

Zitierhinweis:

Wilhelm, O. (2006). Messinstrumente emotionaler Intelligenz: Praxis und Standards In R. Schulze, P. A. Freund & R. D. Roberts (Hrsg.), *Emotionale Intelligenz. Ein internationales Handbuch* (S. 141-163). Göttingen: Hogrefe.

7

Messinstrumente emotionaler Intelligenz: Praxis und Standards

Oliver Wilhelm
Humboldt-Universität Berlin, Deutschland

Zusammenfassung

In diesem Kapitel wird emotionale Intelligenz (EI) aus einer psychometrischen Perspektive diskutiert, wobei sich der Fokus insbesondere auf EI-Fähigkeitstests richtet. Vorliegende Forschungsergebnisse demonstrieren, dass in der EI-Forschung, ähnlich wie in vielen anderen psychologischen Bereichen auch, Instrumente, die zwar das gleiche Konstrukt adressieren, aber entweder auf in Tests erbrachter Leistung oder auf Selbstberichten basieren, wenig oder keine Konvergenz zeigen. Es wird argumentiert, dass leistungsbasierte Tests angemessenere Indikatoren für EI sind. Der Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) wird als ein solches Messinstrument vorgestellt und die Validität des Verfahrens wird diskutiert. Aus mehreren Perspektiven wird beleuchtet, inwiefern Leistungsaufgaben zur EI als Intelligenzaufgaben kategorisiert werden können. Im Rahmen der Kritik der EI-Instrumente wird versucht, ausstehende Forschungsfragen aufzuwerfen, denen es sich in der Zukunft zuzuwenden gilt. Die Empfehlungen beinhalten a) verstärkt Aufgaben auf emotionspsychologischem Hintergrund zu entwickeln, b) eine größere Aufgabenvariation in multivariaten Studien zu verwenden und c) bei der Validierung von EI geeignete Kriterien zu verwenden.

7.1 Einleitung

Wenn neue Konstrukte in der Psychologie vorgeschlagen werden, sind Wissenschaftler dazu aufgefordert, skeptisch, kritisch und konservativ zu reagieren. Wenn zusätzlich noch neue Messinstrumente für diese neuen Konstrukte entworfen werden, stellt sich die Situation noch schwieriger dar. Dafür mögen zweierlei Gründe verantwortlich sein: Zum einen haben Laien, historisch gesehen, keine lohnenswerten Konstrukte und Messinstrumente zur differenziellen Psychologie beigesteuert, und obwohl es Psychologen waren, die das Konstrukt „Emotionale Intelligenz“ als erste untersucht haben (Mayer, DiPaolo & Salovey, 1990), wurde es doch, auch innerhalb der Psychologie, durch Laien (Goleman, 1995, 1998) popularisiert. Zum anderen sehen sich Psychologen zur Rechtfertigung gezwungen, warum sie so großes Aufhebens um *ihre* Messinstrumente zur Erfassung individueller Unterschiede machen (d. h. warum diese sich von Tests, die etwa in der *Brigitte* erscheinen, unterscheiden). Diese Routinereaktion ist äußerst hilfreich bei der Vermeidung von falsch-positiven Entscheidungen, wenn es um den Einsatz neuer Konstrukte und Instrumente geht. Auf der anderen Seite besteht die Gefahr, es mit der Vorsicht zu übertreiben und neue Ideen und Konstrukte vorschnell abzulehnen, obwohl es vielleicht angezeigt wäre, sie genauer zu untersuchen, weiterzuentwickeln und ggf. in der Praxis einzusetzen. Auf diese Weise kann eine inakzeptabel hohe Quote falsch-negativer Beurteilungen entstehen.

Historisch vereinfachend kann man sagen, dass das Konstrukt EI erstmals in den frühen neunziger Jahren (Mayer et al., 1990; Mayer & Salovey, 1993; Salovey & Mayer, 1990) auf der Bildfläche der differenzialpsychologischen Forschung und Diagnostik auftauchte und seither großes öffentliches Interesse (Goleman, 1995, 1998) erregt hat (für eine ausführliche Beschreibung siehe Matthews, Zeidner & Roberts, 2002). Dieses öffentliche Interesse kann man durch die Entwicklung wissenschaftlicher Messinstrumente als mehr oder weniger befriedigt betrachten. Einige Wissenschaftler untersuchen das Konstrukt nach bestem Wissen und Gewissen, während andere es zu diversen anderen unbrauchbaren Ideen in der Psychologie zählen.

In diesem Kapitel wird es zunächst um den wesentlichsten Unterschied verschiedener Messinstrumente gehen, nämlich um die Erfassung maximalen vs. typischen Verhaltens. Einer kurzen Evaluation derzeit gebräuchlicher EI-Instrumente zur Erfassung typischen Verhaltens folgt eine detaillierte Diskussion der Tests für maximales emotional intelligentes Verhalten. Letztere beginnt mit einer Beschreibung des MSCEIT (MSCEIT V.2; Mayer, Salovey, Caruso und Sitarenios, 2002; Mayer, Salovey, Caruso und Sitarenios, 2003), fährt fort mit der Erläuterung von Anforderungen an ein Messinstrument, die es zu erfüllen gilt, um als Intelligenztest eingestuft zu werden, und schließt mit einer Kritik und Empfehlungen an zukünftige Forschung.

7.2 Typisches und maximales Verhalten

Die Unterscheidung von typischem und maximalem Verhalten geht auf Cronbach (1949) zurück und ist insbesondere mit dem Iteminhalt assoziiert. Typisches Verhalten wird üblicherweise durch Selbstberichte über Präferenzen und Valenzen erfasst. Maximales Verhalten bezieht sich hingegen auf die Messung von Fähigkeiten, Leistungen, Begabungen und deklarativem Wissen. Situationen, in denen maximales Verhalten erfasst wird, lassen sich wie folgt charakterisieren: (a) die betreffende Person ist sich der Tatsache bewusst, dass ihre Leistung beurteilt wird, (b) sie ist willens und in der Lage,

maximales Verhalten zu zeigen und (c) die Bewertungsstandards des Verhaltens sind zur Fällung eines diagnostischen Urteils angemessen (Sackett, Zedeck & Fogli, 1988).

Leistungstests und Instrumente, die auf Selbstberichten beruhen, unterscheiden sich in vielfältiger Weise. Leistungsbasierte Instrumente verlangen maximales Verhalten, sie erfordern eine externe Verhaltensbewertung, sie führen zu minimalem response-bias, ihre Administration ist langwierig und aufwendig und sie sollen eine „Fähigkeit“ erfassen. Bei Instrumenten auf der Grundlage von Selbstberichten verhält es sich genau entgegengesetzt: Sie verlangen Informationen über typisches Verhalten, die Bewertungsinstanz liegt intern und bezieht sich auf Präferenzen, der response-bias kann ganz erheblich ausfallen (insbesondere in Auswahl-situationen), sie sind einfach und schnell durchzuführen und sollen persönlichkeitsähnliche Konstrukte erfassen. Deshalb werden sie vor allem in der Persönlichkeitspsychologie eingesetzt, Instrumente maximalen Verhaltens hingegen zur Ermittlung interindividueller Unterschiede in der Leistungsfähigkeit.

Unterscheidet man zwischen typischem und maximalem Verhalten, so lässt man bisher jedoch die bedeutende Frage außer acht, wie dicht Menschen am jeweils möglichen maximalen Verhalten operieren, wenn sie typisches Verhalten zeigen. Ebenso ist es nicht immer möglich und nützlich, mehr Anstrengung auf maximales Verhalten zu verwenden (Kahneman, 1973). Bemühungen, die Kluft zwischen maximalem und typischem Verhalten zu schließen, können von beiden Seiten unternommen werden: Es ist sowohl denkbar, Persönlichkeitsmerkmale mit Tests maximalen Verhaltens, als auch Fähigkeiten mit Instrumenten typischen Verhaltens zu erfassen (Riemann & Abels, 1994). Zahlreiche Beispiele dokumentieren den zuletzt genannten Ansatz, und genau diesen gilt es zu berücksichtigen, wenn es um die Erfassung Emotionaler Intelligenz geht.

Erstens wird in der Altersforschung regelmäßiger Gebrauch gemacht von Selbstberichten über Gedächtnisprobleme, die mit objektiven Gedächtnistests konkurrieren (Hertzog, Park, Morrell & Martin, 2000). Zweitens wurden in der klinischen Neuropsychologie Fragebögen entwickelt, die einen klaren Mangel an Einsicht in objektiv gemessenen Leistungsabfall in Tests maximalen Verhaltens demonstrieren (McGlynn & Schacter, 1989; Seidenberg, Haltiner, Taylor, Hermann & Wyler, 1994). Drittens wimmelt es in der kognitiven Psychologie geradezu von Selbstberichtsmaßen zur Erfassung von Aufmerksamkeits- und Gedächtnisproblemen (Broadbent, Cooper, Fitzgerald & Parkes, 1982; Herrmann, 1982; Reason, 1993), die wiederum in Beziehung gestellt werden können zu Tests der maximalen Arbeits- und Kurzzeitgedächtnis-Leistung und Aufmerksamkeitsspanne (Oberauer, Süß, Schulze, Wilhelm & Wittmann, 2000). Und letztendlich wurden in der Pädagogischen, der Differenziellen und der Sozialpsychologie zahlreiche Instrumente entwickelt, die der Erfassung typischen intellektuellen Engagements dienen (Cacioppo & Petty, 1982; Epstein, Pacini, Denes-Raj & Heier, 1996; Goff & Ackerman, 1992; McCrae, 1990, 1996; Wilhelm, Schulze, Schmiedek & Süß, 2003) und die in Bezug zu etablierten Intelligenztests stehen können.

In all diesen Gebieten war man bisher jedoch nicht in der Lage, durchgängig substantielle oder hohe Korrelationen zu finden. Tatsächlich bestehen lediglich im letztgenannten Bereich kleine bis moderate Korrelationen zwischen den beiden Ansätzen (d. h. zwischen Maßen des typischen intellektuellen Engagements und Intelligenztests). Immer wieder wurde argumentiert, dass die klassischen Tests zur Erfassung maximalen Verhaltens unter kontrollierten Laborbedingungen durchgeführt werden und dadurch ein Mangel an Vorhersagekraft für relevante Kriterien entsteht (Dennis, Sternberg & Beatty, 2000). Wenn man Alltagsaktivitäten betrachtet, dann ist die Anzahl und Dauer von Situationen, in denen man sein bestmögliches Verhalten zeigt, eher limitiert. Dennoch ist die Vorhersagekraft von Messinstrumenten zur Erfassung maximalen Ver-

haltens einwandfrei belegt und von substantieller Höhe (Ones, Viswesvaran & Dilchert, 2004).

Die derzeitige Forschungslage lässt vermuten, dass das soeben dargestellte Ergebnismuster vergleichbarer Konstrukte auch für EI zutrifft – zumindest unter Vernachlässigung konzeptioneller und empirischer Probleme bei der Messung typischen und maximalen Verhaltens. Trotz einer beachtenswerten konzeptuellen Überschneidung in der Definition von EI gibt es bisher im Kontext maximalen und typischen Verhaltens keine wesentlichen Relationen zwischen Instrumenten der beiden Ansätze (siehe z. B. O'Connor & Little, 2003). Diese Nullkorrelationen lassen keine anderen Interpretationen zu, als dass die beiden Ansätze offensichtlich zwei verschiedene Merkmale erfassen. Methodenartefakte für die mangelnden Korrelationen verantwortlich zu machen kann nicht befriedigend sein, wenn es darum gehen soll, ein neues Konstrukt inklusive neuer Messinstrumente zu etablieren (siehe Kapitel 9 von Pérez, Petrides & Furnham). In Begriffen der Multitrait-Multimethod (MTMM)-Validierung ausgedrückt sind substantielle Korrelationen über verschiedene Methoden hinweg und innerhalb eines Traits erforderlich (monotrait-heteromethod). Betrachtet man Instrumente maximalen und typischen Verhaltens jedoch als unterschiedliche Methoden, so sind deren Korrelationen von nicht ausreichender Größe. Auf der anderen Seite sollten heterotrait-monomethod Korrelationen niedrig oder gleich Null sein, was im Falle von EI-Selbstberichtsmaßen jedoch typischerweise nicht zutrifft. EI-Fähigkeitstests korrelieren moderat und signifikant mit anderen Fähigkeitstests. Ähnliche Ergebnisse wurden im Bereich sozialer Intelligenz gefunden (siehe Kapitel 10 von Weis & Süß).

Wenn Instrumente typischer und maximaler EI nicht in Beziehung zueinander stehen, dann sollten sie auch nicht die gleiche Bezeichnung tragen. EI soll offensichtlich eine Fähigkeit darstellen. Folglich sollten Selbstberichtsmaße nicht die Bezeichnung *Intelligenz* erhalten. Ausweichbegriffe wie „Trait-EI“ (siehe Kapitel 9 von Pérez, Petrides & Furnham) stellen dabei keine Lösung dar, da auch „Fähigkeits-EI“ eine Disposition sein soll.

7.3 Selbstberichtete und selbsteingeschätzte emotionale Intelligenz

Die Versuche, EI zu operationalisieren, gehen in zwei Richtungen. Auf der einen Seite stehen traditionelle Selbstberichtsinstrumente, die typisches Verhalten erfragen sollen. Auf der anderen Seite stehen Maße, die in konzeptueller Nähe zu herkömmlichen Fähigkeitstests stehen. Letztere werden im Folgenden als „Fähigkeitsmodelle“ (abgeleitet von Mayer, Caruso & Salovey, 2000) bezeichnet, obwohl bisher nicht zweifelsfrei nachgewiesen werden konnte, dass sich diese Instrumente tatsächlich als Fähigkeitstests qualifizieren. Dieser Aspekt wird in 7.5 thematisiert. Unter Berücksichtigung der Unabhängigkeit der zwei Messansätze und der Konvention, dass der Ausdruck *Intelligenz* in Verbindung steht mit Instrumenten, die maximales Verhalten erfassen, sollten EI-Instrumente, die auf Selbstberichten basieren, nicht mit dem Ausdruck *Intelligenz* etikettiert werden.

Wesentlicher als diese begrifflichen Unklarheiten ist der Status korrespondierender Selbstberichtsfragebögen. Diese Instrumente wurden auf der Basis diverser unterschiedlicher Definitionen dessen entwickelt, was angeblich emotional intelligentes Verhalten konstituiert. Bar-On (1997, 2000) unterscheidet ganze 15 Komponenten des erfolgreichen emotionalen Funktionierens. Diese 15 Komponenten sind in fünf breiten, interkorrelierenden Dimensionen organisiert. Die fünf Dimensionen lauten intrapersonale EI,

interpersonale EI, adaptive EI, Stress-Management EI und allgemeine Stimmungslage EI. Der entsprechende Test zu diesem Model ist das BarOn Emotional Quotient Inventory, EQ-i (Bar-On, 1997). Jedoch konnten die vorgeschlagenen sowie alternative Strukturen empirisch nicht bestätigt werden (Palmer, Manocha, Gignac & Stough, 2003; Petrides und Furnham, 2000, 2001). Ebenso fanden das Schutte et al. Emotional Intelligence Inventory (Schutte et al., 1998) und seine Erweiterungen (Saklofske, Austin & Minski, 2003) weitläufige Verbreitung, ohne dass bisher eine eindeutige Aussage über seine interne Struktur gemacht werden könnte (siehe Kapitel 9 von Pérez, Petrides & Furnham). Unter Berücksichtigung weiterer verfügbarer Instrumente ist der Trait Emotional Intelligence Questionnaire (TEIQue) der vielversprechendste Kandidat bezüglich vorliegender empirischer Befunde und Volumen der Validierungsbemühungen (siehe Kapitel 9 von Pérez et al.). Der TEIQue besteht aus 144 Items, die zu zehn Skalen zusammengefasst werden: Adaptability, assertiveness, emotion perception, emotion expression, emotion regulation, empathy, low impulsivity, relationship skills, social competence und stress-management.

Ein problematischer Aspekt des TEIQue und diverser anderer Selbstberichtsinstrumente für EI ist, dass ein Großteil seiner Items etablierten Messinstrumenten für beispielsweise Emotionale Empathie (Mehrabian & Epstein, 1970), der Toronto Alexithymie Skala (Bagby, Parker & Taylor, 1994a, 1994) und anderen EI-Selbstberichtsinstrumenten entnommen wurde. Somit ist der TEIQue technisch gesehen größtenteils ein Konglomerat bereits existierender Items und das damit erfasste Konstrukt kann dann wohl kaum als neu bezeichnet werden. Ein weiteres Problem des TEIQue sowie anderer Selbstreport EI-Instrumente ist, dass bisher kein zufrieden stellendes Messmodell, weder auf der Item-, noch auf der Test-Ebene oder für einzelne Skalen, besteht. Ein drittes Problem aller EI-Selbstberichtsinstrumente ist, dass mögliche Redundanzen mit konkurrierenden und etablierten Konstrukten, die bei Selbstberichten entstehen, bisher nicht ausreichend ermittelt wurden. Dieser letzte Punkt ist besonders wichtig, da in der differenzialpsychologischen Forschung geballte Anstrengungen darauf verwendet wurden, die Dimensionalität traditioneller Selbstberichte zu ergründen. Das Fünf-Faktoren-Modell (Costa und McCrae, 1992) ist das prominenteste Beispiel dieser Bemühungen. Innerhalb dieses Modells werden diverse Facetten niedrigerer Ordnung für jeden einzelnen Faktor postuliert und untersucht. Darüber hinaus wurden eine ganze Reihe weiterer Selbstberichtsdimensionen in der Vergangenheit untersucht. Wenn nun ein neues Konstrukt durch Selbstberichtsinstrumente erfasst werden soll, dann muss zunächst zweifelsfrei geklärt werden, ob die interindividuellen (Leistungs-)Unterschiede in diesem neuen Instrument nicht auf interindividuelle Unterschiede in bereits etablierten Instrumenten zurückgeführt werden können. Nachdem eine große Bandbreite konkurrierender Selbstberichtsdimensionen kontrolliert wurde, müssen die Tests des neuen Konstrukts noch immer signifikant und in angemessener Höhe miteinander korrelieren. Zusätzlich soll das neue Instrument inkrementell dazu beitragen, interessante Kriterien vorherzusagen, das heißt, die Vorhersagegüte soll noch über die Verwendung konkurrierender Selbstberichtsdimensionen und etablierter Prädiktoren hinausgehen. Zum jetzigen Zeitpunkt scheint es keine wissenschaftliche Evidenz für die Eindimensionalität, die inkrementelle Validität und die Nützlichkeit von selbstberichteter EI zu geben.

Deshalb sollten die bisher entwickelten Selbstberichtsinstrumente für emotionale Intelligenz nicht das Label EI erhalten. Verfügbare Evidenz weist nicht darauf hin, dass diese Tests etwas Neues erfassen. Diese Instrumente als Indikatoren eines neuen Kon-

strukts zu verstehen – zum Beispiel emotionale Selbstwirksamkeit – erfordert eine elaboriertere Einbettung in verwandte und bereits etablierte nomologische Netzwerke.

Die Selbstbeurteilung von Fähigkeiten fällt in das Niemandland zwischen Tests zu typischem und zu maximalem Verhalten (Stankov, 1999). Es ist nicht ungewöhnlich, Items zur Erfassung einer Selbstberichtsdimension vorzufinden, die die Selbsteinschätzungen von Fähigkeiten repräsentieren. Items wie „Ich kann meine Stimmungen und Gefühle gut ausdrücken“ sind strukturierten Ansätzen zur Messung von Selbstberichts-EI sehr ähnlich. Derartige Items reflektieren weniger eine Präferenz typischerweise emotional intelligent zu handeln, sondern beziehen sich eher auf eine Einschätzung der relativen eigenen Position in der Fähigkeit, Emotionen und Gefühle adäquat auszudrücken. Für einige Fähigkeiten ist es leichter als für andere angemessene Selbsteinschätzungen abzuliefern. Je genauer die Introspektion und das Wissen über eigene Fähigkeiten sind, desto höher ist der Zusammenhang zwischen den Fähigkeiten und ihren Selbstratings. So ist es wichtig zu erwähnen, dass in traditionellen Intelligenzdomänen die Korrelationen zwischen Selbstratings und tatsächlicher Fähigkeit zwischen .20 und .50 liegen (Ackerman, Beier & Bowen, 2002). Obwohl gewisse konvergente und diskriminante Validität zwischen den Selbstberichten von Fähigkeiten und Wissen und ihrer tatsächlichen Messung bestehen, so sind diese Zahlen doch überraschend niedrig in Anbetracht der Tatsache, dass Menschen ständig mit einem Feedback über ihre Leistung in diversen Bereichen versorgt werden. Die Korrelation zwischen Selbstberichts-EI und Fähigkeits-EI ist höchstwahrscheinlich nicht höher. Auf der Basis von Korrelationen, die weit unter .50 liegen, ist es gewiss nicht angemessen, Selbstberichte als Annäherungen für Fähigkeits-EI zu verwenden. Darüber hinaus ist unklar, was EI-Selbstberichtsinstrumente überhaupt reflektieren. Präferenzen, Valenzen, Fähigkeiten, die Neigung, Fähigkeiten zu über- oder unterschätzen und andere Persönlichkeitskonstrukte sind die offenkundigsten Kandidaten, die für Selbstberichts-EI verantwortlich gemacht werden können. Zur Etablierung eines neuen Konstrukts und einer neuen Messprozedur sind Selbstberichte nur von begrenzter Nützlichkeit, sowohl als Kriterien als auch als Prädiktoren. Folglich werden wir und im Rest dieses Kapitels nur noch mit Fähigkeits-EI auseinandersetzen.

7.4 Der MSCEIT: Beschreibung, Struktur und Validität

Der Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) ist eine gekürzte und verbesserte Version der Multi-Factor Emotional Intelligence Scales (MEIS; Mayer et al., 2002; Mayer, Caruso & Salovey, 1999). Die wesentlichen Ziele der Entwicklung des MSCEIT waren es, die MEIS zu verkürzen und die psychometrischen Eigenschaften der Items und Skalen zu verbessern. Dem MSCEIT kommt deshalb hier besondere Beachtung zu, weil er die jüngste und aktuellste Entwicklung der Forschungsgruppe um Mayer, Salovey und Caruso repräsentiert und er der am weitesten verbreitete Test zur Messung von emotionaler Intelligenz ist.

Der MSCEIT besteht aus acht Subskalen. Jeweils zwei dieser Subskalen werden kombiniert und so zu einem der vier „branches“, also Zweige, emotionaler Intelligenz zusammengefasst. Die vier Zweige schließen sich wiederum zu zwei Bereichen zusammen, die zusammen den Testgesamtwert ergeben. Der MSCEIT repräsentiert somit ein hierarchisches EI-Modell. Es besteht aus drei Ebenen, denen Fähigkeitsstatus zugesprochen wird. An der Spitze dieser Hierarchie steht „Emotionale Intelligenz“, es folgen darunter die zwei Bereiche „Emotionales Erleben“ und „Emotionales Verarbeiten“, und

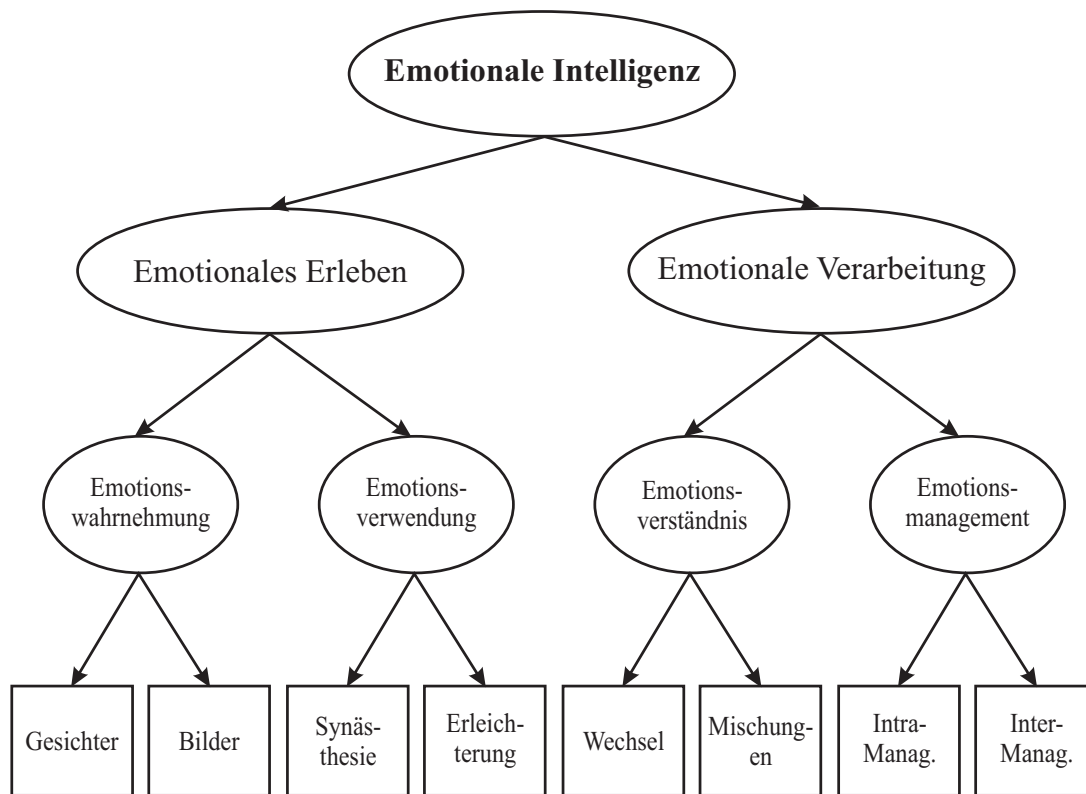


Abbildung 7.1 Struktur und Subskalen des MSCEIT.

wiederum unterhalb dieser Ebene die vier Zweige „Emotionswahrnehmung“, „Emotionsverwendung“, „Emotionsverständnis“ und „Emotionsmanagement“. In Abbildung 7.1 wird noch eine weitere vierte Ebene eingeführt, nämlich die der spezifischen Tests, aus denen der MSCEIT besteht.

Es wird vorgeschlagen, die erzielten Werte bis zu den vier Zweigen zu interpretieren und im Einzelfall auch die Leistung auf der Ebene der Einzelaufgaben einzuschließen. Aus dem hierarchischen Modell ergibt sich, dass im MSCEIT insgesamt sieben Faktoren postuliert werden, denen auch allesamt Fähigkeitsstatus zugesprochen wird. Im Folgenden wird kurz erläutert, wie die Werte für die vier Zweige interpretiert werden sollen.

Emotionswahrnehmung: Probanden mit hohen Werten sind in der Lage, eigene Emotionen und Emotionen anderer richtig zu identifizieren und zu benennen. Zusätzlich sind sie fähig, ihre Gefühle adäquat auszudrücken und falsche bzw. gestellte Gesichtsausdrücke zu erkennen.

Emotionsverwendung: Personen mit hohen Werten auf diesem Faktor sind in der Lage Emotionen so zu generieren, dass sie ihnen in Problemlösesituationen helfen. Sie können außerdem ihre Aufmerksamkeit adäquat auf wichtige Veränderungen ausrichten, mehrere Perspektiven bei der Bewertung von Emotionen einnehmen und Denkprozesse durch verschiedene Stimmungen unterstützen.

Emotionsverständnis: Probanden mit hohen Werten verstehen Ursachen und Wechsel von Emotionen, sowohl abstrakt als auch in Bezug auf Beziehungen. Sie sind au-

ßerdem in der Lage, Ähnlichkeiten zwischen Emotionen unterschiedlicher Intensität zu erkennen und Gefühlsdynamik im interpersonellen Bereich zu deuten.

Emotionsmanagement: Personen mit hohen Werten setzen erfolgreich ihr emotionales Bewusstsein ein, um optimale Entscheidungen zu treffen, wobei sie ihren Gefühlen dabei einen adäquaten Stellenwert einräumen. Diesen Probanden gelingt es, offen für Gefühle zu bleiben und sich situationsgemäß zu engagieren oder zurückzuziehen. Darüber hinaus können sie ihre Stimmungen hinsichtlich ihres typischen Auftretens, ihrer Akzeptanz und Relevanz aus einer übergeordneten Perspektive evaluieren.

Die ersten beiden Fähigkeiten werden zum Faktor „Emotionales Erleben“ zusammengefasst. Diese Fähigkeit soll eine akkurate Wahrnehmung, Reaktion und Manipulation emotionalen Materials reflektieren. Das zweite Fähigkeitspaar wird zum Faktor „Emotionales Verarbeiten“ zusammengefasst. Diese Fähigkeit soll das Verständnis von und den Umgang mit Gefühlen reflektieren. Sie soll auch abbilden, wie akkurat eine Person die Bedeutung der Emotionen erfasst und wie gut sie auf eigene Emotionen und auf die wichtiger anderer Personen Einfluss nimmt.

An der Spitze der vorgeschlagenen Hierarchie befindet sich die allgemeine emotionale Intelligenz (Mayer & Salovey, 1997; Mayer, Salovey & Caruso, 2000). Sie wird bei der Auswertung des Tests als Durchschnitt aus den Unterfaktoren emotionales Erleben und emotionales Verarbeiten berechnet. Interpretiert wird sie als die Fähigkeit, Gefühle wahrzunehmen, sie unterstützend zum Denken einzusetzen, Emotionen zu verstehen und erfolgreich zu regulieren.

Es gibt viele gute Beschreibungen der Aufgaben des MSCEIT (Mayer et al., 2002) und seiner Vorgängerin, der MEIS (siehe Kapitel 2 von Neubauer & Freudenthaler). Das Scoring im MSCEIT erfolgt für alle Untertests nach dem gleichen Schema. Die Häufigkeitsverteilungen der Antwortmöglichkeiten einer Subskala werden herangezogen, um die individuellen Antworten eines Probanden im jeweiligen Test zu gewichten. Im Gesichtersubtest sind die Probanden beispielsweise aufgefordert, auf einer fünfstufigen Skala von „nicht glücklich“ bis „extrem glücklich“ zu beurteilen, wie viel Fröhlichkeit ein Fotoportrait ausdrückt. Angenommen, ein bestimmtes Foto wurde in der Referenzstichprobe folgendermaßen eingeschätzt: 10% der Befragten sagten „extrem unglücklich“, 20% „eher unglücklich“, 40% „weder noch“, 20% „eher glücklich“ und 10% gaben an, die Person auf dem Foto sei „extrem glücklich“. Eine Person, die die Antwort 3 („weder noch“) gegeben hat, würde dann den Wert 0.40 erhalten, für die individuelle Antwort 5 („extrem glücklich“) würde ein Wert von 0.10 vergeben. Diese Prozedur wird für alle Einzelantworten wiederholt. Die so erhaltenen Einzelwerte werden für jeden Untertest gemittelt, um die Gesamtleistung im jeweiligen Untertest zu quantifizieren. Das dieser Scoringprozedur zu Grunde liegende Rationale ist, dass in vielen Fähigkeits- und Wissensbereichen keine universellen, unanfechtbaren Korrektheitsstandards existieren (siehe Kapitel 8 von Legree, Psotka, Tremble & Bourne). Diese Scoringprozedur wird als konsensbasiertes Scoring bezeichnet. Sie wird häufig und erfolgreich bei Situational Judgment Tests (SJTs; McDaniel, Morgeson, Finnegan, Campion & Braverman, 2001) und Tests für implizites oder prozedurales Wissen eingesetzt. Für den MSCEIT wurde ein empirischer Vergleich des konsensbasierten Scorings auf der Grundlage der Standardisierungsstichprobe und der auf der Grundlage einer Expertengruppe, die aus 21 Mitgliedern der Internationalen Gesellschaft für Emotionsforschung bestand, durchgeführt. Die Korrelationen fielen sehr hoch aus – doch steht die Replikation solcher Befunde noch aus. Auf die Validität der MEIS beziehungsweise des MSCEIT richten sich intensive Forschungsbemühungen, es können jedoch derzeit noch keine klaren Schlussfolgerungen gezogen werden. Anstatt die vorliegenden Evidenzen erschöpfend

darzulegen, soll hier der Schwerpunkt auf eine Zusammenfassung ausgewählter prototypischer Untersuchungen gelegt werden. Besondere Beachtung wird dabei zum einen der MSCEIT finden, da er als das momentane „state-of-the-art“ Verfahren zur Messung von EI gilt. Zum anderen werden einige Mängel und Standards für die zukünftige Forschung dargestellt.

Die Struktur des MSCEIT scheint auf der Grundlage der bisher verfügbaren Daten im Großen und Ganzen robust zu sein. Dennoch kann das Strukturmodell mit vier korrelierten Faktoren nur dann als gültig angesehen werden, wenn die Kovarianzen zwischen den Faktoren Emotionswahrnehmung und Emotionsverwendung und den Faktoren Emotionsverständnis und Emotionsmanagement gleichgesetzt werden. Explorative Faktorenanalysen bestätigen die vorgeschlagene Faktorenunterscheidung zumeist. Jedoch variieren die Ladungen der Aufgaben sowohl in konfirmatorischen als auch in explorativen Faktorenanalysen sehr stark, was impliziert, dass die Faktoren durch Einzelaufgaben dominiert werden. Beispielsweise lädt die Aufgabe „Synästhesie“ auf dem Faktor Emotionsverwendung viel stärker als die zweite Aufgabe „Erleichterung“, die denselben Faktor erfassen soll. Folglich fehlt es den Faktoren an Inhaltsbreite. Inhaltsvalidität wurde für den MSCEIT noch nicht umfassend demonstriert. Weiterhin besteht im MSCEIT ein Zuordnungsproblem zwischen den Faktorenbezeichnungen und ihren Aufgaben. Zum Beispiel wird der erste Zweig auch als „Wahrnehmung und Ausdruck von Emotionen“ bezeichnet, obwohl offensichtlich nur einzelne Aspekte der Emotionswahrnehmung getestet werden.

Die prädiktive Validität des Tests wurde durch Korrelation mit zahlreichen Kriterien untersucht. Die Korrelationen mit fluider Intelligenz sind im allgemeinen klein, dagegen sind für einige Einzeltests die Korrelationen mit kristalliner Intelligenz deutlich höher. MacCann, Roberts, Matthews und Zeidner (2004) berichten Korrelationen von EI- Einzelaufgaben mit dem Faktor broad visualization (G_v) im Bereich von .20. Eine aktuelle Metaanalyse konnte bestätigen, dass der MEIS – im Gegensatz zu den Selbstberichtsinstrumenten – mit der allgemeinen mentalen Leistungsfähigkeit assoziiert ist ($\rho = .33$; $SD\rho = .093$; Van Rooy & Viswesvaran, 2004). In zahlreichen Stichproben wurden zwischen dem MSCEIT und Fragebögen zur Empathie Korrelationen von .17 bis .52 gefunden. Korrelationen mit anderen Selbstberichtsskalen waren zumeist gering, in manchen Fällen jedoch signifikant. Ciarrochi, Chan und Caputi (2000) fanden eine Korrelation von .31 mit dem Selbstwertgefühl. Korrelationen mit der Lebenszufriedenheit streuen sehr stark, doch liegen die Koeffizienten für größere Studien bei .20 (Mayer et al., 2002). Emotionale Intelligenz ist stark negativ assoziiert mit durch die Peergruppe eingeschätzter Aggressivität und positiv assoziiert mit prosozialem Verhalten (Mayer et al., 2002). EI im Allgemeinen und emotionales Erleben im Besonderen sind negativ assoziiert mit illegalem Drogen- und Alkoholmissbrauch und kriminellem Verhalten. Diese Korrelationen beziehen sich allerdings nur auf die männliche Untergruppe (Brackett, Mayer & Warner, 2004). Korrelationen zwischen den vier Zweigen und dem Gesamtwert für den MSCEIT und Selbstberichtsinstrumenten für EI liegen maximal bei .28 und sind meistens deutlich geringer (Brackett & Mayer, 2003). In dieser Studie konnte der MSCEIT soziale Auffälligkeit vorhersagen – sogar dann, wenn die Big Five und der verbale SAT-Wert kontrolliert wurden. Trotz dieser Evidenzen bestehen nach wie vor vielfältige Probleme und Kontroversen, die in den folgenden Abschnitten diskutiert werden.

7.5 Ist Fähigkeits-EI eine Form von Intelligenz?

Die Verwendung des Wortes *Intelligenz* in der Konstruktbezeichnung „Emotionale Intelligenz“ hat zu beträchtlichem Diskussionsbedarf geführt. Es stellt sich die Frage, welche Gründe dafür sprechen, ein neues Konstrukt in die Rubrik „Intelligenz“ einzuordnen. Erstens führt die Messung von Intelligenz zu guten Erfolgen in vielen praktischen Angelegenheiten (Ones et al., 2004). Intelligenz ist der beste Einzelprädiktor, den die Psychologie jemals hervorgebracht hat. Da Intelligenztests weit verbreitet sind, ermöglicht die Integration eines neuen Konstrukts in diesen Kontext, dass der Ruhm und die Glaubwürdigkeit etablierter Verfahren auf die neuen Tests überspringen.

Neben diesen Marketingüberlegungen kann argumentiert werden, dass Tests zur emotionalen Intelligenz eine effektive Informationsverarbeitung erfordern und dass Menschen dazu mehr oder weniger gut in der Lage sind. Wenn weniger Anstrengung auf die Informationsverarbeitung verwendet wird, so wird, *ceteris paribus*, auch eine schlechtere Leistung erzielt. Die Levels of Emotional Awareness Scale (LEAS; Lane, Quinlan, Schwartz, Walker & Zeitlin, 1990) ist ein kognitiv-entwicklungspsychologisches Instrument zur Messung von Emotionen. Es unterscheidet fünf Komplexitätsebenen, von einfacherer bis anspruchsvollerer Informationsverarbeitung. Diese fünf Ebenen lauten: Körperliche Empfindungen, Handlungsneigungen, einzelne Emotionen, gemischte Emotionen und die Kombinationen aus gemischten Emotionen (Lane & Schwartz, 1987). Die Probanden werden aufgefordert, zu 20 Schablonen zwei Fragen schriftlich zu beantworten: „Wie würden Sie sich fühlen?“ und „Wie würde sich die andere Person fühlen?“. Die LEAS wurde aufwändig entwickelt und validiert (Ciarrochi, Scott, Deane & Heaven, 2003; Lane et al., 1998). Da Informationsverarbeitungsprozesse bei der Testentwicklung explizit berücksichtigt wurden, repräsentiert die LEAS einen interessanten und möglicherweise besseren Ansatz, emotionale Intelligenz zu erfassen als Verfahren, die auf Konsens- oder Expertenscoring basieren. Momentan wird Informationsverarbeitung im MSCEIT nicht explizit berücksichtigt. Es wäre sehr interessant, alternative und zusätzliche Messmethoden zur Erfassung des allgemeinen emotionalen Bewusstseins auf der Grundlage dieses 5-Ebenen-Modells der emotionalen Komplexität zu entwickeln. Der Einfluss der Informationsverarbeitung wird noch stärker bei experimentellen Prozeduren zur Messung von Emotionen, wie zum Beispiel dem emotionalen Stroop Test, betont (Coffey, Berenbaum & Kerns, 2003; Matthews et al., 2002). Zurzeit ist jedoch unklar, wie diese Verfahren der emotionalen Informationsverarbeitung mit anderen Fähigkeitstests der emotionalen Intelligenz zusammenhängen und ob diese Messmethoden kohärente Persönlichkeitsmerkmale abbilden.

Ein dritter Grund, EI als „Intelligenz“ zu bezeichnen, könnte darin bestehen, dass ihre Tests, abgesehen von EI-Selbstberichtsmaßen, wie alle anderen Fähigkeitstests auch, anhand der Übereinstimmung mit externen Korrektheitsstandards ausgewertet werden können (Guttman, 1965). Solche Standards ermöglichen in der Regel die Beurteilung individueller Antworten als richtig oder falsch. In manchen Fällen werden auch Punkte für den Grad der Richtigkeit vergeben. Diese Korrektheitsstandards können nur bei Leistungstests eingesetzt werden. Bei den meisten Verfahren zur Messung von Einstellungen, beispielsweise, werden die Probanden gebeten, Auskunft darüber zu geben, wie wohl gesonnen sie gegenüber einem bestimmten Objekt sind. Bei der Erfassung von Persönlichkeitseigenschaften sollen die Personen angeben, wie typisch bestimmte Gedanken, Verhaltensmuster und ähnliches für sie sind. Bei Leistungstests können die Antworten anhand einer expliziten Regel als mehr oder weniger richtig eingestuft werden. Bei der Klassifizierung von Antwortverhalten können zahlreiche Bewertungs-

standards verwendet werden (Nevo, 1993). Die Leistung kann durch die Anzahl der Richtigantworten, die Latenzzeiten der Richtigantworten, die Vielfalt der generierten Antworten oder als Übereinstimmung mit einer Autorität ermittelt werden. Bei der Erfassung emotionaler Intelligenz werden zur Klassifikation der individuellen Antworten hauptsächlich die Methoden Zielpersonen-Scoring, allgemeines konsensbasiertes Scoring und expertenkonsensbasiertes Scoring verwendet – allesamt Variationen des Autoritätsansatzes. Die Antwortnormen der EI-Tests können als Korrektheitsstandards konzeptualisiert werden. Obwohl nicht wünschenswert, ist es dennoch üblich, dass sich diverse Instrumente eines spezifischen Konstrukts auf einen einzigen Korrektheitsstandard beschränken. Beispielsweise werden bei Aufgaben zum Schlussfolgernden Denken logische Standards angewendet, bei Aufgaben zur mentalen Geschwindigkeit die Arbeitsraten, und bei Messungen der emotionalen und praktischen Intelligenz kommt zumeist ein Konsensstandard mit einer Vergleichsgruppe zum Einsatz.

Ein vierter Grund für die Klassifikation der EI-Messinstrumente als Intelligenzaufgaben ist deren Zusammenhang zu anderen Fähigkeitstests. Ausgehend von den positiven Zusammenhängen, die zwischen verschiedenen Intelligenztests gefunden wurden, müssen EI-Verfahren mit anderen Intelligenzmessungen korrelieren, wenn sie ebenfalls eine Intelligenz repräsentieren sollen. Es gibt jedoch auch andere mit Intelligenz assoziierte Indikatoren, die Wissenschaftler nicht als Intelligenzaufgaben zu klassifizieren bereit sind. So könnte zum Beispiel die Bildung der Eltern mit der Intelligenz des Nachwuchses korrelieren, aber es wäre sehr unüblich, sie als einen Indikator für die Intelligenz der Kinder zu verwenden. Geht man von einem Generalfaktor der Intelligenz (g) aus, so lässt sich die Korrelation zwischen zwei etablierten Intelligenztests primär oder ausschließlich auf deren Korrelation mit diesem Generalfaktor zurückführen. Dennoch ist die positive Vielfalt an sich nicht der Grund für den beobachteten Zusammenhang zwischen den Intelligenztests. Diese korrelieren positiv miteinander, weil in ihnen interindividuelle Unterschiede auf die gleichen Fähigkeiten zurückzuführen sind. Es wurde mehrfach argumentiert, dass EI-Methoden mit g korrelieren sollten (Mayer, Salovey, Caruso & Sitarenios, 2001; Roberts, Zeidner & Matthews, 2001; Zeidner, Matthews & Roberts, 2001). Der Grund hierfür ist der mehrfach replizierte Befund, dass zwischen allen als Intelligenztests bezeichneten Verfahren positive Zusammenhänge gefunden werden. Daher müssen EI-Indikatoren mit traditionellen Intelligenzindikatoren positiv korrelieren, falls EI tatsächlich einen Aspekt der allgemeinen Intelligenz repräsentiert.

Es finden sich Erweiterungen und Vertiefungen dieses Arguments, die zum Teil weit über statistische Argumente hinausgehen. Bei einigen spezifischen EI-Instrumenten erwartete und fand man Korrelationen zu diversen Intelligenzaspekten, bei anderen nicht (MacCann et al., 2004). Aus psychologischer Sicht ist es kein Problem, wenn spezifische EI-Instrumente unabhängig von spezifischer oder allgemeiner Intelligenz sind. Wenn es keine Überlappung hinsichtlich der Ursachen für interindividuelle Unterschiede gibt, müssen zwei Tests nicht notwendigerweise korrelieren. In den meisten Fällen findet man jedoch eine gewisse Überlappung. So ist zum Beispiel die Fähigkeit des Schlussfolgernden Denkens in den Aufgaben zum „Emotionsverständnis“ des MSCEIT enthalten. Ähnlich erfordern manche Aufgaben visuelle Verarbeitung und könnten daher mit dem Faktor „broad visualization“ zusammenhängen. Bei anderen EI-Aufgaben wird eine gewisse Allgemeinbildung vorausgesetzt. Diese Aufgaben könnten folglich mit kristalliner Intelligenz assoziiert sein. Andererseits mag ein Teil der gefundenen Zusammenhänge zwischen einigen Aufgaben Artefakte repräsentieren. Wenn zum Beispiel eine Aufgabe zur Messung der emotionalen Intelligenz zum Großteil auf verbalen Beschreibungen

beruht, könnte Leseverständnis eine notwendige, aber nicht hinreichende Bedingung dafür sein, sich mit dem emotionalen Inhalt eines Items auseinandersetzen zu können. Folglich könnte ein künstlicher Zusammenhang entstehen zwischen der Leistung in der EI-Aufgabe und Leseverständnis, verbaler Intelligenz oder sogar allgemeiner Intelligenz.

Im Gegensatz dazu wurde argumentiert, dass die Validität der EI-Fähigkeitstests gestützt wird, wenn keine Zusammenhänge mit Persönlichkeitsskalen vorliegen. Auch wenn hohe Korrelationen ein Grund zur Beunruhigung wären, könnten moderate Korrelationen sehr wohl von Bedeutung sein. Wenn zum Beispiel „Offenheit für Ästhetik“, eine Facette des Big Five Faktors „Offenheit für Erfahrungen“, mit der Leistung in der Aufgabe „Bilder“ aus dem MSCEIT korreliert, könnte dieser Zusammenhang ein gewisses ästhetische Empfinden der Probanden reflektieren, das eine bessere Leistung im Bildertest begünstigt. Allein durch lange Exposition oder intellektuelle Elaboration könnten Personen mit hohen Werten in Offenheit für Ästhetik bessere Leistungen bei Aufgaben wie „Bilder“ zeigen, da sie mit dem Reizmaterial vertrauter sind und auf ein umfassendes Wissen über verschiedene Designs zurückgreifen können. Auch andere Persönlichkeitskonstrukte könnten sinnvoll mit Leistungen in EI-Verfahren korrelieren. Daher reicht es nicht aus, die Fähigkeits-Modelle (oder auch die gemischten Modelle) der emotionalen Intelligenz durch das Aufzeigen konvergenter positiver Zusammenhänge zu anderen Leistungstests und gleichzeitig diskriminanter Nullkorrelationen zu Persönlichkeitskonstrukten zu validieren. Ob und warum konvergente und diskriminante Befunde als hinreichend bezeichnet werden können, ist eine psychologische Fragestellung, die in jedem Fall neu überdacht werden muss.

Ob nun Aufgaben wie die des MSCEIT als Intelligenztests bezeichnet werden sollten, ist eine Entscheidung, die unter Berücksichtigung diverser konzeptueller Aspekte getroffen werden muss. Der Begriff „Intelligenz“ ist selbst so ungenau definiert, dass keine klare Trennungslinie existiert, die ein Verfahren als Intelligenztest zulässt oder eben nicht. Es bleibt die Option bestehen, die MSCEIT- und ähnliche Aufgaben als kognitive Fähigkeitstests anzuerkennen. Ein kognitiver Fähigkeitstest ist sicherlich durch mehrere Eigenschaften gekennzeichnet. Zum Beispiel sollte sich die Leistung verringern, wenn weniger Zeit für die Bearbeitung der Aufgaben zu Verfügung steht. Die Leistungen in solchen Aufgaben – Wissenstests ausgenommen – sollten *ceteris paribus* außerdem schlechter ausfallen, wenn von den Probanden weniger Anstrengung gefordert wird. Das heißt, wenn Teilnehmer gebeten werden, typisches Verhalten zu zeigen, sollten sie schlechtere Leistungen erbringen, als wenn sie zu maximalem Verhalten angehalten werden. In ähnlicher Weise ist die Simulation besserer Leistungen bei Fähigkeitstests bzw. Leistungstests normalerweise nicht möglich. Wenn Probanden die Möglichkeit haben, einen traditionellen IQ-Test erneut zu bearbeiten, werden sie sich signifikant verbessern. Ein profundes Verständnis der Fähigkeit, die in einem solchen Leistungstest erfasst wird, schließt ein, dass Ideen zur Variation der Problemschwierigkeit vorliegen. Die Frage, ob diese Eigenschaften auch für die EI-Tests vorliegen, ist bisher nicht adressiert worden. Caruso, Mayer und Salovey (2002) berichten beim Retest über Leistungsabnahmen bei neun der zwölf Einzeltests im MEIS und gleich bleibende Ergebnisse in den restlichen drei Aufgaben. Die derzeitige Befundlage erlaubt bisher keine Entscheidung darüber, ob sich die Fähigkeitstests zur Messung emotionaler Intelligenz als kognitive Leistungstests qualifizieren.

7.6 Kritik und Empfehlungen

Es sei hier noch einmal ausdrücklich betont, dass frühere Erfahrungen aus der differenzialpsychologischen Forschung, aktuelle Befunde zu *emotionaler Intelligenz* und die Hoffnung auf eine florierende Zukunft dieses Konstrukts allesamt darauf hinweisen, dass EI nicht auf der Grundlage von Selbstberichten oder Selbstbeurteilungen untersucht werden sollte. Eine einfache begriffliche Erweiterung des EI-Konstrukts reicht nicht aus, um klarzustellen, dass Selbstberichts- und Fähigkeits-EI fundamental konzeptuell verschieden und empirisch im Großen und Ganzen voneinander unabhängig sind. Darüber hinaus können Selbstberichtsfragebögen leicht entwickelt und Daten auf einfache Art gesammelt werden. Daher existieren eine Vielzahl von Selbstberichtsmaßen und Studien zur Untersuchung der Struktur individueller Unterschiede in solchen Verfahren. Bei jedem Versuch, ein neues Konstrukt zu etablieren, das ausschließlich durch Fragebögen erfasst wird, muss seine Abgrenzbarkeit gegenüber den anerkannten und etablierten Verfahren nachgewiesen werden. Bei einem breit definierten Konstrukt wie emotionaler Intelligenz wird es auch nötig sein, die interne Struktur der vorgeschlagenen Indikatoren zu untersuchen und zu prüfen, ob die Indikatoren kohärent genug sind, um unter einem Begriff zusammengefasst zu werden. Eine Sammlung von Indikatoren, von denen der eine am ehesten zu selbstberichteter Extraversion, der andere zu selbstberichteter Verträglichkeit gezählt werden kann, ist nicht befriedigend. Es ist erstrebenswert, zumindest ein gewisses Maß an Methodenunabhängigkeit zu demonstrieren. Beispielsweise ist ein Zusammenhang zwischen biographischen Daten und entsprechenden Selbstberichtsangaben wünschenswert. Substantielle Konvergenzen zwischen Selbstbericht und Informationen der Peergruppe über dieselbe Person sind ein weiteres Beispiel für den Nachweis methodenunabhängiger Merkmalsvarianz. Um die ganzen Forschungsanstrengungen überhaupt zu rechtfertigen, ist es schließlich notwendig, eine nicht-triviale inkrementelle Validität des neuen Verfahrens nachzuweisen. All diese Schritte sind essentiell, um das Konstrukt EI, konzeptualisiert als typisches Verhalten, zu etablieren. Letztendlich muss dieses vermeintliche Konstrukt jedoch einen anderen Namen als das auf den EI-Fähigkeitsmessungen basierende Konstrukt erhalten. Ein Begriff wie *typisches emotionales Engagement* könnte eine geeignete Bezeichnung für solch ein prospektives und schwer fassbares Konstrukt sein.

7.6.1 Scoring

Wird EI als Fähigkeit konzeptualisiert, so wird theoretisch postuliert, dass alle Mitglieder der angestrebten Anwendungspopulation diese Fähigkeit besitzen, und zwar in verschieden starker Ausprägung, und dass sie über eine gewisse Zeit stabil bleibt. Die psychometrischen Messeigenschaften von EI sollten etablierten Standards gerecht werden. Dabei ist bei der Verfolgung dieser Standards zu beachten, dass die psychometrische Evaluation eines Verfahrens in der Regel nach der Zuordnung von Werten zu bestimmten Antwortkategorien erfolgt. Dieser Prozess der Zuordnung – das Testscoring – muss ebenfalls gerechtfertigt werden (für eine detaillierte Beschreibung des konsensbasierten Scorings siehe Kapitel 8 von Legree et al.). Das Scoring auf Konsensbasis ist nur ein Weg, den Antwortkategorien bestimmte Werte zuzuordnen. Ihr Einsatz ist bei Tests für implizites oder prozedurales Wissen gerechtfertigt (Legree et al., Kapitel 8). Die Befürworter von EI scheinen jedoch nicht der Überzeugung zu sein, dass EI solch implizites Wissen erfasst. Interessanterweise wird konsensbasiertes Scoring in zwei weiteren Domänen der Psychologie verwendet. Situational Judgment Tests stel-

len eine Methodik dar, implizites berufsrelevantes Wissen zu erfassen (McDaniel et al., 2001; McDaniel & Nguyen, 2001). Praktische Intelligenz (PI) ist ein höchst umstrittenes Konstrukt (Gottfredson, 2003; Sternberg, 2003), das den Erfolg im Alltagskontext messen soll (Wagner, 1987). Die vorherrschende kritische Frage in der Forschung zu PI und SJT ist die, ob PI nicht ohne Rückgriff auf konsensbasiertes Scoring gemessen werden könne. Die gleiche Frage kann auch bezüglich EI gestellt werden: Gibt es ein einheitliches Konstrukt EI und können die Validitätsbefunde des Konstrukts und seiner Messinstrumente auch auf andere, nicht konsensbasierte EI-Instrumente übertragen werden? Zurzeit kann diese Frage nicht zufrieden stellend beantwortet werden.

Ein mehr technisches aber möglicherweise kritisches Problem ist, dass die zahlreichen konsensbasierten Scoringprozeduren nicht ausreichend konvergieren (MacCann et al., 2004). So stellt sich die Frage, welche Prozedur die am besten geeignete ist. Dieses Problem sollte nicht gelöst werden, indem auf die psychometrischen Ergebnisse zurückgegriffen und diejenige Prozedur ausgewählt wird, die die reliabelsten oder die konsistentesten Werte produziert. Welche Scoringprozedur man bei Fähigkeitstests einsetzt, muss auch rational begründet sein. Die Prozeduren, die MacCann et al. miteinander verglichen, waren letztendlich bezüglich ihres Rationales sehr ähnlich. Es gibt jedoch konkurrierende und nicht konvergierende Scoringprozeduren für Tests wie die MEIS oder den MSCEIT. Lediglich befriedigende Konvergenzen zwischen experten- und konsensbasiertem Scoring sind nicht hinreichend, um die Scoringprozedur des MSCEIT zu rechtfertigen.

7.6.2 Vorliegende Validitätsstudien

Es kann ziemlich schwierig sein, bei psychometrischen Messinstrumenten, seien es Selbstberichte oder Leistungstests, korrelative Befunde richtig einzuschätzen. Dies liegt vor allem daran, dass die Zusammenhänge relativ willkürlich interpretiert werden können. Dieses Problem ist nicht neu. Wann immer ein neuer IQ-Test konstruiert wird, wird er validiert, indem Korrelationen zu bestehenden Messinstrumenten erhoben werden. Üblicherweise wird erwartet, dass die Korrelationen hoch, aber nicht perfekt sind. Im Falle perfekter Korrelationen wäre es wenig sinnvoll, ein neues Messinstrument zu etablieren. Ist die Korrelation hoch, so besteht die Möglichkeit, dass der neue Test besser als bereits existierende Tests ist. Dennoch kann eine hohe aber nicht perfekte Korrelation auch völlig andere Szenarien implizieren. So ist ebenfalls möglich, dass der neue Test erhebliche psychometrische Mängel aufweist, und dass, wenn diese Mängel nicht bestünden, die Korrelation mit dem etablierten Instrument perfekt wäre. Bei den Instrumenten zur Messung von EI sieht die Situation ganz ähnlich aus. Wenn zum Beispiel ein neuer EI-Fragebogen mit einem Fragebogen zur Lebenszufriedenheit zu $.70$ und mit Neurotizismus zu $-.50$ korreliert, kann man das dann als Validitätshinweise des neuen Messinstruments verstehen? Man könnte den Befund als ein starkes Indiz für konvergente Validität des EI-Fragebogens deuten. Es könnte jedoch ebenso argumentiert werden, dass dieses Resultat wenig Raum für die Einzigartigkeit des EI-Selbstberichts lässt und dass das neue Verfahren und das Konstrukt, das es erfasst, daher redundant sind.

Nicht wesentlich anders verhält es sich mit kleinen bis moderaten Korrelationen zwischen einem Test emotionaler Fähigkeiten und einem etablierten kognitiven Fähigkeitstest, zum Beispiel zur verbalen Intelligenz. Einerseits können diese als Nachweis für diskriminante Validität verstanden werden, da das EI-Messinstrument wahrscheinlich etwas anderes als der Test zur verbalen Intelligenz erfasst. Andererseits könnte man einwenden, dass die gefundenen Zusammenhänge ein Artefakt darstellen und ein-

zig und allein auf das Testmedium zurückgehen, so dass Probanden mit hoher verbaler Intelligenz bei der Bearbeitung des EI-Tests einen Vorteil haben. Und schließlich könnte man behaupten, dass die kleine Korrelation eine gemeinsame Varianz ausdrückt, die auf die allgemeine Intelligenz zurückgeführt werden kann. Für ein und denselben Befund gibt es also mehrere Erklärungen, und so wird die Interpretation immer in gewissem Maße willkürlich sein. Um ganz offen zu sein, sieht die Situation für jeglichen spezifischen Intelligenztest und dessen Zusammenhänge mit anderen etablierten Messinstrumenten nicht viel anders aus. Dennoch liegen für die meisten Intelligenztests weitere replizierte Befunde vor, die ihre Einbettung in ein nomologisches Netz, ihren inkrementellen Praxisnutzen sowie ihre theoretisch vorhergesagte Redundanz mit ähnlichen und unähnlichen Testformen u. v. m. demonstrieren. Obwohl die meisten Untersuchungen traditioneller Intelligenztests lediglich konzeptuelle Replikationen vorangegangener Studien sind, verbleibt dennoch eine Vielzahl von sicheren Ergebnissen, die alternative Erklärungen für korrelative Befunde ausschließen und so die Interpretation und Validität der Ergebnisse untermauern.

7.6.3 Nicht vorliegende Validitätsstudien

Trotz der zahlreichen Validitätsbefunde findet man überraschenderweise sehr wenige Studien, die die Beziehung der MEIS oder des MSCEIT zu verwandten Aufgaben explorieren. Es wurden z. B. selten Zusammenhänge mit Indikatoren zur Erfassung der sozialen Intelligenz betrachtet. Ebenso wenig liegen Daten unter Verwendung experimenteller Paradigmen vor, zum Beispiel standardisierte Prozeduren zur Erforschung der Gesichtererkennung oder die Emotional Stroop Aufgabe. Auch die nahverwandten methodischen Ansätze zur Erforschung individueller Unterschiede emotionaler Fähigkeiten – zum Beispiel die oben beschriebene LEAS – wurden nur selten als Korrelate herangezogen. Eine umfassende Untersuchung, die neben den üblichen MSCEIT-Aufgaben eine große Bandbreite weiterer emotionaler Aufgaben, inklusive unterschiedlicher Scoringprozeduren, enthält, würde viele richtungsweisende Informationen für die zukünftige Forschung liefern. Bezüglich der Konstruktvalidität ist es sehr wichtig, mehr über die Einbettung von EI in das nomologische Netz verwandter Konstrukte und Instrumente zu erfahren. Neben den etablierten kognitiven Fähigkeitskonstrukten (Carroll, 1993) ist auch die Diskussion wichtig, wie EI mit sozialer Intelligenz, empathischer Akkuratheit, praktischer Intelligenz, inter- und intrapersonellen Fähigkeiten und emotionalem Bewusstsein zusammenhängt. Bedauerlicherweise sind die meisten der eben genannten Konstrukte selbst von zweifelhaftem Wert und Status.

7.6.4 Alternative Modelle

Möglichen Alternativmodellen für die gefundenen Daten wurde bisher nicht ausreichend Beachtung geschenkt. In Abbildung 7.2 sind nur drei mögliche Alternativmodelle dargestellt (für eine Diskussion verschiedener Modellarchitekturen siehe Schulze, 2005).

Die in Abbildung 7.2 dargestellten Modelle weisen eine ähnliche Struktur auf wie das im MSCEIT angenommene Modell, dennoch existieren zwischen ihnen auch wichtige Unterschiede. Manche dieser Modelle kommen ohne Generalfaktor aus (Darstellungen A und B). Mit anderen Worten: Sind die Modellpassungen akzeptabel, können die Kovarianzen zwischen den Aufgaben auch ohne Annahme eines allgemeinen Faktors der emotionalen Intelligenz adäquat erklärt werden. Die Struktur der individuellen Unterschiede in den verfügbaren EI-Messinstrumenten ist nicht gut fundiert. Umfassende

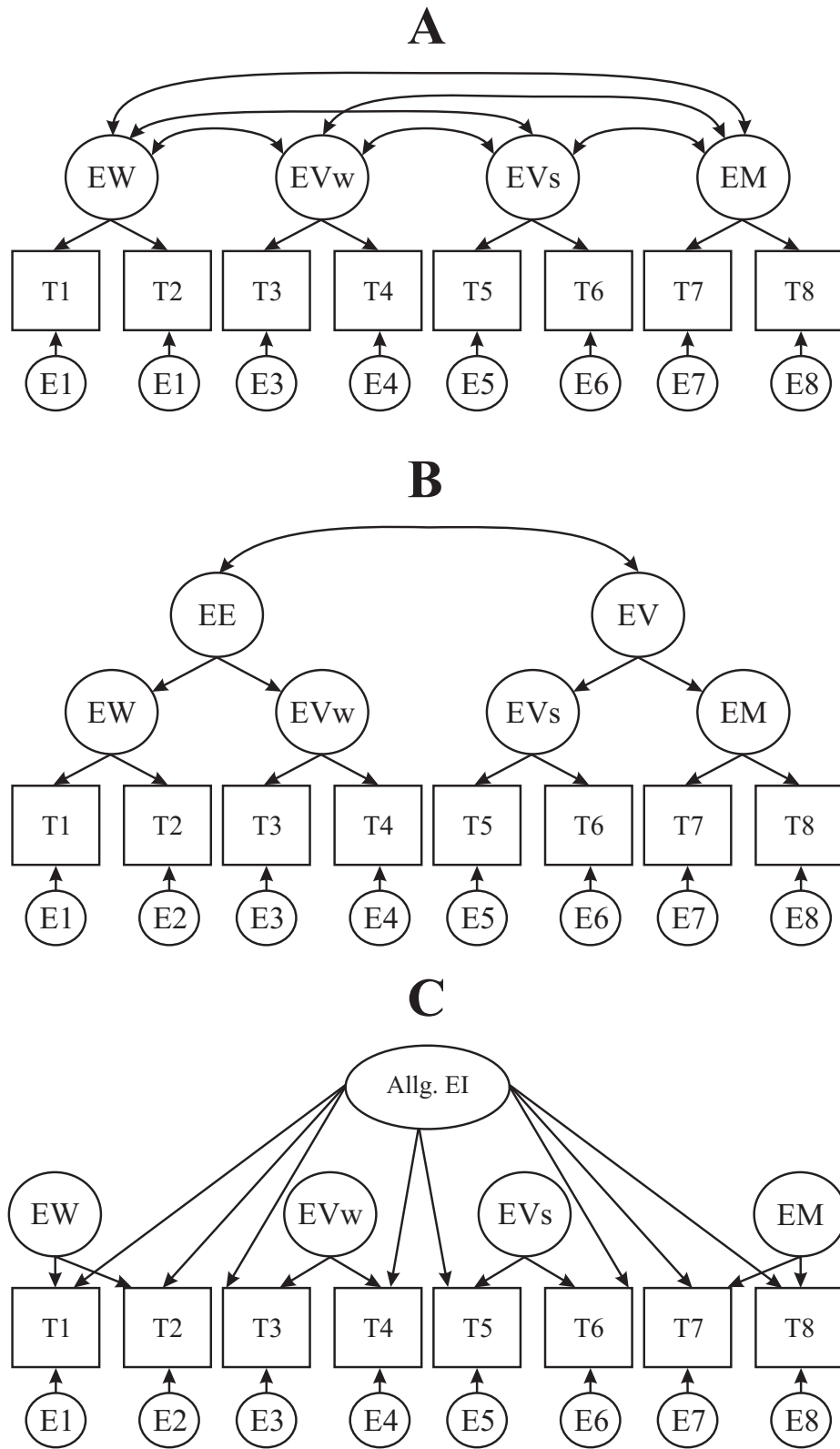


Abbildung 7.2 Alternative Modelle für den MSCEIT.

und vielseitige Forschung unter Verwendung einer großen Bandbreite an verfügbaren Messinstrumenten ist notwendig, um EI-Strukturmodelle miteinander vergleichen zu können.

Mit dem MSCEIT soll eine sehr große Anzahl an Fähigkeiten erfasst werden. Auf der Grundlage von nur acht Aufgabentypen erhalten die Probanden ein Feedback über sieben Fähigkeiten, da die Werte für die Einzeltests insgesamt je dreimal verwendet werden. Die erste Verwendung erfolgt bei der Berechnung der Werte für die vier Zweige, die zweite Verwendung, wenn die Werte der vier Zweige ihrerseits in zwei Werte überführt werden und die dritte Verwendung erfolgt bei der Bildung des Gesamtwerts des MSCEIT aus den beiden Bereichswerten. Das Versagen in oder Probleme mit einer Einzelaufgabe wird daher dreifach bestraft. Diese Redundanz bei der Informationsverwendung wird den Probanden nicht ausreichend verdeutlicht. Außerdem wird den Folgen, die durch die hierarchische Struktur des MSCEIT entstehen, nicht hinreichend Beachtung geschenkt. Um derartige Redundanzen bei der Datenanalyse zu vermeiden, könnte die Varianz der Indikatoren aufgeteilt werden in Varianz aufgrund des Generalfaktors EI und solche aufgrund der Faktoren niedrigerer Ebenen.

7.6.5 Anzahl der Aufgaben pro Faktor und die Allgemeingültigkeit der Faktoren

Die im MSCEIT verwendeten Aufgaben reichen nicht aus, um die EI-Zweige zu erfassen. In die Gesamtwerte der Zweige gehen jeweils nur zwei Aufgabentypen ein. Die Interpretation der Faktoren geht weit über das hinaus, was im Test inhaltlich tatsächlich enthalten ist. Es sollten weit vielseitigere Aufgabentypen eingesetzt werden, bevor Fähigkeiten postuliert werden. Zum Beispiel wurde seit Jahrzehnten immer wieder demonstriert, dass Matrizenaufgaben für zuverlässige Messungen des Konstrukts „fluide Intelligenz“ eingesetzt werden können. Es besteht also ausreichende Evidenz für die Verwendung dieser Aufgaben als Indikator dieses Konstrukts. Natürlich sollte sich ein guter psychologischer Test nicht auf einen einzigen Aufgabentyp verlassen. Andererseits muss jeder verwendete Indikator eine zuverlässige Erfassung des Konstrukts gewährleisten. Es lässt sich einfach überprüfen, ob ein spezifischer Indikator in ein Messmodell der jeweiligen Fähigkeit passt. Es werden lediglich vier oder mehr Indikatoren benötigt, die die gleiche Fähigkeit erfassen sollen. Mit nur drei Faktoren wird das Modell lediglich identifiziert – es ist jedoch keine adäquate Modelltestung möglich. Bei nur zwei Aufgaben ist es nötig, das Modell um weitere Faktoren und Variablen zu erweitern, damit es überhaupt erst bedeutungsvoll ist. Der Status der vier Zweige des MSCEIT kann also nicht adäquat getestet werden, denn dazu sind mehr als zwei Aufgaben für jeden Zweig notwendig.

7.6.6 Alternative Aufgaben (Anzahl und Variationsbreite)

Bei der Ableitung von Einzelaufgaben aus der Beschreibung eines Konstrukts waltet ein erhebliches Maß an Willkür. Technisch gesehen lassen die Definitionen der meisten differenzialpsychologischen Konstrukte die Ableitung einer unbegrenzten Vielzahl von Aufgabentypen zu. Ein hohes Maß an Spezifität der Indikatoren für die zu messende Fähigkeit wird als irrelevant betrachtet. Der Versuchsleiter, das Testmedium, die spezifischen Stimuli (z. B. welche Gesichter in der „Gesichteraufgabe“ präsentiert werden) – das alles wird als irrelevant für das zu messende Konstrukt betrachtet. Variation der Testinstruktionen und Antwortskalen sollte keinen signifikanten Einfluss darauf haben,

was mit dem spezifischen Test gemessen wird. Die Beschreibung der EI-Fähigkeiten lässt noch viele weitere Variationen zu. Zum Beispiel kann die Wahrnehmung von Emotionen durch Musik, Prosa oder Videos von Gesichtsausdrücken, mit künstlerischen Reizen etc. gemessen werden. Ohne derartige Variationen muss Vorsicht vor Übergeneralisierungen der Ergebnisse geboten sein. Im Falle der MEIS und des MSCEIT sind mehr und vielseitigere Indikatoren notwendig, um zu schlussfolgern, dass das postulierte Modell mit Faktoren für die vier Zweige, Faktoren für die zwei Bereiche und schließlich dem Generalfaktor ein angemessenes Modell für EI ist. In anderen Worten liefert der MSCEIT unter Verwendung sehr spezifischer Aufgaben eine sehr allgemeine Interpretation der emotionalen Intelligenz.

7.6.7 Testkonstruktion

Bei der Testkonstruktion sollte nach Möglichkeit deduktiv und theoriegeleitet vorgegangen werden. Die Messintention der EI-Instrumente sollte möglichst durch experimentelle und neuropsychologische Forschung inspiriert werden. Ein kritisches und wichtiges Anliegen ist der Aufbau und die Aufrechterhaltung eines engen Zusammenhangs zwischen der Messung differenzieller Unterschiede und allgemeinen Emotionstheorien. Der Gebrauch eines Instruments sollte durch das gerechtfertigt sein, was er misst. Neben einer präzisen Beschreibung der Messintention und der Operationalisierung sind die psychometrischen Kriterien wichtig, aber der Testinhalt ist entscheidend. Obwohl die Aufgaben in der MEIS und im MSCEIT gute Indikatoren für die vorgeschlagenen Zweige zu sein scheinen, ist nur wenig über Alternativen oder den emotionstheoretischen Hintergrund bekannt. Es wäre wünschenswert, stärkere Assoziationen zwischen den Einzelindikatoren und den Konstrukten, die sie erfassen sollen, zu etablieren.

7.7 Schlussfolgerungen

Trotz der oben dargestellten kritischen Aspekte repräsentiert der MSCEIT den ambitioniertesten und bis heute besten Zugang zur umfassenden Messung der emotionalen Fähigkeiten. Es gilt sich jedoch der methodischen und auch der psychologischen Herausforderungen anzunehmen. Da sich die Forschung zur Validität und zum Nutzen des EI-Konstrukts sowie der vorgeschlagenen Messinstrumente noch in einem sehr frühen Stadium befindet, ist bei der Anwendung der Instrumente Vorsicht geboten. Zusammenfassend lässt sich feststellen, dass sich die Befürworter der als Fähigkeit definierten EI bisher anständig geschlagen haben. Während es noch zu früh ist, ein einfaches Modell anzuerkennen, wurden in diesem Forschungsfeld in den letzten zehn Jahren große Fortschritte errungen. Neben den nachteiligen Folgen aufgrund des verfrühten Gebrauchs der Instrumente und des inflationären Charakters des Gesamtkonzepts brachte die starke öffentliche Aufmerksamkeit für EI auch einen wichtigen positiven Effekt: Sie führte dazu, dass dieser wichtige und bisher ignorierte Bereich menschlicher Fähigkeiten wissenschaftlich intensiver untersucht wird.

Andererseits hat die enthusiastische Annahme der frühen Vorschläge zum EI-Konstrukt den Blick für zeitgemäße und sachgerechte Forschungsprinzipien verstellt. Eine erschreckende Konsequenz davon ist, dass Tests zur Erfassung der emotionalen Intelligenz in der Praxis verwendet wurden, bevor kritische Fragen empirisch beantwortet werden konnten. Tatsächlich sind wir derzeit noch nicht einmal in der Lage, die richtigen Fragen überhaupt zu stellen.

Um dieses Kapitel optimistisch abzuschließen soll ein viel versprechender Forschungsansatz vorgestellt werden, der durch neuropsychologische und experimentelle Arbeiten zur Gesichtererkennung inspiriert wurde. Gute Gesichterwahrnehmung und Gesichtererkennung erlauben es dem Menschen, Informationen über das Alter, das Geschlecht, die Stimmung und die Identität einer Person zu erhalten. Folglich kann die Gesichtererkennung ein limitierender Faktor für einige EI-Aspekte sein. Interindividuelle Unterschiede in diesen Bereichen der EI könnten dann auf die individuellen Unterschiede bei der Gesichtererkennung zurückgeführt werden. Es liegen physiologische Befunde vor, die besagen, dass das Erkennen von Gesichtern aus zwei verschiedenen Komponenten besteht. Die erste Komponente bezieht sich auf die Enkodierung von Gesichtern. Erfolgreiches und nicht erfolgreiches Lernen von unbekanntem Gesichtern ist mit neurophysiologischen Unterschieden verbunden (Schweinberger, Pfütze & Sommer, 1995; Sommer, Komoss & Schweinberger, 1997). Die zweite Komponente hat mit dem Abruf von bekannten und unbekanntem Gesichtern aus dem Gedächtnis zu tun. Der so genannte frühe Wiederholungseffekt unterscheidet sich für mit privaten Bekanntschaften, bekannten öffentlichen Personen und unbekanntem Personen assoziierte Gesichter (Herzmann, Schweinberger, Sommer & Jentsch, 2004). Von diesen und ähnlichen Forschungszugängen kann man sich überzeugende Befunde für die Aufdeckung individueller Unterschiede, die sich auf Enkodierung und Abruf von Gesichtern zurückführen lassen versprechen (Pfütze, Sommer & Schweinberger, 2002). Dieses Gebiet ist sehr inspirierend. Wie wäre es, zum Beispiel, Verfahren zur Erfassung der Gesichtswahrnehmung und zur Erkennung von Änderungen im Gesichtsausdruck zu entwickeln oder Instrumente zur Erfassung minimaler Veränderungen des Gesichtsausdrucks?

Literatur

- Ackerman, P. L., Beier, M. E. & Bowen, K. R. (2002). What we really know about our abilities and our knowledge. *Personality and Individual Differences*, 33, 587–605.
- Bagby, R. M., Parker, J. D. A. & Taylor, G. J. (1994a). The Twenty-Item Toronto Alexithymia Scale-II: Convergent, discriminant, and concurrent validity. *Journal of Psychosomatic Research*, 38, 33–40.
- Bagby, R. M., Parker, J. D. A. & Taylor, G. J. (1994b). The Twenty-Item Toronto Alexithymia Scale-I: Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, 38, 23–32.
- Bar-On, R. (1997). *BarOn Emotional Quotient Inventory (EQ-i): Technical manual*. Toronto, Canada: Multi-Health Systems.
- Bar-On, R. (2000). Emotional and social intelligence: Insights from the emotional quotient inventory. In R. Bar-On & J. D. A. Parker (Eds.), *The handbook of emotional intelligence: Theory, development, assessment, and application at home, school, and in the workplace* (pp. 363–388). San Francisco, CA: Jossey-Bass.
- Brackett, M. A. & Mayer, J. D. (2003). Convergent, discriminant, and incremental validity of competing measures of emotional intelligence. *Personality and Social Psychology Bulletin*, 29, 1147–1158.
- Brackett, M. A., Mayer, J. D. & Warner, R. M. (2004). Emotional intelligence and its relation to everyday behaviour. *Personality and Individual Differences*, 36, 1387–1402.
- Broadbent, D. E., Cooper, P. J., Fitzgerald, P. F. & Parkes, K. R. (1982). The Cognitive Failures Questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology*, 21, 1–16.

- Cacioppo, J. T. & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*, 116–131.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Caruso, D. R., Mayer, J. D. & Salovey, P. (2002). Relation of an ability measure of emotional intelligence to personality. *Journal of Personality Assessment*, *79*, 306–320.
- Ciarrochi, J., Chan, A. Y. & Caputi, P. (2000). A critical evaluation of the emotional intelligence construct. *Personality and Individual Differences*, *28*, 539–561.
- Ciarrochi, J., Scott, G., Deane, F. P. & Heaven, P. C. L. (2003). Relations between social and emotional competence and mental health: a construct validation study. *Personality and Individual Differences*, *35*, 1947–1963.
- Coffey, E., Berenbaum, H. & Kerns, J. G. (2003). The dimensions of emotional intelligence, alexithymia, and mood awareness: Associations with personality and performance on an emotional stroop task. *Cognition & Emotion*, *17*, 671–679.
- Costa, P. T., Jr. & McCrae, R. R. (1992). *Revised NEO personality and five factor inventory professional manual*. Odessa, FL: Psychological Assessment Resources.
- Cronbach, L. J. (1949). *Essentials of psychological testing*. New York: Harper & Row.
- Dennis, M. J., Sternberg, R. J. & Beatty, P. (2000). The construction of “user-friendly” tests of cognitive functioning: A synthesis of maximal- and typical-performance measurement philosophies. *Intelligence*, *28*, 193–211.
- Epstein, S., Pacini, R., Denes-Raj, V. & Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology*, *71*, 390–405.
- Goff, M. & Ackerman, P. L. (1992). Personality-intelligence relations: Assessment of typical intellectual engagement. *Journal of Educational Psychology*, *84*, 537–552.
- Goleman, D. (1995). *Emotional intelligence: Why it can matter more than IQ*. New York: Bantam Books.
- Goleman, D. (1998). *Working with emotional intelligence*. New York: Bantam.
- Gottfredson, L. S. (2003). Dissecting practical intelligence theory: Its claims and evidence. *Intelligence*, *31*, 343–397.
- Guttman, L. (1965). A faceted definition of intelligence. *Scripta Hierosolymitana*, *14*, 166–181.
- Herrmann, D. J. (1982). Know thy memory: The use of questionnaires to assess and study memory. *Psychological Bulletin*, *92*, 434–452.
- Hertzog, C., Park, D. C., Morrell, R. W. & Martin, M. (2000). Ask and ye shall receive: Behavioral specificity in the accuracy of subjective memory complaints. *Applied Cognitive Psychology*, *14*, 257–275.
- Herzmann, G., Schweinberger, S. R., Sommer, W. & Jentzsch, I. (2004). What’s special about personally familiar faces? A multimodal approach. *Psychophysiology*, *41*, 688–701.
- Kahneman, D. (1973). *Attention and effort*. Hillsdale, NJ: Prentice-Hall.
- Lane, R., Quinlan, D., Schwartz, G. E., Walker, P. & Zeitlin, S. (1990). The levels of emotional awareness scale: A cognitive-developmental measure of emotion. *Journal of Personality Assessment*, *55*, 124–134.
- Lane, R., Reiman, E., Axelrod, B., Yun, L.-S., Holmes, A. H. & Schwartz, G. E. (1998). Neural correlates of levels of emotional awareness: Evidence of an interaction between emotion and attention in the anterior cingulate cortex. *Journal of Cognitive Neuroscience*, *10*, 225–235.
- Lane, R. & Schwartz, G. E. (1987). Levels of emotional awareness: A cognitive-developmental theory and its application to psychopathology. *American Journal of Psychiatry*, *144*, 133–143.

- MacCann, C., Roberts, R. D., Matthews, G. & Zeidner, M. (2004). Consensus scoring and empirical option weighting of performance-based emotional intelligence (EI) tests. *Personality and Individual Differences*, *36*, 645–662.
- Matthews, G., Zeidner, M. & Roberts, R. D. (2002). *Emotional intelligence: Science and myth*. Cambridge, MA: MIT Press.
- Mayer, J. D., Caruso, D. R. & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence*, *27*, 267–298.
- Mayer, J. D., Caruso, D. R. & Salovey, P. (2000). Selecting a measure of emotional intelligence: The case for ability testing. In R. Bar-On & J. D. A. Parker (Eds.), *The handbook of emotional intelligence: Theory, development, assessment, and application at home, school, and in the workplace* (pp. 320–342). San Francisco: Jossey-Bass.
- Mayer, J. D., DiPaolo, M. & Salovey, P. (1990). Perceiving affective content in ambiguous visual stimuli: A component of emotional intelligence. *Journal of Personality Assessment*, *54*, 772–781.
- Mayer, J. D. & Salovey, P. (1993). The intelligence of emotional intelligence. *Intelligence*, *17*, 433–442.
- Mayer, J. D. & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. J. Sluyter (Eds.), *Emotional development and emotional intelligence: Educational implications* (pp. 3–31). New York: Basic Books.
- Mayer, J. D., Salovey, P. & Caruso, D. R. (2000). Models of emotional intelligence. In R. J. Sternberg (Ed.), *The handbook of intelligence* (pp. 396–420). New York: Cambridge University Press.
- Mayer, J. D., Salovey, P. & Caruso, D. R. (2002). *The Mayer, Salovey, and Caruso Emotional Intelligence Test: Technical manual*. Toronto: Multi-Health Systems.
- Mayer, J. D., Salovey, P., Caruso, D. R. & Sitarenios, G. (2001). Emotional intelligence as a standard intelligence. *Emotion*, *1*, 232–242.
- Mayer, J. D., Salovey, P., Caruso, D. R. & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. *Emotion*, *3*, 97–105.
- McCrae, R. R. (1990). Traits and trait names: How well is openness represented in natural languages. *European Journal of Personality*, *4*, 119–129.
- McCrae, R. R. (1996). Social consequences of experiential openness. *Psychological Bulletin*, *120*, 333–337.
- McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A. & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology*, *86*, 730–740.
- McDaniel, M. A. & Nguyen, N. T. (2001). Situational judgment tests: A review of practice and constructs assessed. *International Journal of Selection & Assessment*, *9*, 103–113.
- McGlynn, S. M. & Schacter, D. L. (1989). Unawareness of deficits in neuropsychological syndromes. *Journal of Clinical and Experimental Neuropsychology*, *11*, 143–205.
- Mehrabian, A. & Epstein, S. (1970). A measure of emotional empathy. *Journal of Personality*, *40*, 525–543.
- Nevo, B. (1993). In search of a correctness typology for intelligence. *New Ideas in Psychology*, *11*, 391–397.
- Oberauer, K., Süß, H.-M., Schulze, R., Wilhelm, O. & Wittmann, W. W. (2000). Working memory capacity—Facets of a cognitive ability construct. *Personality and Individual Differences*, *29*, 1017–1045.
- O'Connor, R. M. & Little, I. S. (2003). Revisiting the predictive validity of emotional intelligence: self-report versus ability-based measures. *Personality and Individual Differences*, *35*, 1893–1902.

- Ones, D. S., Viswesvaran, C. & Dilchert, S. (2004). Cognitive ability in selection decisions. In O. Wilhelm & R. W. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 431–463). London: Sage.
- Palmer, B. R., Manocha, R., Gignac, G. & Stough, C. (2003). Examining the factor structure of the Bar-On Emotional Quotient Inventory with an Australian general population sample. *Personality and Individual Differences, 35*, 1191–1210.
- Petrides, K. V. & Furnham, A. (2000). On the dimensional structure of emotional intelligence. *Personality and Individual Differences, 29*, 313–320.
- Petrides, K. V. & Furnham, A. (2001). Trait emotional intelligence: Psychometric investigation with reference to established trait taxonomies. *European Journal of Personality, 15*, 425–448.
- Pfütze, E.-M., Sommer, W. & Schweinberger, S. R. (2002). Age-related slowing in face and name recognition: Evidence from event-related brain potentials. *Psychology and Aging, 17*, 140–160.
- Reason, J. T. (1993). Self-report questionnaires in cognitive psychology: Have they delivered the goods? In A. D. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control* (pp. 152–170). Oxford: Clarendon Press.
- Riemann, R. & Abels, D. (1994). Personality abilities: Construct validation. In B. de Raad, W. K. B. Hofstee & G. L. M. van Heck (Eds.), *Personality psychology in europe* (Vol. 5, pp. 201–215). Tilburg: Tilburg University Press.
- Roberts, R. D., Zeidner, M. & Matthews, G. (2001). Does emotional intelligence meet traditional standards for an intelligence? Some new data and conclusions. *Emotion, 1*, 196–231.
- Sackett, P. R., Zedeck, S. & Fogli, L. (1988). Relations between measures of typical and maximal job performance. *Journal of Applied Psychology, 73*, 482–486.
- Saklofske, D. H., Austin, E. J. & Minski, P. S. (2003). Self-reported emotional intelligence: Factor structure and evidence for construct validity. *Personality and Individual Differences, 34*, 707–721.
- Salovey, P. & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality, 9*, 185–211.
- Schulze, R. (2005). Modeling structures of intelligence. In O. Wilhelm & R. W. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 241–263). London: Sage.
- Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D. J., Cooper, J. T., Golden, C. J. et al. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences, 25*, 167–177.
- Schweinberger, S. R., Pfütze, E.-M. & Sommer, W. (1995). Repetition priming and associative priming of face recognition: Evidence from event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 722–736.
- Seidenberg, M., Haltiner, A., Taylor, M. A., Hermann, B. B. & Wyler, A. (1994). Development and validation of a multiple ability self-report questionnaire. *Journal of Clinical and Experimental Neuropsychology, 14*, 93–104.
- Sommer, W., Komoss, E. & Schweinberger, S. R. (1997). Differential localization of brain systems subserving memory for names and faces in normal subjects with event-related potentials. *Electroencephalography and Clinical Neurophysiology, 102*, 192–199.
- Stankov, L. (1999). Mining on the “no man’s land” between intelligence and personality. In P. L. Ackerman, P. C. Kyllonen & R. D. Roberts (Eds.), *Learning and individual differences: Process, trait, and content determinants* (pp. 315–337). Washington, DC: American Psychological Association.
- Sternberg, R. J. (2003). Our research program validating the triarchic theory of successful intelligence: Reply to Gottfredson. *Intelligence, 31*, 399–413.

- Van Rooy, D. L. & Viswesvaran, C. (2004). Emotional intelligence: A meta-analytic investigation of predictive validity and nomological net. *Journal of Vocational Behavior*, 65, 71–95.
- Wagner, R. K. (1987). Tacit knowledge in everyday intelligent behavior. *Journal of Personality and Social Psychology*, 52, 1236–1247.
- Wilhelm, O., Schulze, R., Schmiedek, F. & Süß, H.-M. (2003). Interindividuelle Unterschiede im typischen intellektuellen Engagement. *Diagnostica*, 49, 49–60.
- Zeidner, M., Matthews, G. & Roberts, R. D. (2001). Slow down, you move too fast: Emotional intelligence remains an “elusive” intelligence. *Emotion*, 1, 265–275.