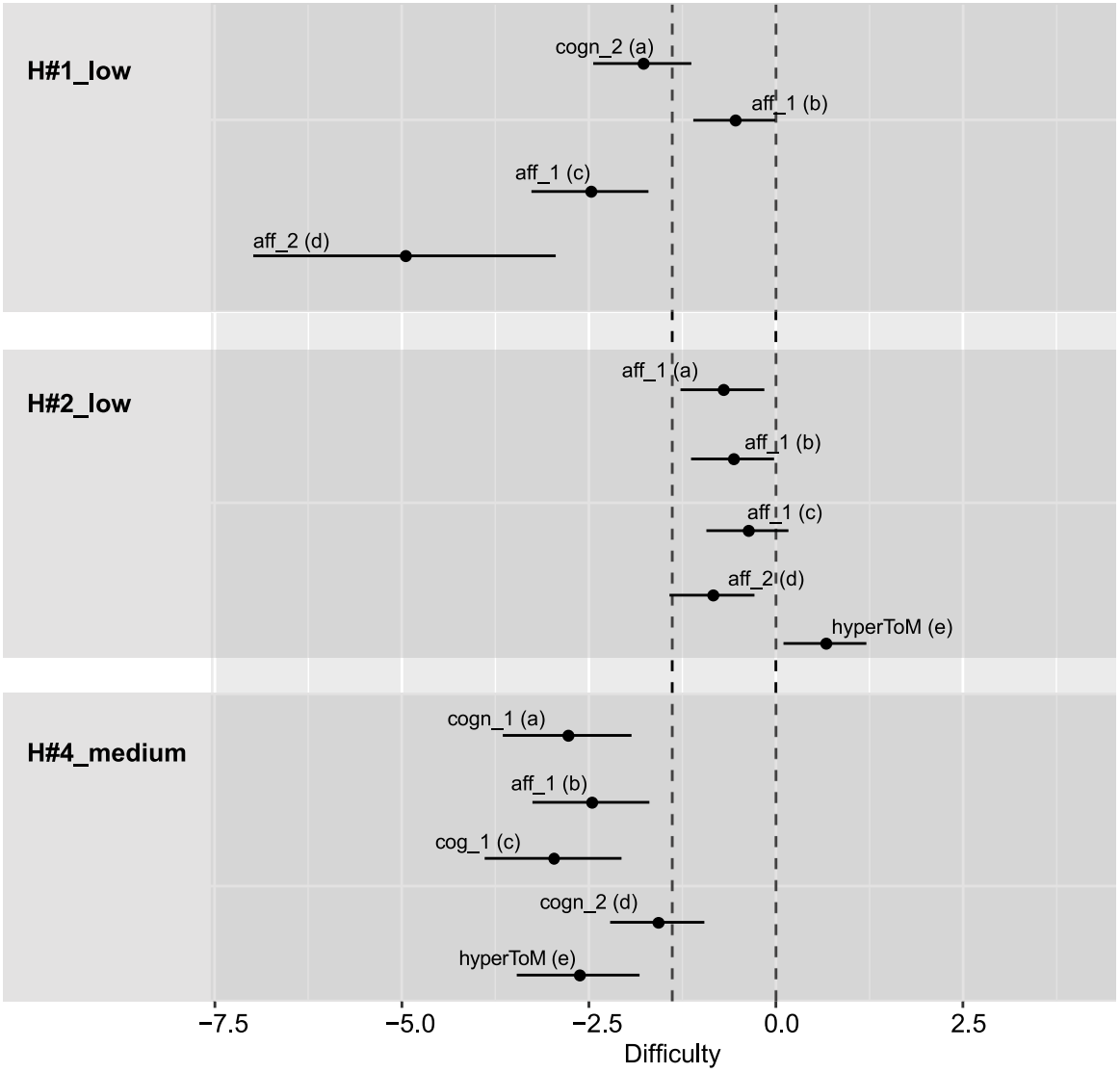
Types and Orders of Theory of Mind for Each Question of the Ten Selected Scenes

Scene	ToM type and order				
	a	b	c	d	e
H#1	cogn_2nd	aff_1st	aff_1st	aff_2nd	hyper
H#2	aff_1st	aff_1st	aff_1st	aff_2nd	hyper
H#4	cogn_1st	aff_1st	cogn_1st	cogn_2nd	hyper
sS#1	aff_2nd	aff_1st	cogn_1st	cogn_2nd	hyper
sS#2	cogn_1st	aff_1st	cogn_1st	cogn_2nd	hyper
sS#3	aff_2nd	aff_1st	cogn_1st	cogn_2nd	hyper
pS#1	cogn_1st	aff_1st	cogn_1st	aff_2nd	hyper
pS#2	cogn_1st	aff_1st	cogn_1st	aff_2nd	hyper
pS#4	aff_2nd	aff_1st	aff_1st	aff_2nd	hyper
Practice(pS)	cogn_1st	aff_1st	cogn_1st	aff_2nd	hyper

Note. ToM=Theory of Mind; Scene code consists of letters (i.e., H=honesty, sS=simple sarcasm, pS= paradoxical sarcasm) and the respective scene numbers; Question ID=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.

Figure 4

95% Confidence Intervals of the Estimated Difficulties of the Selected BASIT-ToM Scenes for Message Type Honesty

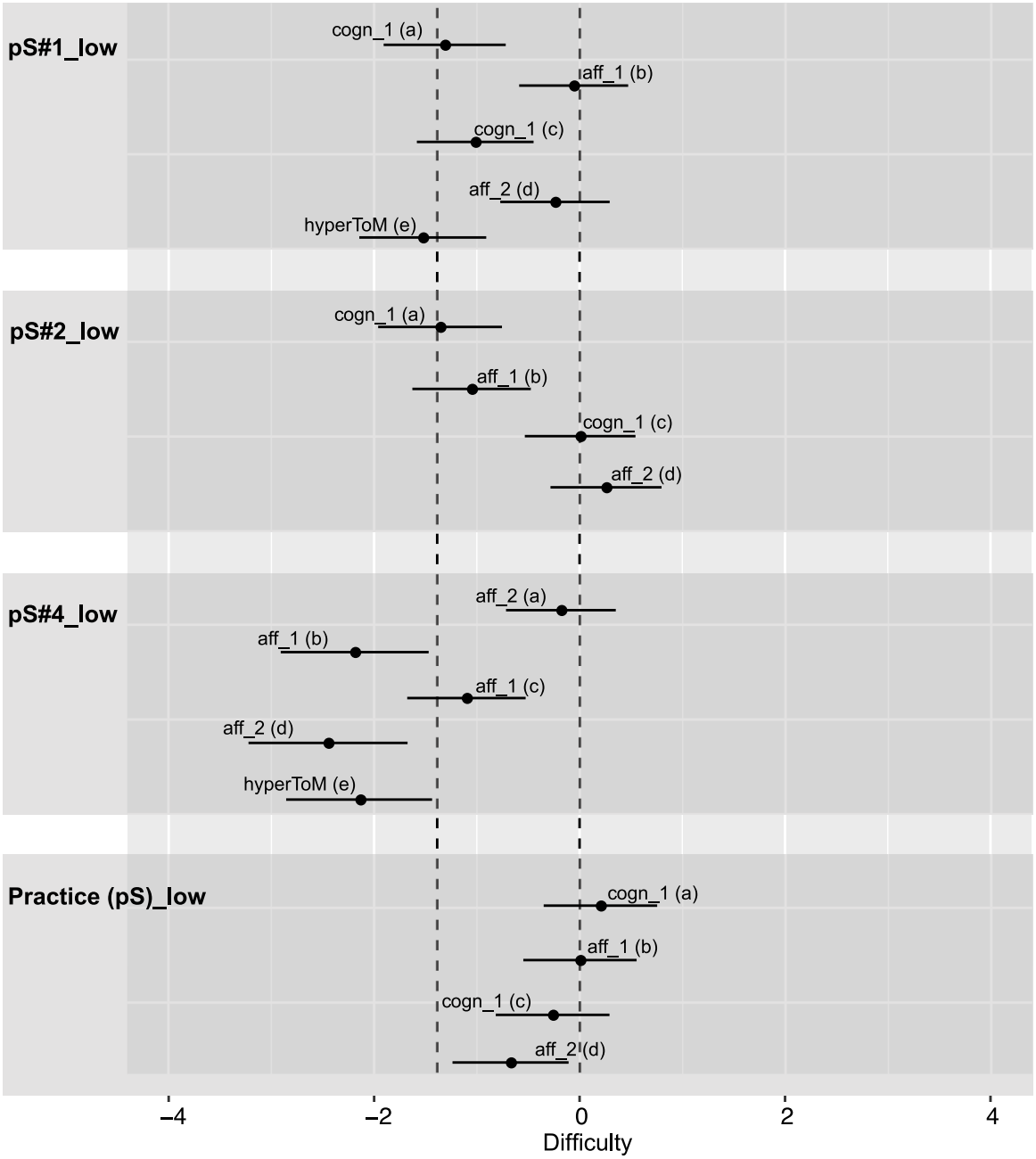


Note. 95% confidence intervals (CI) of the estimated difficulties of the Theory of Mind (ToM) and hypermentalization questions (hyperToM) of the three selected BASIT-ToM scenes for message type honesty (i.e., A: H#1_low, B: H#2_low, C: H#4_medium). CI of the hyperToM question is only depicted in presence 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=Basel Version of The Awareness of Social Inference Test - Theory of Mind.

Figure 5

95% Confidence Intervals of the Estimated Difficulties of the Selected BASIT-ToM Scenes for Message Type Paradoxical Sarcasm

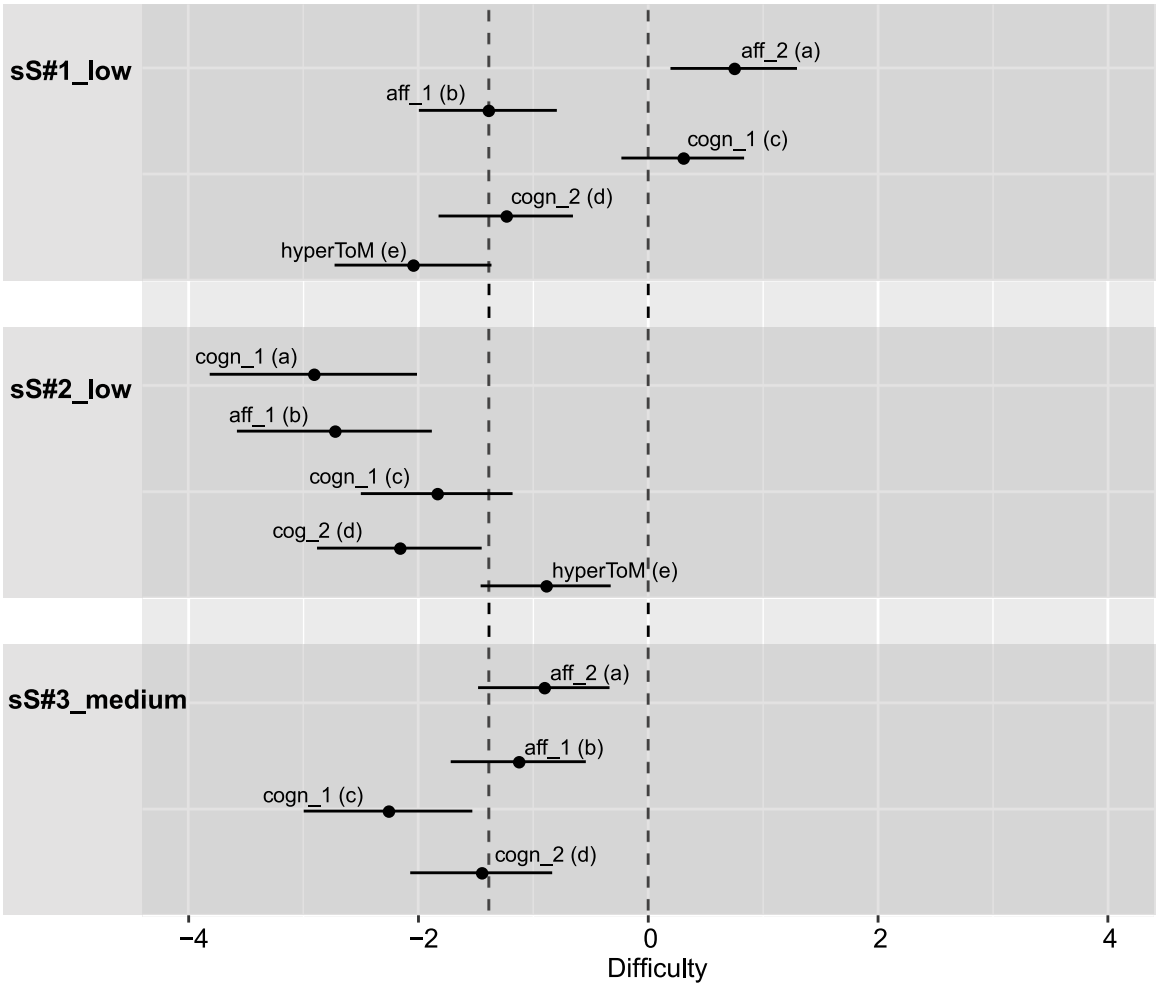


Note. 95% confidence intervals (CI) of the estimated difficulties of the Theory of Mind (ToM) and hypermentalization questions (hyperToM) of the four selected BASIT-ToM scenes for message type paradoxical sarcasm (i.e., A: scene pS#1_low, B: scene pS#2_low, C: scene pS#4_low, D: Practice (pS)_low). CI 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.

Figure 6

95% Confidence Intervals of the Estimated Difficulties of the Selected BASIT-ToM Scenes for Message Type Simple Sarcasm



Note. 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.

Number of Times the Selected Scenes were Watched

The majority of the selected test scenes (87%) were watched only once. About 13% were watched at least twice, i.e., 12% were watched twice, 1% three times, less than 1% four times and less than 1% five times.

Discussion

In this study, we developed and provided preliminary validation of the BASIT-ToM, the first German-language TASIT-SIM adaption in a sample of 240 healthy adults with a large age range. We adapted TASIT-SIM to address 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 analyze its response probalities in 6 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. However, for example, in schizophrenic patients with positive symptoms, incorrect answers may be a sign of hypermentalization (Canty et al., 2017).

Notably, we found no systematic differences in difficulty between first- and second-order ToM questions implying no larger differences in cognitive demand between these two types of ToM. Similarly, we found no systematic differences in difficulty between cognitive and affective ToM questions.

Participants had the option to re-watch the scenes in case they did not have understood the content of the scene. Notably, most of our participants did not use this option, which speaks in favor of the scenes's content. However, in patients with neuropsychological disorders, who make up the target population for whom the BASIT-ToM was designed, the opportunity to rewatch the scene before answering the questions to be sure about the scene's content is critical. This way, we minimize the risk that low ToM test scores are due to cognitive dysfunction other than the ability of ToM. It may be argued that in real life, situations are only experienced once and the option of multiple watching may make the task less realistic. However, real life situations generally last longer and include more context than our scenes.

Finally, our study should not be limited to the development of the BASIT-ToM. We have generated a well-characterized set of dynamic ToM stimuli of varying difficulty from a sample of 240 healthy people with a large age range, evenly balanced in age and sex. Depending on the research question, type of participants and time available for testing, researchers can choose the stimuli and paradigms they favor.

Our study contains some limitations. First, due to the removal of three scenes as part of the scene selection procedure, we could not achieve a complete balance of the different ToM types within the selected scenes (i.e., first-order affective ToM: n = 14, first-order cognitive ToM: n = 12, second-order affective ToM: n=9, second-order cognitive ToM: n = 5). Nevertheless, there should be enough questions of the different types of ToM to analyze these types separately in the subsequent validation study. Second, in 6 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 behavioral-variant frontotemporal dementia (e.g., Bora et al., 2015; Henry et al., 2014) and healthy individuals. Lastly, in 4 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. Along the same line, each of these four questions seemed unclear, showing percentages of correct answers between 11% and 50%. Interestingly, all four questions refered to protagonists' affects ("Is he in love with her?", "Can they not stand Simone?"), whereas the hyperToM questions with an adequate model fir refered 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.

As a next step, we will validate the BASIT-ToM in a clinical sample to examine its sensitivity, reliability, construct and ecological validity and the potential influence of impaired cognitive functions 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. Lastly, the study will guide us in the final selection of BASIT-ToM scenes (and potentially questions) for later use in clinical settings

In conclusion, we have developed and preliminary validated the first German-language adaption of TASIT-SIM, called the BASIT-ToM, in healthy adults. The ultimate aim is for the BASIT-ToM, ideally in a shortened version, to become a core part of any comprehensive German-language neuropsychological exam in clinical routine.

References

Abu-Akel, A., & Bailey, A. L. (2000). Correspondence. *Psychological Medicine*, *30*(3), 735–738. https://doi.org/10.1017/S0033291799002123

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*. American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

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

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. https://doi.org/10.1037/amp0000191

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, and Psychiatry*, *86*(7), 714–719. https://doi.org/10.1136/jnnp-2014-309445

Canty, A. L., Neumann, D. L., Fleming, J., & Shum, D. H. K. (2015). *Neuropsychological Rehabilitation An International Journal Evaluation of a newly developed measure of theory of mind: The virtual assessment of mentalising ability*. https://doi.org/10.1080/09602011.2015.1052820

Canty, A. L., Neumann, D. L., & Shum, D. H. K. (2017). Using virtual reality to assess theory of mind subprocesses and error types in early and chronic schizophrenia. *Schizophrenia Research: Cognition*, *10*, 15–19. https://doi.org/10.1016/J.SCOG.2017.09.001

Castelli, I., Pini, A., Alberoni, M., Liverta-Sempio, O., Baglio, F., Massaro, D., Marchetti, A., & Nemni, R. (2011). Aging & Mental Health Mapping levels of theory of mind in Alzheimer's disease: A preliminary study Mapping levels of theory of mind in

Alzheimer's disease: A preliminary study. *Aging & Mental Health*, *15*(2), 157–168. https://doi.org/10.1080/13607863.2010.513038

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. https://doi.org/10.1016/J.CORTEX.2020.03.019

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

Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., Kessler, J., Woike, J. K., Wolf, O. T., & Convit, A. (2006). Introducing MASC: a movie for the assessment of social cognition. *Springer*, *36*, 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. https://doi.org/10.1016/J.NEUROL.2017.07.013

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

Frith, C. D. (2004). Schizophrenia and theory of mind. *Psychological Medicine*, *34*(3), 385–389. https://doi.org/10.1017/S0033291703001326

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*, 1–6. https://doi.org/10.1017/S1355617719001371

Ghosh, B. C. P., 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.

https://doi.org/10.1093/brain/aws128

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

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. https://doi.org/10.1038/nrneurol.2015.229

Jarsch, M., Bartsch, K., Berres, M., Ryf, I., Jurisic, K., Sluka, C., Monsch, A. U., Piguet, O., & Sollberger, M. (n.d.). The Basel Version of the Awareness of Social Inference Test – Emotion Recognition (BASIT-ER): Preliminary Validation Analyses in Healthy Adults. *Neuropsychology*.

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

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

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*(08), 896–900. https://doi.org/10.1017/S1355617719000584

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. https://doi.org/10.1111/j.1742-9544.2011.00054.x

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. https://doi.org/10.1111/jnp.12086

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.

Meijering, B., Van Rijn, H., Taatgen, N. A., & Verbrugge, R. (2011). *I Do Know What You Think I Think: Second-Order Theory Of Mind In Strategic Games Is Not That Difficult*.

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., … Yarkoni, T. (2015). SCIENTIFIC STANDARDS. Promoting an open research culture. *Science (New York, N.Y.)*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

Patel, G. H., Arkin, S. C., Ruiz-Betancourt, D. R., DeBaun, H. M., Strauss, N. E., Bartel, L. P., Grinband, J., Martinez, A., Berman, R. A., Leopold, D. A., & Javitt, D. C. (2020). What you see is what you get: visual scanning failures of naturalistic social scenes in schizophrenia. *Psychological Medicine*, 1–10. https://doi.org/10.1017/S0033291720001646

Peirce, J. W. (2007). PsychoPy--Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*, 10. https://doi.org/10.3389/neuro.11.010.2008

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

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. https://doi.org/10.1002/pchj.322

Quinting, J., Jonas, K., Wendt, C., & Stenneken, P. (2020). Social cognition in cognitive-communication disorders following traumatic brain injury: a video-based assessment. *Aphasie Und Verwandte Gebiete*, *48*(2).

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. https://doi.org/10.1016/J.NEUROIMAGE.2009.05.077

Rasch, G. (1960). Probabilistic Models for some Intelligence and Attainment Tests. In *Probabilistic Models for some Intelligence and Attainment Tests*. Paedagogike Institut, Copenhagen.

Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., van Swieten, J. C., Seelaar, H., Dopper, E. G. P., 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. https://doi.org/10.1093/brain/awr179

Rizopoulos, D. (2006). **ltm** : An *R* Package for Latent Variable Modeling and Item Response Theory Analyses. *Journal of Statistical Software*, *17*(5), 1–25. https://doi.org/10.18637/jss.v017.i05

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. https://doi.org/10.1080/87565641.2018.1529175

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. https://doi.org/10.1093/scan/nsx066

Sebastian, C. L., Fontaine, N. M. G., Bird, G., Blakemore, S.-J., De Brito, S. A., McCrory, E. J. P., & Viding, E. (2012). Neural processing associated with cognitive and affective Theory of Mind in adolescents and adults. *Social Cognitive and Affective Neuroscience*, *7*(1), 53–63. https://doi.org/10.1093/scan/nsr023

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., Harari, H., Aharon-Peretz, J., & Levkovitz, Y. (2010). The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies. *Cortex*, *46*(5), 668–677. https://doi.org/10.1016/J.CORTEX.2009.04.008

Sofologi, M., Koutsouraki, E., Tsolaki, M., Tsolaki, A., Koukoulidis, T., Theofilidis, A., Papantoniou, G., & Moraitou, D. (2019). Analyzing social cognition and understanding of social inferences in patients with multiple sclerosis. A comparative study. *Hellenic Journal of Nuclear Medicine*, *22 Suppl 2*, 15–26. http://www.ncbi.nlm.nih.gov/pubmed/31802039

Wang, Y., Wang, Y., Zou, Y., Ni, K., Tian, X., Sun, H., 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. https://doi.org/10.1016/J.SCHRES.2017.11.003

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. https://doi.org/10.1080/13803395.2014.966661

Yildirim, E., Soncu Buyukiscan, E., Demirtas-Tatlidede, A., Bilgiç, B., & Gurvit, H. (2020). An investigation of affective theory of mind ability and its relation to neuropsychological functions in Alzheimer's disease. *Journal of Neuropsychology*, *14*(3), 399–415. https://doi.org/10.1111/jnp.12207

Teil 3

5 Zusammenfassung der Erkenntnisse und Allgemeine Diskussion

Diese Dissertationsthesis zeigt, dass sowohl in der Forschung als auch in der klinischen Anwendung ein Konsens darüber besteht, dass die Untersuchung der EE- und ToM-Fähigkeiten sowohl für die bvFTD als auch für die AD Diagnostik relevant sein könnte. Informationen über individuelle sozial-kognitive Fähigkeiten könnten nicht nur für die Früh-, Differenzial- und Verlaufsdiagnostik von Nutzen sein, sondern auch wichtige Impulse für die Beratung und Behandlung von Betroffenen und ihren Betreuungspersonen liefern. Eine Voraussetzung für die standardmässige Untersuchung von EE- und ToM-Fähigkeiten in der Demenzdiagnostik ist das Vorhandensein von reliablen, validen, sensitiven und normierten Testverfahren, die zusätzlich von den MK/GS als zeitlich praktikabel und als passend für ihre Patient:innen eingeschätzt werden. Die explorative Onlineumfrage aus Manuskript 1 konnte erstmals zeigen, dass an deutschsprachigen MC/GS kaum sozial-kognitive Tests durchgeführt werden, obwohl die Diagnostiker:innen die Untersuchung sozial-kognitiver Fähigkeiten als wichtig erachteten. Die MC/GS erhielten die Möglichkeit Kritikpunkte an durchgeführten sozial-kognitiven Tests zu nennen oder zu begründen, warum sie keine sozial-kognitiven Tests an ihrer Einrichtung einsetzen. Diese Antworten bildeten die Basis für die Ausformulierung von neun Anforderungen, die von deutschsprachigen MC/GS an sozial-kognitive Tests gestellt werden. Die Anforderungen 1 bis 6 wurden in der aktuellen Phase der Entwicklung des BASIT-ER und des BASIT-ToM berücksichtigt, um dem deutschsprachigen Raum Testinstrumente zur Verfügung stellen zu können, die mit guten psychometrischen Eigenschaften an den MC/GS eingesetzt werden können. Abschliessend stellt sich die Frage, ob beide Tests die geforderten Anforderungen tatsächlich erfüllen oder ob weitere Entwicklungsschritte zur Erfüllung notwendig wären.

5.1 Anforderung (1): Untersuchung der EE-Fähigkeit von Basis- und Komplexen Emotionen

Der BASIT-ER wurde zur Untersuchung der menschlichen Fähigkeit Emotionen bei anderen Personen zu erkennen entwickelt. In der Entwicklungsphase ermöglichte ein kombiniertes Antwortformat aus geschlossenen und freien Antwortoptionen sowohl die Erfassung der Fähigkeit Basisemotionen (Multiple Choice Auswahl) als auch der Fähigkeit komplexe Emotionen (Freitextfeld) zu erkennen. Die Auswertung der Freitextfelder zeigte jedoch, dass sich der BASIT-ER auf die Untersuchung der Basisemotionen nach Paul Ekman beschränkt, da komplexe Emotionen äusserst selten genannt wurden (durchschnittlich 5% aller erkannten Emotionen, siehe Supplemental Material L des Manuskripts 2). Dieser Befund ist bei genauer Betrachtung jedoch wenig erstaunlich, da die BASIT-ER Filmszenen auf den TASIT-Filmszenen basieren, die explizit zur Untersuchung der EE-Fähigkeit von Basisemotionen entwickelt wurden.

Der Nachweis von Defiziten beim Erkennen von Basisemotionen würde gemäss Zinck & Newen (2008) auf eine verminderte EE-Fähigkeit hinweisen, die sich auf einer ontogenetisch früheren Entwicklungsstufe entwickelt hat. Demnach würden Personen die Defizite in der Erkennung von Basisemotionen zeigen, auch Defizite beim Erkennen komplexer Emotionen aufweisen. Defizite beim Erkennen von Basisemotionen wären demnach tiefgreifender und könnten im Gegensatz zu Defiziten beim Erkennen komplexer Emotionen als kognitiver Marker deutlich veränderter sozial-kognitiver Prozesse angesehen werden.

Wie valide und reliabel der BASIT-ER tatsächlich die EE-Fähigkeit von Testteilnehmer erfassen kann, wird sich allerdings erst in zukünftigen Validierungsstudien mit klinischen Populationen zeigen. Da der BASIT-ER vor allem für den Einsatz in der Demenzdiagnostik entwickelt wurde, empfehle ich Validierungsstudien unter Einbezug von Patientengruppen mit bvFTD und AD durchzuführen.

Nachfolgendend werden Empfehlungen zum weiteren Vorgehen bezüglich des Antwortformates und der Beurteilung der Validität sowie der Reliabilität des BASIT-ER beschrieben.

5.1.1 Empfehlungen zum Weiteren Vorgehen in Zukünftigen Validierungsstudien des BASIT-ER

5.1.1.1 Änderung des Antwortformats

In zukünftigen Validierungsstudien des BASIT-ER könnte aufgrund der niedrigen Benennungsrate komplexer Emotionen auf das freie Antwortformat verzichtet werden. Diese Änderung würde die Auswertung des BASIT für den klinischen Alltag deutlich praktikabler gestalten, da keine Auswertung eines Freitextfeldes erfolgen müsste. Man könnte allerdings argumentieren, dass eine Änderung des Antwortformates Einfluss auf die Schätzung der Schwierigkeiten der Testszenen mittels Rasch-Analysen haben könnte und sich nach Wegfall des freien Antwortformates die Szenenauswahl für den BASIT-ER verändern könnte. Meiner Meinung nach kann jedoch aufgrund der niedrigen Benennungsrate komplexer Emotionen davon ausgegangen werden, dass diese Änderung keinen Einfluss auf die Szenenauswahl haben wird.

Die Mehrfachauswahl der Basisemotionen sollte jedoch weiterhin ermöglicht werden, da diese zu einem zusätzlichen Informationsgewinn über die EE-Leistung der Testpersonen führen könnte. Aufgrund der Mehrfachauswahl könnte z. B. nicht nur festgestellt werden, ob Personen ein Defizit beim Erkennen der jeweils dargestellten Emotion aufweist, sondern auch ob es zu systematischen Verwechslungen von Emotionen kommt (z. B. systematische Verwechslung von Angst und Wut). Derartige Befunde eröffnen die Möglichkeit festzustellen, ob diese systematischen EE-Defizite und Verwechslungen für bestimmte demenzielle Syndrome spezifisch sind und somit für die Differenzialdiagnostik nutzbar gemacht werden könnten. Als Beispiel kann der Vergleich der die EE-Leistungen von bvFTD Patient:innen mit Patent:innen mit einer schweren unipolaren Depression (MDD) angeführt werden. Es konnte gezeigt werden, dass bvFTD

Patient:innen dazu neigen negative Emotionen zu verwechseln, während dieser Verwechslungsfehler bei den MDD Patient:innen nicht vorlag (Chiu et al., 2018).

5.1.1.2 Überprüfung der Konstruktvalidität des BASIT-ER

Die Beurteilung der Validität des BASIT-ER hat eine hohe Priorität, denn sie beantwortet die Frage, ob anhand des BASIT-ER tatsächlich die EE-Fähigkeit von Testteilnehmern erfasst werden kann. Im Folgenden werden konkrete Vorschläge zur Beurteilung verschiedener Aspekte der Konstruktvalidität genannt.

5.1.1.2.1 Konvergente Validität

Zur Überprüfung der konvergenten Validität des BASIT sollten Korrelationen zwischen den BASIT-ER Skalen und den Skalen von anderen EE-Tests berechnet werden. Beispielsweise könnten Korrelationsanalysen zwischen dem BASIT-ER und dem EE-Test des *Social Cognition and Emotional Assessment* (SEA; Bertoux et al., 2012) oder dem filmszenenbasierten Geneva Emotion Recognition Test erfolgen (GERT; Schlegel et al., 2014).

5.1.1.2.2 Diskriminante Validität

Die Überprüfung der diskriminanten Validität sollte anhand der Ermittlung von Korrelationen zwischen dem BASIT-ER und Test mit anderen Gültigkeitsbereichen stattfinden. Da bei der EE- und ToM-Fähigkeit weitgehend von unterscheidbaren Konstrukten ausgegangen wird, könnten Korrelationsanalysen zwischen den Skalen des BASIT-ER und des BASIT-ToM durchgeführt werden. Zusätzlich könnten die Zusammenhänge zwischen den Skalen des BASIT-ER und der ToM-Tests (z. B. MASC; Dziobek et al., 2006), die zur Untersuchung der konvergenten Validität des BASIT-ToM genutzt wurden, analysiert werden.

5.1.1.2.3 Faktorielle Validität

Zur Überprüfung der faktoriellen Validität sollte eine konfirmatorische Faktorenanalyse durchgeführt werden, um zu evaluieren, ob eine Unterscheidung verschiedener Basisemotionen anhand einer Faktorenstruktur gezeigt werden kann.

5.1.1.3 Überprüfung der Reliabilität des BASIT-ER

Die Überprüfung der Reliabilität des BASIT-ER soll der Frage nachgehen, wie hoch der Grad der Genauigkeit ist, mit welcher der BASIT-ER ein bestimmtes Merkmal misst, unabhängig davon, ob er tatsächlich die EE-Fähigkeit von Personen erfasst.

Die Entwicklung des BASIT-ER beruht auf den Prinzipien der Testkonstruktion der Item Response Theorie (IRT)[2]. Die Genauigkeit der Messung und die Trennschärfe der Items[3] wird in der Item Response Theorie über die sogenannte Iteminformationsfunktion[4] geschätzt (Kelava & Moosbrugger, 2020).

Jede Testfrage des BASIT-ER ist grundsätzlich dazu geeignet Unterschiede in den EE-Fähigkeiten der Testpersonen zu erfassen, dennoch kann es Unterschiede im Informationsgehalt der jeweiligen Testfragen geben. Je grösser die Iteminformationsfunktion ist, desto höher ist ihr Gewinn an Information über die EE-Fähigkeit einer Person, bzw. desto genauer misst die Testfrage die EE-Fähigkeit. Die jeweilige Gesamt-Testgenauigkeit des BASIT-ER wird über die Testinformationsfunktion angegeben, die sich aus der Aufsummierung der jeweiligen Iteminformationsfunktionen ergibt. Zusätzlich zu einer hohen Testgenauigkeit ist es wichtig, dass die Testfragen zur Unterscheidung zwischen Personen mit unterschiedlichen EE-Fähigkeiten geeignet sind.

---

[2] Eine ausführliche Beschreibung der Prinzipien der Item Response Theorie kann beispielsweise bei Kelava & Moosbrugger (2020) eingesehen werden.

[3] Ein Item entspricht im BASIT-ER und BASIT-ToM einer Testfrage

[4] Eine Erklärung wie die Iteminformationsfunktion berechnet wird, kann bei Kelava & Moosbrugger (2020) nachgelesen werden.

Deutliche Unterschiede in der Lösungswahrscheinlichkeit der Testfrage sind jedoch nur dann sichtbar, wenn die Schwierigkeiten der Testfragen im Bereich der jeweiligen Personenfähigkeit liegen. In der Entwicklungsphase des BASIT wurden Testszenen für Personen mit einer mittleren ER-Fähigkeit ausgewählt. Es ist jedoch durchaus denkbar, andere Testszenen auszuwählen, um bessere Trennschärfen für Personen mit einem deutlich niedrigeren oder höheren Leistungsniveau zu erreichen.

5.2 Anforderung (2): Untersuchung Verschiedener Konzepte der ToM-Fähigkeiten

Der BASIT-ToM wurde zur Untersuchung der menschlichen ToM-Fähigkeiten entwickelt. Der Test soll nicht nur eine Gesamtanalyse individueller ToM-Fähigkeiten, sondern auch eine differenzierte Beurteilung verschiedener ToM-Konzepte ermöglichen. Entsprechend wurden die Testfragen so konzipiert, dass eine getrennte Untersuchung kognitiver bzw. affektiver ToM-Fähigkeiten in jeweils erster und zweiter Ordnung möglich ist. Die getrennte Analyse verschiedener Aspekte der ToM soll in Zukunft helfen, störungsspezifische Muster der ToM-Defizite verschiedener demenzieller Syndrome zu identifizieren, um diese als kognitiven Marker für die Früh-, Differenzial- und Verlaufsdiagnostik nutzbar zu machen zu können.

Jedoch wird sich erst in den nachfolgenden Validierungsstudien zeigen, wie valide und reliabel der BASIT-ToM die verschiedenen ToM-Fähigkeiten (affektive, kognitive ToM in jeweils erster und zweiter Ordnung) untersucht und ob diese differenzierte Erfassung tatsächlich zur Unterscheidung verschiedener demenzieller Syndrome einen Beitrag leistet. Da der BASIT-ToM entsprechend dem BASIT-ER für den Einsatz in der Demenzdiagnostik entwickelt wurde, empfehle ich auch für die Validierungsstudien des BASIT-ToM den Einbezug von Patientengruppen mit bvFTD und AD.

Im Nachfolgenden werden Empfehlungen zur Beurteilung der Validität sowie der Reliabilität des BASIT-ToM beschrieben.

5.2.1 Empfehlungen zum Weiteren Vorgehen in Zukünftigen Validierungsstudien des BASIT-ToM

6.2.1.1 Überprüfung der Konstruktvalidität des BASIT.

Die Beurteilung der Validität, soll Aufschluss darüber geben, ob der BASIT-ToM tatsächlich die ToM-Fähigkeiten von Testteilnehmern erfasst. Im Folgenden werden konkrete Vorschläge zur Beurteilung verschiedener Aspekte der Konstruktvalidität genannt.

5.2.1.1.1 Konvergente Validität

Zur Überprüfung der konvergenten Validität des BASIT-ToM sollten Korrelationen zwischen den BASIT-ToM Skalen und den Skalen anderer ToM-Tests berechnet werden. Beispielsweise könnten Korrelationsanalysen mit dem BASIT-ToM und dem Movie for the Assesment of Social Cognition (MASC; Dziobek et al., 2006) durchgeführt werden, einem weiteren deutschsprachigen videobasierten ToM-Test. Zusätzlich könnten die Zusammenhänge zwischen dem BASIT-ToM und dem für die deutschsprachige Population normierten Reading the Eyes in the Mind Test (RMET; Kynast et al., 2021), sowie dem ToM Test des SEA (Bertoux et al., 2012) analysiert werden.

5.2.1.1.2 Diskriminante Validität

Die Überprüfung der diskriminanten Validität sollte anhand der Ermittlung von Korrelationen zwischen dem BASIT-ToM und Tests mit anderen Gültigkeitsbereichen stattfinden. Entsprechend dem vorgeschlagen Vorgehen der BASIT-ER Validierungsstudie, wären auch für die Beurteilung der diskriminanten Validität des BASIT-ToM, Korrelationsanalysen zwischen den Skalen des BASIT-ToM und des BASIT-ER sinnvoll. Zusätzlich könnten Korrelationen zwischen den BASIT-ToM Skalen und den Tests, die zur Überprüfung der konvergenten Validität des BASIT-ER genutzt wurden (z. B. GERT; Schlegel et al., 2014), berechnet werden.

5.2.1.1.3 Faktorielle Validität

Es sollte eine konfirmatorische Faktorenanalyse durchgeführt werden, um zu evaluieren, ob anhand einer Faktorenstruktur die Unterscheidung zwischen affektiver und kognitiver ToM gezeigt werden kann.

5.2.1.2 Überprüfung der Reliabilität des BASIT-ToM

Die Überprüfung der Reliabilität und der Trennschärfen des BASIT-ToM erfolgt analog zum Vorgehen beim BASIT-ER (siehe Abschnitt 6.1.1.3).

5.3 Anforderung (3): Hohe Sensitivität zur Detektion Subtiler Defizite in den EE- und ToM-Fähigkeiten

Derzeit stellt vor allem die Frühdiagnostik demenzieller Syndrome die MC/GS vor grosse Herausforderungen, da die standardmässig durchgeführten neuropsychologischen Untersuchungen häufig keine Unterscheidung zwischen den verschiedenen Mild Cognitive Impairment (MCI) - Subtypen zulassen (Dodich et al., 2021) und sich die Symptome möglicher zugrunde liegender Pathologien überschneiden können (Johnen & Bertoux, 2019). Ein potenzieller Grund für diese Problematik könnte sein, dass die etablierten neuropsychologischen Testverfahren nicht alle betroffenen Hirnregionen und neuronale Netzwerke abdecken, die kognitive Prozesse repräsentieren. Insbesondere das „Salience Network“ (Seeley et al., 2007) und das „Semantic Appraisal Network“ (Yeo et al., 2011), die mit sozial-kognitiven Fähigkeiten assoziiert werden (z. B. Christidi et al., 2018; Toller et al., 2018) und beispielsweise bei einer frontotemporalen Lobärdegeneration betroffen sind (Menon, 2011; Seeley et al., 2009), können nicht erfasst werden. Die Detektion subtiler Defizite der EE- und ToM-Fähigkeiten könnte diese Lücke schliessen und so einen wichtigen Beitrag zur Frühdiagnostik demenzieller Erkrankungen liefern. Sowohl subtile ToM- als auch EE-Defizite könnten Hinweise auf den MCI-Subtyp und eine mögliche Progression hin zur AD liefern (Dodich et al., 2021; McCade et al.,

2013; Michaelian et al., 2021; ZhongQuan et al., 2020). Zudem könnte die Erfassung subtiler sozial-kognitiver Defizite dazu beitragen, ein frühes Stadium der bvFTD von einer AD oder psychiatrischen Erkrankung abzugrenzen (Chiu et al., 2018; Cotter et al., 2016; Desmarais et al., 2018; Johnen & Bertoux, 2019). Derzeit sind die Testinstrumente zur Untersuchung der EE- und ToM- Fähigkeiten zwar häufig dazu geeignet deutliche Defizite zu detektieren, jedoch können subtile Defizite nicht erkannt werden (z. B. Deuse et al., 2016; Henry et al., 2015; Kessels et al., 2014). Hierfür könnte ein Grund sein, dass gesunde Probanden eine hohe EE- und ToM-Fähigkeit besitzen, was sich bei der Analyse ihres Antwortverhaltens in Deckeneffekten äussert (z. B. Deuse et al., 2016; Henry et al., 2015; Kessels et al., 2014). Dementsprechend war ein Ziel der BASIT-Testentwicklung Filmszenen mit Testfragen zu entwickeln, deren Beantwortung für gesunde Erwachsene mit durchschnittlichen EE- und ToM-Fähigkeiten weder zu schwer noch zu leicht ist. Indem jede Testszene in drei unterschiedlichen Intensitäten (niedrig, mittel, hoch) aufgenommen wurde, konnte mit Hilfe von Rasch-Analysen Szenen ausgewählt werden, bei denen sich mehrheitlich keine Deckeneffekte zeigten.

5.3.1 Empfehlungen zum Weiteren Vorgehen in Zukünftigen Validierungsstudien des BASIT

Die Validierungsstudien sollen die Sensitivität beider BASIT-Subtests untersuchen und mit der Sensitivität anderer EE- und ToM-Tests [z. B. SEA; (Bertoux et al., 2012)] vergleichen. Als Studienteilnehmer empfehle ich Patient:innen mit bvFTD in einem frühen und fortgeschrittenen Stadium, MCI Patient:innen sowie gesunde Probanden.

5.4 Anforderung (4): Testleistung soll Prädiktiv für die EE- und ToM-Leistungen in Realen Sozialen Interaktionen sein

Zur Beurteilung des Einflusses der EE- und ToM-Fähigkeiten auf soziale Verhaltensweisen und das psychosoziale Funktionsniveau, ist es notwendig, dass die

Leistungen im BASIT prädiktiv für die ToM- und EE-Fähigkeiten der Testpersonen in realen sozialen Interaktionen sind. Zudem könnte das Wissen um systematische Verwechslungen von Emotionen oder um spezifische EE- und ToM-Defizite für eine erfolgreiche Gestaltung der sozialen Beziehungen genutzt werden. Ein geschärftes Bewusstsein für sozial-kognitive Defizite könnte Missverständnissen vorbeugen und das Verständnis für inadäquate soziale Verhaltensweisen erhöhen. Zudem könnten gezielte Trainingsprogramme, sowohl für die Betroffenen als auch für deren Betreuungspersonen angeboten werden, um die zwischenmenschlichen Beziehungen zu stärken und das psychosoziale Funktionsniveau der Betroffenen zu erhalten bzw. zu verbessern.

Es ist wahrscheinlich, dass das Antwortverhalten auf realistische dynamische Teststimuli prädiktiver für die sozial-kognitiven Fähigkeiten in realen Situationen ist, als das Antwortverhalten, das durch statische Stimuli erzeugt wird (Westerhof-Evers et al., 2014). Diese Annahme wird durch Befunde gestützt, die zeigen, dass dynamische Stimuli eine erhöhte Aktivität in Hirnregionen auslösen können, die mit Gesichtserkennung, EE und emotionaler Verarbeitung assoziiert sind (z. B. Arsalidou et al., 2011; Risko et al., 2012). Der BASIT beinhaltet filmszenebasiertes, realitätsnahes Stimulusmaterial, das sowohl die visuellen als auch die auditiven sensorische Kanäle der Testpersonen ansprechen. Das Erkennen der dargestellten Emotionen bzw. das Ausbilden der korrekten ToM-Mentalisierungen im BASIT sollte somit den sozial-kognitiven Prozessen einer in der Realität beobachteten sozialen Interaktion ähneln.

5.4.1 Empfehlungen zum Weiteren Vorgehen in Zukünftigen Validierungsstudien des BASIT

Der prädiktive Nutzen der Leistungen im BASIT-ToM auf Sozialverhalten und das psychosoziale Funktionsniveau sollte untersucht werden. Für die Erfassung des sozialen Verhaltens könnte der speziell für neurodegenerativ bedingte demenzielle Syndrome entwickelte Behavioural Dysfunction Questionnaire (BDQ; Sollberger et al., 2021) und für

das psychosoziale Funktionsniveau die World Health Organization Disability Assessment Schedule 2 (WHODAS 2.0; American Psychiatric Association, 2013) verwendet werden.

5.5 Anforderung (5): Durchführbarkeit mit Personen, die Beeinträchtigungen in Mehreren Kognitiven Domänen Aufweisen

Personen mit neurodegenerativ bedingten dementiellen Syndromen weisen in der Regel in mehreren kognitiven Domänen Defizite auf und haben zusätzlich, aufgrund des meist höheren Lebensalters, körperliche Beeinträchtigungen. Diese kognitiven und physischen Einschränkungen können die Testdurchführung und somit auch die Testergebnisse erheblich beeinflussen.

Bei der Entwicklung des BASIT wurden daher diese möglichen Beeinträchtigungen explizit berücksichtigt. Zunächst wurde als Basis für die Testentwicklung der TASIT (McDonald et al., 2003) gewählt, da er als Stimulusmaterial kurze Filmszenen mit niedriger Komplexität beinhaltet. Zudem wurde eine Auswahl aus den TASIT-Filmszenen getroffen, um die Durchführungsdauer erheblich zu kürzen. Die Testpersonen des BASIT haben darüber hinaus die Möglichkeit die Filmszenen wiederholt zu betrachten, sodass der Einfluss weiterer kognitiver Fähigkeiten möglichst reduziert wird. Zudem bietet eine computerbasierte Testapplikation bei körperlichen Einschränkungen weitere Vorteile, wie beispielsweise die Adaption der Lautstärke oder der Schriftgrösse für Personen mit Hör- bzw. Sehschwäche. Darüber hinaus können sowohl die Testdurchführung des BASIT-ER als auch des BASIT-TOM assistiert erfolgen, z. B. in Form von Unterstützung bei der Navigation durch computerbasierte Applikationen oder mittels Vorlesen der Aufgabenstellung, Fragen und Antwortoptionen.

Neben der Problematik, dass kognitive Beeinträchtigung die Durchführbarkeit des eines Tests negativ beeinflussen könnten, ist die Untersuchung von Zusammenhängen der Testleistungen und der Leistungen in anderen kognitiven Bereichen (z. B. Exekutivfunktionen, Gedächtnisleistungen) für das Verständnis der Ursache eines Defizits

wichtig. Werden beispielsweise deutliche Zusammenhänge zwischen den sozial-kognitiven Fähigkeiten und anderen kognitiven Fähigkeiten (z. B. Exekutivfunktionen) gefunden, könnte dies ein Hinweis darauf sein, dass es sich bei der sozial-kognitiven Dysfunktion eher um ein sekundäres als um ein primäres Defizit handelt [z. B. hoher Einfluss exekutiver Funktionen auf ToM-Leistungen von AD Patient:innen (siehe Fortier et al., 2018)] .

In den nachfolgenden Validierungsstudien sollte überprüft werden, in wie weit die Leistungen der Testpersonen in den Subtests des BASIT mit den Fähigkeiten in anderen kognitiven Bereichen zusammenhängen.

5.5.1 Empfehlungen zum Weiteren Vorgehen in Zukünftigen Validierungsstudien des BASIT

Zur Messung der kognitiven Fähigkeiten und der Analyse möglicher Zusammenhänge mit dem BASIT, empfehle ich die Durchführung der deutschsprachigen CERAD-Plus Testbatterie (Consortium to Establish a Registry for Alzheimer's Disease; Monsch, 2015), die ein reliables, valides, sensitives und normiertes Testinstrument zur differenzierten Untersuchung der fünf neurokognitiven Domänen nach DSM-5 (American Psychiatric Association, 2013) Sprache, Lernen und Gedächtnis, komplexe Aufmerksamkeit, Exekutivfunktion und perzeptuell-motorische Funktion darstellt.

Störungsspezifische Befunde zu den Zusammenhängen zwischen EE- und ToM-Fähigkeiten und anderen kognitiven Fähigkeiten, könnten unter anderem wichtige Hinweise zu einer aktuellen Diskussion im Bereich der neurodegenerativ-bedingten Demenzen beitragen. Während im Bereich der bvFTD vermutet wird, dass es sich bei der sozial-kognitiven Dysfunktion um ein primäres Defizit handelt, wird bei der AD eher ein sekundäres Defizit vermutet (siehe Übersichtsarbeit Fortier et al., 2018).

5.6 Anforderung (6): Praktikable Testdauer

Ein Test, der sehr viel Zeit in Anspruch nimmt, kann nur schwer in die klinische Routine integriert werden. Die explorative Onlineumfrage aus Manuskript 1 bestätigt die Annahme, dass eine angemessene Durchführungsdauer für die MC/GS wichtig ist. Die befragten MC/GS gaben an, dass sie für ein ausführliches Screeninginstrument eine maximale Durchführungsdauer von 20 Minuten bevorzugen würden.

Die Beantwortung der BASIT-Filmszenen nehmen bei gesunden Probanden in etwa 1-2 Minuten in Anspruch. Somit beträgt die geschätzte Testdauer für den BASIT-ER 44 Minuten (22 Szenen x 2 Minuten) und für den BASIT-ToM 20 Minuten (10 Szenen x 2 Minuten). Bei Patient:innen mit demenziellen Syndromen kann davon ausgegangen werden, dass die Durchführungsdauer aufgrund weiterer kognitiver Defizite deutlich erhöht ist, was unter anderem zu mehrmaligem Betrachten der Testszenen führen könnte.

5.6.1 Empfehlungen zum Weiteren Vorgehen in Zukünftigen Validierungsstudien des BASIT

Im Rahmen der Validierungsstudien könnte die Testdauer weiter verkürzt werden, indem eine erneute Selektion der Szenen stattfindet. Als Basis für die Szenenauswahl könnten die Iteminformationsfunktionen, die Testinformationsfunktionen und die Kennzahlen der Diskriminationsfähigkeit der Subtests des BASIT für verschiedene Störungsbilder dienen. Dabei könnte die Auswahl und Anzahl der Testszenen je nach den zu untersuchenden Störungsbildern variiert werden. Darüber hinaus hätte ein solches Vorgehen den Vorteil, dass störungsspezifische Testsets zusammengestellt werden könnten, was sich deutlich positiv auf die Validität des BASIT-ER bzw. des BASIT-ToM auswirken könnte.

5.7 Normierungsstudie

Für die Beurteilung der individuellen EE- und ToM-Leistungen der Testteilnehmer ist es von zentraler Bedeutung, sowohl den BASIT-ER als auch den BASIT-ToM, unter Berücksichtigung demografischer Variablen wie Alter, Geschlecht und Ausbildungsjahre, für die deutschsprachige Population zu normieren.

5.8 Mögliche Weiterentwicklungen des BASIT

Der BASIT, mit seinen Subtests BASIT-ER und BASIT-ToM, ist eine Testbatterie, die mögliche Weiterentwicklungen über die ursprünglich vorgesehenen Einsatzgebiete hinaus zulässt. Im Folgenden werden vier weitere Einsatzmöglichkeiten besprochen.

5.8.1 Erweiterung des BASIT um die Untersuchung Sozialen Verhaltens zu Ermöglichen

In der Literatur existieren verschiedene Modellvorstellungen über einen positiven Zusammenhang von sozialer Kognition, sozialem Verhalten und psychosozialem Funktionsniveau (z. B. Keifer et al., 2020; Kennedy & Adolphs, 2012). In der Demenzforschung haben sich jedoch zwei nahezu getrennte Forschungszweige herauskristallisiert, die soziale Kognition und soziales Verhalten getrennt voneinander untersuchen. Während sich die einen Studien auf Veränderungen sozial-kognitiver Fähigkeiten und deren Bedeutung für die Diagnostik, Beratung und Therapie demenzieller Syndrome konzentrieren (z. B. Christidi et al., 2018; Fortier et al., 2018; Henry et al., 2016), fokussieren sich andere Studien auf Verhaltensauffälligkeiten bei demenziellen Syndromen (z. B. Taragano et al., 2018; Volicer, 2018).

In den letzten Jahren wurden vor allem zwei Bereiche ermittelt, für die eine Untersuchung von Verhaltensauffälligkeiten einen Nutzen haben könnte. Bereich 1 beinhaltet die Frühdiagnostik demenzieller Erkrankungen: Die Beobachtung, dass Verhaltensauffälligkeiten sich häufig schon manifestieren bevor sich die ersten kognitiven

Defizite bei Personen mit demenziellen Syndromen zeigen (Liu et al., 2004), hat zu der Überlegung geführt, Verhaltensauffälligkeiten als frühen Marker für kognitiven Abbau zu nutzen (Ismail et al., 2016; Taragano et al., 2018). Daher wurde das Syndrom des Mild behavioral Impairment (MBI) beschrieben, das erworbene, anhaltende und bedeutsame neuropsychologische Störungen im späteren Leben, von beliebigem Schweregrad beinhaltet, aber nicht notwendigerweise zusammen mit kognitiven Defiziten auftreten muss (Ismail et al., 2016; Mortby et al., 2018). Befunde zeigen, dass das Vorliegen eines MBI als Risikofaktor für die Entstehung eines demenziellen Syndroms angesehen werden kann (z. B. Ismail et al., 2017; Johansson et al., 2021; Taragano et al., 2018).
Bereich zwei stellt das Vorliegen von Verhaltensauffälligkeiten bei demenzerkrankten Personen und die daraus resultierenden Auswirkungen auf die Betreuungspersonen in den Vordergrund. Demenzerkrankte Personen leben häufig viele Jahre zu Hause und benötigen von ihren Betreuungspersonen kontinuierliche Unterstützung. Die Verhaltensauffälligkeiten demenzerkrankter Personen werden als Hauptrisikofaktor für eine hohe psychische Belastung der Betreuungspersonen demenzerkrankter Personen betrachtet (z. B. Deardorff & Grossberg, 2019; Poveda et al., 2017).

Es stellt sich nun die Frage wie der BASIT für eine gemeinsame Untersuchung sozial-kognitiver Fähigkeiten und Sozialverhalten nutzbar gemacht werden kann. Die Kriterien eines MBI anhand des BASIT zu überprüfen wird kaum möglich sein. Jedoch könnte untersucht werden, wie sich die EE- und ToM-Fähigkeiten, die im BASIT gemessen werden, unmittelbar auf Verhaltensentscheidungen der Testteilnehmer auswirken. Beispielsweise könnten hierfür mehrere zusätzliche Aussagen pro Testszene entwickelt werden, die sich auf das weitere mögliche Verhalten der Protagonisten der Filmszenen beziehen und entweder angemessenes oder unangemessenes Sozialverhalten widerspiegeln. Nach jeder Aussage würden die Testteilnehmer dann aufgefordert werden sich zu entscheiden, ob ein adäquates oder inadäquates Verhalten dargestellt wurde.

Die Ergebnisse könnten genutzt werden, um ein gezieltes soziales Kompetenztraining und kognitives Training vorzunehmen, das diejenigen sozialen Verhaltensweisen trainiert, die von den Betroffenen nur noch eingeschränkt gezeigt werden können. Das differenzierte Wissen um spezifische sozial-kognitive Defizite und daraus resultierenden Defiziten im Sozialverhalten könnte darüber hinaus den Betreuungspersonen helfen Missverständnisse zu vermeiden und die Akzeptanz für inadäquates Sozialverhalten zu erhöhen. Zudem könnten auch für die Betreuungspersonen Schulungseinheiten angeboten werden, die gezielt den Umgang mit den vorliegenden sozial-kognitiven Defiziten und inadäquaten Verhaltensweisen adressiert.

5.8.2 Einsatz für Verschieden Fähigkeitsniveaus

In der Entwicklungsphase des BASIT wurden für jede Szene drei verschiedene Versionen generiert, in denen die jeweils dargestellte Emotion bzw. die Art der Botschaft in unterschiedlicher Intensität (niedrig, mittel, hoch) dargestellt wurde. Alle Szenen wurden an einer Stichprobe von 240 gesunden Teilnehmern mit einer grossen Altersspanne, die in Bezug auf Alter und Geschlecht ausbalanciert wurde, getestet. Anschliessend wurde für jede Szene Rasch-Analysen durchgeführt, die Aussagen über die Schwierigkeit der jeweiligen Testszene zuliessen. Durch dieses Vorgehen ist ein umfangreiches Set mit Testszenen in unterschiedlichen Schwierigkeiten entstanden. Abhängig von der jeweiligen Forschungsfrage, Art der Testteilnehmer und zur Verfügung stehender Zeit, können Forscher:innen und Kliniker:innen diejenigen BASIT-Szenen auswählen, die sie für ihre Untersuchungen favorisieren. Es könnte zum Beispiel sinnvoll sein, Personen mit niedrigen ToM- und EE-Fähigkeiten sehr einfache Szenen vorzulegen, um zu verhindern, dass sich bei schweren Szenen Bodeneffekte zeigen, wodurch eine differenzierte Beurteilung der Leistungen nicht mehr möglich wäre. Ebenso wäre es vorstellbar, Personen mit überdurchschnittlichen Fähigkeiten schwere Testszenen zu zeigen, um Deckeneffekte bei zu leichten Testszenen zu verhindern.

5.8.3 Differenzierte Untersuchung Verschiedener Sinneskanäle

Die Filmszenen des BASIT sprechen mehrere Sinneskanäle der Testpersonen an. Die EE- und ToM-Leistungen der Testpersonen basieren somit auf der gemeinsamen Interpretation der auditiven und visuellen Signale. Die computerbasierte BASIT-Applikation eröffnet deshalb auch die Möglichkeit, die EE- und ToM-Fähigkeiten basierend auf einer getrennten Analyse der sensorischen Signale untersuchen zu können. So ist es beispielsweise möglich Probanden die Filmszenen ohne den zugehörigen Ton zu zeigen, oder ausschliesslich die Tonspur abzuspielen.

5.8.4 Einsatz bei Anderen Erkrankungen

Die Testszenen des BASIT wurden explizit für eine grosse Altersspanne konzipiert (35 bis >75 Jahre) und eigenen sich deshalb nicht nur für die Untersuchung der EE- und ToM-Fähigkeiten von Personen mit demenziellen Erkrankungen, sondern können auch bei anderen Störungsbildern eingesetzt werden, deren Betroffene durchschnittlich jünger sind. Beispielsweise wäre der Einsatz bei Patient:innen mit Schizophrenie gut vorstellbar, bei denen EE- und ToM-Defizite, deren Einfluss auf das individuelle psychosoziale Funktionsniveau und deren prädiktiven Nutzen für den Erkrankungsverlauf ausführlich dokumentiert sind (z. B. Fett et al., 2011; Maat et al., 2012; Mike et al., 2019; Patel et al., 2020; Savla et al., 2013). Zudem sind für diese Patientengruppe Hypermentalisierungstendenzen gut dokumentiert (z. B. siehe Bliksted et al., 2019), die ebenfalls mit dem BASIT-ToM untersucht werden könnten. Im englischsprachigen Raum wird der TASIT häufig zur Untersuchung der sozial-kognitiven Fähigkeiten von Schizophrenie-Patient:innen eingesetzt (Rapp et al., 2018), da er als dynamischer Test eine Untersuchung sozial-kognitiver Fähigkeiten anhand sozial-relevanter, realistisch konzipierter Testszenen ermöglicht (Patel et al., 2020). Der Einsatz des BASIT könnte die Untersuchung der EE- und ToM-Fähigkeiten von Schizophrenie-Patient:innen anhand dynamischer, realistischer Stimuli im deutschsprachigen Raum fördern und zudem die Vergleichbarkeit internationaler Studien begünstigen.

6 Fazit

Das Ziel den BASIT als ein praktikables, sensitives, valides, reliables und normiertes Testinstrument für den Einsatz in der Demenzdiagnostik des deutschsprachigen Raumes zu entwickeln, ist noch nicht erreicht, befindet sich jedoch auf einem guten Weg. Die Testentwicklung hat innovative Ansätze verwendet, wie zum Beispiel die Generierung von verschiedenen Intensitäten bzw. Schwierigkeiten einer Testszene, die Konzeptualisierung als ersten realitätsnahen deutschsprachigen dynamischen EE-Test und die Optimierung aller Komponenten des Original-TASIT, abgestimmt auf die spezifischen Anforderungen der deutschsprachigen MC/GS und der evidenzbasierten Erkenntnisse aus der Demenzforschung. Ich bin der Überzeugung, dass der BASIT nach Abschluss weiterer Validierungs- und Normierungsstudien einen erheblichen Nutzen für die Diagnostik demenzieller Erkrankungen liefern wird.

8 Literaturverzeichnis

Abu-Akel, A., & Bailey, A. L. (2000). Correspondence. *Psychological Medicine*, *30*(3), 735–738. https://doi.org/10.1017/S0033291799002123

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*. American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

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

Arsalidou, M., Morris, D., & Taylor, M. J. (2011). Converging Evidence for the Advantage of Dynamic Facial Expressions. *Brain Topography*, *24*(2), 149–163. https://doi.org/10.1007/s10548-011-0171-4

Bertoux, M., Delavest, M., de Souza, L. C., Funkiewiez, A., Lépine, J.-P., Fossati, P., Dubois, B., & Sarazin, M. (2012). Social Cognition and Emotional Assessment differentiates frontotemporal dementia from depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, *83*(4), 411–416. https://doi.org/10.1136/jnnp-2011-301849

Bertsch, K., & Herpertz, S. C. (2021). Neurobiologische Grundlagen der Borderline-Störung: eine Integration in das ICD-11-Modell der Persönlichkeitsstörungen. *Der Nervenarzt*, *92*(7), 653–659. https://doi.org/10.1007/s00115-021-01133-w

Bliksted, V., Frith, C., Videbech, P., Fagerlund, B., Emborg, C., Simonsen, A., Roepstorff, A., & Campbell-Meiklejohn, D. (2019). Hyper- and Hypomentalizing in Patients with First-Episode Schizophrenia: fMRI and Behavioral Studies. *Schizophrenia Bulletin*, *45*(2), 377–385. https://doi.org/10.1093/schbul/sby027

Braak, H., & Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathologica*, *82*(4), 239–259. https://doi.org/10.1007/BF00308809

Brüne, M., & Juckel, G. (2010). Soziale Kognition bei Schizophrenien Mentalisierung und psychosoziales Funktionsniveau. *Nervenarzt*, *81*, 339–346.

https://doi.org/10.1007/s00115-009-2841-y

Canty, A. L., Neumann, D. L., Fleming, J., & Shum, D. H. K. (2015). *Neuropsychological Rehabilitation An International Journal Evaluation of a newly developed measure of theory of mind: The virtual assessment of mentalising ability*. https://doi.org/10.1080/09602011.2015.1052820

Castelli, I., Pini, A., Alberoni, M., Liverta-Sempio, O., Baglio, F., Massaro, D., Marchetti, A., & Nemni, R. (2011). Aging & Mental Health Mapping levels of theory of mind in Alzheimer's disease: A preliminary study Mapping levels of theory of mind in Alzheimer's disease: A preliminary study. *Aging & Mental Health*, *15*(2), 157–168. https://doi.org/10.1080/13607863.2010.513038

Chiu, I., Piguet, O., Diehl-Schmid, J., Riedl, L., Beck, J., Leyhe, T., Holsboer-Trachsler, E., Kressig, R. W., Berres, M., Monsch, A. U., & Sollberger, M. (2018). Facial Emotion Recognition Performance Differentiates Between Behavioral Variant Frontotemporal Dementia and Major Depressive Disorder. *The Journal of Clinical Psychiatry*, *79*(1), 16m11342. https://doi.org/10.4088/JCP.16m11342

Christidi, F., Migliaccio, R., Santamaría-García, 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., & Drake, R. J. (2016). Social cognition in multiple sclerosis: A systematic review and meta-analysis. *Neurology*, *87*(16), 1727. 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 & Biobehavioral Reviews*, *84*, 92–99. https://doi.org/10.1016/J.NEUBIOREV.2017.11.014

De Lucena, A. T., Bhalla, R. K., Belfort Almeida Dos Santos, T. T., & Dourado, M. C. N. (2020). The relationship between theory of mind and cognition in Alzheimer's

disease: A systematic review. *Journal of Clinical and Experimental Neuropsychology*, *42*(3), 223–239. https://doi.org/10.1080/13803395.2019.1710112

Deardorff, W. J., & Grossberg, G. T. (2019). Behavioral and psychological symptoms in Alzheimer's dementia and vascular dementia. *Handbook of Clinical Neurology*, *165*, 5–32. https://doi.org/10.1016/B978-0-444-64012-3.00002-2

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

Deuse, L., Rademacher, L. M., Winkler, L., Schultz, R. T., Gründer, G., & Lammertz, S. E. (2016). Neural correlates of naturalistic social cognition: brain-behavior relationships in healthy adults. *Social Cognitive and Affective Neuroscience*, *11*(11), 1741–1751. https://doi.org/10.1093/scan/nsw094

Dodich, A., Crespi, C., Santi, G. C., Cappa, S. F., & Cerami, C. (2020). Evaluation of Discriminative Detection Abilities of Social Cognition Measures for the Diagnosis of the Behavioral Variant of Frontotemporal Dementia: a Systematic Review. *Neuropsychology Review*, 1–16. https://doi.org/10.1007/s11065-020-09457-1

Dodich, A., Crespi, C., Santi, G. C., Marcone, A., Iannaccone, S., Perani, D., Cappa, S. F., & Cerami, C. (2021). Emotion Recognition Deficits in the Differenzial Diagnosis of Amnestic Mild Cognitive Impairment: A Cognitive Marker for the Limbic-Predominant Phenotype. *Journal of the International Neuropsychological Society*, 1–7. https://doi.org/10.1017/S1355617721000254

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. *Springer*, *36*, 623–636.

Ekman, P. (1992). Are there basic emotions? *Psychological Review*, *99*(3), 550–553. https://doi.org/10.1037/0033-295X.99.3.550

Ekman, P., & Friesen, W. (1976). *Pictures of facial affect*. Consulting Psychologists Press.

Ferretti, V., & Papaleo, F. (2018). Understanding others: emotion recognition abilities in

humans and other animals. *Genes, Brain and Behavior*, *18*(1), e12544. https://doi.org/10.1111/gbb.12544

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

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. https://doi.org/10.1016/J.NEUROL.2017.07.013

Frith, C. D., & Frith, U. (2007). Social Cognition in Humans. *Current Biology*, *17*(16), R724–R732. https://doi.org/10.1016/J.CUB.2007.05.068

Garcia-Cordero, I., Migeot, J., Fittipaldi, S., Aquino, A., Campo, C. G., García, A., & Ibáñez, A. (2021). Metacognition of emotion recognition across neurodegenerative diseases. *Cortex*, *137*, 93–107. https://doi.org/10.1016/J.CORTEX.2020.12.023

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

Henry, J. D., 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., & 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

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. https://doi.org/10.1038/nrneurol.2015.229

Hoffmann, S., & Akbar, P. (2019). Emotion. In *Konsumentenverhalten* (pp. 53–69).

Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-23567-3_4

Ismail, Z., Agüera-Ortiz, L., Brodaty, H., Cieslak, A., Cummings, J., Fischer, C. E., Gauthier, S., Geda, Y. E., Herrmann, N., Kanji, J., Lanctôt, K. L., Miller, D. S., Mortby, M. E., Onyike, C. U., Rosenberg, P. B., Smith, E. E., Smith, G. S., Sultzer, D. L., Lyketsos, C., & NPS Professional Interest Area of the International Society of to Advance Alzheimer's Research and Treatment (NPS-PIA of ISTAART). (2017). The Mild Behavioral Impairment Checklist (MBI-C): A Rating Scale for Neuropsychiatric Symptoms in Pre-Dementia Populations. *Journal of Alzheimer's Disease : JAD*, *56*(3), 929–938. https://doi.org/10.3233/JAD-160979

Ismail, Z., Smith, E. E., Geda, Y., Sultzer, D., Brodaty, H., Smith, G., Agüera-Ortiz, L., Sweet, R., Miller, D., & Lyketsos, C. G. (2016). Neuropsychiatric symptoms as early manifestations of emergent dementia: Provisional diagnostic criteria for mild behavioral impairment. *Alzheimer's & Dementia*, *12*(2), 195–202. https://doi.org/10.1016/j.jalz.2015.05.017

Johansson, M., Stomrud, E., Insel, P. S., Leuzy, A., Johansson, P. M., Smith, R., Ismail, Z., Janelidze, S., Palmqvist, S., van Westen, D., Mattsson-Carlgren, N., & Hansson, O. (2021). Mild behavioral impairment and its relation to tau pathology in preclinical Alzheimer's disease. *Translational Psychiatry*, *11*(1), 76. https://doi.org/10.1038/s41398-021-01206-z

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

Keifer, C. M., Mikami, A. Y., Morris, J. P., Libsack, E. J., & Lerner, M. D. (2020). Prediction of social behavior in autism spectrum disorders: Explicit versus implicit social cognition. *Autism : The International Journal of Research and Practice*, *24*(7), 1758–1772. https://doi.org/10.1177/1362361320922058

Kelava, A., & Moosbrugger, H. (2020). Einführung in die Item-Response-Theorie (IRT). In

Testtheorie und Fragebogenkonstruktion (pp. 369–409). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-61532-4_16

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

Kessels, R. P. C., Montagne, B., Hendriks, A. W., Perrett, D. I., & de Haan, E. H. F. (2014). Assessment of perception of morphed facial expressions using the Emotion Recognition Task: Normative data from healthy participants aged 8-75. *Journal of Neuropsychology*, *8*(1), 75–93. https://doi.org/10.1111/jnp.12009

Kessels, R. P. C., Waanders-Oude Elferink, M., & Tilborg, I. (2021). Social cognition and social functioning in patients with amnestic mild cognitive impairment or Alzheimer's dementia. *Journal of Neuropsychology*, *15*(2), 186–203. https://doi.org/10.1111/jnp.12223

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

Klein-Koerkamp, Y., Beaudoin, M., Baciu, M., & Hot, P. (2012). Emotional Decoding Abilities in Alzheimer's Disease: A Meta-Analysis. *Journal of Alzheimer's Disease*, *32*(1), 109–125. https://doi.org/10.3233/JAD-2012-120553

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

Kynast, J., Polyakova, M., Quinque, E. M., Hinz, A., Villringer, A., & Schroeter, M. L. (2021). Age- and Sex-Specific Standard Scores for the Reading the Mind in the Eyes Test. *Frontiers in Aging Neuroscience*, *12*, 529. https://doi.org/10.3389/fnagi.2020.607107

Lawrence, K., Campbell, R., & Skuse, D. (2015). Age, gender, and puberty influence the

development of facial emotion recognition. *Frontiers in Psychology*, *6*, 761. https://doi.org/10.3389/fpsyg.2015.00761

Liu, W., Miller, B. L., Kramer, J. H., Rankin, K., Wyss-Coray, C., Gearhart, R., Phengrasamy, L., Weiner, M., & Rosen, H. J. (2004). Behavioral disorders in the frontal and temporal variants of frontotemporal dementia. *Neurology*, *62*(5), 742–748. https://doi.org/10.1212/01.wnl.0000113729.77161.c9

Maat, A., Fett, A.-K., & Derks, E. (2012). Social cognition and quality of life in schizophrenia. *Schizophrenia Research*, *137*(1–3), 212–218. https://doi.org/10.1016/J.SCHRES.2012.02.017

McCade, D., Savage, G., Guastella, A., Hickie, I. B., Lewis, S. J. G., & Naismith, S. L. (2013). Emotion Recognition in Mild Cognitive Impairment: Relationship to Psychosocial Disability and Caregiver Burden. *Journal of Geriatric Psychiatry and Neurology*, *26*(3), 165–173. https://doi.org/10.1177/0891988713491832

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. A., & Verbrugge, R. (2011). *I Do Know What You Think I Think: Second-Order Theory Of Mind In Strategic Games Is Not That Difficult*.

Menon, V. (2011). Large-scale brain networks and psychopathology: a unifying triple network model. *Trends in Cognitive Sciences*, *15*(10), 483–506. https://doi.org/10.1016/J.TICS.2011.08.003

Michaelian, J. C., Duffy, S. L., Mowszowski, L., Guastella, A. J., McCade, D., McKinnon, A. C., & Naismith, S. L. (2021). Poorer Theory of Mind in Amnestic Mild Cognitive Impairment Is Associated with Decreased Functional Connectivity in the Default Mode Network. *Journal of Alzheimer's Disease*, *81*(3), 1079–1091. https://doi.org/10.3233/JAD-201284

Mike, L., Guimond, S., Kelly, S., Thermenos, H., Mesholam-Gately, R., Eack, S., & Keshavan, M. (2019). Social cognition in early course of schizophrenia: Exploratory

factor analysis. *Psychiatry Research*, *272*, 737–743. https://doi.org/10.1016/J.PSYCHRES.2018.12.152

Monsch, A. U. (2015). Case-finding in der hausärztlichen Praxis und neuropsychologische Diagnostik an einer Memory Clinic. *Therapeutische Umschau*, *72*(4), 213–217. https://doi.org/10.1024/0040-5930/a000667

Mortby, M. E., Black, S. E., Gauthier, S., Miller, D., Porsteinsson, A., Smith, E. E., & Ismail, Z. (2018). Dementia clinical trial implications of mild behavioral impairment. *International Psychogeriatrics*, *30*(2), 171–175. https://doi.org/10.1017/S1041610218000042

Nelis, S. M., Clare, L., Martyr, A., Markova, I., Roth, I., Woods, R. T., Whitaker, C. J., & Morris, R. G. (2011). Awareness of social and emotional functioning in people with early-stage dementia and implications for carers. *Aging & Mental Health*, *15*(8), 961–969. https://doi.org/10.1080/13607863.2011.575350

Patel, G. H., Arkin, S. C., Ruiz-Betancourt, D. R., DeBaun, H. M., Strauss, N. E., Bartel, L. P., Grinband, J., Martinez, A., Berman, R. A., Leopold, D. A., & Javitt, D. C. (2020). What you see is what you get: visual scanning failures of naturalistic social scenes in schizophrenia. *Psychological Medicine*, 1–10. https://doi.org/10.1017/S0033291720001646

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*(4), 813–823. https://doi.org/10.1093/schbul/sbt081

Plutchik, R. (1991). *The emotions*. https://books.google.com/books?hl=en&lr=&id=JaQauznPoiEC&oi=fnd&pg=PP5&ots=EFyg9H9kGc&sig=TdeXDwUtkdKucAUgh7ITb6Vkg3U

Plutchik, Robert. (1962). *The emotions: Facts, theories and a new model*. Crown Publishing Group/Random House.

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

Poveda, B., Osborne-Crowley, K., Laidlaw, K., Macleod, F., & Power, K. (2017). Social Cognition, Behaviour and Relationship Continuity in Dementia of the Alzheimer Type. *Brain Impairment*, *18*(2), 175–187. https://doi.org/10.1017/BrImp.2016.35

Rapp, A., Purr, F., & Felsenheimer, A. (2018). T58. SARCASM COMPREHENSION AS A SOCIAL COGNITION MEASURE IN SCHIZOPHRENIA – A SYSTEMATIC LITERATURE SEARCH AND META-ANALYSIS ON THE USE OF THE TASIT. *Schizophrenia Bulletin*, *44*(suppl_1), S136–S136. https://doi.org/10.1093/schbul/sby016.334

Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., van Swieten, J. C., Seelaar, H., Dopper, E. G. P., 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. https://doi.org/10.1093/brain/awr179

Risko, E. F., Laidlaw, K., Freeth, M., Foulsham, T., & Kingstone, A. (2012). Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. *Frontiers in Human Neuroscience*, *6*, 143. https://doi.org/10.3389/fnhum.2012.00143

Savla, G. N., Vella, L., Armstrong, C. C., Penn, D. L., & Twamley, E. W. (2013). Deficits in Domains of Social Cognition in Schizophrenia: A Meta-Analysis of the Empirical Evidence. *Schizophrenia Bulletin*, *39*(5), 979–992. https://doi.org/10.1093/schbul/sbs080

Schlegel, K., Grandjean, D., & Scherer, K. R. (2014). Introducing the Geneva emotion recognition test: an example of Rasch-based test development. *Psychological Assessment*, *26*(2), 666–672. https://doi.org/10.1037/a0035246

Sebastian, C. L., Fontaine, N. M. G., Bird, G., Blakemore, S.-J., De Brito, S. A., McCrory, E. J. P., & Viding, E. (2012). Neural processing associated with cognitive and affective Theory of Mind in adolescents and adults. *Social Cognitive and Affective Neuroscience*, *7*(1), 53–63. https://doi.org/10.1093/scan/nsr023

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*(1), 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., Reiss, A. L., & Greicius, M. D. (2007). *Behavioral/Systems/Cognitive Dissociable Intrinsic Connectivity Networks for Salience Processing and Executive Control*. https://doi.org/10.1523/JNEUROSCI.5587-06.2007

Shamay-Tsoory, S. G., Harari, H., Aharon-Peretz, J., & Levkovitz, Y. (2010). The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies. *Cortex*, *46*(5), 668–677. https://doi.org/10.1016/J.CORTEX.2009.04.008

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology*, *52*, 1061–1086. https://d1wqtxts1xzle7.cloudfront.net/41383936/Emotion_Knowledge_Further_Exploration_of20160121-10163-jqeiwm-with-cover-page-v2.pdf?Expires=1636293166&Signature=CBVfkJBtQYCbqEJ8jP~U~dqPcdVmGu2nPSki5aL8cnEGEwaQaSaovIR3afNnzGOrscHyg0xwvQUsqkvBP4AaxNxng1S5IT

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (2001). Emotional Knowledge: Further Exploration of a Prototype Approach. In G. Parrot (Ed.), *Emotions in social psychology : essential readings* (pp. 26–56). Philadelphia: Taylor & Francis.

Sollberger, M., Warnecke, T., Berres, M., Leyhe, T., Mülhauser, M., Baumgartner, M., Felbecker, A., Hägele-Link, S., & Johnen, A. (2021). *Standardisation and Validation Study of the Behavioural Dysfunction Questionnaire for Use in Early*

Neurodegenerative Disease Syndromes. Alzheimer Schweiz. https://www.alzheimer-schweiz.ch/fileadmin/dam/Alzheimer_Schweiz/Dokumente/Ueber_Demenz/Forschung/ftd_diagnose_projektinformationen.pdf

Taragano, F. E., Allegri, R. F., Heisecke, S. L., Martelli, M. I., Feldman, M. L., Sánchez, V., García, V. A., Tufro, G., Castro, D. M., Leguizamón, P. P., Guelar, V., Ruotolo, E., Zegarra, C., & Dillon, C. (2018). Risk of Conversion to Dementia in a Mild Behavioral Impairment Group Compared to a Psychiatric Group and to a Mild Cognitive Impairment Group. *Journal of Alzheimer's Disease : JAD*, *62*(1), 227–238. https://doi.org/10.3233/JAD-170632

Timmer, M. L., Jacobs, B., Schonherr, M. C., Spikman, J. M., & van der Naalt, J. (2020). The Spectrum of Long-Term Behavioral Disturbances and Provided Care After Traumatic Brain Injury. *Frontiers in Neurology*, *11*, 246. https://doi.org/10.3389/fneur.2020.00246

Toller, G., Brown, J., Sollberger, M., Shdo, S. M., Bouvet, L., Sukhanov, P., Seeley, W. W., Miller, B. L., & 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

Volicer, L. (2018). Behavioral Problems and Dementia. *Clinics in Geriatric Medicine*, *34*(4), 637–651. https://doi.org/10.1016/j.cger.2018.06.009

Wang, Y., Wang, Y., Zou, Y., Ni, K., Tian, X., Sun, H., 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. https://doi.org/10.1016/J.SCHRES.2017.11.003

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.

https://doi.org/10.1080/13803395.2014.966661

World Health Organization. (n.d.). *ICD-11 for Mortality and Morbidity Statistics*. 05/2021. Retrieved November 7, 2021, from https://icd.who.int/browse11/l-m/en#/http%3A%2F%2Fid.who.int%2Ficd%2Fentity%2F546689346

Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fisch, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*(3), 1125–1165. https://doi.org/10.1152/jn.00338.2011

ZhongQuan, Y., PanWen, Z., Hui, Z., YuanYuan, S., HaiChun, S., JianGuo, Z., & PingLei, P. (2020). Theory of mind in Alzheimer's disease and amnestic mild cognitive impairment: a meta-analysis. *Neurological Sciences*, *41*(5), 1027–1039. https://doi.org/10.1007/s10072-019-04215-5

Zinck, A., & Newen, A. (2008). Classifying emotion: a developmental account. *Synthese*, *161*(1), 1–25. https://doi.org/10.1007/s11229-006-9149-2

Anhang

Curriculum Vitae

Persönliche Daten:

Name:	Marianne Somaya Jarsch (geb. Glock)
Geburtsdatum:	26. Januar 1987
Geburtsort:	Friedrichshafen (Deutschland)
Staatsangehörigkeit:	deutsch

Schulische Ausbildung:

1993-1997	Grundschule am Kirchplatz / Sulgen (Deutschland)
1997-2002	Gymnasium Schramberg (Deutschland)
2002-2003	Gymnasium Überlingen (Deutschland)
2003-2006	Justus von Liebig Schule – Ernährungswissenschaftliches Gymnasium / Überlingen (Deutschland)
	Abitur im Juni 2006 (Note: 1,8)

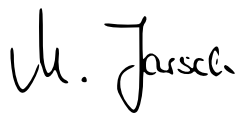
Berufliche Ausbildung / Studium:

09/2006 - 09/2007	Bachelorstudium Soziale Arbeit (Katholische Hochschule Freiburg) (Deutschland)
09/2007 - 09/2010	Ausbildung zur Gesundheits- und Krankenpflegerin (Universitätsklinikum Freiburg) (Deutschland)
	Examen im August 2010
09/2011 - 09/2014	Bachelorstudium Psychologie (Universität Ulm) (Deutschland)
	Bachelor of Science im September 2014 (Note:1,7)

02/2015 – 02/2018	Masterstudium Psychologie (Universität Basel) Vertiefungsrichtung: Klinische Psychologie und Neurowissenschaften Master of Science im Februar 2018 (Note: 5,7)
seit 02/2018	Doktorat an der Fakultät für Psychologie der Universität Basel / Universitäre Altersmedizin FELIX PLATTER, Memory Clinic Basel Betreuung: Prof. Dr. phil. Andreas U. Monsch Thema: Die Entwicklung der Basel Version of The Awareness of Social Inference Test (BASIT): Ein Innovatives Instrument zur Erfassung Sozial-Kognitiver Fähigkeiten

Berufliche Tätigkeit

Seit 03/2019	Wissenschaftliche Mitarbeiterin (Doktorandin) in der Memory Clinic der Universitären Altersmedizin FELIX PLATTER, Basel
01/2011 -06/2011	Gesundheits- und Krankenpflegerin in der Onkologie des Universitätsklinikums Freiburg (Deutschland)
12/2011 -08/2014	Gesundheits- und Krankenpflegerin in der Hämato-Onkologie und in der Psychosomatik des Universitätsklinikums Ulm (Deutschland)



Basel, den 30.11.2021

Universität Basel

Fakultät für Psychologie

Die Entwicklung der Basel Version of The Awareness of Social Inference Test (BASIT): Ein Innovatives Instrument zur Erfassung Sozial-Kognitiver Fähigkeiten

Band II von II

Zusatzmaterial

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

Marianne Somaya Jarsch

aus Friedrichshafen, Deutschland

Basel, 2021

Universität Basel

Fakultät für Psychologie

Genehmigt von der Fakultät für Psychologie auf Antrag von

Prof. Dr. phil. Andreas U. Monsch

Prof. Dr. phil. Klaus Opwis

Datum des Doktoratsexamen: 27.01.2022

Prof. Dr. phil. Jens Gaab

Dekan der Fakultät für Psychologie

Universität Basel

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:

- Jarsch, M., Monsch, A. U., Kressig, R. W., & Sollberger, M. (2021). Eine Lücke, die es zu schließen gilt: Die Untersuchung sozial-kognitiver Fähigkeiten an deutschsprachigen Memory Clinics und Gedächtnissprechstunden. *Zeitschrift für Neuropsychologie*. Manuscript submitted.

- Jarsch, M., Bartsch, K., Berres, M., Ryf, I., Jurisic, K., Sluka, C., Kressig, R. W., Monsch, A. U., McDonald, S., Piguet, O., & Sollberger, M. (2021). The Basel Version of the Awareness of Social Inference Test – Emotion Recognition (BASIT-ER): Preliminary Validation Analyses in Healthy Adults. *Neuropsychology*. https://doi.org/10.1037/neu0000786. Manuskript in press.

- Jarsch, M., Piguet, O., Berres, M., Sluka, C., Kressig, R. W., Monsch, A. U., McDonald, S., & Sollberger, M. (2021). The Basel Version of the Awareness of Social Inference Test – Theory of Mind (BASIT-ToM): Preliminary Validation Analyses in Healthy Adults. *Neuropsychology*. Manuscript submitted.

Basel, den 30.11.2021

M. Jarsch

Marianne Somaya Jarsch

Inhaltsverzeichnis

Manuskript 1: Elektronisches Zusatzmaterial

Eine Lücke die es zu schließen gilt: Die Untersuchung sozial-kognitiver Fähigkeiten an deutschsprachigen Memory Clinics und Gedächtnissprechstunden **1**

Manuskript 2: Supplementary Material:

The Basel Version of The Awareness of Social Inference Test – Emotion Recognition (BASIT – ER): Preliminary Validation Analyses in Healthy Adults **23**

Manuskript 3: Supplementary Material:

The Basel Version of The Awareness of Social Inference Test – Theory of Mind (BASIT – ToM): Preliminary Validation Analyses in Healthy Adults **132**

Elektronisches Zusatzmaterial

Manuskript 1

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

Inhaltsverzeichnis

Elektronisches Zusatzmaterial 1 3

Zusätzliche Informationen über die Auswahl- und Teilnahmebedingungen der Studienteilnehmer 3

Elektronisches Zusatzmaterial 2 4

Druckversion des Fragebogens 4

Elektronisches Zusatzmaterial 3 19

3.1 Einschätzung der diagnostischen Relevanz verschiedener sozial-kognitiven Fähigkeiten 19

3.2 Angabe von Patientengruppen mit unterschiedlichen Erkrankungen, für die eine Untersuchung sozial-kognitiver Fähigkeiten relevant sein könnte 19

3.3 Gründe keine Emotionserkennungstests durchzuführen, die von jeweils einer Institution genannt wurden 19

3.4 Statische Bildmaterial, das zur Untersuchung der Emotionserkennung verwendet wurde 20

3.5 Vor- und Nachteile der durchgeführten Emotionserkennungstests, die von weniger als 20% der Institutionen im Freitextfeld genannt wurden 20

3.6 Gründe keine Theory of Mind-Tests durchzuführen, die von weniger als 20% der Institutionen genannt wurden: 21

3.7 Statische Bildmaterial, das zur Untersuchung der Theory of Mind-Fähigkeiten verwendet wurde 21

3.8 Vor- und Nachteile der durchgeführten Theory of Mind-Tests, die von weniger als 20% der Institutionen im Freitextfeld genannt wurden 22

Elektronisches Zusatzmaterial 1

Zusätzliche Informationen über die Auswahl- und Teilnahmebedingungen der Studienteilnehmer

Für Deutschland wurden die Kontaktdaten über die Webseite der *Deutschen Alzheimer Gesellschaft e.V.* (DAlzG) (https://www.deutsche-alzheimer.de/unser-service/gedaechtnissprechstunden.html) ermittelt. Die Kontaktdaten der schweizerischen MK/GS wurden über die Webseite des Vereins *Swiss Memory Clinics* (https://www.swissmemoryclinics.ch/de/home/mitglieder) akquiriert, die österreichischen Kontaktdaten 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 aufgrund fehlender oder fehlerhafter E-Mail-Kontaktdaten bereinigt.

Alle Teilnehmenden gaben im Vorfeld eine informierte Einwilligung zur freiwilligen Teilnahme an der Umfrage. Die Institutionen erhielten keine Aufwandsentschädigung.

Elektronisches Zusatzmaterial 2

In dieser Umfrage sind 23 Fragen enthalten.

Herzlich Willkommen!

Bevor Sie an dieser Umfrage teilnehmen, möchten wir Sie bitten die nachfolgenden Informationen aufmerksam zu lesen.

Mit Ihrer Teilnahme an unserem Fragebogen haben Sie die Möglichkeit uns Ihre Erfahrungen mit Tests zur Untersuchung sozial kognitiver Fähigkeiten mitzuteilen. Sie können uns darüber informieren welche Testanforderung erfüllt sein sollten, damit Sie einen neu entwickelten Test in Ihre klinische Routine integrieren würden. Mit Ihrer Teilnahme leisten Sie einen substanziellen Beitrag zur Entwicklung eines validen und praktikablen Tests zur Untersuchung sozial kognitiver Fähigkeiten.

Ablauf des Fragebogens:

Abhängig von Ihrem Antwortverhalten werden Ihnen maximal 22 Fragen mit unterschiedlichen Antwortformaten präsentiert. Dem blauen Hinweistext können Sie entnehmen, ob Sie eine oder mehrere Antwortoptionen auswählen können.

Falls Sie eine Antwort abändern möchten, können Sie jederzeit zu einer vorherigen Frage zurückkehren, indem Sie auf „zurück“ klicken. Die Fragen können jedoch nicht übersprungen werden, mit Ausnahme der Fragen mit offenem Antwortformat.

Möchten Sie zu einem späteren Zeitpunkt mit der Beantwortung des Fragebogens fortfahren, haben Sie die Möglichkeit zur Zwischenspeicherung.

Die Beantwortung des Fragebogens dauert etwa 15 Minuten.

Freiwilligkeit der Teilnahme und Abbruchsrecht:

Die Teilnahme an der Umfrage ist freiwillig. Sie können die Teilnahme jederzeit ohne Angabe von Gründen beenden. Durch einen Abbruch entstehen Ihnen keine Nachteile. Sie haben Anspruch darauf über die Ergebnisse der Umfrage informiert zu werden. Kontaktieren Sie hierfür bitte die Studienleiterin Marianne Jarsch (marianne.jarsch@unibas.ch). Sie erhalten die Kontaktinformationen erneut am Ende des Fragebogens.

Aufklärung über den Datenschutz:

Alle im Verlauf der Studie erhobenen Daten werden streng vertraulich behandelt, anonymisiert erhoben und anonymisiert weiterverarbeitet. Es sind keine Rückschlüsse auf Ihre Institution möglich. Die Daten werden in anonymisierter Form auf einem geschützten Server des psychologischen Instituts der Universität Basel für 10 Jahre gespeichert (DNS: "survey.psychologie.unibas.ch"). Die Beachtung des Datenschutzstandards der Europäischen Union ist in vollem Umfang gewährleistet. Die Ergebnisse der Umfrage werden ohne jede Bezugsmöglichkeit auf Ihre Institution voraussichtlich in medizinisch-psychologischen Fachzeitschriften und einer open science Plattform veröffentlicht. Wir weisen Sie darauf hin, dass aufgrund der Durchführung der Studie in einem Onlineportal keine Garantie für den Schutz personengebundener Daten gemäß DSGVO gegeben werden kann. Durch die Onlineerhebung ergeben sich Risiken, dass Daten durch Drittanbieter abgegriffen werden könnten. Durch den Einsatz von Verschlüsselungstechniken und die Speicherung auf einem geschützten Server reduziert sich dieses Risiko, ein Restrisiko bleibt jedoch bestehen.

Risiken und Nebenwirkungen

Es werden keine Risiken und Belastungen für die Teilnehmer erwartet, die durch das Ausfüllen der der Umfrage entstehen.

Aufwandsentschädigung

Für Ihre Teilnahme erhalten Sie keine Aufwandsentschädigung.

*****

❗ Bitte wählen Sie eine der folgenden Antworten:
Bitte wählen Sie nur eine der folgenden Antworten aus:

◯ Ich habe die Teilnahmeinformationen der Umfrage vollständig gelesen und verstanden. Mit einer Teilnahme an der Umfrage erkläre ich mich mit den oben stehenden Informationen einverstanden.

Wie relevant finden Sie die Untersuchung verschiedener sozial kognitiver Fähigkeiten im Rahmen der Diagnostik kognitiver Störungen?

Eine Definition der aufgelisteten sozial kognitiven Domänen finden Sie im untenstehenden Hilfetext.

*

Bitte wählen Sie die zutreffende Antwort für jeden Punkt aus:

	nicht relevant	eher nicht relevant	eher relevant	relevant
Emotionserkennung	◯	◯	◯	◯
Gezielter Umgang mit wahrgenommenen Emotionen	◯	◯	◯	◯
Soziale Wahrnehmung	◯	◯	◯	◯
Theory of Mind (Perspektivübernahme)	◯	◯	◯	◯
Attributionsstil	◯	◯	◯	◯

Unter sozial kognitiven Fähigkeiten kann man alle informationsverarbeitenden Prozesse verstehen, die notwendig sind, um angemessen mit anderen Personen zu interagieren und zu kommunizieren (Vogeley, Schilbach & Newen, 2014).

Emotionserkennung: Emotionen wahrnehmen und identifizieren.

Gezielter Umgang mit wahrgenommenen Emotionen: Umgang, der das persönliche Wachstum und soziale Beziehungen fördert.

Soziale Wahrnehmung: Verarbeitung von sozialem Kontext und sozialem Wissen, das definiert werden kann als das Kennen, Verstehen und Nutzen sozialer Regeln, Rollen und Zielen.

Theory of Mind: (Perspektivübernahme): Fähigkeit, die mentalen Zustände anderer zu repräsentieren und sie von den eigenen mentalen Zuständen abzugrenzen.

Attributionsstil: Die Art und Weise, wie Individuen die Ursachen sozialer Ereignisse oder Interaktionen erklären oder ihnen Sinn geben.

Gibt es weitere sozial kognitive Fähigkeiten, deren Untersuchung Sie für relevant erachten?

Falls "ja", bennnen Sie diese bitte im Freitextfeld.

Bitte wählen Sie nur eine der folgenden Antworten aus:

◯ ja

◯ nein

Bitte schreiben Sie einen Kommentar zu Ihrer Auswahl

Werden an Ihrer Institution Emotionserkennungstests durchgeführt?

Falls ja, benennen Sie diese(n) bitte im Freitextfeld.

Bitte wählen Sie nur eine der folgenden Antworten aus:

◯ Ja

◯ Nein

Bitte schreiben Sie einen Kommentar zu Ihrer Auswahl

Falls an Ihrer Institution keine quantitativen Tests durchgeführt, sondern subjektive Befragungen vorgenommen werden (z.B. Patientenbefragung oder Angehörigenbefragung), beantworten Sie die Frage bitte mit "nein".

Wie zufrieden sind Sie mit den Emotionserkennungstests, die an Ihrer Institution durchgeführt werden?

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:
Antwort war 'Ja' bei Frage '4 (Werden an Ihrer Institution Emotionserkennungstests durchgeführt? Falls ja, benennen Sie diese(n) bitte im Freitextfeld.)

Bitte wählen Sie die zutreffende Antwort für jeden Punkt aus:

	unzufrieden	eher unzufrieden	neutral	eher zufrieden	zufrieden
	◯	◯	◯	◯	◯

Bitte listen Sie die Vorteile und Nachteile der Emotionserkennungstests auf, die sie an Ihrer Institution durchführen.

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:
Antwort war 'Ja' bei Frage '4 (Werden an Ihrer Institution Emotionserkennungstests durchgeführt? Falls ja, benennen Sie diese(n) bitte im Freitextfeld.)

Bitte geben Sie Ihre Antwort hier ein:

Vor- und Nachteile können sich auch auf praktische Aspekte beziehen (z.B. Patientenpassung, Art der Durchführung, Durchführungsdauer etc.)

Bitte geben Sie Gründe an, warum an Ihrer Institution keine Emotionerkennungstests durchgeführt werden.

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:
Antwort war 'Nein' bei Frage '4 (Werden an Ihrer Institution Emotionserkennungstests durchgeführt? Falls ja, benennen Sie diese(n) bitte im Freitextfeld.)

Bitte wählen Sie alle zutreffenden Antworten aus:

- [ ] Die uns bekannten Tests sind für Patienten mit Störungen in mehreren kognitiven Domänen oftmals ungeeignet.
- [ ] Die Durchführung uns bekannter Emotionserkennungstests dauert zu lange.
- [ ] Wir würden gerne subtile Defizite in der Fähigkeit Emotionen zu erkennen detektieren. Die uns bekannten Tests sind jedoch wenig sensitiv.
- [ ] Die Testergebnisse haben keine Relevanz für die Therapie.
- [ ] Wir haben es bis jetzt nicht in Erwägung gezogen, die Fähigkeit zur Emotionserkennung zu untersuchen.
- [ ] Sonstiges: _____

Werden an Ihrer Institution Fähigkeiten zur Emotionserkennung subjektiv erfasst? *

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:
Antwort war 'Nein' bei Frage '4 (Werden an Ihrer Institution Emotionserkennungstests durchgeführt? Falls ja, benennen Sie diese(n) bitte im Freitextfeld.)

Bitte wählen Sie alle zutreffenden Antworten aus:

- [ ] Nein
- [ ] Ja, durch Patientenbefragung
- [ ] Ja, durch Angehörigenbefragung
- [ ] Sonstiges: _____

Werden an Ihrer Institution Theory of Mind (Perspektivübernahme) Tests durchgeführt? Falls ja, benennen Sie diese(n) bitte im Freitextfeld.

Bitte wählen Sie nur eine der folgenden Antworten aus:

◯ ja

◯ nein

Bitte schreiben Sie einen Kommentar zu Ihrer Auswahl

Falls an Ihrer Institution keine quantitativen Tests durchgeführt, sondern subjektive Befragungen vorngenommen werden (z.B. Patientenbefragung oder Angehörigenbefragung), beantworten Sie die Frage bitte mit "nein".

Wie zufrieden sind Sie mit den Theory of Mind Tests, die an Ihrer Institution durchgeführt werden? *

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:
Antwort war 'ja' bei Frage '9 (Werden an Ihrer Institution Theory of Mind (Perspektivübernahme) Tests durchgeführt? Falls ja, benennen Sie diese(n) bitte im Freitextfeld.)

Bitte wählen Sie die zutreffende Antwort für jeden Punkt aus:

	unzufrieden	**eher unzufrieden**	**neutral**	**eher zufrieden**	**zufrieden**
	◯	◯	◯	◯	◯

Bitte listen Sie die Vorteile und Nachteile der Theory of Mind Tests auf, die Sie an Ihrer Institution durchführen.

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:
Antwort war 'ja' bei Frage '9 (Werden an Ihrer Institution Theory of Mind (Perspektivübernahme) Tests durchgeführt? Falls ja, benennen Sie diese(n) bitte im Freitextfeld.)

Bitte geben Sie Ihre Antwort hier ein:

Vor- und Nachteile können sich auch auf praktische Aspekte beziehen (z.B. Patientenpassung, Art der Durchführung, Durchführungsdauer etc.)

Bitte geben Sie Gründe an, warum an Ihrer Institution keine Theory of Mind Tests durchführt werden.

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:
Antwort war 'nein' bei Frage '9 (Werden an Ihrer Institution Theory of Mind (Perspektivübernahme) Tests durchgeführt? Falls ja, benennen Sie diese(n) bitte im Freitextfeld.)

Bitte wählen Sie alle zutreffenden Antworten aus:

- [ ] Die uns bekannten Tests sind sind für Patienten mit Störungen in mehreren kognitiven Domänen oftmals ungeeignet.
- [ ] Die Durchführung uns bekannter Theory of Mind Tests dauert zu lange.
- [ ] Wir würden gerne subtile Defizite in der Fähigkeit eine Theory of Mind auszubilden detektieren. Die uns bekannten Tests sind jedoch wenig sensitiv.
- [ ] Die Testergebnisse haben keine Relevanz für die Therapie.
- [ ] Wir haben es bis jetzt nicht in Erwägung gezogen, Theory of Mind Fähigkeiten zu untersuchen.
- [ ] Sonstiges:

Werden an Ihrer Institution Theory of Mind Fähigkeiten subjektiv erfasst?

Beantworten Sie diese Frage nur, wenn folgende Bedingungen erfüllt sind:
Antwort war 'nein' bei Frage '9 (Werden an Ihrer Institution Theory of Mind (Perspektivübernahme) Tests durchgeführt? Falls ja, benennen Sie diese(n) bitte im Freitextfeld.)

Bitte wählen Sie alle zutreffenden Antworten aus:

- [ ] Nein
- [ ] Ja, durch Patientenbefragung
- [ ] Ja, durch Angehörigenbefragung
- [ ] Sonstiges:

Welche maximale Durchführungsdauer eines Tests zur Untersuchung sozial kognitiver Fähigkeiten halten Sie für praktikabel?

Bitte wählen Sie die zutreffende Antwort für jeden Punkt aus:

	max. 10 min	max. 20 min	max. 30 min	max. 40 min	max. 50 min	max. 1h	über 1h
Screening-Tests	◯	◯	◯	◯	◯	◯	◯
Ausführliche diagnostische Tests	◯	◯	◯	◯	◯	◯	◯

Wie wichtig finden Sie es, dass die Testleitung einem Patienten bei der Durchführung eines Tests assisstieren kann (z.B. Vorlesen der Fragen, Hilfe bei der Navigation durch eine computerbasierte Applikation)?

Bitte wählen Sie die zutreffende Antwort für jeden Punkt aus:

	unwichtig	eher unwichtig	eher wichtig	wichtig
	◯	◯	◯	◯

Wie wichtig finden Sie es, dass die Stimuli eines Tests (z.B. Bilder oder Filmszenen), der sozial kognitive Fähigkeiten untersucht, möglichst realitätsnah sind?

Bitte wählen Sie die zutreffende Antwort für jeden Punkt aus:

	unwichtig	eher unwichtig	eher wichtig	wichtig
	◯	◯	◯	◯

Welches Antwortformat würden Sie für einen Test, der sozial kognitive Fähigkeiten untersucht, bevorzugen?

Bitte wählen Sie nur eine der folgenden Antworten aus:

◯ offen (Freitextfeld)

◯ geschlossen, nur eine Antwort kann ausgewählt werden

◯ geschlossen, mehrere Antworten können ausgewählt werden

◯ Kombination aus offenem und geschlossenem Antwortformat

Wie wichtig finden Sie es, dass ein Test nicht nur deutliche sondern auch leichte Defizite in sozial kognitiven Fähigkeiten detektieren kann?

Bitte wählen Sie die zutreffende Antwort für jeden Punkt aus:

	unwichtig	eher unwichtig	eher wichtig	wichtig
	◯	◯	◯	◯

Bitte geben Sie alle Patientengruppen an, bei denen Sie gerne sozial kognitive Fähigkeiten im Rahmen der neurokognitiven Diagnostik untersuchen würden.

Bitte wählen Sie alle zutreffenden Antworten aus:

☐ Patienten mit neurodegenerativen Erkrankungen (z.B. Frontotemporale Demenz, Alzheimer Demenz, Amyotrophe Lateralsklerose)

☐ Patienten mit neuroinflammatorischen Erkrankungen (z.B. Multiple Sklerose)

☐ Patienten mit traumatisch verursachten Hirnleistungsstörungen (z.B. Schädelhirntrauma)

☐ Patienten mit vaskulär verursachten Hirnleistungsstörungen (z.B. Schlaganfall)

☐ Patienten mit Entwicklungsstörungen (z.B. Autismus)

☐ Patienten mit psychiatrischen Erkrankungen (z.B. Schizophrenie, Depression)

☐ Bei keiner Patientengruppe

☐ Sonstiges: _____

An wie vielen Patienten pro Monat werden an Ihrer Institution neuropsychologische Untersuchungen durchgeführt?

Bitte wählen Sie nur eine der folgenden Antworten aus:

- ◯ wir führen keine neuropsychologischen Untersuchungen durch
- ◯ 1 - 4 pro Monat
- ◯ 5 - 20 pro Monat
- ◯ 20 - 40 pro Monat
- ◯ 40 - 60 pro Monat
- ◯ über 60 pro Monat
- ◯ keine Angabe

Bitte nennen Sie uns Ihre Berufsgruppe. Falls mehrere Personen an der Beantwortung des Fragebogens beteiligt waren, geben Sie bitte alle Berufsgruppen an.

Bitte wählen Sie alle zutreffenden Antworten aus:

- ☐ Neuropsychologe / Neuropsychologin
- ☐ Psychologe / Psychologin
- ☐ Psychologischer Psychotherapeut / Psychologische Psychotherapeutin
- ☐ Facharzt / Fachärztin für Neurologie
- ☐ Facharzt / Fachärztin für Psychiatrie und Psychotherapie
- ☐ Arzt / Ärztin in Weiterbildung
- ☐ Sonstiges: _____

Wie hoch schätzen Sie die Wahrscheinlichkeit ein, dass Sie an Ihrer Institution einen neu entwickelten Test, der sozial kognitive Fähigkeiten praktikabel, valide, reliabel und realitätsnah untersucht, einsetzen?

Bitte wählen Sie die zutreffende Antwort für jeden Punkt aus:

	unwahrscheinlich	eher unwahrscheinlich	eher wahrscheinlich	wahrscheinlich
	◯	◯	◯	◯

Möchten Sie uns gerne noch etwas mitteilen? Wir freuen uns über Ihre Kommentare und Anregungen.

Bitte geben Sie Ihre Antwort hier ein:

Herzlichen Dank für Ihre Teilnahme!

Bei Interesse an den Umfrageergebnissen oder Fragen und Problemen freuen wir uns über Ihre Kontaktaufnahme mit der Studienleiterin Marianne Jarsch.

marianne.jarsch@unibas.ch

Vielen Dank für die Beantwortung des Fragebogens.

Elektronisches Zusatzmaterial 3

3.1 Einschätzung der diagnostischen Relevanz verschiedener sozial-kognitiven Fähigkeiten

Auswertung des Freitextfeldes

11.5% der Institutionen (n=10) gaben die Untersuchung von verhaltensbezogenem Outcome sozial-kognitiver Fähigkeiten (z.B. Sozialverhalten, psychosoziale Funktionsniveau, Compliance bei Unterstützungsangeboten, Delinquenz) an und 4.6% (n=4) die Untersuchung der Empathiefähigkeit.

3.2 Angabe von Patientengruppen mit unterschiedlichen Erkrankungen, für die eine Untersuchung sozial-kognitiver Fähigkeiten relevant sein könnte

72% (n=59) der Institutionen befürworteten sozial-kognitive Untersuchungen bei Patient:innen mit vaskulär verursachten Hirnleistungsstörungen, 67.1% (n=55) bei psychiatrischen Erkrankungen, 58.5% (n=48) bei traumatisch verursachten Hirnleistungsstörungen, 47.6% (n=39) bei Entwicklungsstörungen und 33.0% (n=27) bei neuroinflammatorischen Erkrankungen. Im Freitextfeld wurde von den Institutionen keine weitere Störungsgruppen benannt, für die sie sozial-kognitive Untersuchungen als relevant erachteten.

3.3 Gründe keine Emotionserkennungstests durchzuführen, die von jeweils einer Institution genannt wurden

Von einer Institution (4.0%) ausgewähltes Optionsfeld:

- Die Testergebnisse haben keine Relevanz für die Therapie

Gründe, die im Freitextfeld von jeweils einer Institution (4.0%) benannt wurden:

- Vorhandene Tests sind nicht valide genug
- Fehlende Normierung und Validierung der vorhandenen EE-Tests für Einzelfalldiagnostik
- Vorhandene Tests liefern zu wenig diagnostische Informationen
- Es existieren zu wenige Tests und es gibt keine Bezugsadressen
- Standardisierte Testverfahren sind nicht notwendig, eine subjektive Beurteilung ist genauer,

- Die Untersuchung der EE-Fähigkeit spielt für die Demenzdiagnostik eine untergeordnete Rolle
- Der Schwerpunkt der Institution liegt auf psychosozialen Interventionen.

3.4 Statische Bildmaterial, das zur Untersuchung der Emotionserkennung verwendet wurde

41.7% (n=5) der Institutionen gaben zur Untersuchung der EE an, Ekman Faces (Ekman & Friesen, 1976) zu verwenden. Eine Institution (8.3%) gab an Fotos mit emotionalen Gesichtsausdrücken zu verwenden, die sie jedoch nicht näher bezeichnete.

Literatur:

Ekman, P., & Friesen, W. V. (1976). Pictures of facial affect. *Consulting Psychologists Press*.

3.5 Vor- und Nachteile der durchgeführten Emotionserkennungstests, die von weniger als 20% der Institutionen im Freitextfeld genannt wurden

Vorteile:

- EckmanFaces/statische Bildmaterial eignen sich im klinischen Kontext als Screeninginstrument [2 Institutionen (16.7%)]
- EckmanFaces/statische Bildmaterial führen im Forschungskontext zu einer guten Einschätzung der EE-Fähigkeit [1 Institution(8.3%)]
- Akzeptable Durchführungsdauer der Eckman Faces im klinischen Kontext [1Institution (8.3%)].

Nachteile:

- Keine Validierung für spezifische Patientengruppen [1 Institution (8.3%)]
- Fragwürdige Validität [1 Institution (8.3%)]
- Sprachlich zu komplex / Patienten müssen lesen können [1 Institution (8.3%)]

- Kein Scoring/ Bewertung möglich von Verwechslung ähnlicher emotionaler Ausdrücke möglich [1 Institution (8.3%)]
- Limitierte Auswahl an Tests [1 Institution (8.3%)].

3.6 Gründe keine Theory of Mind-Tests durchzuführen, die von weniger als 20% der Institutionen genannt wurden:

Ausgewähltes Optionsfeld:

- Testergebnisse haben keine Relevanz für die Therapie [15.2%(n=5)]

Freitextfeld:

- Stimulusmaterial der Tests sind eher für Jüngere als für Patient:innen in höherem Lebensalter geeignet [3.0% (n=1)].

3.7 Statische Bildmaterial, das zur Untersuchung der Theory of Mind-Fähigkeiten verwendet wurde

- Faux Pas Geschichten, nicht näher bezeichnet: 41.7% (n=5)
- Reading the Mind in the Eyes (Baron-Cohen et al., 2001): 25% (n=3)
- Sally & Anne Test (Baron-Cohen et al., 1985): 25% (n=3)
- Geneva Social Cognition Scale (Martory et al., 2015): 8.3% (n=1)
- ToM Cartoons von Brüne (Brüne, 2005): 8.3% (n=1)
- Bildergeschichten nicht näher bezeichnet: 8.3% (n=1)

Literatur:

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. J. Child Psychol. Psychiatry 42, 241–251.

Baron-Cohen S, Leslie AM, Frith U. Does the autistic child have a "theory of mind"? Cognition 1985; 21: 37–46.

Brüne, M. (2005). Emotion recognition,'theory of mind,'and social behavior in schizophrenia. Psychiatry research, 133(2-3), 135-147.

Martory, M. D., Pegna, A. J., Sheybani, L., Métral, M., Pertusio, F. B., & Annoni, J. M. (2015). Assessment of social cognition and theory of mind: initial validation of the Geneva Social Cognition Scale. European neurology, 74(5-6), 288-295.

3.8 Vor- und Nachteile der durchgeführten Theory of Mind-Tests, die von weniger als 20% der Institutionen im Freitextfeld genannt wurden

Vorteile:

- Die Vielfältigkeit und Durchführungsdauer der Geneva Social Cognition Scale (GeSocS; Martory et al., 2015) wird positiv bewertet [7.7% (n=19)].

Nachteile:

- Ungenaue Angaben zur Auswertung [15.4% (n=2)]
- Erschwerte Durchführung bei Patient:innen mit Sprach- und Leseschwierigkeiten [15.4% (n=2)]

Literatur:

Martory, M. D., Pegna, A. J., Sheybani, L., Métral, M., Pertusio, F. B., & Annoni, J. M. (2015). Assessment of social cognition and theory of mind: initial validation of the Geneva Social Cognition Scale. *European neurology*, *74*(5-6), 288-295.

Supplementary Material

Manuskript 2

The Basel Version of the Awareness of Social Inference Test – Emotion Recognition (BASIT-ER): Preliminary Validation Analyses in Healthy Adults

Table of Contents

Supplementary Material A **29**

Table A1 ..... **29**

Evaluation criteria of the TASIT-EET scenes for scene selection

Supplementary Material B **30**

Table B1.1 ..... **30**

Evaluation and selection of the Anger scenes from the TASIT-EET form A and form B

Table B1.2 ..... **32**

Evaluation and selection of the Disgust scenes from the TASIT-EET form A and form B

Table B1.3 ..... **34**

Evaluation and selection of the Fear scenes from the TASIT-EET form A and form B

Table B1.4 ..... **36**

Evaluation and selection of the Happiness scenes from the TASIT-EET form A and form B

Table B1.5 ..... **38**

Evaluation and selection of the Sadness scenes from the TASIT-EET form A and form B

Table B1.6 ..... **40**

Evaluation and selection of the Surprise scenes from the TASIT-EET form A and form B

Table B1.7 ..... **42**

Evaluation and selection of the Neutral scenes from the TASIT-EET form A and form B

Table B2 ..... **44**

Selected scenes for the BASIT-ER film project

Supplementary Material C **45**

Document C1 ..... **45**

BASIT-ER screenplays

Document C2 ..... **55**

Examples of the major script changes

Supplementary Material D **59**

Table D1 ..... **59**

Distribution of the actors on the scenes of the BASIT-ER and the corresponding scenes of the TASIT-EET

Supplementary Material E **61**

Document E1 ..... **61**

Background and intensity information of the BASIT-ER scenes for the actors

Table E2 ..... **84**

Differences between the BASIT-ER and the TASIT-EET

Supplementary Material F **86**

Document F1 ..... **86**

Details of the programming and data storage

Figure F2 ..... **88**

Exemplary representation of the BASIT-ER computer based application process

Supplementary Material G **89**

Figure G1 ..... **89**

Algorithm to decide whether the named term can be classified as an emotion or not as an emotion

Supplementary Material H **90**

Figure H1 ..... **90**

Algorithm to define whether the perceived “Other Emotions” are within the spectrum of the portrayed emotion

Supplementary Material I **91**

Document I1 ..... **91**

Geneva Affect Label Coder (Scherer, 2005)

Figure I2 ..... **94**

Algorithm to classify whether participants' answers are correct or not

Supplementary Material J **95**

Figures J1a-h **95**

95% Confidence intervals of the estimated difficulties of the selected intensity version scenes and the neutral scenes

Supplementary Material K **99**

Figure K1 99

Item Characteristic Curves (ICC) of the 60 intensity version scenes and the neutral scenes

Supplementary Material L **103**

Figure L1 ..... **103**

Distribution of the frequencies of the "other emotion" terms

Figure L2 ..... **104**

Distribution of the percentages of participants whose answers were classified as correct based on having named an "other emotion" term

Table L3.1 ..... **105**

Scene Anger#1_high: Percentages of emotions

Table L3.2 ..... **107**

Scene Anger#2_medium: Percentages of emotions

Table L3.3 ..... **108**

Scene Anger#3_medium: Percentages of emotions

Table L3.4 ..... **109**

Scene Disgust#1_high: Percentages of emotions

Table L3.5 ..... **110**

Scene Disgust#2_low: Percentages of emotions

Table L3.6 ..... **111**

Scene Disgust#3_medium: Percentages of emotions

Table L3.7 ..... **112**

Scene Fear#1_low: Percentages of emotions

Table L3.8 ..... **113**

Scene Fear#2_low: Percentages of emotions

Table L3.9 ..... **114**

Scene Fear#3_low: Percentages of emotions

Table L3.10 ..... **115**

Scene Happiness#1_low: Percentages of emotions

Table L3.11 ..... **116**

Scene Happiness#2_medium: Percentages of emotions

Table L3.12 ..... **117**

Scene Happiness#3_low: Percentages of emotions

Table L3.13 ..... **118**

Scene Sadness#1_high: Percentages of emotions

Table L3.14 ..... **120**

Scene Sadness#2_low: Percentages of emotions

Table L3.15 ..... **121**

Scene Sadness#3_low: Percentages of emotions

Table L3.16 ..... **122**

Scene Surprise#1_medium: Percentages of emotions

Table L3.17 ..... **123**

Scene Surprise#2_low: Percentages of emotions

Table L3.18 ..... **124**

Scene Surprise#3_medium: Percentages of emotions

Table L3.19 ..... **126**

Scene Neutral_#1: Percentages of emotions

Table L3.20 ..... **128**

Scene Neutral_#2: Percentages of emotions

Table L3.21 ..... **130**

Scene Neutral_#3: Percentages of emotions

Supplementary Material A

Table A1

Evaluation criteria of the TASIT-EET 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
4)	The content of the spoken text does not indicate the emotions portrayed
5)	The emotion of the film scene can be portrayed in different intensities*
6)	The semantics of the spoken text appears congruent with the non-verbal emotional cues (e.g., there is no indication of sarcasm)
7)	The emotion is portrayed by different modalities (e.g. gestures, facial expressions, linguistic characteristics)*
8)	Only one emotion seems to be portrayed

Note. TASIT-EET = The Awareness of Social Inference Test – Emotion Evaluation Test.

*Criterion does not apply to the neutral film scenes.

Supplementary Material B

Table B1.1

Evaluation and selection of the Anger scenes from the TASIT-EET form A and form B

	Scene ID							
Evaluation criteria	**A1**	A12	A18	A23	B4	B13	**B23**	**B28**
The topic of the film scene is suitable for the age groups over 60	1	1	1	1	1	1	1	1
The topic of the film scene is suitable for people from German-speaking cultural area	1	1	1	1	1	1	1	1
The film scene appears realistic	1	1	1	0(b)	1	0(b)	1	1
The verbal content of the film scene does not indicate the portrayed emotions	1	0	0	0	0	0	1	1
The emotion of the film scene can be portrayed in different intensities	1	1	1	1	1	1	1	1

	Scene ID							
Evaluation criteria	**A1**	A12	A18	A23	B4	B13	**B23**	**B28**
The semantic content of the film scene is congruent with the non-verbal emotional cues (e.g., no sarcasm)	1	0(d)	0(d)	1	1	1	0(d)	0(d)
The emotion is portrayed by different modalities (e.g. gestures, facial expressions, linguistic characteristics)*	1	1	1	1	1	1	1	1
Only one emotion seems to be portrayed*	1	1	1	1	1	1	1	1
Total score	6/6	4/6	4/6	4/6	5/6	4/6	5/6	5/6

Note. TASIT-EET = The Awareness of Social Inference Test-Emotion Evaluation Test; A = TASIT- EET form A, B = TASIT-EET form B. Scoring: 0 = criterion was not fulfilled; 1 = criterion was fulfilled. (a) content not changeable; (b) presentation not changeable; (c) inconsistency between verbal content and emotion; (d) sarcasm. Selected scenes are in bold. *Criterion is not part of the total score.

Table B1.2

Evaluation and selection of the Disgust scenes from the TASIT-EET form A and form B

	Scene ID							
Evaluation criteria	A4	**A9**	A15	**A22**	**B7**	B14	B20	B25
The topic of the film scene is suitable for the age groups over 60	1	1	1	1	1	1	1	1
The topic of the film scene is suitable for people from the German-speaking cultural area	1	1	1	1	1	1	1	1
The film scene appears realistic	1	1	0(b)	1	1	1	1	0(b)
The verbal content of the film scene does not indicate the portrayed emotions	0	1	0	1	1	0	0	0
The emotion of the film scene can be portrayed in different intensities	1	1	1	1	1	1	1	1

Evaluation criteria	Scene ID							
	A4	**A9**	A15	**A22**	**B7**	B14	B20	B25
The semantic content of the film scene is congruent with the non-verbal emotional cues (e.g., no sarcasm)	1	0(c)	1	1	1	1	1	1
Emotion is portrayed by different modalities (e.g. gestures, facial expressions, linguistic characteristics)*	1	1	1	1	1	1	1	1
Only one emotion seems to be portrayed*	1	0	1	0	1	1	1	1
Total score	5/6	5/6	4/6	6/6	6/6	5/6	5/6	4/6

Note. TASIT-EET = The Awareness of Social Inference Test-Emotion Evaluation Test; A = TASIT- EET form A, B = TASIT-EET form B. Scoring: 0 = criterion was not fulfilled; 1 = criterion was fulfilled. (a) content not changeable; (b) presentation not changeable; (c) inconsistency between verbal content and emotion; (d) sarcasm. Selected scenes are in bold. *Criterion is not part of the total score.

Table B1.3

Evaluation and selection of the Fear scenes from the TASIT-EET form A and form B

	Scene ID							
Evaluation criteria	A7	**A14**	**A20**	A25	B6	B9	B15	**B24**
The topic of the film scene is suitable for the age groups over 60	1	1	1	1	1	1	1	1
The topic of the film scene is suitable for people from the German-speaking cultural area	1	1	1	1	1	1	1	1
The film scene appears realistic	0(a)	1	1	1	1	1	0(b)	1
The verbal content of the film scene does not indicate the portrayed emotions	1	1	1	1	1	1	1	1
The emotion of the film scene can be portrayed in different intensities	1	1	1	1	1	1	1	1

Evaluation criteria	Scene ID							
	A7	**A14**	**A20**	A25	B6	B9	B15	**B24**
The semantic content of the film scene is congruent with the non-verbal emotional cues (e.g., no sarcasm)	1	1	1	1	1	1	1	1
Emotion is portrayed by different modalities (e.g. gestures, facial expressions, linguistic characteristics)*	1	1	1	1	1	1	0	1
Only one emotion seems to be portrayed*	0	0	1	1	1	1	1	1
Total score	5/6	6/6	6/6	6/6	6/6	6/6	5/6	6/6

Note. TASIT-EET = The Awareness of Social Inference Test-Emotion Evaluation Test; A = TASIT- EET form A, B = TASIT-EET form B. Scoring: 0 = criterion was not fulfilled; 1 = criterion was fulfilled. (a) content not changeable; (b) presentation not changeable; (c) inconsistency between verbal content and emotion; (d) sarcasm. Selected scenes are in bold. *Criterion is not part of the total score.

Table B1.4

Evaluation and selection of the Happiness scenes from the TASIT-EET form A and form B

	Scene ID							
Evaluation criteria	A3	**A8**	A17	**A26**	B2	B11	B17	**B22**
The topic of the film scene is suitable for the age groups over 60	1	1	1	1	1	1	1	1
The topic of the film scene is suitable for people from the German-speaking cultural area	1	1	1	1	1	1	1	1
The film scene appears realistic	0(b)	1	1	1	1	1	0(b)	1
The verbal content of the film scene does not indicate the portrayed emotions	1	1	1	1	0	1	1	1
The emotion of the film scene can be portrayed in different intensities	1	1	1	1	1	1	1	1

	Scene ID							
Evaluation criteria	A3	**A8**	A17	**A26**	B2	B11	B17	**B22**
The semantic content of the film scene is congruent with the non-verbal emotional cues (e.g., no sarcasm)	1	1	1	1	1	0(c)	1	1
Emotion is portrayed by different modalities (e.g. gestures, facial expressions, linguistic characteristics)*	0	1	1	1	1	1	1	1
Only one emotion seems to be portrayed*	1	1	0	1	1	0	1	1
Total score	5/6	6/6	6/6	6/6	5/6	5/6	5/6	6/6

Note. TASIT-EET = The Awareness of Social Inference Test-Emotion Evaluation Test; A = TASIT- EET form A, B = TASIT-EET form B. Scoring: 0 = criterion was not fulfilled; 1 = criterion was fulfilled. (a) content not changeable; (b) presentation not changeable; (c) inconsistency between verbal content and emotion; (d) sarcasm. Selected scenes are in bold. *Criterion is not part of the total score.

Table B1.5

Evaluation and selection of the Sadness scenes from the TASIT-EET form A and form B

	Scene ID							
Evaluation criteria	A5	A10	A16	A27	B1	**B12**	**B19**	**B26**
The topic of the film scene is suitable for the age groups over 60	1	1	1	1	1	1	1	1
The topic of the film scene is suitable for people from the German-speaking cultural area	1	1	1	1	1	1	1	1
The film scene appears realistic	1	1	1	0(b)	0(b)	1	1	1
The verbal content of the film scene does <u>not</u> indicate the portrayed emotions	1	0	0	1	0	1	1	1
The emotion of the film scene can be portrayed in different intensities	1	1	1	1	1	1	1	1

Evaluation criteria	Scene ID							
	A5	A10	A16	A27	B1	**B12**	**B19**	**B26**
The semantic content of the film scene is congruent with the non-verbal emotional cues (e.g., no sarcasm)	1	1	1	1	1	1	1	1
Emotion is portrayed by different modalities (e.g. gestures, facial expressions, linguistic characteristics)*	1	1	1	1	1	1	1	1
Only one emotion seems to be portrayed*	1	0	1	1	1	0	0	1
Total score	6/6	5/6	5/6	5/6	4/6	6/6	6/6	6/6

Note. TASIT-EET = The Awareness of Social Inference Test-Emotion Evaluation Test; A = TASIT- EET form A, B = TASIT-EET form B. Scoring: 0 = criterion was not fulfilled; 1 = criterion was fulfilled. (a) content not changeable; (b) presentation not changeable; (c) inconsistency between verbal content and emotion; (d) sarcasm. Selected scenes are in bold. *Criterion is not part of the total score.

Table B1.6

Evaluation and selection of the Surprise scenes from the TASIT-EET form A and form B

	Scene ID							
Evaluation criteria	A2	A11	A21	**A24**	**B5**	B10	**B16**	**B27**
The topic of the film scene is suitable for the age groups over 60	1	1	1	1	1	1	1	1
The topic of the film scene is suitable for people from the German-speaking cultural area	1	1	1	1	1	1	1	1
The film scene appears realistic	1	0(b)	1	1	1	1	1	1
The verbal content of the film scene does not indicate the portrayed emotions	0	0	1	1	1	0	1	0
The emotion of the film scene can be portrayed in different intensities	1	1	1	1	1	1	1	1

Evaluation criteria	Scene ID							
	A2	A11	A21	**A24**	**B5**	B10	**B16**	**B27**
The semantic content of the film scene is congruent with the non-verbal emotional cues (e.g., no sarcasm)	1	1	1	1	1	1	1	1
Emotion is portrayed by different modalities (e.g. gestures, facial expressions, linguistic characteristics)*	1	1	1	1	1	1	1	1
Only one emotion seems to be portrayed*	1	1	1	1	0	1	1	1
Total score	5/6	4/6	6/6	6/6	6/6	5/6	6/6	5/6

Note. TASIT-EET = The Awareness of Social Inference Test-Emotion Evaluation Test; A = TASIT- EET form A, B = TASIT-EET form B. Scoring: 0 = criterion was not fulfilled; 1 = criterion was fulfilled. (a) content not changeable; (b) presentation not changeable; (c) inconsistency between verbal content and emotion; (d) sarcasm. Selected scenes are in bold. *Criterion is not part of the total score.

Table B1.7

Evaluation and selection of the Neutral scenes from the TASIT-EET form A and form B

	Scene ID							
Evaluation criteria	**A6**	A13	A19	A28	B3	B8	**B18**	**B21**
The topic of the film scene is suitable for the age groups over 60	1	1	1	1	1	1	1	1
The topic of the film scene is suitable for people from the German-speaking cultural area	1	1	1	1	1	1	1	1
The film scene appears realistic	1	1	0(b)	1	0(b)	1	1	1
The verbal content of the film scene does not indicate the portrayed emotions	1	0	0	0	0	0	0	1
The emotion of the film scene can be portrayed in different intensities*	-	-	-	-	-	-	-	-

Evaluation criteria	Scene ID							
	A6	A13	A19	A28	B3	B8	**B18**	**B21**
The semantic content of the film scene is congruent with the non-verbal emotional cues (e.g., no sarcasm)	1	1	1	1	1	1	1	1
Emotion is portrayed by different modalities (e.g. gestures, facial expressions, linguistic characteristics)*+	-	-	-	-	-	-	-	-
Only one emotion seems to be portrayed+	1	1	1	1	1	1	1	1
Total score	5/5	4/5	4/5	4/5	3/5	5/5	5/5	5/5

Note. TASIT-EET = The Awareness of Social Inference Test-Emotion Evaluation Test; A = TASIT- EET form A, B = TASIT-EET form B. Scoring: 0 = criterion was not fulfilled; 1 = criterion was fulfilled. (a) content not changeable; (b) presentation not changeable; (c) inconsistency between verbal content and emotion; (d) sarcasm. Selected scenes are in bold. *Criterion is not part of the total score.

Table B2

Selected scenes for the BASIT-ER film project

Emotion	Scene 1	Scene 2	Scene 3
Practice scene (surprise)	B27		
Anger	A1	B23	B28
Disgust	A22	B7	A9
Fear	A14	A20	B24
Happiness	A8	A26	B22
Sadness	B19	B26	B12
Surprise	B16	A24	B5
Neutral	A6	B18	B21

Note. BASIT-ER = Basel Version of the Awareness of Social Inference Test – Emotion Recognition. Scene description: The letter indicates the form of the TASIT (The Awareness of Social Inference Test) and the number indicates the scene number.

Supplementary Material C

Document C1

BASIT-ER screenplays

TASIT - Spielszenen

(deutsch)

V 4.2

1 ÜBUNGSSZENE 1 - ÜBERRASCHUNG ("GESCHENK" B27) - WOZI 1

M4 und F3 sind im Wohnzimmer. M4 gibt F3 eine Kartonkiste.

M4
Hier, für dich!

F3
Für mich?

M4
Ja, nimm's!

M4 nimmt die Kiste, öffnet sie und schaut rein.

F3
Oh, das kann ich doch nicht...

M4
Wieso nicht? Du hast es dir verdient.

2 ANGST 1 ("AUFFLIEGENDER FEHLER", A14) - BÜRO 2

M1 sitzt auf einem Stuhl und spricht mit jemandem am Telefon.

M1 hört zu.

M1
In Ordnung.

M1 (CONT'D)
Ja.

M1 (CONT'D)
Nein.

M1 (CONT'D)
Woher weißt du das?

M1 (CONT'D)
Okay.

M1 (CONT'D)
Okay.

3 ANGST 2 ("EINBRECHER", A20) - WOZI 3

F1 sitzt auf einem Stuhl und spricht mit jemandem am Telefon. Im Hintergrund ist eine Tür zu sehen.

F1
Ja... Ja.

Sie hört zu und dreht sich zur Tür um.

F1 (CONT'D)
Nein... Nein, ich hab nur gedacht...

Sie hört.

F1 (CONT'D)
Könnte sein!

F1 (CONT'D)
Oh! Oh nein!

4 ANGST 3 ("STALKER", B24) - WOZI 4

F2 sitzt auf einem Stuhl und hält das Telefon.

F2
Ja.

F2 (CONT'D)
Ja.

F2 (CONT'D)
Okay.

F2 (CONT'D)
Okay... Ohh!

F2 (CONT'D)
Hmm, ja.

F2 (CONT'D)
Okay!

5 EKEL 1 ("KOLLEGIN", A22) - BÜRO 5

M2 sitzt auf einem Stuhl und spricht mit jemandem am Telefon.

M2
Aah, aah... ah nein! Ah...

Seine Stimme bricht in ein leichtes Husten. Er hört weiter zu.

M2 (CONT'D)
Oh.

M2 (CONT'D)
J... Ich... ja ja, ja, ich, ich..

M2 (CONT'D)
Ja ich verstehe, ja...

M2 (CONT'D)
ooh... ahh, ohh, hör auf!

6 EKEL 2 ("VERDORBENE ESSWAREN", B7) - WOZI 6

F2 und M1 stehen sich gegenüber. Sie halten zusammen eine verschlossene Kartonkiste.

M1
Mach es auf und schau rein!

F2 weicht vom Paket weg.

F2
Nein, ich will nicht.

M1
Soll ich es aufmachen?

F2
Nein.

M1
Komm schon.

M1 beginnt vorsichtig, das Paket ein bisschen zu öffnen, aber F2 hält es zu.

M1 (CONT'D)
Dann mach ich es auf!

F2
Nein! Lass mich in Ruhe.

F2 drückt ihm das Paket entgegen und stützt die Hände in die Seiten.

7 EKEL 3 ("SAUEREI", A9) - BÜRO 7

F4 und M3 stehen nebeneinander und schauen sich zunächst nicht an. F4 zeigt M3 etwas, das sich vor ihnen befindet. F4 spricht während der ganzen Szene relativ neutral.

F4
Siehst du, was sie gemacht haben?

Er schaut zu M3, der weiterhin auf das vor ihnen liegende fixiert ist.

M3
Ja.

F4 wendet den Blick wieder auf das, was sich vor ihnen befindet.

F4
Was hälst du davon?

M3
Das haben sie wirklich gut hingekriegt.

F4
Ich glaube, ich habe noch nie etwas dergleichen gesehen.

M3
Nein.

M3 lacht kurz auf.

M3 (CONT'D)
Das übertrifft sicherlich alles, was ich bis jetzt gesehen habe.

8 TRAUER 1 ("TODESFALL", B19) - WOZI 8

F4 und M3 befinden sich im Wohnzimmer, F4 schaut ins Leere.

M3
Willst du hingehen?

F4
Ja, natürlich.

F4 schaut dabei kurz zu M3, dann wieder ins Leere.

M3
Es werden viele Leute dort sein.

F4 schaut zu M3.

F4
Hm ja, ich weiß.

F4 wendet den Blick wieder ab ins Leere.

M3
Wie wär's, wenn ich dich abhole?

F4
Danke.

F4 (CONT'D)
Das wäre nett.

9 TRAUER 2 ("INSOLVENZ", B26) - BÜRO 9

M4 sitzt und spricht am Telefon.

M4
Ja gut.

M4 (CONT'D)
Ja, ich weiß... ich weiß.

M4 (CONT'D)
Ja.

M4 (CONT'D)
Ja.

M4 (CONT'D)
Ich weiß. Danke Dir!

10 TRAUER 3 ("DIAGNOSE", B12) - WOHN-ESSZIMMER 10

F1 sitzt auf einem Stuhl und spricht mit jemandem am Telefon.

F1
Ich weiß, Liebes, ich weiß.

F1 (CONT'D)
Ja, sicher.

F1 (CONT'D)
Nein, nein, es wird schon gehen.

F1 (CONT'D)
Das werde ich.

F1 (CONT'D)
Ich verspreche es.

11 WUT 1 ("NERVIGER KUNDE", A1) - BÜRO 11

M2 steht im Büro, hält einen Block und liest etwas. In der anderen Hand hält er einen Stift.

F3 tritt ins Büro ein

F3
Da ist jemand, der dich sehen will.

M2 schlägt mit dem Stift aufs Papier. Schaut kurz auf, setzt zum Schreiben an und schaut wieder auf seine Blätter.

M2
Wer denn?

F3
Ich weiß nicht, er wollte es nicht sagen.

M2
Wie sieht er denn aus?

F3
Groß, blonde Haare, mit Brille.

M2 wedelt abweisend mit der Hand und wendet sich wieder seinen Blättern zu.

M2
Sag ihm, ich sei nicht da.

F3
Oh, das geht nicht, ich habe ihm schon gesagt, dass du hier bist.

M2
Dann sag ihm, ich habe keine Zeit!

12 WUT 2 ("LONDON", B23) - WOHN-/ESSZIMMER 12

M1 hat es sich im Wohnzimmer gemütlich gemacht.

F4 betritt die Wohnung.

F4
Rate mal...

M1
Was?

F4
Ich werde nach London gehen.

M1
Was?

Kamera fährt langsam näher auf das Gesicht von M1.

F4
Ja! Der Chef ist gerade zu mir gekommen und hat mich gefragt, ob ich nicht nach London gehen möchte.

M1
Wann?

F4
In zwei Wochen.

Die Kamera ist nun ganz auf dem Gesicht von M1.

M1
Zwei Wochen! Wie soll das denn gehen?

F4
Weiß ich auch noch nicht, aber ich werde hin gehen!n

M1
Du Glückspilz!

13 WUT 3 ("BAUPROJEKT", B28) - BÜRO 13

F2 sitzt auf einem Stuhl und spricht mit jemandem am Telefon.

F2 setzt an, um etwas zu sagen aber hört nochmals kurz zu.

F2
Ich... Ich muss jetzt wieder zurück, okay?

F2 will wieder etwas sagen aber kommt nicht zu Wort.

F2 (CONT'D)
Ah, vielen Dank!

F2 (CONT'D)
Ach, ja? Hat sie? Na prima... Das ist doch einfach grossartig!

14 FREUDE 1 ("TREFFEN", A8) - BÜRO 14

F3 spricht mit jemandem am Telefon und sitzt dazu am Schreibtisch.

F3
Nein. Das hätte ich nicht gedacht.

Sie lacht einmal kurz auf.

F3 (CONT'D)
Ja, im zehnten Stock.

F3 (CONT'D)
Ja!

F3 (CONT'D)
Ich weiss, ich weiss.

F3 (CONT'D)
Das wird sie.

15 FREUDE 2 ("NEUES AUTO", A26) - WOHN-ESSZIMMER 15

M3 sitzt auf einem Stuhl und spricht mit jemandem am Telefon. Ein Bein leger über das andere geschlagen. Er hört lächelnd zu.

M3
Danke. Danke der Nachfrage.

Er hört zu und lacht kurz auf.

M3 (CONT'D)
Ja, J... Ja.

M3 (CONT'D)
In Ordnung, ja, ich werd's ihr sagen.

M3 (CONT'D)
Ja okay... in Ordnung. Dann, nächsten Mittwoch um halb drei.

M3 (CONT'D)
Ja, genau, und welche Farbe hat es?

M3 (CONT'D)
Aha. In Ordnung, ich werd's ihr ausrichten.

M3 (CONT'D)
(lächelnd)
Okay!

16 FREUDE 3 ("URLAUB", B22) - BÜRO 16

M4 sitzt im Büro und spricht mit jemandem am Telefon.

M4
Das werden wir sicherlich tun.

M4 (CONT'D)
Ja!

Er hört kurz zu und lacht dann.

M4 (CONT'D)
Ich wüsste nicht, was dagegen spricht.

M4 (CONT'D)
Und das auch, ja.

M4 (CONT'D)
Oh ich weiss, ich weiss, das wird schon klappen!

M4 (CONT'D)
Ja!

17 ÜBERRASCHUNG 1 ("KOLLISION", B16) - BÜRO 17

F4 sitzt im Büro und spricht mit jemandem am Telefon.

F4
Genau...

F4 (CONT'D)
Ooh!

Sie setzt zum Sprechen an aber hört dann nochmal zu.

F4 (CONT'D)
Beide?

F4 hört aufmerksam zu und kann es fast nicht glauben, was sie hört.

F4 (CONT'D)
Ich...

Sie lacht erstaunt auf.

F4
Also ich... ich weiss!

18 ÜBERRASCHUNG 2 ("AFFÄRE EINES FREUNDES", A24) - WOZI 18

M3 sitzt auf einem Stuhl und spricht mit jemandem am Telefon.

M3
Wann?

M3 (CONT'D)
Letztes Jahr?!

M3 (CONT'D)
Die ganze Zeit?!

M3 (CONT'D)
Wie konnte er...?

19 ÜBERRASCHUNG 3 ("AUTOPANNE", B5) - WOHN-/ESSZIMMER 19

F1 steht mit der Schulter am Fenster und schaut hinaus.

F1
Da hat ein Auto angehalten.

Sie beobachtet etwas draußen auf der Straße.

F1 (CONT'D)
Jemand steigt aus.

Sie beobachtet weiter, schüttelt leicht den Kopf.

F1 (CONT'D)
Na sowas!

Sie beobachtet weiter mit offenem Mund und einem Ansatz eines ungläubigen Lächelns.

F1 (CONT'D)
Und jetzt kommt er zum Haus!

F1 (CONT'D)
Das gibt's doch nicht!

Schüttelt ungläubig den Kopf.

20 NEUTRAL 1 ("VERABREDUNG", A6) - WOHN-/ESSZIMMER 20

M1 sitzt auf einem Stuhl und spricht mit jemandem am Telefon.

M1
Ja.

Er nickt nachdenklich, während er zuhört.

M1 (CONT'D)
Naja, ich könnte dort hinkommen, aber...

Er hört kurz zu.

M1 (CONT'D)
J... Ja aber... wärst du dann bereit?

M1 (CONT'D)
Ja.

M1 (CONT'D)
Mhm, okay.

M1 (CONT'D)
Ja. Also...

M1 (CONT'D)
Ja, lass... lass uns dann treffen... in Ordnung?

Er hört kurz zu.

M1 (CONT'D)
Gut.

21 NEUTRAL 2 ("PAKET", B18) - BÜRO 21

M2 sitzt an einem Tisch, auf dem sich eine geschlossene Kartonkiste befindet.

M2
Ist das für mich?

Schaut sich das Paket an.

M2 (CONT'D)
Vielen Dank.

22 NEUTRAL 3 ("FEUERLÖSCHÜBUNG", B21) - BÜRO 22

M4 und F3 stehen nebeneinander. F3 hält ein Buch in der einen und ein Stift in der anderen Hand und liest im Buch. M4 hält eine Liste und ein Bleistift.

M4
Du bist als nächstes dran.

F3 schaut nicht von ihrem Buch hoch.

F3
Also, ich bin mir nicht sicher ob ich das machen will.

Sie schreibt etwas in das Buch.

M4 schaut irritiert zur Seite in die Richtung wo F3 jetzt hingehen müsste und wieder zurück zu ihr.

M4
Naja, es ist etwas spät um jetzt einen Rückzieher zu machen.

Er schaut wieder in die Richtung und zeigt jetzt auch dort hin. F3 blickt zu M4 hoch und hört kurz auf zu schreiben.

F3
Es tut mir Leid, aber ich habe zu tun.

Sie wendet sich wieder dem Buch zu und kritzelt etwas.

M4
Ich glaube nicht, dass das geht.

Wie aus Gedanken gerissen schaut F3 hoch zu M4.

F3
Was?

M4
Ähm...

M4 (CONT'D)
Du musst dort hingehen.

Document C2

Examples of the major script changes:

In the TASIT-EET, actors were given names, which they kept throughout all scenes. In contrast, in the BASIT-ER, no names were given to the actors. To dispatch actors to each scene, numbers were given to each actor, i.e., M1 to M4 for the four male actors and F1 to F4 for the four female actors.

1. Simplification of the script:

Practice scene (surprise): B27

TASIT-EET script:

Voice: Here is a present for you

Mick: Me?

Voice: Yes! Take it! *(Mick opens the box)*

Mick: Ooh, I couldn't possibly

Voice: Why not? You deserve it

Mick: You really mean I can have it?

Voice: Of course, it's yours.

Mick: Thank you

BASIT-ER script:

M4: Hier, für dich!

F3: Für mich?

M4: Ja, nimm's (Fx nimmt die Kiste, öffnet sie und schaut rein)

F3: Oh, das kann ich doch nicht...

M4: Wieso nicht? Du hast es dir verdient.

Disgust: A9

TASIT-EET script:

A: Did you see what they did?

B: Yeah

A: What did you make of it?

B: They've done a really good job of it.

I don't think I've ever seen anything quite like it.

A: No, it sure beats anything I've seen before.

B: Shall we call in the others?

A: No, let's keep this one to ourselves.

BASIT-ER Script:

F4: Siehst du, was sie gemacht haben?

M3: Ja.

F4: Was hälst du davon?

M3: Das haben sie wirklich gut hingekriegt.

F4: Ich glaube, ich habe noch nie etwas dergleichen gesehen.

M3: Nein.

F4: Das übertrifft sicherlich alles, was ich bis jetzt gesehen habe.

---

2. Changes of the text suggested by the actors:

Disgust: B7

BASIT-ER script:

M1: Mach es auf und schau rein!

F2: Nein, ich will nicht.

M1: Soll ich es aufmachen?

F2: Nein.

M1: Komm schon. Dann mach ich es auf!

F2: Nein! Lass mich in Ruhe!

Adapted version:

M1: Mach es auf und schau rein!

F2: Nein, ich will nicht.

M1: Soll ich es aufmachen?

F2: Nein.

M1: Komm schon. Dann mach ich es auf!

F2: Nein, bitte!

Anger: B28

BASIT-ER script:

F2: Ich...Ich muss jetzt wieder zurück, okay?

F2: Ah, vielen Dank!

F2: Ach, ja? Hat sie? Na prima...

F2: Das ist doch einfach grossartig!

Adapted version:

F2: Ich...Ich muss jetzt weitermachen, okay?

F2: Ah, vielen Dank!

F2: Ach ja? Hat sie? Na prima...

F2: Das ist doch einfach grossartig!

Surprise: A24

BASIT-ER script:

M3: Wann?

M3: Letztes Jahr?

M3: Die ganze Zeit?

M3: Wie konnte er...?

Adapted version:

M3: Ah, das Übliche, du weisst schon.

M3: Ich war schon mal besser in Form.

M3: Wann?

M3: Letztes Jahr?

M3: Die ganze Zeit?

Neutral: B21

BASIT-ER script:

M4: Du bist als nächstes dran.

F3: Also, ich bin mir nicht sicher ob ich das machen will.

M4: Naja, es ist etwas spät um jetzt einen Rückzieher zu machen.

F3: Es tut mir Leid, aber ich habe zu tun.

M4: Ich glaube nicht, dass das geht.

F3: Was?

M4: Ähm... Du musst dort hingehen.

Adapted version:

M4: Du bist als nächstes dran.

F3: Also, ich bin mir nicht sicher ob ich das machen will.

M4: Naja, es ist etwas spät um jetzt einen Rückzieher zu machen.

F3: Es tut mir Leid, aber ich habe zu tun.

M4: Ich glaube nicht, dass das geht.

F3: Was?

M4: Ähm... Du musst dort hingehen.

F3: Ich hab leider keine Zeit.

Supplementary Material D

Table D1

Distribution of the actors on the scenes of the BASIT-ER and the corresponding scenes of the TASIT-EET

Scene	BASIT-ER	TASIT-EET
Practice scene (surprise)		
B27	**F3**+M4	**M+**F
Anger		
A1	**M2**+F3	**M+**F
B23	**M1**+F4	**M+**F
B28	**F2**	**F**
Disgust		
A22	**M2**	**M**
B7	**F2**+M1	**F**+M
A9	**M3**+F4	**M**+M
Fear		
A14	**M1**	**M**
A20	**F1**	**F**
B24	**F2**	**F**
Happiness		
A8	**F3**	**F**
A26	**M3**	**M**
B22	**M4**	**M**
Sadness		
B19	**F4+**M3	**F**+M
B26	**M4**	**M**
B12	**F1**	**F**
Surprise		
B16	**F4**	**F**
A24	**M3**	**F**
B5	**F1**	**F**

Scene	BASIT-ER	TASIT-EET
Neutral		
A6	**M1**	**M**
B18	**M2**	**M**
B21	**F3**+M4	**M+**M

Note. The four female actors are labelled F1 to F4 and the four male actors are labelled M1 to M4. The actors portraying the emotion is printed in bold. BASIT-ER = Basel Version of the Awareness of Social Inference Test – Emotion Recognition; TASIT-EET = The Awareness of Social Inference Test – Emotion Evaluation Test.

Supplementary Material E

Document E1

Information on background and intensity of each BASIT-ER scene for the actors

Practice scene: Surprise (B27) – "gift"

Actor portraying surprise: **F3**	Communication partner: M4

Background information

- F3 is a newspaper editor and has no family
- M4 is a salesman and has a good position in this profession. He is happily married and has two children together with his wife.
- F3 and M4 have been good friends for many years and have a warm relationship.
- M4 is visiting F3.

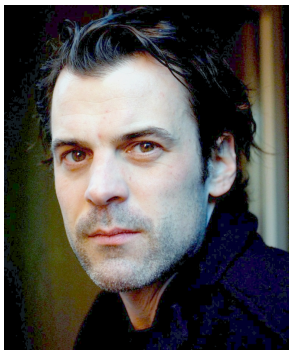
Intensities

- **Low intensity:** It is F3's birthday today. M4 is visiting F3 and is giving her a birthday present. F3 is not very surprised about the present.

- **Medium intensity:** F3 was promoted. She has been working towards this promotion for years. She told M4 about it a few days ago with full of joy. However, she has not expected a gift from her friend and is therefore surprised when she's receiving a gift from M4.

- **High intensity:** F3 knows M4 as a person who does not like to give presents. F3 has never received a gift from M4 before. M4 knows that F3 feels bad and alone after the separation from her longtime boyfriend. Therefore, he wants to visit her to cheer her up and make her happy with a small gift. She has not expected this at all and is therefore very surprised when she is receiving a gift.

Fear 1 (A14) – "Fear of detecting an error"

Actor portraying fear: **M1**	Communication partner is on the phone
	

Background information

- M1 is an overworked investment broker and a single father.
- M1 made a mistake and tried to cover it up.
- M1 is on the phone with a colleague who knows about the mistake.
- The mistake is about to be discovered.

Intensities

- **Low intensity:** There is a mistake in his work, which can be corrected with a lot of effort. The colleague is explaining him the extent of the mistake and is giving him ideas how to correct the mistake. M1 is afraid of the negative reaction of his boss, but is quite sure that he will not lose his job.

- **Medium intensity:** The mistake turns out to be greater than M1 assumed. The mistake may be corrected with the help of the colleague, but there is a risk that M1 will lose his job. This is a frightening prospect for a single father.

- **High intensity:** M1 has embezzled more than hundreds of thousands of dollars. Should this be discovered, he will be prosecuted and face a long prison sentence. The colleague now is blackmailing him and giving him instructions on what to do so that he does not betray him.

Fear 2 (A20) – "Fear of burglars"

Actor portraying fear: **F1**	Communication partner is on the phone

Background information

- F1 is an employee of a boutique, married and has one child.
- In the evening she is alone at home in the living room and is hearing noises. She is on the phone with her husband.

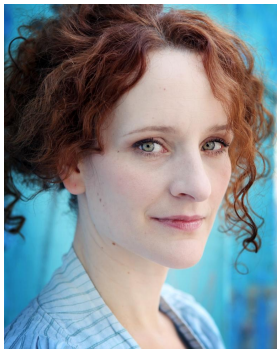
Intensities

- **Low intensity:** F1 is talking to her husband on the phone. They are talking about unimportant matters. While talking on the phone, she suddenly is hearing noises outside the house. The noises could have come from a burglar or it could has been the wind blowing branches against the house. F1 is a bit worried about the noises and is feeling a bit uneasy.

- **Medium intensity:** While F1 is watching TV she is hearing an unusual noise in front of the house. She is afraid of a burglar and therefore she is calling her husband, who is trying to calm her down.

- **High intensity:** F1 heard on the radio this afternoon about a burglar in their neighborhood. The burglars would enter the house through the basement. Now, F1 is hearing noises from the basement, she is calling her husband in panic because she doesn't know what to do.

Fear 3 (B24) – "Fear of stalkers"

Actor portraying fear: **F2**	Communication partner is on the phone
	

Background information

- F2 is a fun-loving art teacher.
- Until a few weeks ago, she was being stalked. But since she moved, she hasn't seen the stalker.
- Her mother is calling and telling her that she saw the stalker.

Intensities

- **Low intensity:** Her mother saw the stalker in the next town today.

- **Medium intensity:** Her mother saw the stalker in the neighborhood of F2 an hour ago.

- **High intensity:** Her mother, who lives in the same apartment building, is seeing the stalker in front of the entrance of the building at this moment. The stalker seems to be waiting for someone coming out so he can enter the building.

Disgust 1 (A22) – "Affair with a disgusting colleague"

Actor portraying fear: **M2**	Communication partner is on the phone

Background information

- M2 is a clerk in the accounting department. He has no family. He is on the phone with a colleague.
- They are talking about the company's last Christmas party, where M2 got drunk and harassed a female colleague. M2 has a memory gap and cannot remember what he did at the Christmas party. The colleague is telling M2 on the phone what happened.

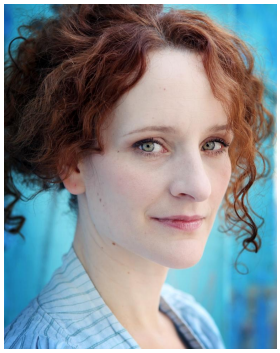
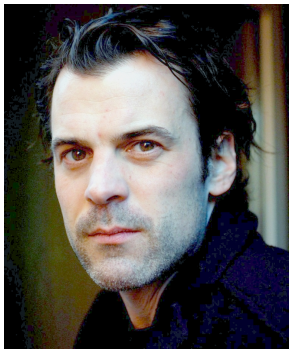
Intensities

- **Low intensity:** M2 only made light advances to her colleague and flirted a little with her but did nothing indecent.

- **Medium intensity:** M2 was quite offensive and danced very close to his colleague.

- **High intensity:** M2 was very offensive and kissed her very intensely.

Disgust 2 (B7) – "Rotten food"

Actor portraying disgust: **F2**	Communication partner: M1
	

Background information

- F2 is a housewife and sometimes a bit clumsy.
- M1 is a shop owner and humorous.
- M1 is a good friend of F2. She received a package that got stuck at the custom. She knows that it contains food which is probably rotten.
- M1 picked up the package. He knows also what is in it and wants F2 to have a look inside too.

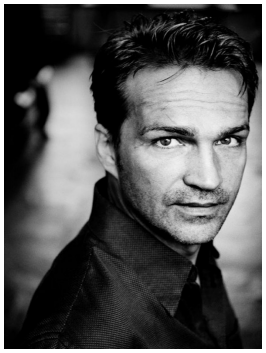
Intensities

- **Low intensity:** There is probably not very rotten food (meat, vegetables, fruit) in the package, as it was only stuck at customs for a few days.

- **Medium intensity:** There is most likely rotten food (eggs, meat, cheese) in the package, as it was sent over a month ago. The package was exposed to high temperature (summer) most of the time.

- **High intensity:** It already smells horribly of rotten food (eggs, meat, cheese) from the closed package. F2 actually doesn't want to open it at all but throw it away directly.

Disgust 3 (A9) – "Dirt in day room/toilet"

Actor portraying disgust: **M3**	Communication partner: F4
	

Background information

- M3 and F4 are owners in an advertising agency. They have a friendly relationship.
- The night before a party was held in the premises of the advertising agency. The next morning M3 is making a disgusting discovery and is showing it to F4.

Intensities

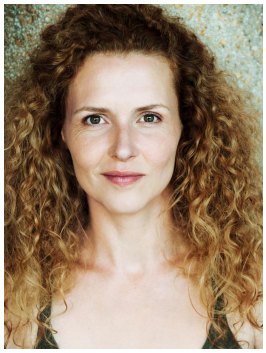
- **Low intensity:** They are looking in the lounge, where food leftovers and dirty dishes are on the table. The table is smeared with traces of ketchup and there are food leftovers on the floor.

- **Medium intensity:** They are looking in the bath room. The toilet is flooded. It stinks of urine and toilet paper is floating around.

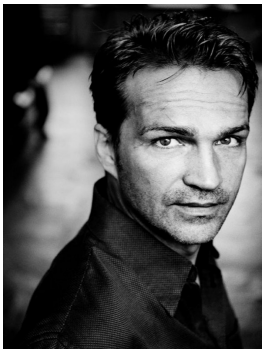
- **High intensity:** They are looking in the bath room. The room is flooded with a brownish, stinking broth, toilet paper floating around. It stinks very strongly of urine and excrement.

Sadness 1 (B19) – "Death of childhood friend / mother / husband"

Actor portraying sadness: **F4**



Communication partner: M3



Background information

- F4 is a kindergarten teacher, has a toddler and two children who go to school.
- M3 is a secondary school teacher.
- F4 and M3 know each other from the past and are very good friends
- They are meeting because of the death of a mutual childhood friend (low intensity), respective the death of F4's mother, respective F4's husband (high intensity).

Intensities

- **Low intensity:** F4 has had no contact with her childhood friend over the years, but is saddened by his early death. F4 is not sure whether she should attend the funeral, because she is worried that it might be too emotional for her, as she will meet many colleagues and friends from her childhood there and this might bring back memories of the time spent together with her childhood friend.

- **Medium intensity:** F4 had a very close relationship with her mother. They were talking on the phone every day. A few days ago the mother had a fatal accident while she was hiking in the mountains. F4 wants to hold the funeral in a private setting so that it does not become too stressful for her. However, her father insists on holding the funeral in public, as her mother was very well known and popular in her hometown and there are

certainly many people in town who would like to attend the funeral. F4 is worried, however, that all this might be too emotional for her.

- **High intensity:** The husband of F4, an aspiring young scientist, recently died in a traffic accident. In honor of him, a symposium is being held, to which F4, his young widow, has also been invited. M3 is asking her if she has the strength to attend this event, which will take place shortly after the funeral. F4 and M3 assume that many friends and colleagues of her deceased husband will be present and want to express their condolences to her.

  M3 is varying his empathic behaviour depending on the intensity.

Sadness 2 (B26) – "Insolvency of own company"

Actor portraying sadness: **M4**	Communication partner is on the phone

Background information

- M4 was managing director of his own company, which meant a lot to him
- The company went bankrupt
- He is on the phone with a colleague who wants to encourage him.

Intensities

- **Low intensity:** It has long been foreseeable that parts of the company would go bankrupt. However, the company will continue to operate at a high financial level with corresponding compensation from M4. For M4 it is a setback, especially as he himself built up these parts of the company years ago.

- **Medium intensity:** It became apparent that core parts of M4's company would go bankrupt only 2-3 month ago. This has happened despite all his efforts during the last weeks. M4 is sad and depressed that this has happend– to him who has defined himself through the company throughout his life. He doesn't know exactly how things will go on and whether what little is left of the company will survive financially for the next few months.

- **High intensity:** M4 has no family and has defined himself throughout his life through his company and his work. He had almost no time for a private life. The company was his life's work and he invested all his money and all his time in it. He is stunned and very sad that his beloved company has gone bankrupt.

Sadness 3 (B12) – "Dealing with the diagnosis of a stroke of a loved one"

Actor portraying sadness: **F1**	Communication partner is on the phone

Background information

- F1 is a mother of three children.
- She is on the phone with her closest friend. They are discussing the medical diagnosis of a person close to F1.
- The friend on the phone is trying to comfort F1.

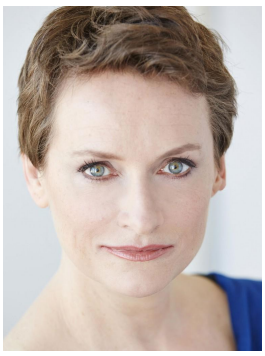
Intensities

- **Low intensity:** F1's father, a 78-year-old man, suffered a stroke a week ago, which has led to a speech disorder and paralysis of the right side of his body. Now, the father is feeling better. He can walk again with help and it is possible to talk to him. He will most likely be able to go home after rehabilitation, albeit with permanent limitations. The father has difficulty coping with the unusual situation. F1 suffers from seeing her father like this.

- **Medium intensity:** F1's father suffered a serious stroke yesterday and can only say a few words at the moment. He is completely paralysed on the right side. The doctors say that he will get better again, but that severe limitations will certainly remain. F1 still can hardly believe what has happened.

- **High intensity:** F1's husband suffered a stroke yesterday. The husband's condition has deteriorated considerably. The doctors are not sure if F1's husband will survive.

Anger 1 (A1) – "Annoying customer and annoying secretary"

Actor portraying anger: **M2**	Communication partner: F3
	

Background information

- M2 is head of a real estate company.
- F3 is a secretary and has been working for 8 months in her new job, which she likes.
- F3 does not like her boss very much, because he acts out stress on employess
- An annoying customer that M2 knows is at the reception desk. F3 has already let him in.

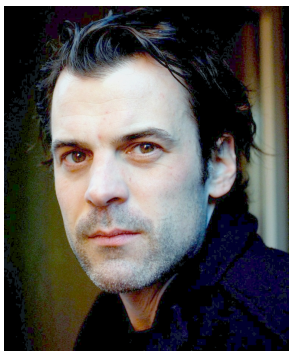
Intensities

- **Low intensity:** M2 is in a good mood, but reacts slightly annoyed at the secretary's information that the annoying customer is there and the secretary has given the customer some information without consulting him.

- **Medium intensity:** M2 is immersed in his work. He reads notes for his presentation, which is planned shortly, has a lot to do and does not want to be disturbed.

- **High intensity:** M2 is very stressed and cannot cope well with the unexpected. He has many appointments. At the moment he is reading his notes just before the next very important meeting starts. He has asked his secretary in the morning not to disturb him.

Anger 2 (B23) – "New job in London"

Actor portraying anger: **M1**	Communication partner: F4
	

Background information

- M1 is a doctor.
- F4 is a management consultant, ambitious and successful in her job.
- F4 got a job offer in London and wants to take it.
- M1 and F4 have been married for several years.

Intensities

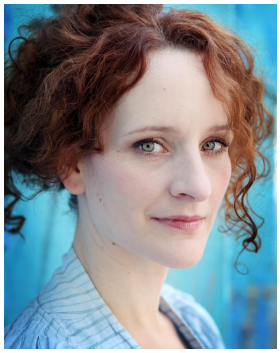
Low intensity: M1 is slightly annoyed that his wife, without asking him, has already taken the job. On the other hand, he is not really that bothered, because he is often abroad and spends quite little time with his wife.

- **Medium intensity:** M1 is angry and upset because his wife has accepted without his consent. Years ago, he turned down a job offer in Asia at her request. F4 is not aware of this. She is happy and wants to share her joy with her husband.

- **High intensity:** M1 is angry and upset because his wife has accepted the job offer without his consent. He thinks that she is deliberately sharing the news so happily to annoy him. F4 does not feel taken seriously by her husband. He has smiled at her in

her professional life and has not taken her seriously. Now she is retaliating, which annoys him tremendously.

Anger 3 (B28) – "Construction project"

Actor portraying anger: **F2**	Communication partner is on the phone
	

Background information

- F2 is an architect in a large architectural office, she is ambitious and does her job well.
- F2 is on the phone with a colleague. They are talking about another colleague taking over her work project.

Intensities

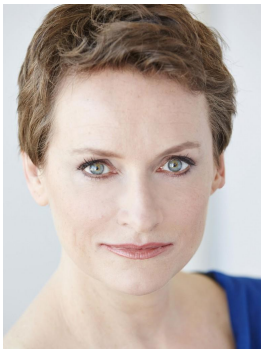
- **Low intensity**: F2 has been reassigned to this project and was already looking forward to it. The fact that another colleague will take over the project annoys F2. Fortunately, she hasn't invested any work in the new project yet, so she only little annoyed.

- **Medium intensity:** F2 has already invested some work in the project. She enjoys the project. The fact that another colleague will take over the project without consulting her makes F2 angry.

- **High intensity:** F2 has been working on this prestigious project of the company in a leading position for more than a year. She has worked overtime and is extremely proud

to lead the project. Now she is learning that a colleague, who she strongly dislikes, has snatched the project away from her in a perfidious way. This makes F2 furious.

Happiness 1 (A8) – "Organizing a meeting of a good colleague with the boss"

Actor portraying happiness: **F3**	Communication partner is on the phone
	

Background information

- F3 is a paralegal and mother of a toddler.
- F3 is talking on the phone with a former colleague who wants to make an appointment with her boss (the lawyer).

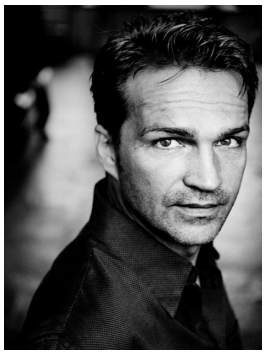
Intensities

- **Low intensity:** F3 always got along well with her former colleague. After he has quit his job at the office, she has no longer any contact with him. She is glad that he comes by the office and she will probably see him.

- **Medium intensity**: F3 always got on well with her former colleague. After he has quit his job at the office, she met him from time to time for lunch. She is happy to see him again.

- **High intensity:** F3 and her colleague are friends. For business reasons the colleague was abroad for several months. F3 is very happy that he returns.

Happiness 2 (A26) – "New car for the partner"

Actor portraying happiness: **M3**	Communication partner is on the phone
	

Background information

- M3 is a concert musician.
- He is in a steady relationship and lives together with his partner in an apartment.
- His partner is looking for a certain car model that can be visited.

Intensities

- **Low intensity:** M3 is not interested in cars, but knows that his partner is looking for a certain car model. M3 is taking a call from a car dealership. He learns that the model his partner is looking has arrived.

- **Medium intensity:** M3 wants to buy a new car together with his partner. As he is not interested in cars, his partner has chosen a quite rare car model. M3 gets a call from a car dealership that actually sells this model. He is happy for this partner as they she has looked for this model for quite some time.

- **High intensity:** M3 and his partner are both big car fans and want to buy a new car. They have been looking for a very rare model for a long time. M3 gets a call from a car dealership that actually sells this rare model. M3 is very happy about the great news and imagines how they will soon drive their dream car. He is very eager to inform his partner about the great news.

Happiness 3 (B22) – "Vacation"

Actor portraying happiness: **M4**	Communication partner is on the phone

Background information

- M4 is a production manager and has a family with three children.
- His vacation request for the period of his children's school holidays was just accepted.
- He is telling a colleague about their holiday planning on the phone.

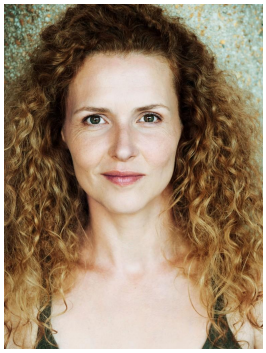
Intensities

- **Low intensity:** M4 is happy about the date of his vacation. If his request had been rejected, he would have need to postpone holiday with his family for a later date.

- **Medium intensity:** M4 is happy that he has been granted four consecutive weeks of leave. In general, holidays not longer than three weeks are granted. For years, he wanted to travel with his family to the West Coast of the USA for four weeks. He is telling a friend on the phone what they are going to see.

- **High intensity:** M4's boss told him yesterday that he would probably have to cancel his vacation due to sick leave in the company. M4 thought that the long awaited four-week family vacation to the USA would likely not take place. However, today his boss reassured him that he could take his holidays as planned as M4's colleagues will be back at work the next days. Full of joy M4 is telling a friend on the phone that he can go on vacation with the best wishes of his boss.

Surprise 1 (B16) – "Collision while skiing"

Actor portraying surprise: **F4**	Communication partner is on the phone
	

Background information

- F4 works as a sales representative in a large pharmaceutical company.
- She is talking to a colleague on the phone about two other colleagues who both had a skiing accident during a conference weekend in their spare time.

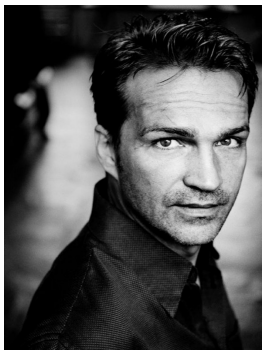
Intensities

- **Low intensity:** F4 knows the two colleagues who had a skiing accident only vaguely. The colleague on the phone is telling F4 that the other two colleagues felt while skiing because they collided. The accident happened at a place in the skiing area that is difficult to ski on. But they had been lucky: one of them had bruised his wrist, the other was unharmed.

- **Medium intensity:** The colleague on the phone is telling F4 that two of her colleagues had fallen because they had had a head-on collision. This had looked funny and had become the topic of conversation during the trip. Fortunately, both of them had no injuries except for a few bruises. The accident had happened at a place that is easy to ski. It was therefore surprising that the two colleagues had collided there.

- **High intensity:** The colleague on the phone is telling F4 that two of her colleagues had fallen because they had had a head-on collision. This had looked funny and had become the topic of conversation during the trip. Fortunately, both of them had no injuries except for a few bruises. The accident had happened very surprisingly, because the weather had been very good and both colleagues claimed to be very good skiers, which they had repeatedly told their colleagues with fervour.

Surprise 2 (A24) – "Affair of a mutual friend"

Actor portraying surprise: **M3**	Communication partner is on the phone
	

Background information

- M3 is a sports coach, he is single.
- M3 is talking to a friend on the phone about a mutual friend who is having a love affair.
- Both know the wife of their mutual friend well.

Intensities

- **Low intensity:** The two has thought that their mutual friend has cheated on his wife with another woman only once. Now M3 is learning on the phone that their mutual friend has been having an affair with this woman ever since. However, it doesn't surprise him much, as their friend is known as a "womanizer".

- **Medium intensity:** The two have known that their mutual friend and his wife have been having marriage problems for some time. However, M3 has not known about an affair with another woman and is surprised. He is surprised even more, because I would not have expected that of their mutual friend.

- **High intensity:** The two thought that their mutual friend has been having a happy marriage. They both know his wife, an attractive, young woman and get along well with her. They have also been invited to the couple's parties several times. Now M3 is learning on the phone that their mutual friend has been cheating on his wife with another woman for several years. M3 would never has expected this from their mutual friend, whom he considered to be absolutely faithful. He is very surprised and can hardly believe it.

Surprise 3 (B5) – "No assistance for car breakdown"

Actor portraying surprise: **F1**	Communication partner is on the phone

Background information

- F1 is a housewife and has been married for a long time.
- She is observing a person on the street with a car breakdown who is trying to ask people for help.

Intensities

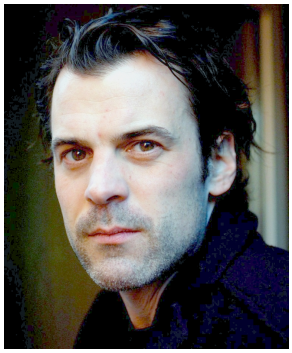
- **Low intensity:** F1 is watching a car on the road that has suddenly stopped and turned on the hazard lights. The driver is getting out of the car and is trying to talk to passers-by to ask for help. He is pointing to his car. However, the people he is talking to are passing by. Then he is walking towards the block of flats where F1 lives.

- **Medium intensity:** F1 is watching a car on the road that has suddenly stopped. The driver is getting out of the car, opening the bonnet and taking a look under it. Then he is trying to stop passers-by on the sidewalk several times, but they all are passing by. Then he is approaching the block of flats where F1 lives.

- **High intensity:** F1 is watching a car on the road that has suddenly stopped . The driver is getting out of the car, opening the bonnet from which there is a lot of smoke. Then

he is trying to stop passers-by, but nobody is stopping to help him. Desperate and with dirty hands he is approaching the block of flats where F1 lives.

Neutral 1 (A6) – "Agreement for a meeting"

Actor portraying neutral: **M1**	Communication partner is on the phone
	

Background information

- M1 is a computer specialist.
- He has to drive to another city for work and picks up a colleague halfway. On the phone they are deciding when M1 should pick up the colleague.

Neutral 2 (B18) – "Opening a package"

Actor portraying neutral: **M2**	Communication partner is on the phone

Background information

- M2 is a journalist and expects documents that he has ordered to write an article.

Neutral 3 (B21) – "Fire drill in the company"

Actor portraying neutral: **F3**	Communication partner: M4

Background information

- M4 is an office employee of a health insurance company. It is his job to make sure that all employees of the company take part in a fire drill, which is taking place today. The participation is compulsory and takes place every four years. He has a list of the people registered with him.

- F3 is also an office employee, but in a higher position than M4. M4 is addressing F3 because she is also supposed to participate in the fire drill. F3 does not want to take part in the drill. She is very busy with reviewing documents. Apart from that, she has already participated in the drill four years ago and considers another drill not necessary anymore.

Table E2

Differences between the BASIT-ER and the TASIT-EET

	BASIT-ER	TASIT-EET
Test length		
Test versions	One test form (scenes from the TASIT-EET 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	21 (7x3) scenes and one practice scene	Form A: 28 (7x4) scenes Form B: 28 (7x4) scenes In total 56 scenes
Scenes		
Distribution of actors	Each actor portrays three different emotions, apart from M4 and F4, who are portraying two emotions	Unevenly distributed across emotions
Gender distribution	Unevenly across emotions (as there are three scenes per emotion). Always one male actor and / or one female actor portrays the emotion. 50% of the main roles are casted by women	Unevenly across emotions
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 emotional 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-ER	TASIT-EET
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; partly close-ups of actors talking to the camera); camera perspective and setting dimensions do not convey realism (e.g., mostly long shots)

Note. BASIT-ER = Basel Version of the Awareness of Social Inference Test – Emotion Recognition; TASIT-EET = The Awareness of Social Inference Test – Emotion Evaluation Test

Supplementary Material F

Document F1

Details of the programming and data storage

The BASIT-ER was 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 each basic emotion and neutral exactly three times was created, with the additional constraint that no successive emotions were equal. Then for each participant ID, the order of the three scenes per emotion was randomised. Finally, the intensities were randomly distributed: The 240 first scenes of each emotion 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. Anger#1_low) the second scene of the given emotion was then randomly splitted into two parts of 40 scenes with the two other intensities (Anger#2_medium and Anger#2_high). The intensity of the third scene was then fixed by the intensities of the first two scenes of each emotion by the constraint that each participant should see each emotion in all three intensities. Since the neutral scenes do not have different intensities, their assigned intensity was removed. 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 is shown to 80 participants.

This resulted in:

- Each participant is shown 21 scenes (3 per emotion).
- Consecutive scenes show different emotions.
- Each participant is shown one low intensity, one medium intensity, and one high intensity scene per basic emotion.
- Each participant is shown all three neutral scenes.
- Each scene and intensity is shown to 80 participants.
- The first scene shown is the practice scene, shown in each intensity to 80 participants, respectively.

The sequence of 22 scenes (1 practice scene, three scenes of 6 basic emotions and three neutral scenes) for a specific participant was loaded into the application at start based on the participant's ID that the examiner had entered.

After the final scene was answered, the results were saved as .csv file (i.e., 240 separate files containing the different sequences of scenes shown to the individual participants) 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

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

Figure F2

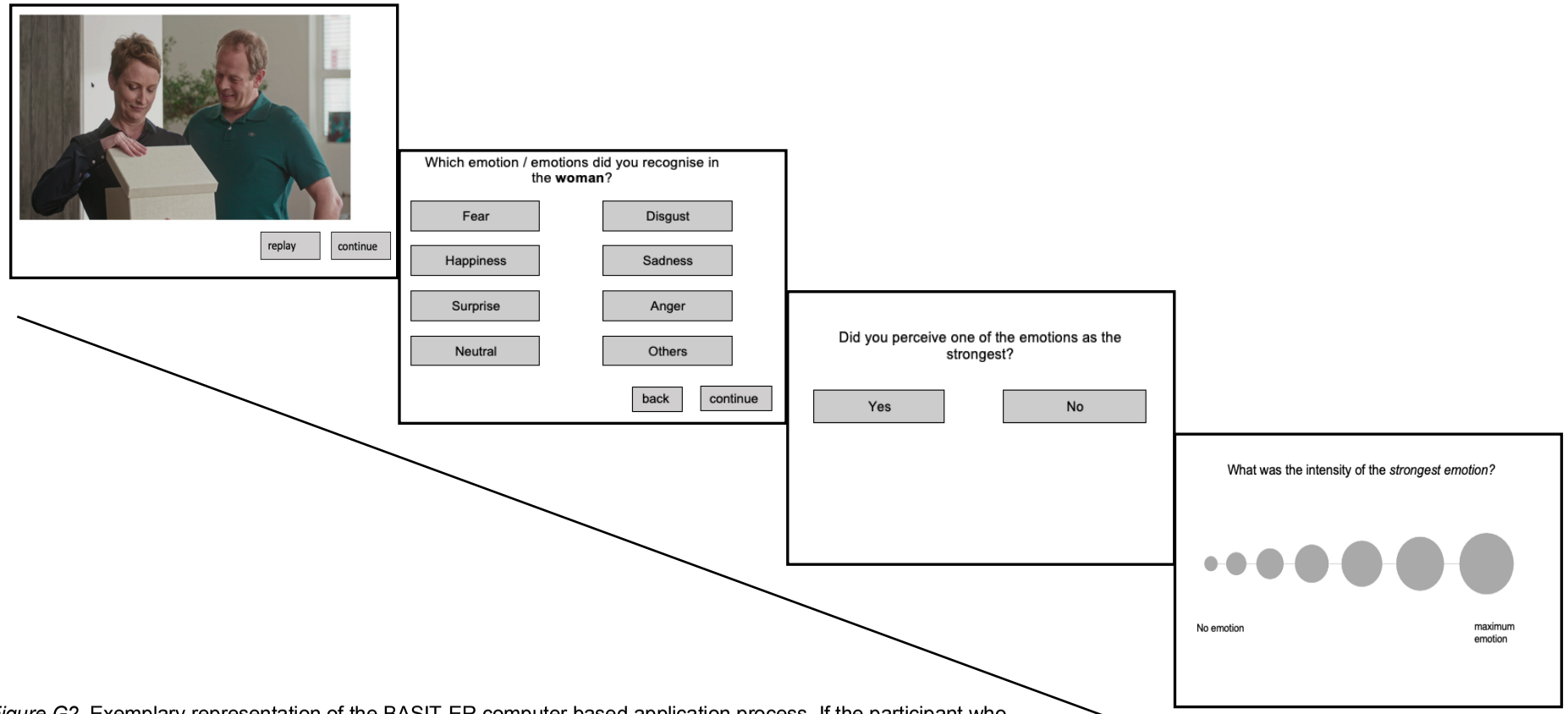
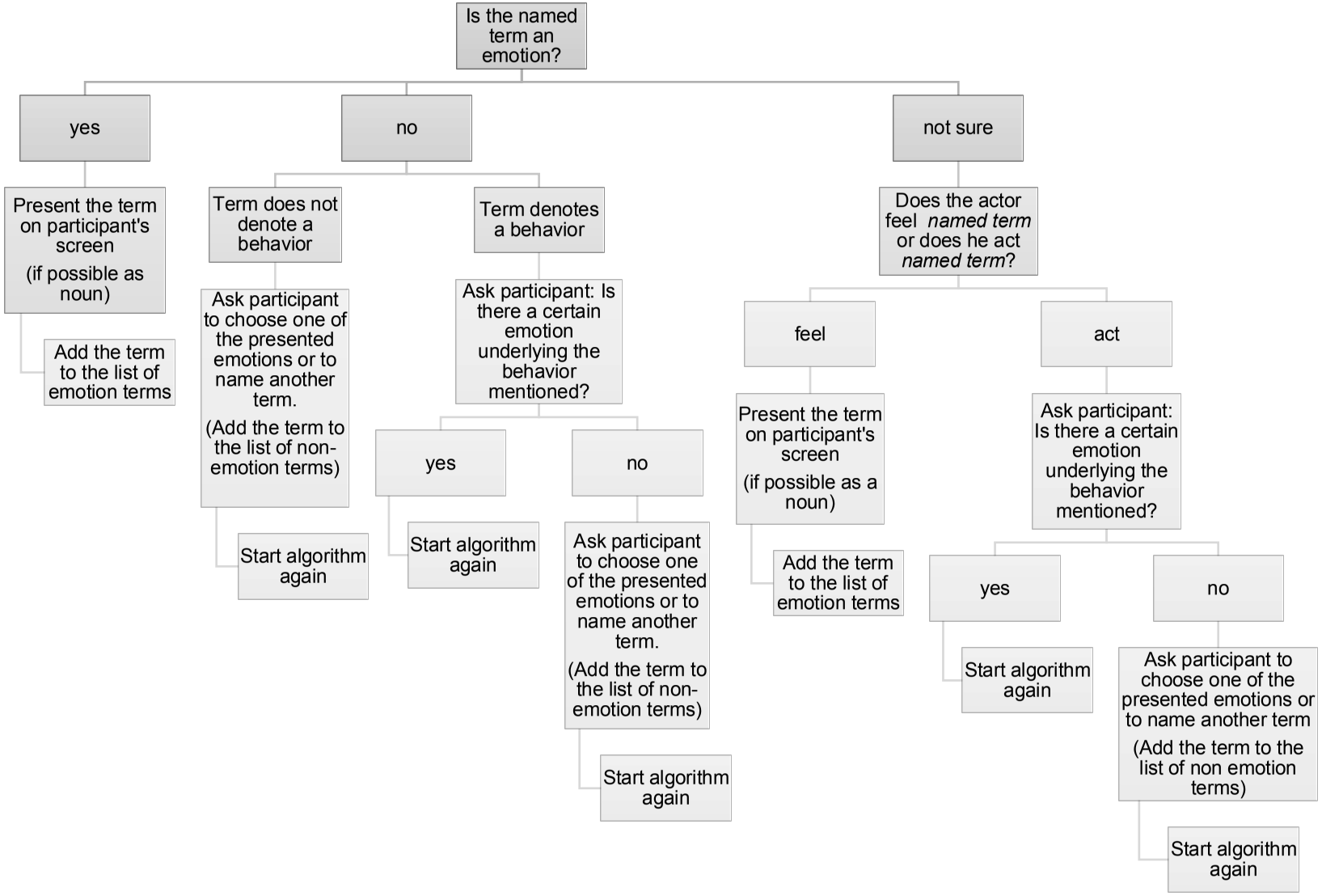


Figure G2. Exemplary representation of the BASIT-ER computer based application process. If the participant who labelled and/or named more than one emotion, he/she was asked “Did you perceive one of the emotions as the strongest” (screen 3). If the answer was no, the next scene started. If the answer was yes, he/she was asked to choose the one emotion and then to indicate its intensity (screen 4). Notably, instead of “emotion” (screen 4), the name of the emotion the participant has chosen either as the only or the strongest emotion appeared on the screen.

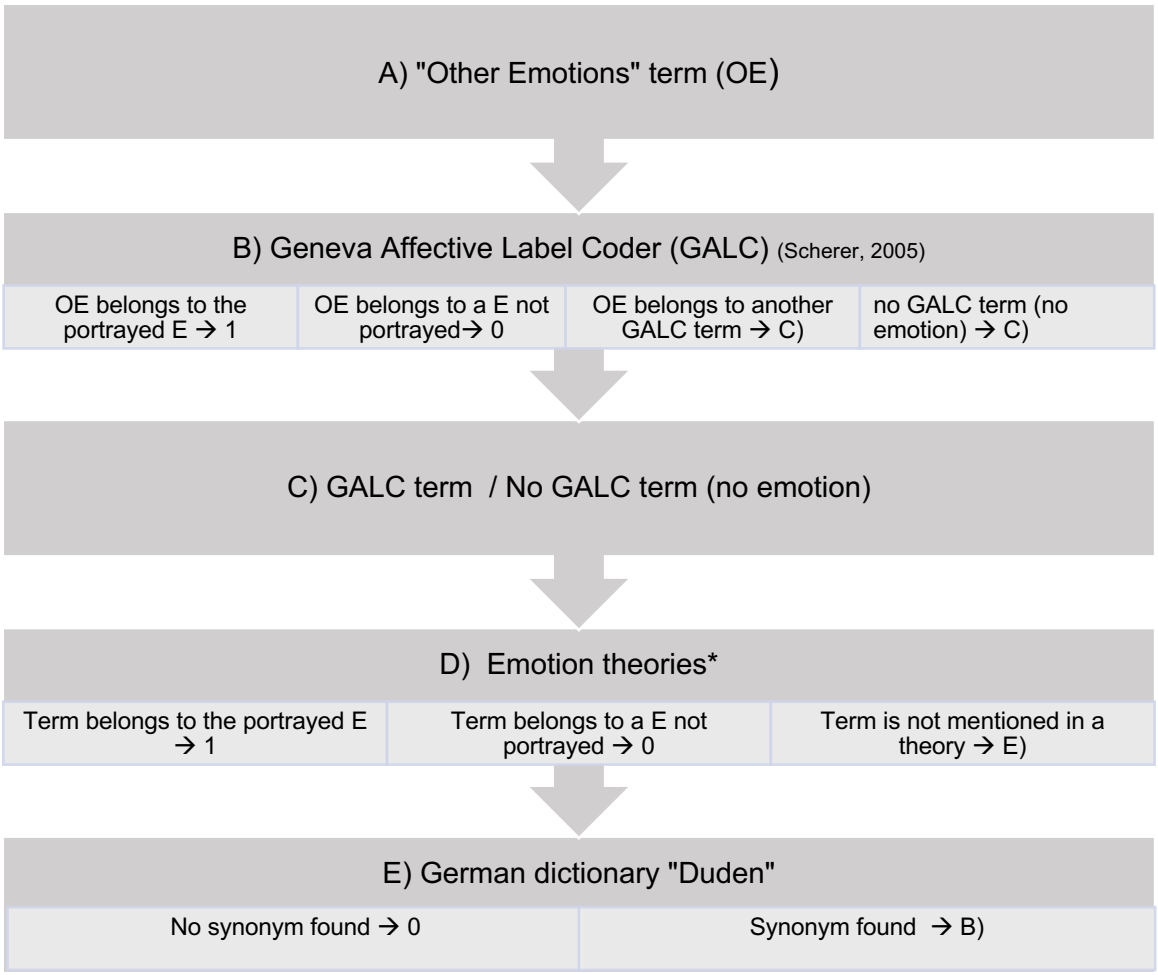
Supplementary Material G

Figure G1: Algorithm to decide whether the named term can be classified as an emotion or not as an emotion



Supplementary Material H

Figure H1: Algorithm to define whether the perceived "Other Emotions" are within the spectrum of the portrayed emotion.



Note. OE = named other emotion term; E = basis emotion (i.e. anger, disgust, fear, happiness, sadness, surprise, or neutral)

References:

*Plutchik, R. (1962). *The emotions: Facts, theories and a new model*. New York: Crown Publishing Group/Random House.

Plutchik, R. (1980). *Emotion, a psychoevolutionary synthesis*. Harper & Row.

Plutchik, R. (2003). *Emotions and life : perspectives from psychology, biology, and evolution*.

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, *44*(4), 695–729. doi.org/10.1177/0539018405058216

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (2001). Emotional Knowledge: Further Exploration of a Prototype Approach. In W. G. Parrott (Ed.), *Emotions in Social Psychology: Essential Readings* (pp. 26–56). Philadelphia: Taylor & Francis.

Zinck, A., & Newen, A. (2008). Classifying emotion: a developmental account. *Synthese*, *161*(1), 1–25. doi.org/10.1007/s11229-006-9149-2

Supplementary Material I

Document I1

Geneva Affect Label Coder (GALC)

The Geneva Affect Label Coder (GALC) is an Excel file that assigns freely named emotional terms to one or two 36 predefined affective categories and/or to so-called “positive” and “negative” categories through text analysis by use of a macro program. The program works for terms in German, English or French. The GALC can be freely downloaded for research use on https://www.unige.ch/cisa/research/materials-and-online-research/research-material/.

In the table below, you find the list of the 36 affective categories and the “positive” and “negative” categories with their synonyms and word stems in German language.

Affective category	Synonyms / Word stems
1. Angst	angst, ängst, bang, beklemm, beklomm, besorg, beunruh, sorge
2. Ärger	ärger, böse, empör, entrüst, erbost, erzürn, grimm, rage, verärger, wut, wüt, zorn
3. Bestürzung	bestürz, betroffen, entgeist, entsetz, erschrock, fassungslos, schock, schreck, verdatter, verstör
4. Bewunderung	bewund, ehrfurcht, ehrfürchtig, verehr
5. Dankbarkeit	dankbar
6. Eifersucht	eifersucht, eifersücht
7. Ekel	abscheu, angewidert, ekel, übel, verleide, widerwill
8. Entspanntheit	ausgeglichen, cool, entspann, gelass, gelöst, locker, ruhe, ruhig, unbeschwert, unbesorgt
9. Enttäuschung	enttäusch, ernüchter, frust
10. Erleichterung	befrei, erleichter, erlös, ermutig
11. Freude	erfreut, freude, freudig, froh, frohlock, jubel, zück
12. Furcht	furcht, fürcht, panik
13. Glück	beglückt, glück, glücklich, glückselig, selig, überglücklich, wonne
14. Hass	hass, verhass

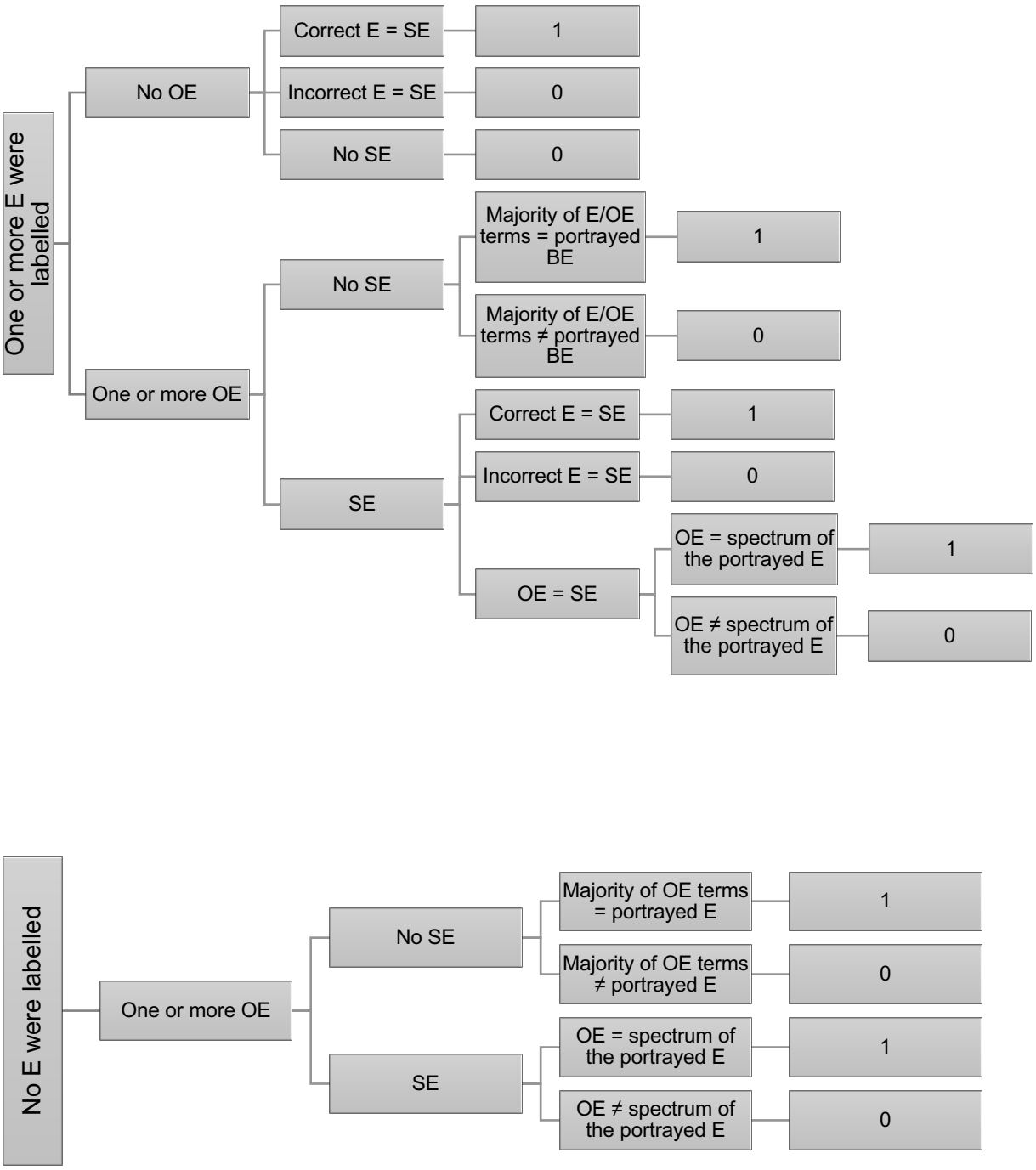
15. Heiterkeit	ausgelassen, belustig, frohsinn, fröhlich, gutgelaunt, heiter, lustig, spass, vergnüg
16. Hoffnung	hoffnung, hoffnungsvoll, optimis, zuversicht
17. Interesse	aufgestellt, begeist, enthusias, interess, neugier
18. Irritation	gereizt, indig, irrit, unmut, verstimm
19. Langeweile	langeweile, langweil
20. Liebe	liebe, verliebt, zärtlich, zuneigung
21. Lust	eksta, leidenschaft, leidenschaftlich, lust, lustvoll
22. Mitleid	anteilnahm, empathie, mitempfind, mitgefühl, mitleid, mitleidgefühl, mitleidig, teilnahm
23. Neid	missgunst, missgünst, neid, neidisch
24. Rührung	bewegt, ergriffen, gerührt, rührung
25. Scham	geschämt, scham, schämen, verlegen, verschämt
26. Schuld	gewissensbiss, reue, schlechtesgewissen, schuld, zerknirsch
27. Sehnsucht	fernweh, heimweh, nostalg, sehnen, sehnsucht, sehnsüchtig, verlang
28. Spannung/Stress	angespannt, anspannung, belast, erreg, nerv, spannung, stress, unruh
29. Stolz	erfolg, stolz
30. Traurigkeit	bedrückt, depress, gram, kummer, kümmer, melanchol, niedergeschlag, resign, schwermut, schwermüt, trauer, traurig, trüb, verzagt, weh, wehm
31. Überraschung	erstaun, konsternier, staun, überrasch, verblüff, verdutz, verwund
32. Unzufriedenheit	missmut, missmüt, unzufried, verdrossen
33. Verachtung	veracht

34. Verzweiflung	entmutig, hilflos, hoffnungslos, machtlos, ohnmacht, ohnmächt, verzweifl
35. Wohlbefinden	wohlbefinden, wohlfüh, wohlgefühl
36. Zufriedenheit	befriedig, genugtu, zufrieden
37. Positiv	angenehm, gut, positiv, schön, toll
38. Negativ	lustlos, mies, negativ, sauer, schlecht, ungut, unlust

References:

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, *44*(4), 695–729. doi:10.1177/0539018405058216

Figure I2: Algorithm to classify whether participants' answers are correct or not

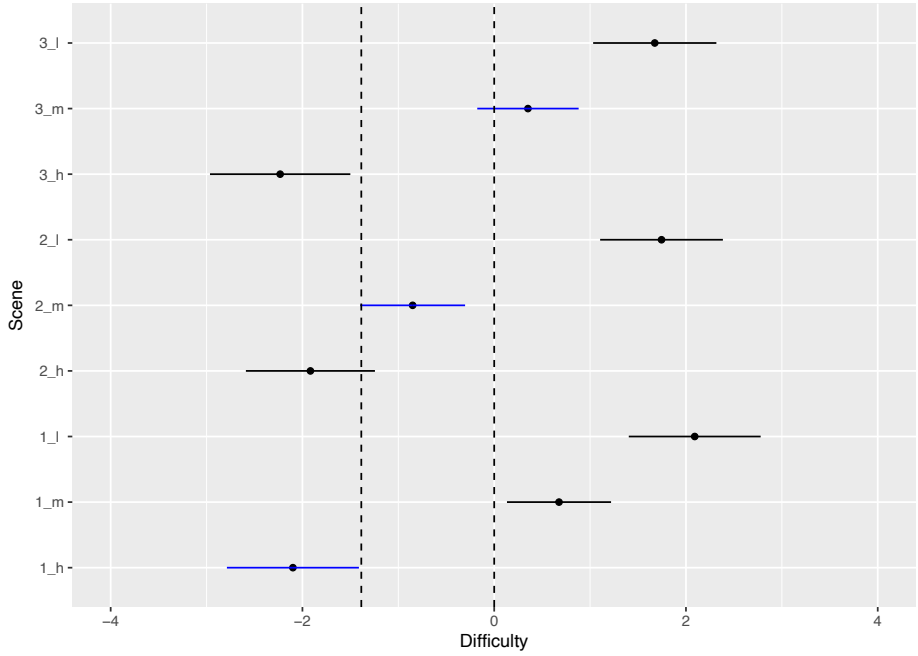


Note: E = Emotion (i.e., anger, disgust, fear, happiness, sadness, surprise, or neutral); OE = named other emotion terms; SE = indicated strongest emotion; Score of 1 = participant recognised the portrayed emotion. Score of 0 = participant did not recognise the portrayed emotion.

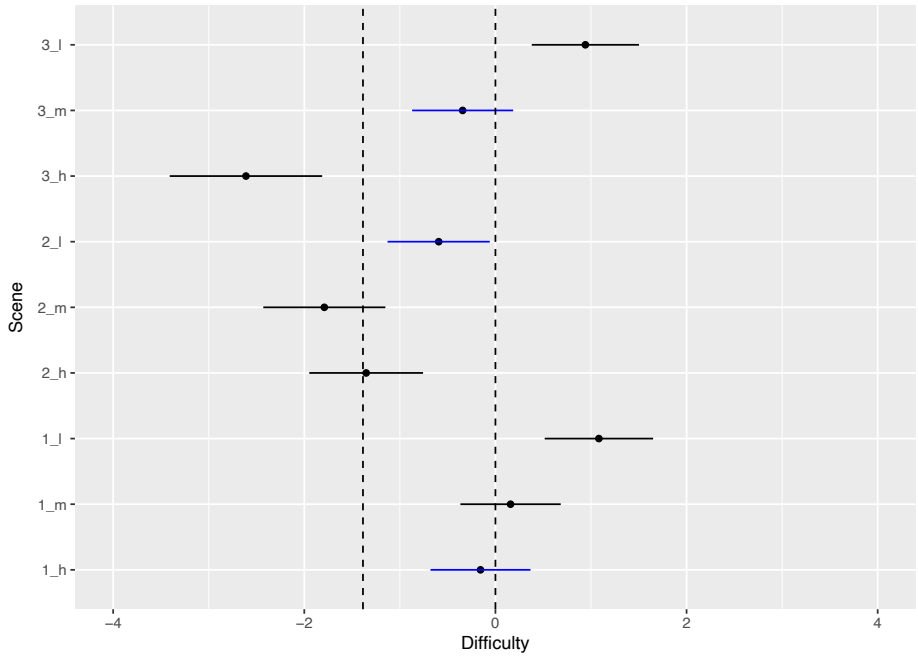
Supplementary Material J

Figures J1

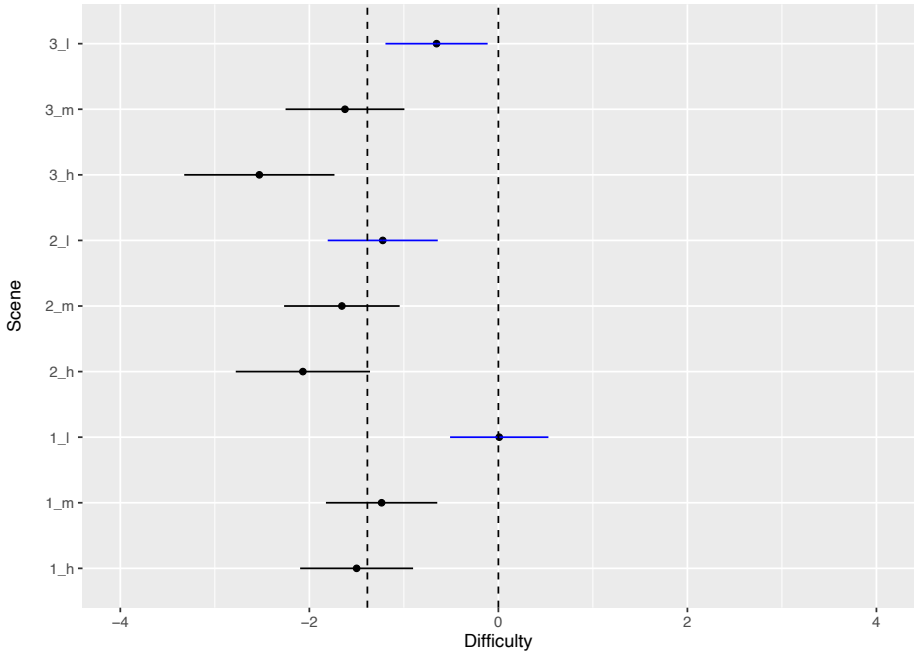
a) 95% Confidence intervals of the estimated difficulties in recognising anger in intensity version scenes of anger



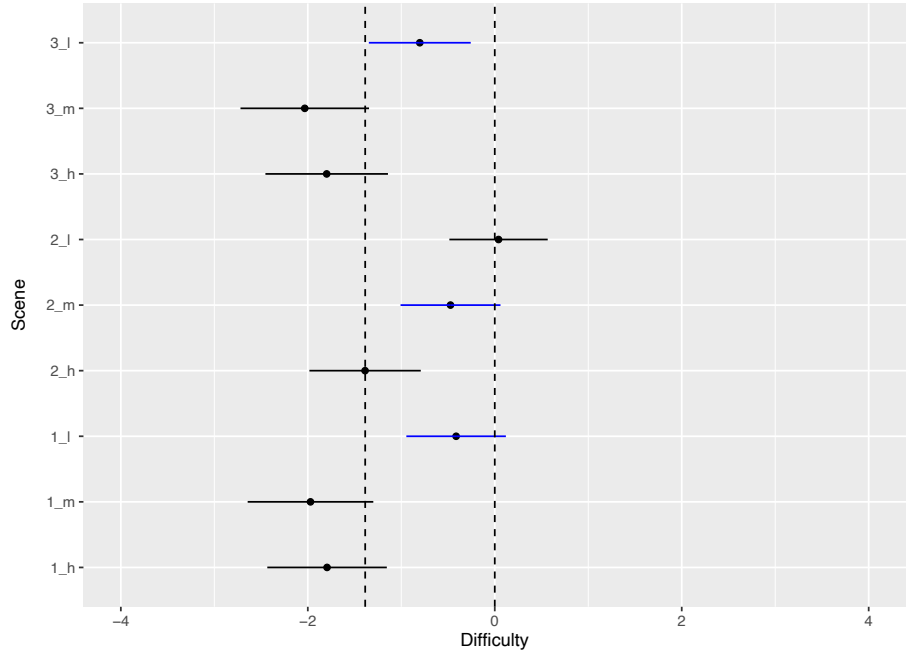
b) 95% Confidence intervals of the estimated difficulties in recognising disgust in intensity version scenes of disgust



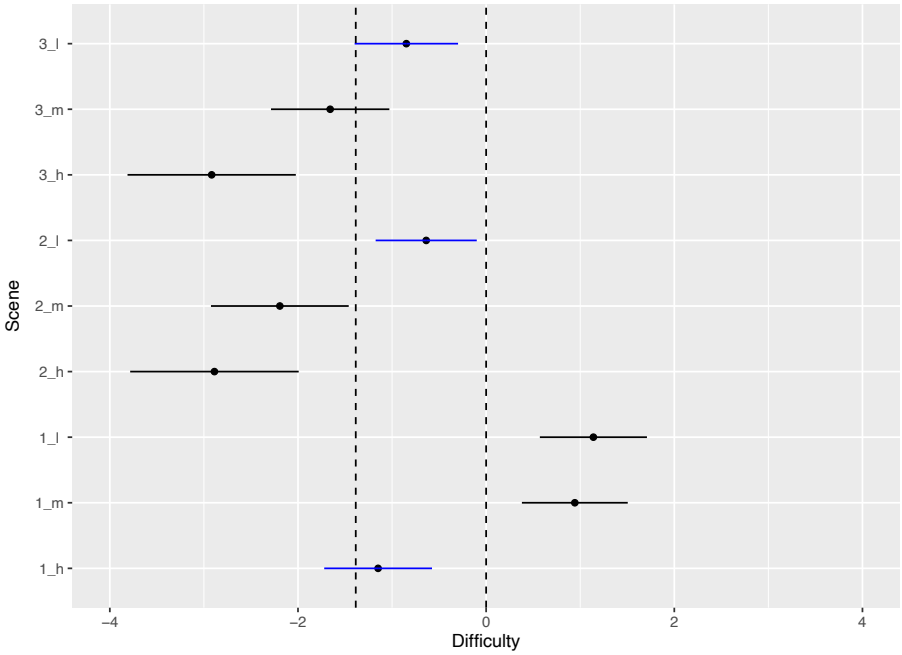
c) 95% Confidence intervals of the estimated difficulties in recognising fear in intensity version scenes of fear



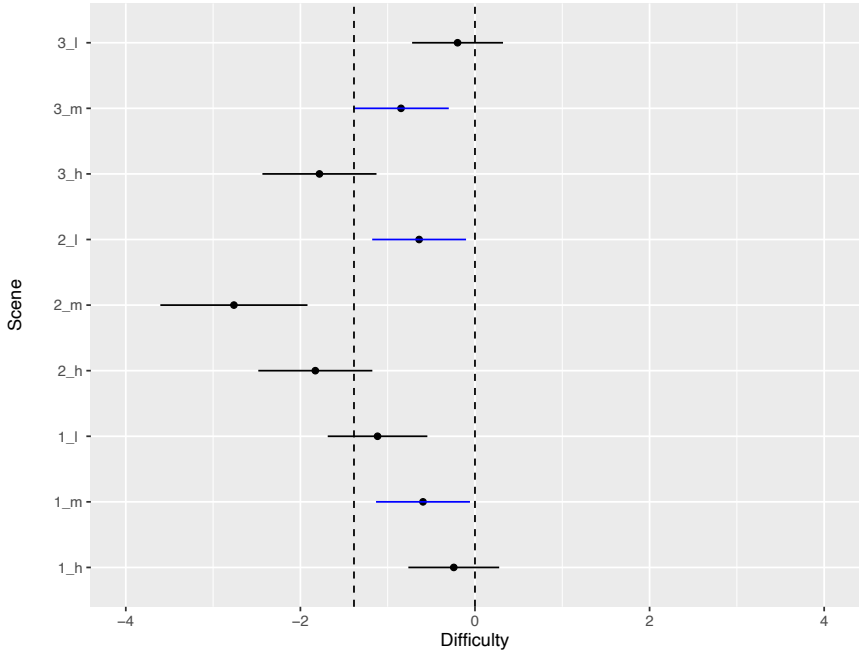
d) 95% Confidence intervals of the estimated difficulties in recognising happiness in intensity version scenes of happiness



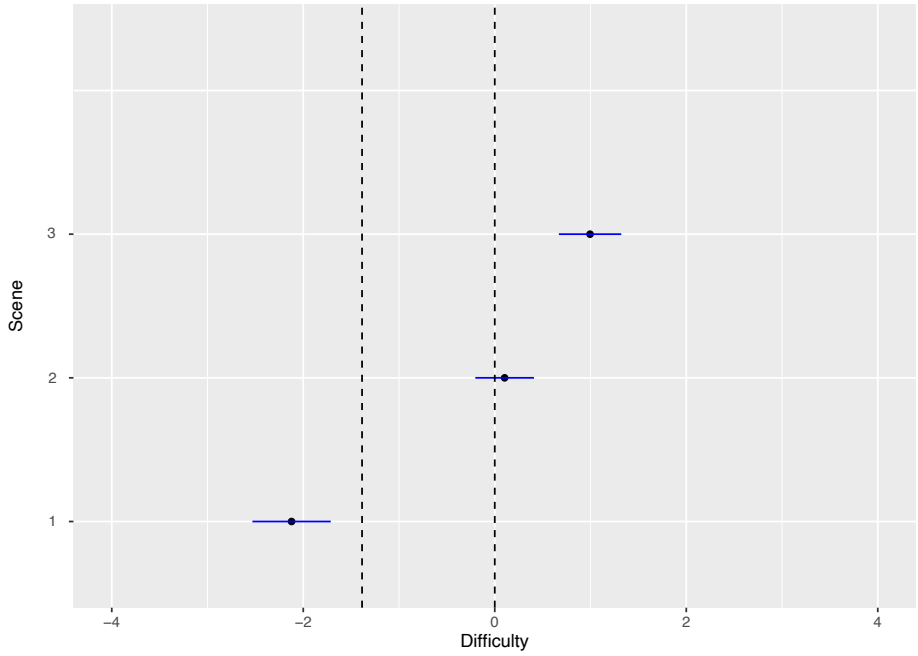
e) 95% Confidence intervals of the estimated difficulties in recognising sadness in intensity version scenes of sadness



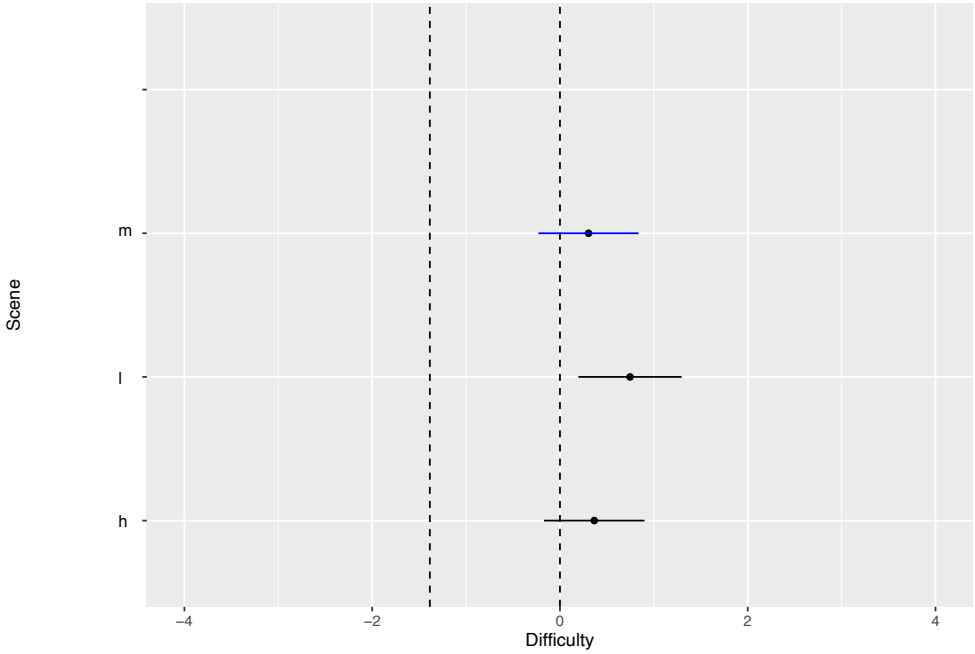
f) 95% Confidence intervals of the estimated difficulties in recognising surprise in intensity version scenes of surprise



g) 95% Confidence intervals of the estimated difficulties in recognising neutral in scenes of neutral



h) 95% Confidence intervals of the estimated difficulties in recognising surprise in intensity version scenes of the practice scene

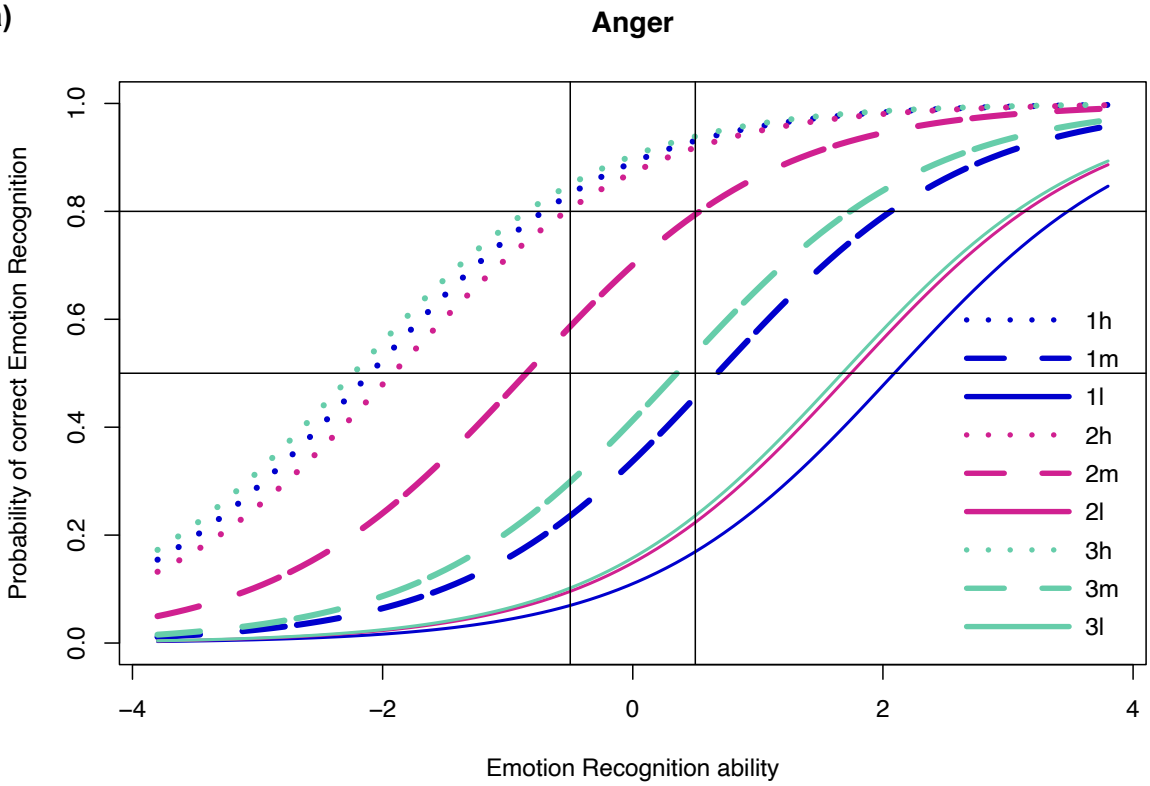


95% confidence intervals (CI) of the estimated difficulties in emotion recognition in the 54 basic emotion intensity version scenes (Fig. J1a-f) and the three neutral scenes (Fig. J1g). Scene codes consist of a number (scene) and a letter (intensity of the portrayed emotion); l = low intensity; m = medium intensity; h = high intensity. The CI of the selected scenes are colored in blue. The dashed lines illustrate the required range of difficulty [-1.386,0].

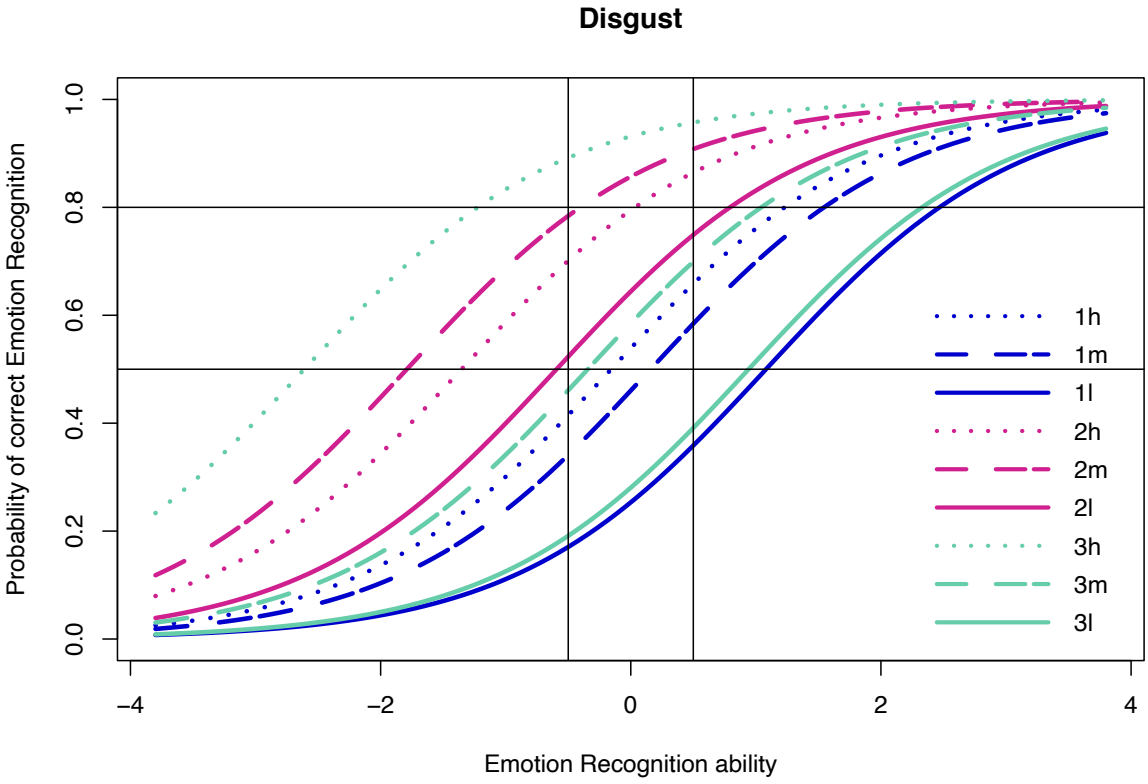
Supplementary Material K

Figure K1

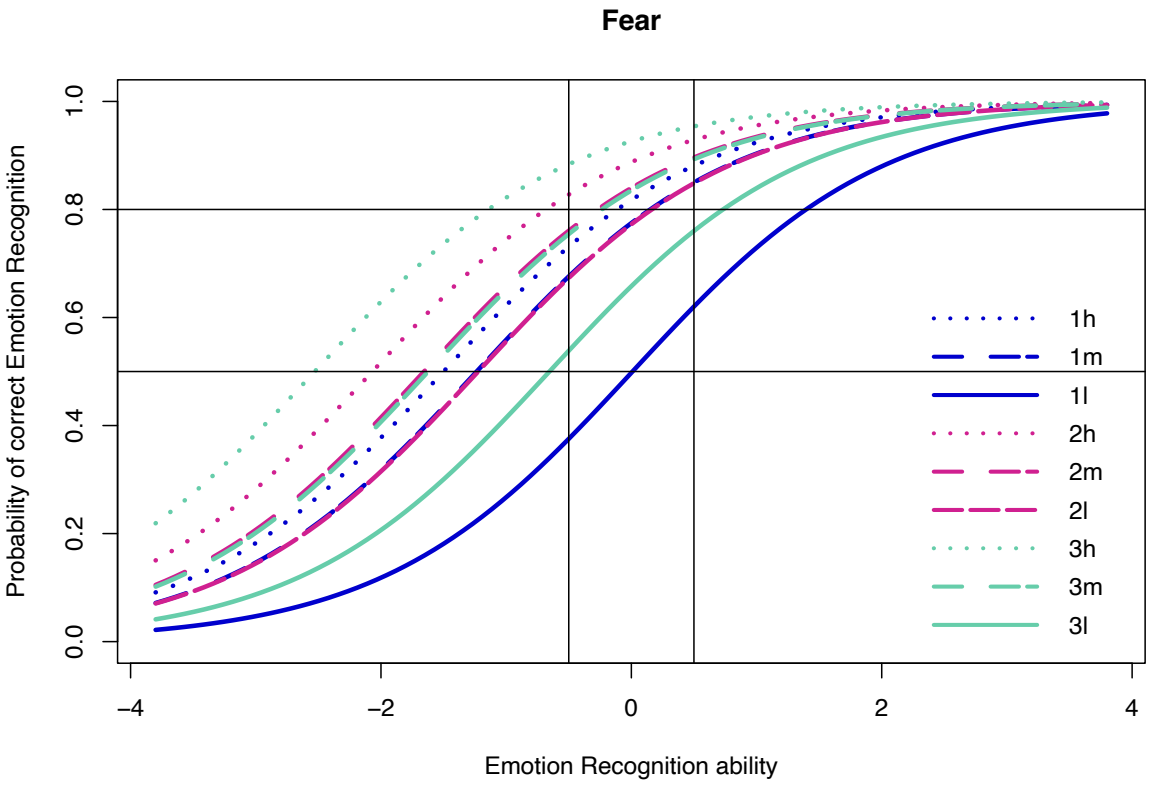
a)



b)



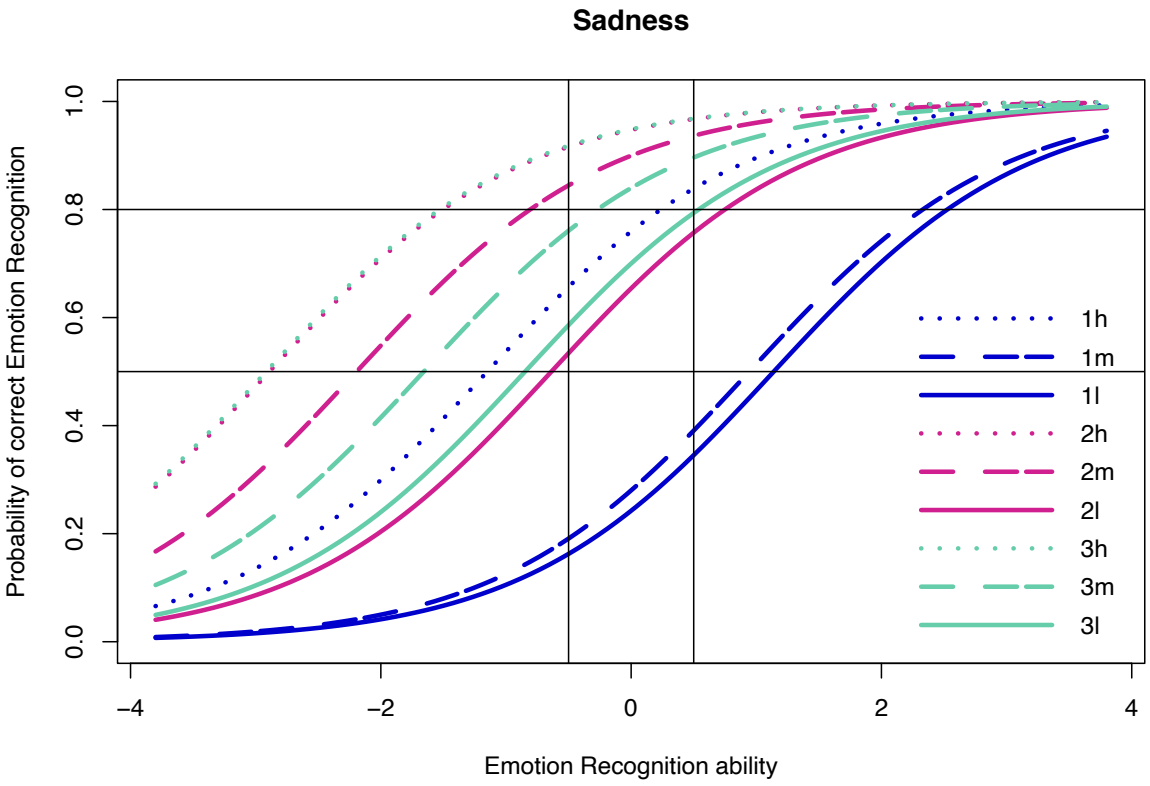
c)


Fear
Probability of correct Emotion Recognition
0.0
0.2
0.4
0.6
0.8
1.0
−4
−2
0
2
4
Emotion Recognition ability
1h
1m
1l
2h
2m
2l
3h
3m
3l

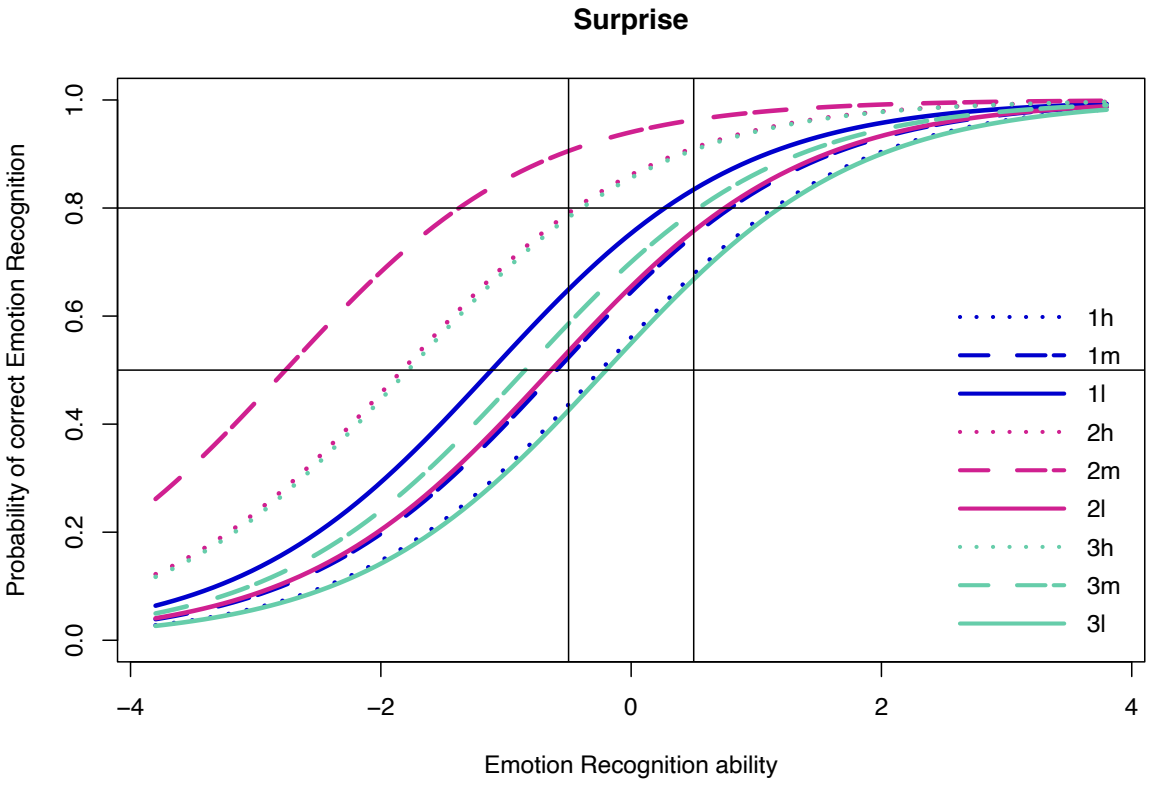
d)

Happiness
Probability of correct Emotion Recognition
0.0
0.2
0.4
0.6
0.8
1.0
−4
−2
0
2
4
Emotion Recognition ability
1h
1m
1l
2h
2m
2l
3h
3m
3l

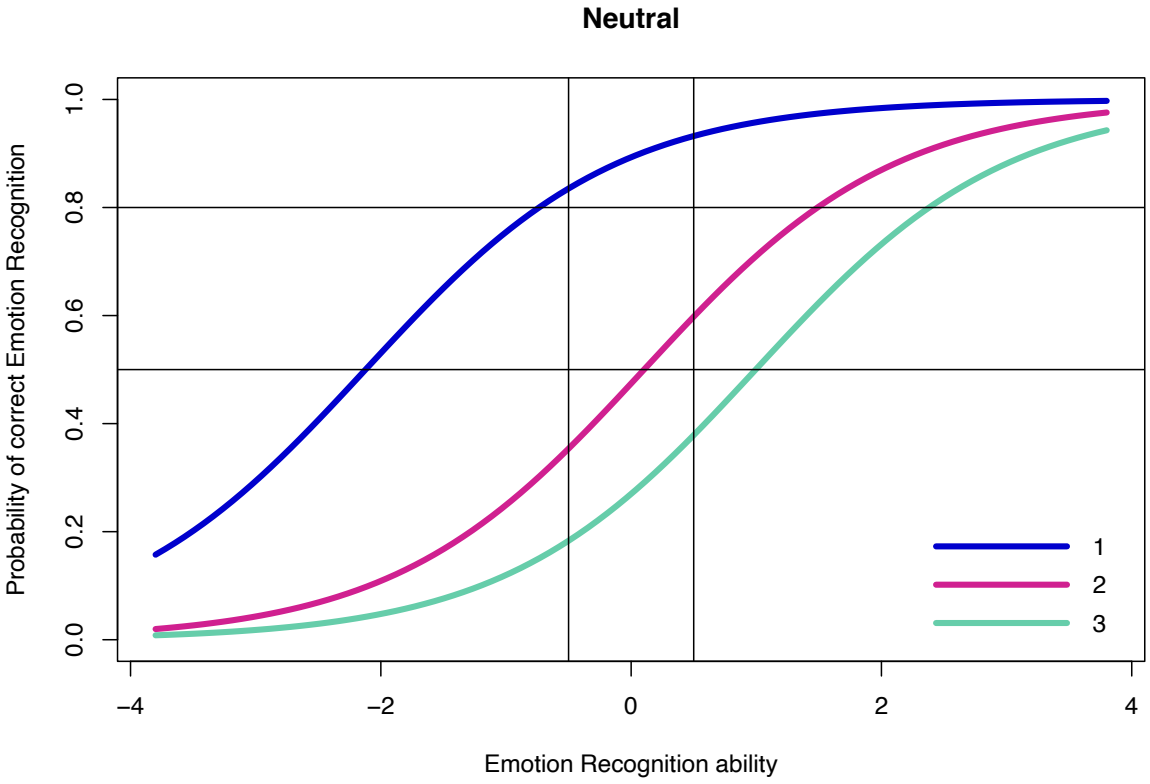
e)


Sadness
Probability of correct Emotion Recognition
0.0
0.2
0.4
0.6
0.8
1.0
−4
−2
0
2
4
Emotion Recognition ability
1h
1m
1l
2h
2m
2l
3h
3m
3l

f)


Surprise
Probability of correct Emotion Recognition
0.0
0.2
0.4
0.6
0.8
1.0
−4
−2
0
2
4
Emotion Recognition ability
1h
1m
1l
2h
2m
2l
3h
3m
3l

g)



h)

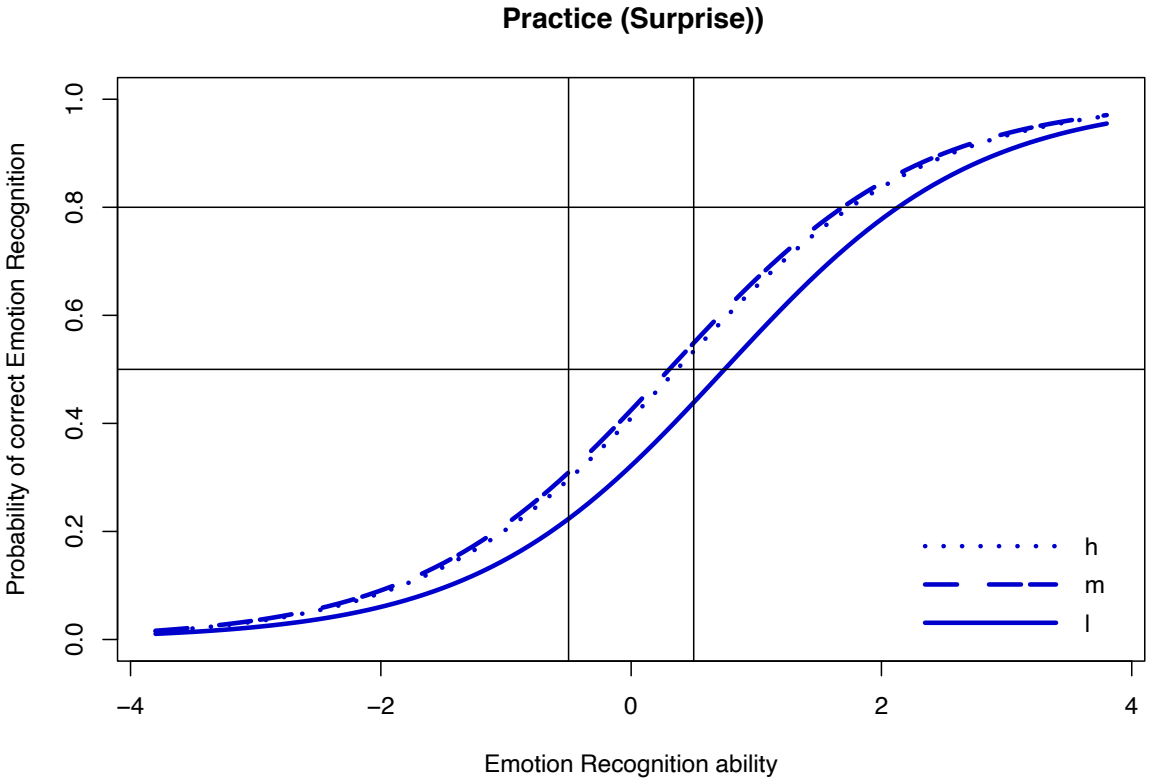
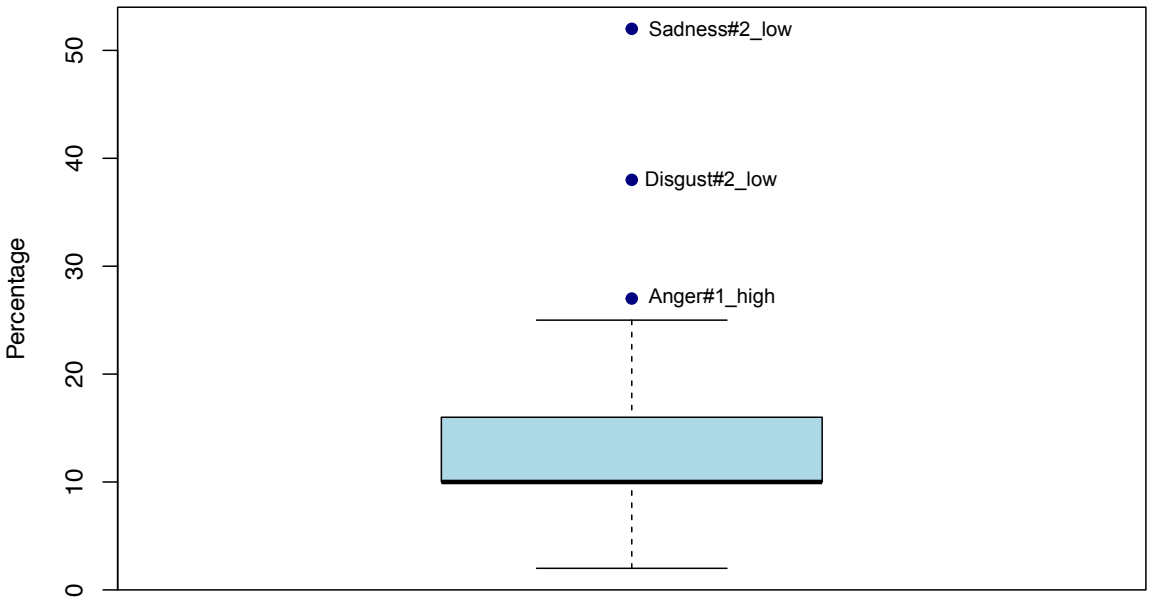


Figure K1a-h. Item Characteristic Curves (ICC) of the 60 intensity versions for the 18 basis emotion scenes (anger, disgust, fear, happiness, sadness, surprise) three neutral scenes and the practice (surprise) scene. Scene code consists of a number (scene) and a letter (emotional intensity). l = low intensity; m = medium intensity; h = high intensity.

Supplementary Material L

Figure L1

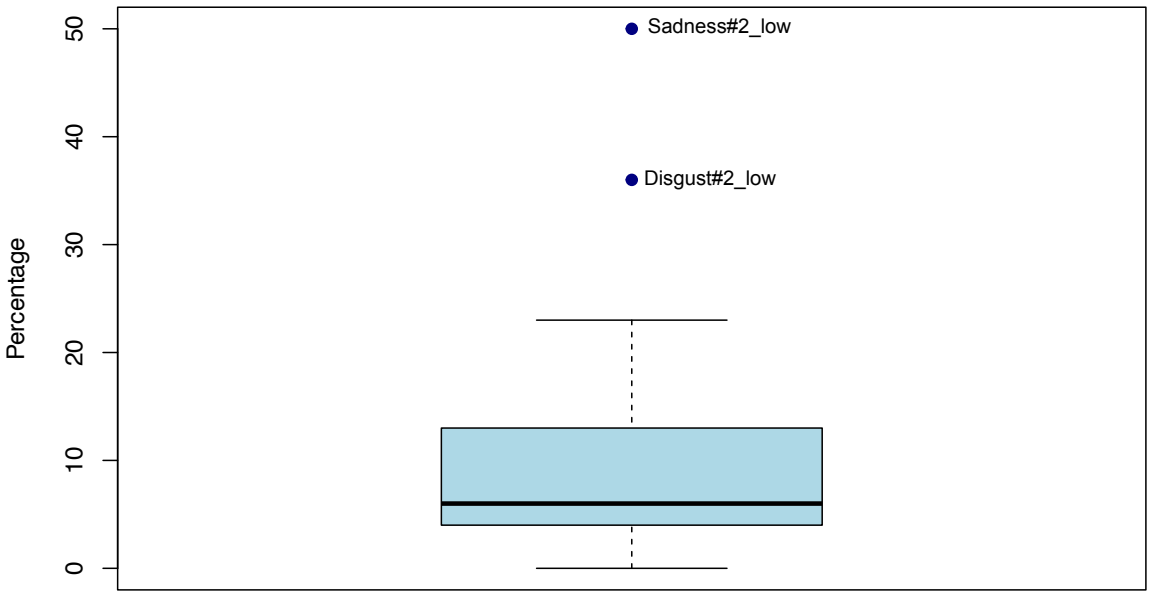
Distribution of the percentages of the "other emotion" terms named by the participants whose answers were classified as correct in the selected scenes



Note. The scene code of the outliers consists of the emotion portrayed, scene number and intensity version.

Figure L2

Distribution of the percentages of participants whose answers were classified as correct based on having named an "other emotion" term in the selected scenes



Note. The scene code of the outliers consists of the emotion portrayed, scene number and intensity version.

Table L3.1

Scene Anger#1_high: Percentages of the number of emotions (n=90), labelled and/or named by participants whose answers were classified as correct [68(85%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=90)	Percentage relative to the number of participants with correct answers (n=68)
Anger	**71%**	**94%**
Named emotions in the spectrum of anger	**12%**	**16%**
Disgust	1%	1%
Named emotions in the spectrum of disgust	4%	6%
Fear	4%	6%
Named emotions in the spectrum of fear	0%	0%
Happiness	0%	0%
Named emotions in the spectrum of happiness	0%	0%
Sadness	0%	0%
Named emotions in the spectrum of sadness	0%	0%
Surprise	3%	4%
Named emotions in the spectrum of surprise	0%	0%
Neutral	0%	0%
Named emotions in the spectrum of neutral	0%	0%
Disappointment (secondary emotion)	1%	1%

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=90)	Percentage relative to the number of participants with correct answers (n=68)
Tension (secondary emotion)	2%	3%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.2

Scene Anger#2_medium: Percentages of the number of emotions (n=99), labelled and/or named by participants whose answers were classified as correct [53(66%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=99)	Percentage relative to the number of participants with correct answers (n=53)
Anger	**51%**	**94%**
Named emotions in the spectrum of anger	**5%**	**9%**
Disgust	0%	0%
Named emotions in the spectrum of disgust	0%	0%
Fear	4%	8%
Named emotions in the spectrum of fear	1%	2%
Happiness	5%	9%
Named emotions in the spectrum of happiness	0%	0%
Sadness	1%	2%
Named emotions in the spectrum of sadness	1%	2%
Surprise	30%	57%
Named emotions in the spectrum of surprise	0%	0%
Neutral	2%	4%
Named emotions in the spectrum of neutral	0%	0%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.3

Scene Anger#3_medium: Percentages of the number of emotions (n=61), labelled and/or named by participants whose answers were classified as correct [35(44%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=61)	Percentage relative to the number of participants with correct answers (n=35)
Anger	**52%**	**91%**
Named emotions in the spectrum of anger	**5%**	**9%**
Disgust	2%	3%
Named emotions in the spectrum of disgust	0%	0%
Fear	0%	0%
Named emotions in the spectrum of fear	0%	0%
Happiness	3%	6%
Named emotions in the spectrum of happiness	0%	0%
Sadness	3%	6%
Named emotions in the spectrum of sadness	0%	0%
Surprise	33%	57%
Named emotions in the spectrum of surprise	0%	0%
Neutral	0%	0%
Named emotions in the spectrum of neutral	0%	0%
Tension (secondary emotion)	2%	3%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.4

Scene Disgust#1_high: Percentages of the number of emotions (n=62), labelled and/or named by participants whose answers were classified as correct [43(54%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=62)	Percentage relative to the number of participants with correct answers (n=43)
Anger	0%	0%
Named emotions in the spectrum of anger	0%	0%
Disgust	**66%**	**95%**
Named emotions in the spectrum of disgust	**6%**	**9%**
Fear	13%	19%
Named emotions in the spectrum of fear	0%	0%
Happiness	0%	0%
Named emotions in the spectrum of happiness	0%	0%
Sadness	0%	0%
Named emotions in the spectrum of sadness	0%	0%
Surprise	11%	16%
Named emotions in the spectrum of surprise	0%	0%
Neutral	2%	2%
Named emotions in the spectrum of neutral	0%	0%
Compassion (secondary emotion)	2%	2%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.5

Scene Disgust#2_low: Percentages of the number of emotions (n=67), labelled and/or named by participants whose answers were classified as correct [50(63%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=67)	Percentage relative to the number of participants with correct answers (n=50)
Anger	0%	0%
Named emotions in the spectrum of anger	0%	0%
Disgust	**51%**	**68%**
Named emotions in the spectrum of disgust	**28%**	**38%**
Fear	19%	26%
Named emotions in the spectrum of fear	0%	0%
Happiness	0%	0%
Named emotions in the spectrum of happiness	0%	0%
Sadness	0%	0%
Named emotions in the spectrum of sadness	0%	0%
Surprise	0%	0%
Named emotions in the spectrum of surprise	0%	0%
Neutral	1%	2%
Named emotions in the spectrum of neutral	0%	0%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.6

Scene Disgust#3_medium: Percentages of the number of emotions (n=68), labelled and/or named by participants whose answers were classified as correct [47(59%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=68)	Percentage relative to the number of participants with correct answers (n=47)
Anger	9%	13%
Named emotions in the spectrum of anger	0%	0%
Disgust	**69%**	**100%**
Named emotions in the spectrum of disgust	**1%**	**2%**
Fear	0%	0%
Named emotions in the spectrum of fear	0%	0%
Happiness	0%	0%
Named emotions in the spectrum of happiness	0%	0%
Sadness	0%	0%
Named emotions in the spectrum of sadness	0%	0%
Surprise	19%	28%
Named emotions in the spectrum of surprise	0%	0%
Neutral	1%	2%
Named emotions in the spectrum of neutral	0%	0%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.7

Scene Fear#1_low: Percentages of the number of emotions (n=71), labelled and/or named by participants whose answers were classified as correct [41(51%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=71)	Percentage relative to the number of participants with correct answers (n=41)
Anger	1%	2%
Named emotions in the spectrum of anger	0%	0%
Disgust	3%	5%
Named emotions in the spectrum of disgust	0%	0%
Fear	**56%**	**98%**
Named emotions in the spectrum of fear	**6%**	**10%**
Happiness	0%	0%
Named emotions in the spectrum of happiness	0%	0%
Sadness	3%	5%
Named emotions in the spectrum of sadness	0%	0%
Surprise	31%	54%
Named emotions in the spectrum of surprise	0%	0%
Neutral	0%	0%
Named emotions in the spectrum of neutral	0%	0%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.8

Scene Fear#2_low: Percentages of the number of emotions (n=80), labelled and/or named by participants whose answers were classified as correct [59(74%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=80)	Percentage relative to the number of participants with correct answers (n=59)
Anger	0%	0%
Named emotions in the spectrum of anger	0%	0%
Disgust	0%	0%
Named emotions in the spectrum of disgust	0%	0%
Fear	**74%**	**100%**
Named emotions in the spectrum of fear	**4%**	**5%**
Happiness	0%	0%
Named emotions in the spectrum of happiness	0%	0%
Sadness	3%	3%
Named emotions in the spectrum of sadness	0%	0%
Surprise	14%	19%
Named emotions in the spectrum of surprise	0%	0%
Neutral	3%	3%
Named emotions in the spectrum of neutral	0%	0%
Insecurity (secondary emotion)	4%	5%

No Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.9

Scene Fear#3_low: Percentages of the number of emotions (n=86), labelled and/or named by participants whose answers were classified as correct [49(61%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=86)	Percentage relative to the number of participants with correct answers (n=49)
Anger	0%	0%
Named emotions in the spectrum of anger	0%	0%
Disgust	0%	0%
Named emotions in the spectrum of disgust	0%	0%
Fear	**55%**	**96%**
Named emotions in the spectrum of fear	**5%**	**8%**
Happiness	2%	4%
Named emotions in the spectrum of happiness	0%	0%
Sadness	10%	18%
Named emotions in the spectrum of sadness	0%	0%
Surprise	27%	47%
Named emotions in the spectrum of surprise	0%	0%
Neutral	0%	0%
Named emotions in the spectrum of neutral	0%	0%
Tension (secondary emotion)	1%	2%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.10

Scene Happiness#1_low: Percentages of the number of emotions (n=63), labelled and/or named by participants whose answers were classified as correct [49(61%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=63)	Percentage relative to the number of participants with correct answers (n=49)
Anger	0%	0%
Named emotions in the spectrum of anger	0%	0%
Disgust	0%	0%
Named emotions in the spectrum of disgust	0%	0%
Fear	0%	0%
Named emotions in the spectrum of fear	0%	0%
Happiness	**73%**	**94%**
Named emotions in the spectrum of happiness	**13%**	**16%**
Sadness	0%	0%
Named emotions in the spectrum of sadness	0%	0%
Surprise	8%	10%
Named emotions in the spectrum of surprise	0%	0%
Neutral	6%	8%
Named emotions in the spectrum of neutral	0%	0%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.11

Scene Happiness#2_medium: Percentages of the number of emotions (n=79), labelled and/or named by participants whose answers were classified as correct [48(60%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=79)	Percentage relative to the number of participants with correct answers (n=48)
Anger	0%	0%
Named emotions in the spectrum of anger	0%	0%
Disgust	0%	0%
Named emotions in the spectrum of disgust	0%	0%
Fear	0%	0%
Named emotions in the spectrum of fear	0%	0%
Happiness	**59%**	**98%**
Named emotions in the spectrum of happiness	**3%**	**4%**
Sadness	0%	0%
Named emotions in the spectrum of sadness	0%	0%
Surprise	32%	52%
Named emotions in the spectrum of surprise	1%	2%
Neutral	4%	6%
Named emotions in the spectrum of neutral	0%	0%
Curiosity	1%	2%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.12

Scene Happiness#3_low: Percentages of the number of emotions (n=68), labelled and/or named by participants whose answers were classified as correct [52(65%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=68)	Percentage relative to the number of participants with correct answers (n=52)
Anger	0%	0%
Named emotions in the spectrum of anger	0%	0%
Disgust	0%	0%
Named emotions in the spectrum of disgust	0%	0%
Fear	0%	0%
Named emotions in the spectrum of fear	0%	0%
Happiness	**66%**	**87%**
Named emotions in the spectrum of happiness	**21%**	**27%**
Sadness	0%	0%
Named emotions in the spectrum of sadness	0%	0%
Surprise	9%	12%
Named emotions in the spectrum of surprise	0%	0%
Neutral	4%	6%
Named emotions in the spectrum of neutral	0%	0%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.13

Scene Sadness#1_high : Percentages of the number of emotions (n=91), labelled and/or named by participants whose answers were classified as correct [58(73%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=91)	Percentage relative to the number of participants with correct answers (n=58)
Anger	3%	5%
Named emotions in the spectrum of anger	0%	0%
Disgust	0%	0%
Named emotions in the spectrum of disgust	1%	2%
Fear	23%	36%
Named emotions in the spectrum of fear	0%	0%
Happiness	2%	3%
Named emotions in the spectrum of happiness	0%	0%
Sadness	**62%**	**97%**
Named emotions in the spectrum of sadness	**2%**	**3%**
Surprise	1%	2%
Named emotions in the spectrum of surprise	0%	0%
Neutral	1%	2%
Named emotions in the spectrum of neutral	0%	0%
Optimism (secondary emotion)	1%	2%

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=91)	Percentage relative to the number of participants with correct answers (n=58)
Determination (secondary emotion)	1%	2%
Guilt (secondary emotion)	1%	2%
Pessimism	1%	2%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.14

Scene Sadness#2_low : Percentages of the number of emotions (n=60), labelled and/or named by participants whose answers were classified as correct [48(60%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=60)	Percentage relative to the number of participants with correct answers (n=48)
Anger	0%	0%
Named emotions in the spectrum of anger	0%	0%
Disgust	0%	0%
Named emotions in the spectrum of disgust	0%	0%
Fear	0%	0%
Named emotions in the spectrum of fear	0%	0%
Happiness	0%	0%
Named emotions in the spectrum of happiness	0%	0%
Sadness	**52%**	**65%**
Named emotions in the spectrum of sadness	**43%**	**54%**
Surprise	2%	2%
Named emotions in the spectrum of surprise	0%	0%
Neutral	3%	4%
Named emotions in the spectrum of neutral	0%	0%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.15

Scene Sadness#3_low : Percentages of the number of emotions (n=74), labelled and/or named by participants whose answers were classified as correct [54(68%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=74)	Percentage relative to the number of participants with correct answers (n=54)
Anger	0%	0%
Named emotions in the spectrum of anger	0%	0%
Disgust	0%	0%
Named emotions in the spectrum of disgust	0%	0%
Fear	15%	20%
Named emotions in the spectrum of fear	0%	0%
Happiness	0%	0%
Named emotions in the spectrum of happiness	0%	0%
Sadness	**65%**	**89%**
Named emotions in the spectrum of sadness	**18%**	**24%**
Surprise	3%	4%
Named emotions in the spectrum of surprise	0%	0%
Neutral	0%	0%
Named emotions in the spectrum of surprise	0%	0%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.16

Scene Surprise#1_medium: Percentages of the number of emotions (n=78), labelled and/or named by participants whose answers were classified as correct [51(64%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=78)	Percentage relative to the number of participants with correct answers (n=51)
Anger	0%	0%
Named emotions in the spectrum of anger	0%	0%
Disgust	0%	0%
Named emotions in the spectrum of disgust	0%	0%
Fear	0%	0%
Named emotions in the spectrum of fear	0%	0%
Happiness	29%	45%
Named emotions in the spectrum of happiness	0%	0%
Sadness	1%	2%
Named emotions in the spectrum of sadness	0%	0%
Surprise	**64%**	**98%**
Named emotions in the spectrum of surprise	**4%**	**6%**
Neutral	1%	2%
Named emotions in the spectrum of surprise	0%	0%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.17

Scene Surprise#2_low : Percentages of the number of emotions (n=66), labelled and/or named by participants whose answers were classified as correct [50(63%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=66)	Percentage relative to the number of participants with correct answers (n=50)
Anger	2%	2%
Named emotions in the spectrum of anger	0%	0%
Disgust	0%	0%
Named emotions in the spectrum of disgust	2%	2%
Fear	0%	0%
Named emotions in the spectrum of fear	0%	0%
Happiness	2%	2%
Named emotions in the spectrum of happiness	0%	0%
Sadness	6%	8%
Named emotions in the spectrum of sadness	3%	4%
Surprise	**74%**	**98%**
Named emotions in the spectrum of surprise	**5%**	**6%**
Neutral	8%	10%
Named emotions in the spectrum of surprise	0%	0%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.18

Scene Surprise#3_medium : Percentages of the number of emotions (n=77), labelled and/or named by participants whose answers were classified as correct [53(66%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=77)	Percentage relative to the number of participants with correct answers (n=53)
Anger	0%	0%
Named emotions in the spectrum of anger	0%	0%
Disgust	0%	0%
Named emotions in the spectrum of disgust	1%	2%
Fear	19%	28%
Named emotions in the spectrum of fear	0%	0%
Happiness	1%	2%
Named emotions in the spectrum of happiness	0%	0%
Sadness	0%	0%
Named emotions in the spectrum of sadness	0%	0%
Surprise	**66%**	**98%**
Named emotions in the spectrum of surprise	**3%**	**4%**
Neutral	3%	4%
Named emotions in the spectrum of surprise	0%	0%

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=77)	Percentage relative to the number of participants with correct answers (n=53)
Insecurity (secondary emotion)	3%	4%
Tension (secondary emotion)	1%	2%
Curiosity	1%	2%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.19

Scene Neutral#1 : Percentages of the number of emotions (n=216), labelled and/or named by participants whose answers were classified as correct [206(86%) out of 240 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=216)	Percentage relative to the number of participants with correct answers (n=206)
Anger	0%	0%
Named emotions in the spectrum of anger	0%	0%
Disgust	0%	0%
Named emotions in the spectrum of disgust	0%	0%
Fear	0%	0%
Named emotions in the spectrum of fear	0%	0%
Happiness	2%	2%
Named emotions in the spectrum of happiness	0%	0%
Sadness	0%	0%
Named emotions in the spectrum of sadness	0%	0%
Surprise	1%	1%
Named emotions in the spectrum of surprise	0%	0%
Neutral	**95%**	**100%**
Named emotions in the spectrum of surprise	**0%**	**0%**

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=216)	Percentage relative to the number of participants with correct answers (n=206)
Satisfaction	0%	0%
Tension (secondary emotion)	0%	0%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.20

Scene Neutral#2: Percentages of the number of emotions (n=142), labelled and/or named by participants whose answers were classified as correct [115(48%) out of 80 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=142)	Percentage relative to the number of participants with correct answers (n=115)
Anger	0%	0%
Named emotions in the spectrum of anger	0%	0%
Disgust	0%	0%
Named emotions in the spectrum of disgust	0%	0%
Fear	0%	0%
Named emotions in the spectrum of fear	0%	0%
Happiness	2%	3%
Named emotions in the spectrum of happiness	0%	0%
Sadness	0%	0%
Named emotions in the spectrum of sadness	0%	0%
Surprise	15%	19%
Named emotions in the spectrum of surprise	0%	0%
Neutral	**81%**	**100%**
Named emotions in the spectrum of surprise	**0%**	**0%**

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=142)	Percentage relative to the number of participants with correct answers (n=115)
Curiosity	1%	1%
Interest	1%	1%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Table L3.21

Scene Neutral#3: Percentages of emotions (n=79), labelled and/or named by participants whose answers were classified as correct [73(30%) out of 240 participants]

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=79)	Percentage relative to the number of participants with correct answers (n=73)
Anger	3%	3%
Named emotions in the spectrum of anger	1%	1%
Disgust	0%	0%
Named emotions in the spectrum of disgust	1%	1%
Fear	1%	1%
Named emotions in the spectrum of fear	0%	0%
Happiness	0%	0%
Named emotions in the spectrum of happiness	0%	0%
Sadness	0%	0%
Named emotions in the spectrum of sadness	1%	1%
Surprise	0%	0%
Named emotions in the spectrum of surprise	0%	0%
Neutral	**91%**	**99%**
Named emotions in the spectrum of surprise	**0%**	**0%**

Emotion	*Percentage relative to the number of labelled and/or named emotions (n=79)	Percentage relative to the number of participants with correct answers (n=73)
Tension (secondary emotion)	1%	1%

Note. Bold = Basic emotion portrayed in the scene.

* Rounded percentages might not necessarily add up to 100%.

Supplementary Material

Manuskript 3

The Basel Version of the Awareness of Social Inference Test – Theory of Mind (BASIT-ToM):

Preliminary Validation Analyses in Healthy Adults

Table of Contents

Supplementary Material A **137**

Table A1 ..... *137*

Evaluation criteria of the TASIT-SIM scenes for scene selection 137

Supplementary Material B **138**

Table B1.1 ..... *138*

Evaluation and selection of the Honest scenes from TASIT-SIM form A and form B 138

Table B1.2 ..... *140*

Evaluation and selection of the Simple Sarcasm scenes from TASIT-SIM form A and form B 140

Table B1.3 ..... *142*

Evaluation and selection of the Paradoxical sarcasm scenes from TASIT-SIM form A and form B 142

Table B2 ..... *144*

Selected scenes for the BASIT-ToM film project 144

Supplementary Material C **145**

Document C1 ..... *145*

Screenplays of the BASIT-ToM Scenes 145

Table C1 ..... *159*

Main text changes of the BASIT-ToM screenplay 159

Supplementary Material D **160**

Table D1 ..... *160*

Distribution of the actors to the scenes of the BASIT-ToM and the corresponding scenes of TASIT-SIM . 160

Supplementary Material E **161**

Document E1 ..... *161*

Background and intensity information of each BASIT-ToM scene for the actors 161

Supplementary Material F 180

Table F1: Scene H#1 180

Test questions of the BASIT-ToM and TASIT-SIM with their corresponding correct answers, type and order of Theory of Mind 180

Table F2: Scene H#2 182

Test questions of the BASIT-ToM and TASIT-SIM with their corresponding correct answers, type and order of Theory of Mind 182

Table F3: Scene H#3 184

Test questions of the BASIT-ToM and TASIT-SIM with their corresponding correct answers, type and order of Theory of Mind 184

Table F4: Scene H#4 186

Test questions of the BASIT-ToM and TASIT-SIM with their corresponding correct answers, type and order of Theory of Mind 186

Table F5: Scene sS#1 188

Test questions of the BASIT-ToM and TASIT-SIM with their corresponding correct answers, type and order of Theory of Mind 188

Table F6: Scene sS#2 190

Test questions of the BASIT-ToM and TASIT-SIM with their corresponding correct answers, type and order of Theory of Mind 190

Table F7: Scene sS#3 192

Test questions of the BASIT-ToM and TASIT-SIM with their corresponding correct answers, type and order of Theory of Mind 192

Table F8: Scene sS#4 194

Test questions of the BASIT-ToM and TASIT-SIM with their corresponding correct answers, type and order of Theory of Mind 194

Table F9: Scene pS#1 196

Test questions of the BASIT-ToM and TASIT-SIM with their corresponding correct answers, type and order of Theory of Mind 196

Table F10: Scene pS#2 ..... *198*

Test questions of the BASIT-ToM and TASIT-SIM with their corresponding correct answers, type and order of Theory of Mind 198

Table F11: Scene pS#3 ..... *200*

Test questions of the BASIT-ToM and TASIT-SIM with their corresponding correct answers, type and order of Theory of Mind 200

Table F12: Scene pS#4 ..... *202*

Test questions of the BASIT-ToM and TASIT-SIM with their corresponding correct answers, type and order of Theory of Mind 202

Supplementary Material G 206

Table G1 ..... *206*

Differences between the BASIT-ToM and TASIT-SIM 206

Supplementary Material H 208

Document H1 ..... *208*

Details of the programming an data storage 208

Supplementary Material I 212

Figure I1 ..... *212*

Exemplary representation of the BASIT-ToM computer based application process 212

Supplementary Material J 213

Table J.1: Honesty scenes *213*

Overview of the intensity versions fitting the Rasch model and the number of the respective ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities 213

Table J.2: Paradoxical sarcasm scenes *217*

Overview of the intensity versions fitting the Rasch model and the number of the respective ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities 217

Table J.3: Simple Sarcasm scenes *222*

Overview of the intensity versions fitting the Rasch model and the number of the respective ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities 222

Supplementary Material 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.

Supplementary Material B

Table B1.1

Evaluation and selection of the Honest scenes from TASIT-SIM form A and form B

	Scene ID									
Evaluation criteria	A1	A4	A7	**A11**	A14	**B1**	B4	**B7**	**B11**	B14
The topic of the film scene is suitable for the age groups over 60	1	1	1	1	1	1	1	1	1	0
The topic of the film scene is suitable for people from German-speaking cultural area	1	1	1	1	1	1	1	1	1	1
The film scene appears realistic and comprehensible	1	1	1	1	1	0	1	1	1	1
The content of the film scene is contemporary	1	1	1	1	1	1	1	1	1	1
Set design appears realistic	1	1	1	1	0	1	1	1	0	0
The message type portrayed appears appropriate for the story of the scene	0	0	0	1	0	1	0	1	1	1
The message type is portrayed by different modalities (e.g. gestures, facial expressions, linguistic characteristics)	1	1	1	1	1	1	1	1	1	1
Office or living room setting is possible	1	1	1	1	1	1	1	1	1	1

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-EET form A, B = TASIT-EET 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 B1.2

Evaluation and selection of the Simple Sarcasm scenes from TASIT-SIM form A and form B

	Scene ID									
Evaluation criteria	A2	**A6**	A9	A10	A13	**B2**	B6	B9	**B10**	**B13**
The topic of the film scene is suitable for the age groups over 60	1	1	1	1	1	1	1	1	1	1
The topic of the film scene is suitable for people from German-speaking cultural area	1	1	1	1	1	1	1	1	1	1
The film scene appears realistic and comprehensible	1	1	0	1	1	1	1	1	1	1
The content of the film scene is contemporary	1	1	1	1	1	1	1	1	1	1
Set design appears realistic	1	1	1	0	1	1	0	1	1	1
The message type portrayed appears appropriate for the story of the scene	0	1	0	0	1	1	1	0	1	0

Evaluation criteria	Scene ID									
	A2	**A6**	A9	A10	A13	**B2**	B6	B9	**B10**	**B13**
The message type is portrayed by different modalities (e.g. gestures, facial expressions, linguistic characteristics)	1	1	1	1	1	1	1	1	1	1
Office or living room setting is possible	1	1	1	1	1	1	1	1	1	1
9) 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	9/9	7/9	7/9	9/9	9/9	8/9	8/9	9/9	8/9

Note. TASIT-SIM= The Awareness of Social Inference Test-Social Inference Minimal; A = TASIT-EET form A, B = TASIT-EET 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 B1.3

Evaluation and selection of the Paradoxical sarcasm scenes from TASIT-SIM form A and form B

Evaluation criteria	Scene ID									
	A3	A5	A8	A12	**A15**	B3	**B5**	**B8**	B12	**B15**
The topic of the film scene is suitable for the age groups over 60	1	1	1	1	1	1	1	1	1	1
The topic of the film scene is suitable for people from German-speaking cultural area	1	1	1	1	1	1	1	1	1	1
The film scene appears realistic and comprehensible	1	1	1	1	1	1	1	1	0	1
The content of the film scene is contemporary	1	1	1	1	1	1	1	1	1	1
Set design appears realistic	1	1	0	1	1	0	1	1	0	1
The message type portrayed appears appropriate for the story of the scene	0	0	1	1	1	1	1	1	1	1

Evaluation criteria	Scene ID									
	A3	A5	A8	A12	**A15**	B3	**B5**	**B8**	B12	**B15**
The message type is portrayed by different modalities (e.g. gestures, facial expressions, linguistic characteristics)	1	1	1	1	1	1	1	1	1	1
Office or living room setting is possible	1	1	1	1	1	1	1	1	1	1
9) 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	9/9	8/9	9/9	9/9	7/9	9/9

Note. TASIT-SIM= The Awareness of Social Inference Test-Social Inference Minimal; A = TASIT-EET form A, B = TASIT-EET 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 Infenerence Test – Theory of Mind. Scene code consists of a letter indicating the form of The Awareness of Social Inference Test (i.e. form A or form B) and a number indicating the respective scene number.

Supplementary Material C

Document C1

Screenplays of the BASIT-ToM Scenes

23 ÜBUNGSZENE 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 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

(CONTINUED)

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!

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!

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.

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.

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!

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.

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.

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.

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.

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.

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 Schuften.

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.

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 text changes of the BASIT-ToM screenplay

Scene	Deleted text that was not transferred from the TASIT-SIM script to the BASIT-ToM screenplay	Text added to the BASIT-ToM screenplay
H#1 (B11)	The Actor repeats what the actress say 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. M=male actor, F= female actor.

Supplementary Material D

Table D1

Distribution of the actors to the scenes of the BASIT-ToM and the corresponding scenes of TASIT-SIM

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 communicate the respective message type (i.e., honesty, simple sarcasm, 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.

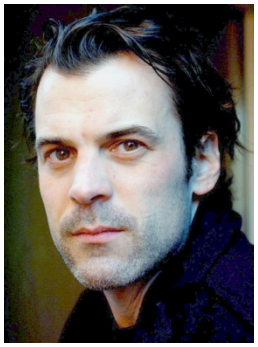
Supplementary Material E

Document E1

Background and intensity information of each BASIT-ToM scene for the actors

Honesty 1 (B11) – Theatre

Actor with honest communication: **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 M2 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.

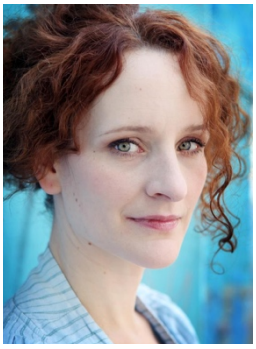
- **Medium intensity:** M1 and F1 have seen a play the night before. They are sitting on the sofa, drinking wine and talking. They rarely go to the theatre, but found the play really good.

- **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

Actor with honest communication: **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 also takes on private organisational tasks, such as booking tickets.

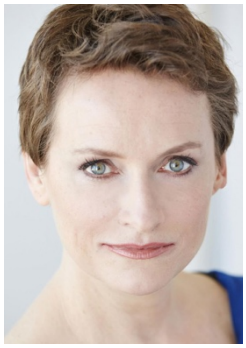
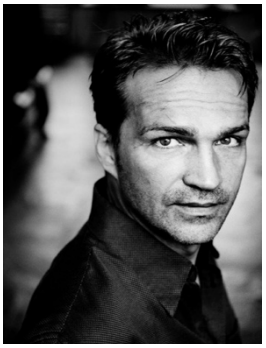
Intensities

- **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 the biggest 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 them for the last Rolling Stones concert and proudly presents him with the tickets. M2 feels excited anticipation and is full of gratitude that F2 organized the tickets for him.

Honesty 3 (B1) – Hard Day

Actor with honest communication: **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 clean-up. 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 out 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 out. 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 organise 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

Actor with honest communication: **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

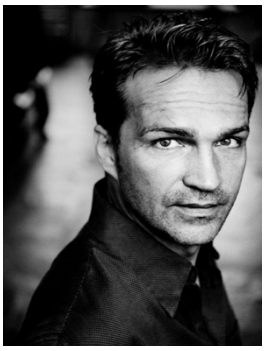
- **Low intensity:** The conversation takes place while they are working together. Meanwhile, M4 wants to talk to F4 for a bit and therefore brings up the subject of the

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”

Actor with simple sarcastic communication: **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 2 children. F1 can only reconcile work and family life with a great deal of organisation 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

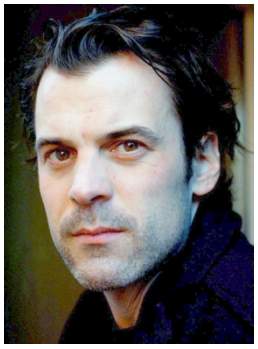
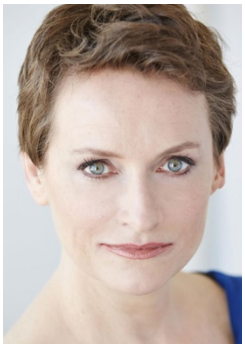
annoys her because she relied on his commitment. If she does not find a replacement, she will have to hold the seminar herself.

- **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 also 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

Actor with simple sarcastic communication: **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

Actor with simple sarcastic communication: **M2**	Communication partner: F4

Background information

- F4 is a gallery owner. She is a confident, content person.

- M2 is an artist. He is confident, articulate and good with people.

Intensities

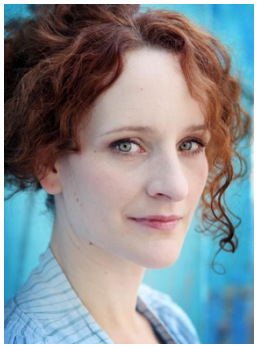
- **Low intensity:** F4 and M2 have been a happy couple for many years. They harmonize very well. Their conversations are usually full of small but affectionate banter. They like to tease each other. They don't have to give their partner any linguistic or visual clues that what they say is meant ironically. 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, 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 who 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 also reacts with irony.

Simple sarcasm 4 (B13) – Shirt

Actor with simple sarcastic communication: **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.

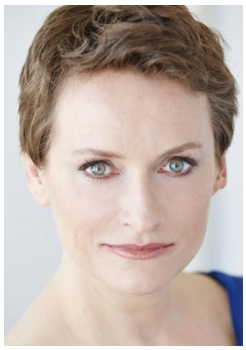
Intensities

- **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

Actor with paradox sarcastic communication: **F3**	Communication partner: M2
	

Background information

- F3 is a management consultant, she works efficiently and reliably and is very 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.

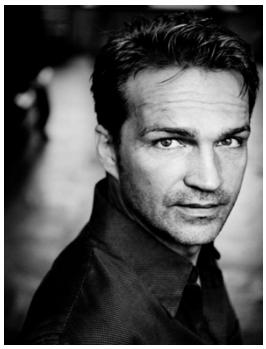
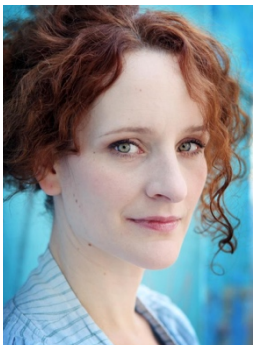
- **Low intensity:** M2 and F3 appreciate each other and have a collegial relationship. F3 is used to writing long reports for her boss at short notice. She has no problem with this, yet she would have liked to spend the weekend doing something else. M2 wants to convey appreciation and thanks to her 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 realises 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

Actor with paradox sarcastic communication: **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

Actor with paradox sarcastic communication: **M4**	Communication partner: F1

Background information

- F1 is a journalist

- M4 is a computer scientist

Intensities

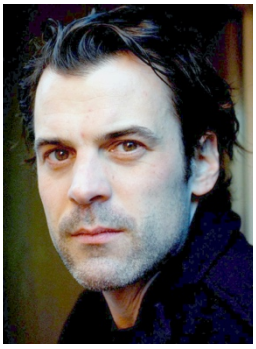
- **Low intensity:** F1 and M4 have been a happy couple for many years. They harmonize very well and understand each other blindly. Their conversations are usually full of small but affectionate asides. They like to tease each other. They don't have to give their partner any linguistic or visual clues 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

Actor with paradox sarcastic communication: **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

Actor with paradox sarcastic communication: **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 color.

- **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 color for her house. However, M2 finds his wife's gossiping excessive. He is not really 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.

Supplementary Material F

Table F1: Scene H#1

Test questions of the BASIT-ToM and TASIT-SIM with their corresponding correct answers, type and order 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	2nd	No	Is Michael agreeing with Ruth about the movie?	cogn	2nd	yes	a
War er von dem Theaterstück gelangweilt?	Was he bored by the play?	aff	1st	No	Is he openly pleased that he saw the movie?	aff	1st	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	1st	No	Is he trying to say he thought the actors were good?	aff	1st	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	2nd	yes	Does he think the movie was bad?	aff	1st	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; 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, type and order 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	1st	yes	Is Michael trying to show he appreciates Gary getting the tickets?	aff	1st	yes	a
Freut er sich, dass sie ihm die Eintrittskarten organisiert hat?	Is he happy that she organised the tickets for him	aff	1st	yes	Is Michael annoyed Gary got him the tickets?	aff	1st	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	1st	yes	Is he trying to say he's pleased about the tickets?	aff	1st	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	2nd	no	By the end of the scene, does Gary think Michael wants to go?	cogn	2nd	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; 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, type and order 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	2nd	yes	Is Ruth trying to make Michael feel appreciated?	aff	2nd	yes	a
Ärgert sie sich über ihn?	Is she annoyed with him?	aff	1st	no	Is she annoyed with him?	aff	1st	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	1st	no	Is he trying to say that he has been a big help?	cogn	1st	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	2nd	yes	Is she annoyed with him?	aff	1st	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; 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, type and order 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	1st	no	Is Ruth sending Michael up about his chances of a promotion?	cogn	1st	yes	a
Würde sie ihm die Beförderung missgönnen?	Would she begrudge him the promotion?	aff	1st	no	Would she like him to get the promotion?	aff	1st	yes	b
Versucht sie ihm zu sagen, dass er gut gearbeitet hat?	Is she trying to say he has worked really well?	cogn	1st	yes	Is she trying to say he has worked really well?	cogn	1st	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	2nd	yes	Does she think he deserves a promotion?	aff	1st	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; 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, type and order 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	2nd	no	Is Ruth trying to make Michael feel OK?	aff	2nd	no	a
Ärgert sie sich über seine Absage?	Is she annoyed by his cancellation?	aff	1st	yes	Is she annoyed with him?	aff	1st	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	1st	yes	Is she trying to say that he is causing a big problem?	cogn	1st	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	2nd	no	Does she believe he is too busy to take the class?	cogn	2nd	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; 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, type and order 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	1st	no	Is he seriously suggesting they inivite other family members?	cogn	1st	no	a
Ärgert er sich über sie?	Is he annoyed with her?	aff	1st	yes	Is he annoyed with her?	aff	1st	yes	b
Versucht er ihr zu sagen, dass ihre Verwandten mitkommen sollen?	Is he trying to say he want her relatives to come?	cogn	1st	no	Is he trying to say he doesn't want her relatives to come?	cogn	1st	yes	c
Denkt sie, er möchte, dass ihre Verwandten mitkommen?	Does she think he wants her relatives to come?	cogn	2nd	no	Does she think he wants her relatives to come?	cogn	2nd	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; 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, type and order 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	2nd	no	Is Gary trying to make Ruth feel bad?	aff	2nd	no	a
Ist er wütend auf sie?	Is he angry with her?	aff	1st	no	Is he angry with her?	aff	1st	no	b
Versucht er ihr zu sagen, dass sie gut aussieht?	Is he trying to say that she looks good?	cogn	1st	yes	Is he trying to say that she looks ok?	cogn	1st	yes	c
Denkt sie, dass er scherzt?	Does she think he is joking with her?	cogn	2nd	yes	Does she think he is joking with her?	cogn	2nd	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; 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, type and order 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	1st	yes	Is Ruth reassuring Gary that the shirt is nice?	cogn	1st	no	a
Ärgert sie sich darüber, dass er das Hemd trägt?	Is she annoyed that he is wearing the shirt?	aff	1st	yes	Is she happy for him to wear the shirt?	aff	1st	no	b
Versucht sie ihm zu sagen, dass sie das Hemd schön findet?	Is she trying to say the shirt is beautiful?	cogn	1st	no	Is she trying to say the shirt is awful?	cogn	1st	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	2nd	yes	Does she think the shirt's OK?	cogn	1st	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; 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, type and order 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	1st	yes	Is Ruth denying the report took a lot of work?	cogn	1st	no	a
Ist sie unzufrieden, weil sie am Wochenende arbeiten musste?	Does Ruth seem unhappy about working all weekend?	aff	1st	yes	Does Ruth seem happy about working all weekend?	aff	1st	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	1st	yes	Is she trying to say she had a lazy, relaxing weekend?	cogn	1st	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	2nd	yes / no*	Does Michael think she took it easy on the weekend?	cogn	1st	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; cogn=cognitive Theory of Mind; aff=affective Theory of Mind, hyper=hypermentalization.

*correct response = yes (low intensity), correyt 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, type and order 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 Ohrringe äussern?	Is he commenting negatively on the earrings?	cogn	1st	yes	Is Michael being complimentary about the dress?	cogn	1st	no	a
Amüsiert er sich über die Qualität der Ohrringe?	Is she making fun of the quality of the earrings?	aff	1st	yes	Does he like the dress?	aff	1st	no	b
Versucht er ihr zu sagen, dass die Ohrringe hochwertig aussehen?	Is he trying to tell her that the earrings look high quality?	cogn	1st	no	Is he trying to say the dress looks cheap?	cogn	1st	yes	c
Denkt er, dass ihr die Ohrringe gefallen?	Does he think she likes the earrings?	aff	2nd	no	Does he think Ruth's sister paid a lot for the dress?	cogn	1st	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; 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, type and order 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	1st	no	Is Gary trying to compliment Ruth on the meal?	cogn	1st	yes	a
Hat er das Essen genossen?	Did he enjoy the meal?	aff	1st	yes	Did he enjoy the meal?	aff	1st	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	1st	no	Is he trying to say that the meal tasted awful?	aff	1st	no	c
Denkt sie, dass ihm das Essen schlecht geschmeckt hat?	Does she think he did not like the meal?	cogn	2st	no	Does he think she is a good cook?	cogn	1st	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; 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, type and order 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	2nd	yes	Is Ruth trying to make Gary feel OK about cancelling?	aff	2nd	no	a
Ist sie von seiner Absage enttäuscht?	Is she disappointed by his cancellation?	aff	1st	yes	Is she happy to cancel?	aff	1st	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	1st	no	Is she trying to say she wanted to go out that night?	cogn	1st	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	2nd	no	*By the end of the scene:* Does he think she's upset about cancelling dinner?	aff	2nd	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; 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, type and order 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	1st	no	Is Ruth being complementary about her friend's house?	cogn	1st	no	a
Gefällt ihr die Farbgebung des Hauses?	Does she like the color scheme of the house?	aff	1st	no	Does she disapprove of the colour scheme?	aff	1st	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	1st	yes	Is she trying to say the colour scheme is dreadful?	cogn	1st	yes	c
Denkt er, dass ihr Simones neues Haus gefällt?	Does he think she likes Simone's new house?	aff	2nd	no	Does she think her friend has good taste?	cogn	1st	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; cogn=cognitive Theory of Mind; aff=affective Theory of Mind, hyper=hypermentalization.

Supplementary Material G

Table G1

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, apart from M3 and F1, who portray paradoxical sarcasm a second time in the practice scene	Unevenly distributed across message types
Gender distribution	Evenly across message types. Always one male actor and one female actor portrays the message types. 50% of the main roles of the test scenes are casted by women	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 emotional manner	Communication partners act in a neutral emotional state

	BASIT-ToM	TASIT-SIM
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
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; partly close-ups of actors talking to the camera); 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., 1st order affective ToM: *n*=17, 1st order cognitive ToM: *n*=17, 2nd order affective ToM: *n*=10, 2nd order cognitive ToM: *n*=8, hypermentalization: *n*=13)	4 test questions per each scene divided based on content (i.e., 1st order affective ToM: *n*=18, 1st order cognitive ToM: *n*=24, 2nd order affective ToM: *n*=5, 2nd order cognitive ToM: *n*=5)
	Questions refer to both actors of the scenes	Questions refer to one actor of each scene

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

Supplementary Material H

Document H1

Details of the programming an data storage

The BASIT-ToM was implemented as part of a suit of four Emotion recognition tests (Emotional Evaluation Test, Social Inference, Congruent Facial Emotional Perception Task, Congruent and Incongruent Facial Emotion Perception Task) and a Gender Discri- mination Task. 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 additional constraint that no successive imessage type were equal. Then for each participant ID the order of the four scenes per message type was randomized.

Then, the intensities were randomly distributed: The 240 first scenes of each message type were randomly splitted in 3 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 2 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 3 parts of 80 scenes with low, medium, and high intensity, for each intensity of the first message type the second message type was randomly splitted into 2 parts of 40 scenes with the two other intensities, and 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 exercise scene shown to each participant at the beginning of the test was randomly assigned, such that each intensity of the exercise scene is shown to 80 participants.

This resulted in:

- Each participant is shown 12 scences (4 per message type),
- Consecutive scenes show different message types,
- Each participant is shown at least one low intensity, one medium intensity, and one high intensity scene per message type
- Each participant is shown 4 low intensity, 4 medium intensity, and 4 high intensity scenes in total.
- Each scene and intensity is shown to 80 participants.
- The first scene shown is the exercise scence, shown in each intensity to 80 participants, respectively.

  The sequence of questions for the five type of minds (“doing”, “saying”, “thinking”, “feeling”, “hypermentalization”) phrased for each scene, and whether the question was phrased with a “Yes” or a “No” as correct answer, was also pseudo-randomly fixed during the development of the test using R: For each message type (“Honesty”, “Simple Sarcasm”, “Paradoxical Sarcasm”), the correct answers for the questions for “doing”, “saying”, “thinking” and “feeling” were randomly fixed to 50% “Yes” and 50% “No” answers over all four scenes and four type of minds per message type. However, for 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 these five questions was randomly determinend for each scene (independent of the intensity).

Note: The final list of correct Yes/No answers “2018-08-22_Endfassung_Fragen_mit_- ToM.docx”, received on 23.08.2017 (!), deviates from my list me15SollbergerAN_Ran- domOrderQuestions.pdf, sent on 27.07.2017 for “paradox sarcasm” scenes “Ohrringe” (PS#2) und “Essen” (PS#3).

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 13 scenes (1 exercise scene, four scenes of three interaction type) for a specific participant was loaded into the application at start, after a Welcome screen was shown and the examinator entered the participant ID, based on this ID.

The test started with a screen summarizing the instructions to the participant. On click of “next”, a test run was performed using the Exercise scene. Afterwards, the participant is asked whether he/she understood the task and encouraged to aks the experimentator for clarification if not. Otherwise the participant can click on “next” to start the task. 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, as described in the main text.

After the final scene was answered, the 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 Departement of the UHBS.

The participants were shown a message that the task was completed, before the application started the next test.

References :

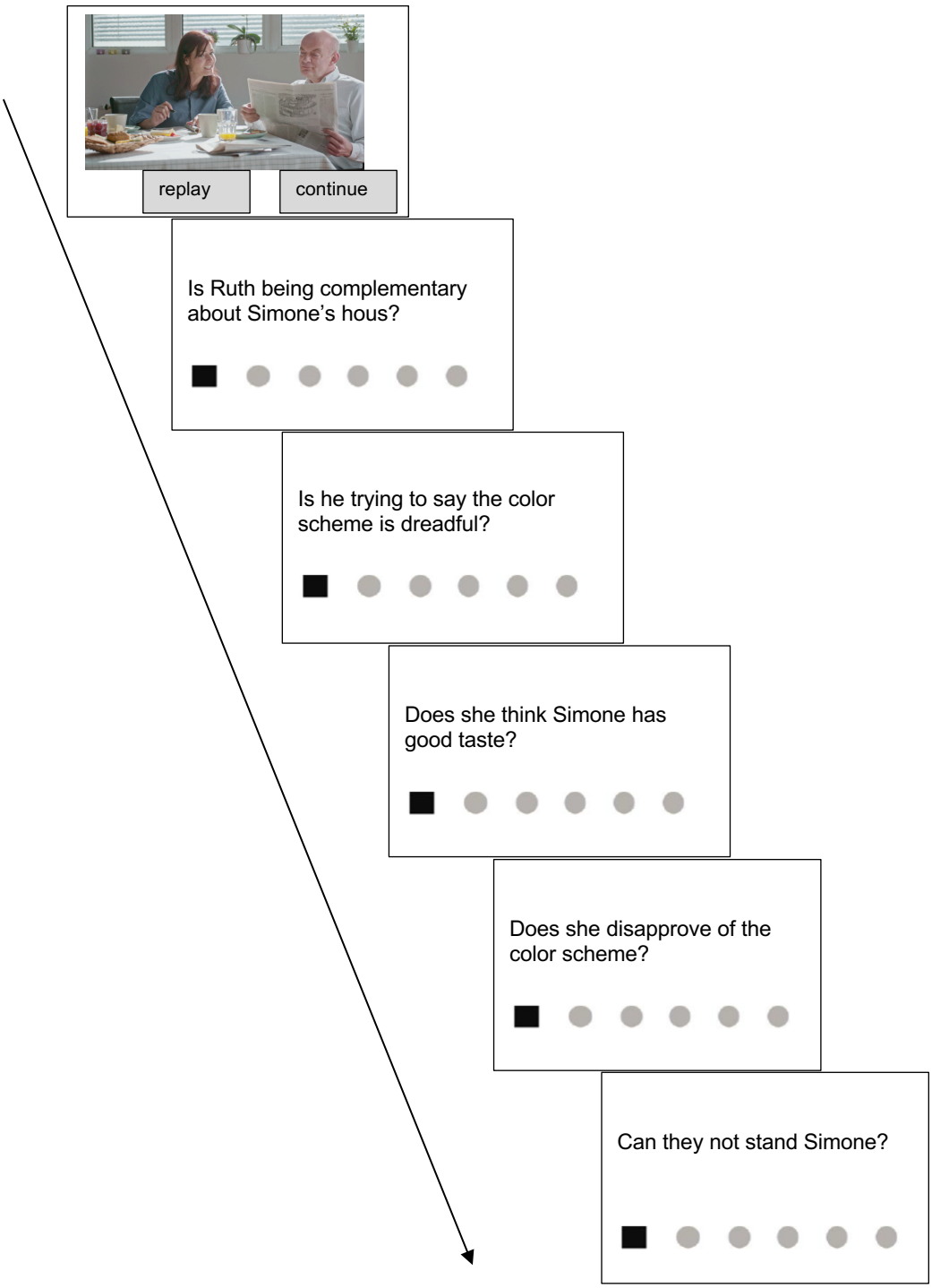
Peirce J. W. (2009). Generating stimuli for neuroscience using PsychoPy. Frontiers in Neuroinformatics, 2 (10), 1-8. doi:10.3389/neuro.11.010.2008

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

Supplementary Material I

Figure I1

Exemplary representation of the BASIT-ToM computer based application process

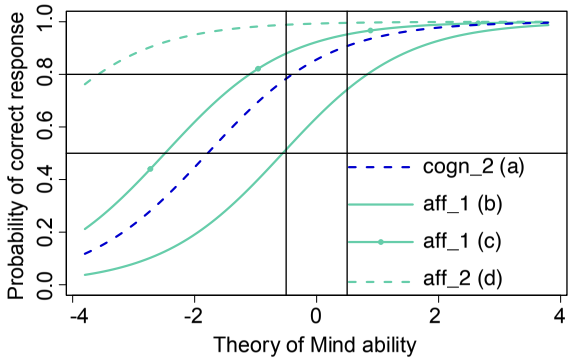


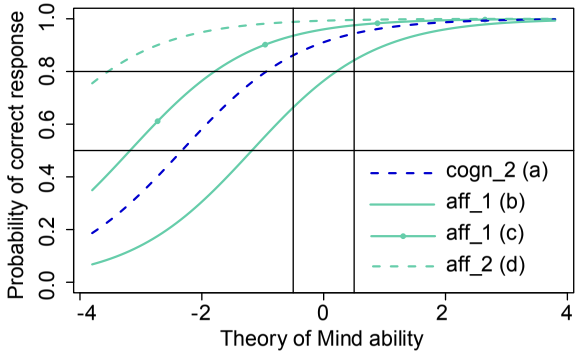
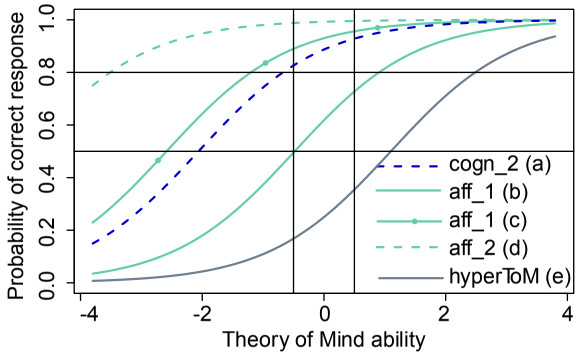
Note. By clicking the square, participants can return to the video to watch it again. Clicking on the respective circles allows participants to return to a previous question to modify their answer, if necessary.

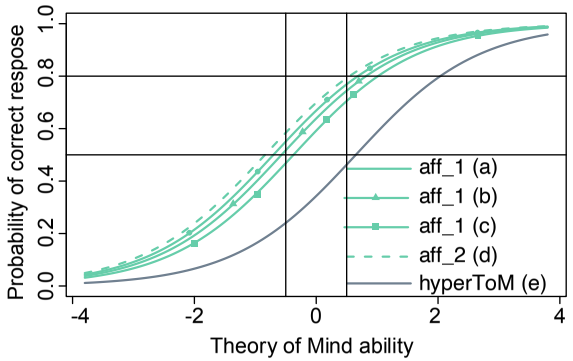
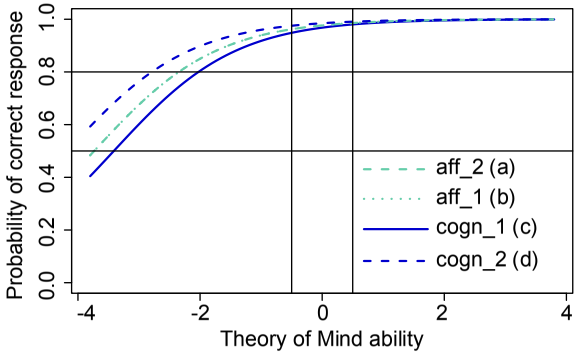
Supplementary Material J

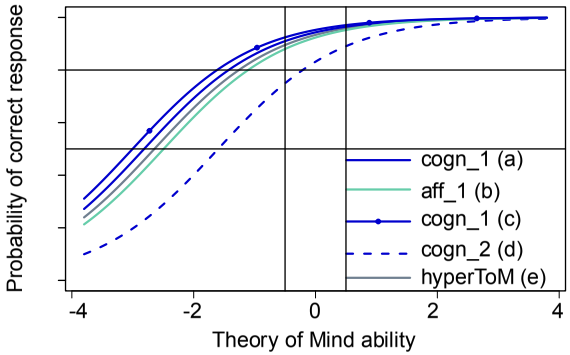
Table J.1: Honesty scenes

Overview of the intensity versions fitting the Rasch model and the number of the respective ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities

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 for the four ToM questions of the respective 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 for the four ToM questions of the respective intensity version fitting the Rasch model
H#1_medium	1	
H#1_high	1	

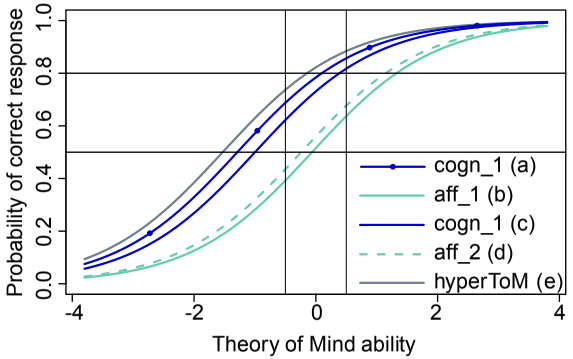
Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves for the four ToM questions of the respective intensity version fitting the Rasch model
H#2_low	**4**	
H#3_low	0	

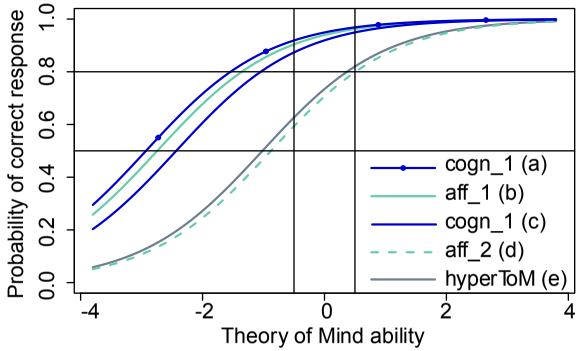
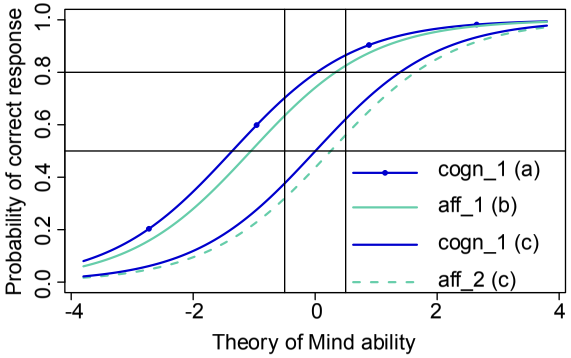
Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves for the four ToM questions of the respective intensity version fitting the Rasch model
H#4_medium	1	

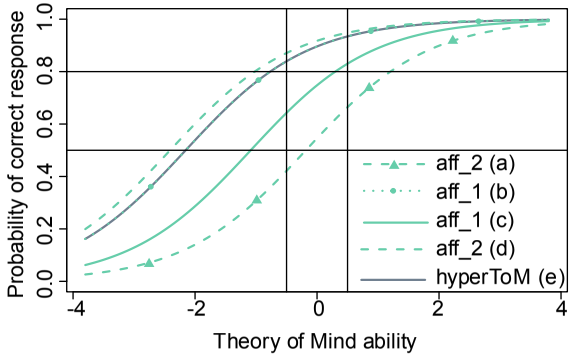
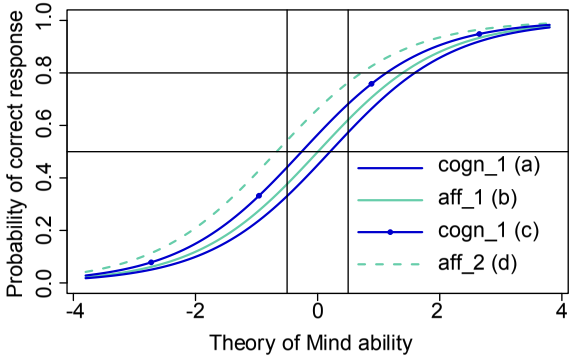
Note. ToM=Theory of Mind; Scene code consists of a scene number, a letter (message type: H=Honesty) and the portrayed intensity (i.e., low, medium, high); Question code consists of an abbreviation for the respective ToM type (i.e., cogn=cognitive ToM, aff=affective ToM, hyperToM), the ToM order (i.e., 1=first order ToM, 2=second order ToM) and a question ID (a-e). Intensities versions selected for the BASIT-ToM are printed in bold.

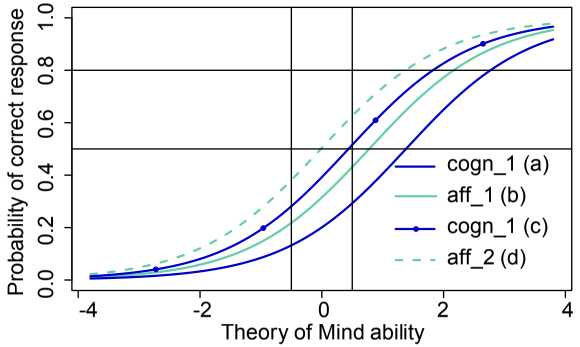
Table J.2: Paradoxical sarcasm scenes

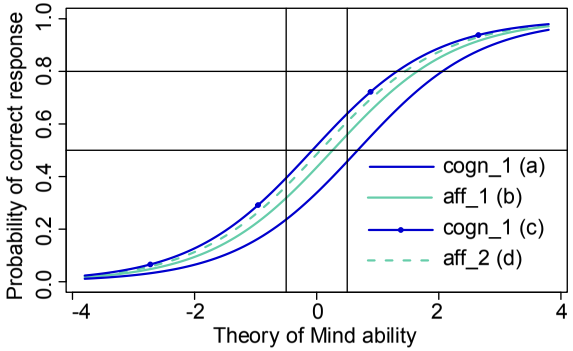
Overview of the intensity versions fitting the Rasch model and the number of the respective ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities

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 for the four ToM questions of the respective 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 for the four ToM questions of the respective intensity version fitting the Rasch model
pS#1_medium	1	
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 for the four ToM questions of the respective intensity version fitting the Rasch model
pS#4_low*	2	
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 for the four ToM questions of the respective 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 for the four ToM questions of the respective intensity version fitting the Rasch model
Practice(pS)_high	3	

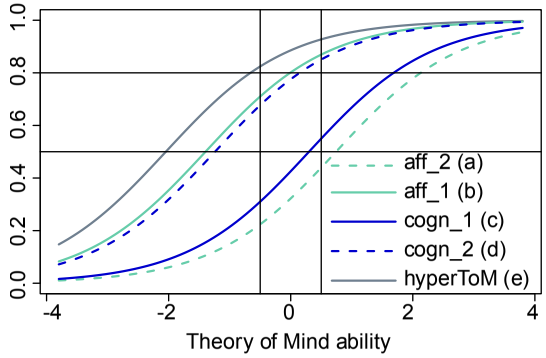
Note. ToM=Theory of Mind; Scene code consists of a scene number, a letter (message type: H=Honesty) and the portrayed intensity (i.e., low, medium, high); Question code consists of an abbreviation for the respective ToM type (i.e., cogn=cognitive ToM, aff=affective ToM, hyperToM), the ToM order (i.e., 1=first order ToM, 2=second order ToM) and a question ID (a-e). Intensities versions selected for the BASIT-ToM are printed in bold.

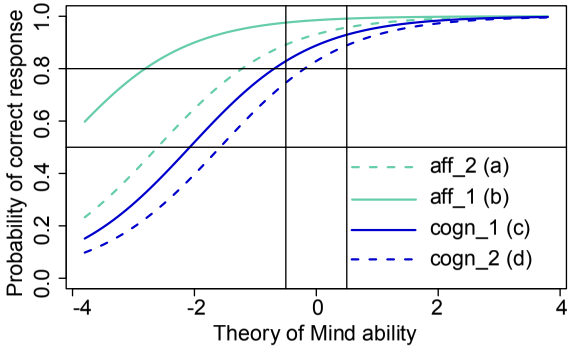
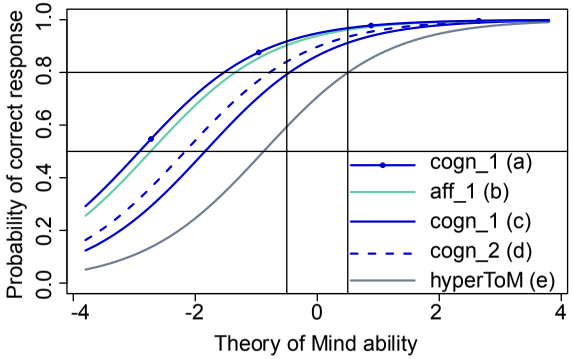
* pS#4_low will be used as practice scene of the Basel Version of The Awareness of Social Inference Test-Theory of Mind (BASIT-ToM)

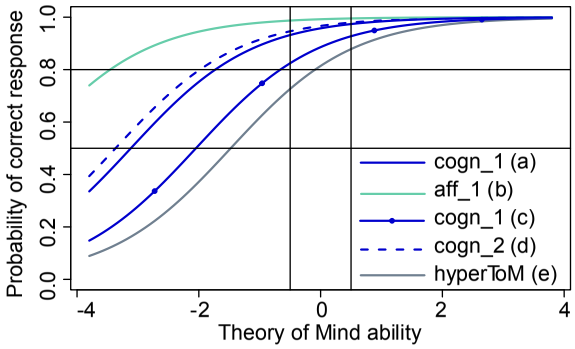
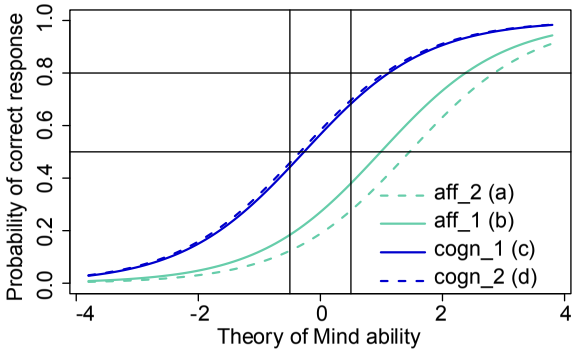
†Practice(pS)_low will be used as a test scene for paradoxical sarcasm of the BASIT-ToM

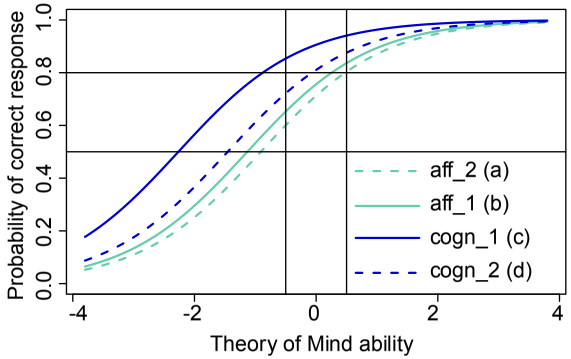
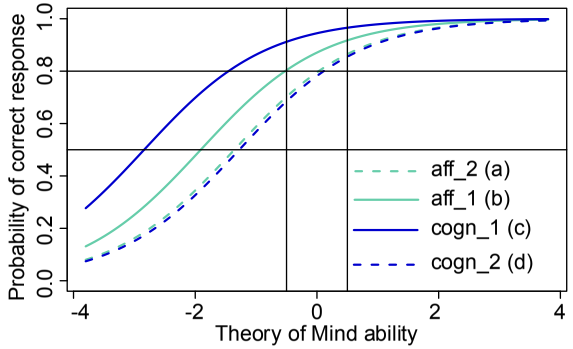
Table J.3: Simple Sarcasm scenes

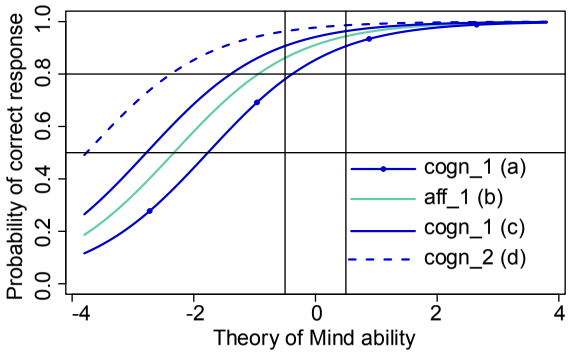
Overview of the intensity versions fitting the Rasch model and the number of the respective ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities

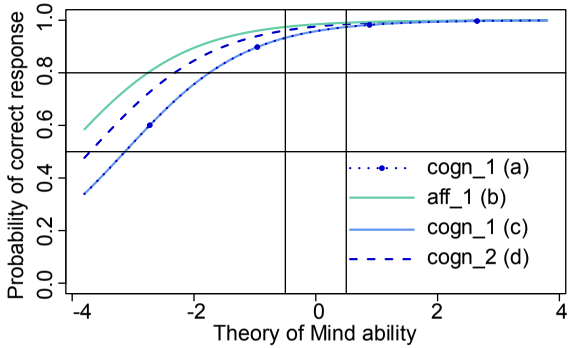
Intensity versions fitting the Rasch model	Number of ToM questions for which the correct response probability is in the required range (0.5-0.8) in subjects with medium ToM abilities	Item characteristic curves for the four ToM questions of the respective intensity version fitting the Rasch model
sS#1_low	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 for the four ToM questions of the respective intensity version fitting the Rasch model
sS#1_high	1	
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 for the four ToM questions of the respective intensity version fitting the Rasch model
sS#2_high	0	
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 for the four ToM questions of the respective intensity version fitting the Rasch model
sS#3_medium	3	
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 for the four ToM questions of the respective 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 for the four ToM questions of the respective intensity version fitting the Rasch model
sS#4_high	0	

Note. ToM=Theory of Mind; Scene code consists of a scene number, a letter (message type: H=Honesty) and the portrayed intensity (i.e., low, medium, high); Question code consists of an abbreviation for the respective ToM type (i.e., cogn=cognitive ToM, aff=affective ToM, hyperToM), the ToM order (i.e., 1=first order ToM, 2=second order ToM) and a question ID (a-e). Intensities versions selected for the BASIT-ToM are printed in bold.