Die Natur der Persönlichkeit:

Validität, Kontinuität und Interaktion von Anlage und Umwelt

Synopse und Publikationen der kumulativen Dissertation
zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)
an der Fakultät für Psychologie und Sportwissenschaft
der Universität Bielefeld

vorgelegt von

Christian Kandler

Bielefeld, März 2010

Abteilung für Psychologie, Universität Bielefeld, Bielefeld, Deutschland

Betreuer und Erstgutachter: Prof. Dr. Rainer Riemann, Universität Bielefeld

Zweitgutachter: Prof. Dr. Peter Borkenau, Martin-Luther-Universität Halle-Wittenberg

Christian Kandler
Universität Bielefeld
Telefon: 0521/106 4540
E-Mail: christian.kandler@uni-bielefeld.de

Danksagung

Es ist mir eine Freude, den vielen Menschen danken zu können, die mich bei meiner Arbeit gefördert und unterstützt haben. Besonders dankbar bin ich meiner Frau Anja, welche mir eine zeitlose und kraftspendende Insel war und ist, die mich bei all meinen Vorhaben uneingeschränkt unterstützte und mir die nötige Ruhe schenkte.

Besonderen Dank möchte ich auch all den Wissenschaftlern aussprechen, die ich in diesem Buch zitiere. Ganz besonders danke ich Rainer Riemann, der mir die Möglichkeit, als Wissenschaftler tätig zu sein, erst eröffnete, die nötigen Freiräume gab und stets mit Rat und Zeit zur Seite stand. Nicht weniger danke ich Wiebke Bleidorn für eine fruchtbare Zusammenarbeit.

Ich möchte mich auch bei Rainer Riemann, Alois Angleitner und Frank M. Spinath für die Initiierung und Bereitstellung unvergleichlicher Datenschätze danken und Wolfgang Thiel für deren Verwaltung. Damit danke ich auch allen Zwillingen, die an der Jena Twin Study of Social Attitudes (JeTSSA) und an der Bielefeld Longitudinal Study of Adult Twins (BiLSAT) teilgenommen haben. Nicht zu vernachlässigen ist der Beitrag von Fritz Ostendorf und Peter Borkenau für die Fertigung erstklassiger Messinstrumente.

Darüber hinaus danke ich allen Koautoren für ihre Unterstützung bei den drei Artikeln. Außerdem möchte ich Martina Diez, Kevin Elsmann und Bärbel Kandler für wertvolle Rückmeldungen zu einer ersten Version der Synopse danken.

“What nature proposes, environment disposes.”

(S. Scarr, 1993, S.1333)

Inhaltsverzeichnis

Liste der Schriften

Diese kumulative Dissertation basiert auf den folgenden Studien, welche in anerkannten wissenschaftlichen Zeitschriften (mit Peer-Review) veröffentlicht sind und werden.

I. Kandler, C., Riemann, R., & Kämpfe, N. (2009). **Genetic and environmental mediation between measures of personality and family environment in twins reared together.** *Behavior Genetics*, *39*, 24-35. DOI: 10.1007/s10519-008-9238-8.

II. Kandler, C., Riemann, R., Spinath, F. M., & Angleitner, A. (2010). **Sources of variance in personality facets: A multiple-rater twin study of self-peer, peer-peer, and self-self (dis-) agreement.** *Journal of Personality*, *78:5.*[1]

III. Kandler, C., Bleidorn, W., Riemann, R., Spinath, F. M., Thiel, W., & Angleitner, A. (2010). **Sources of cumulative continuity in personality: A longitudinal multiple-rater twin study.** *Journal of Personality and Social Psychology*.[2]

Auf diese Schriften wird in der folgenden Synopse unter Nennung der römischen Ziffern Bezug genommen werden. Die vollständigen Schriften sind angehängt.

---

[1] The manuscript (JOPY-09-0115.R2) is scheduled to go to the publisher at the end of April and proofs will be sent to authors via email in June. Print publication is scheduled for the JP October 2010 issue, 78:5, with on-line early publication 1-2 months prior to that (Aug. or Sept.). Page numbers of course, will not be available for some time.

Ellen M. Malaspina
Editorial Manager
Journal of Personality
University of Connecticut Health Center
Department of Community Medicine and Health Care
263 Farmington Avenue
Farmington, CT 06030-6325
Tel 860-679-1594 Fax 860-679-5464
Email: malaspina@nso2.uchc.edu

[2] The manuscript (2009-0097-RR) is accepted for publication at JPSP.

Lisa Jensen
Editorial Manager
Journal of Personality and Social Psychology
610 W. Stewart Road
Columbia, MO 65203
Email: jensenl@missouri.edu

Die Natur der Persönlichkeit

1. Einleitung

Spricht man über Persönlichkeit, ist es wichtig, zunächst einmal zu klären, worüber man eigentlich spricht. Nicht nur zwischen Alltag und Wissenschaft, auch innerhalb der Wissenschaft und sogar innerhalb der Psychologie herrscht große Uneinigkeit. Ein Persönlichkeitspsychologe versteht unter Persönlichkeit die einzigartige Individualität eines Menschen in physischer und psychischer Erscheinung, im Verhalten und Erleben, durch die sich Individuen im Vergleich zu anderen Menschen unterscheiden. Im Vergleich zu Persönlichkeit *im weiten Sinne* (Asendorpf, 2009) beziehe ich mich auf den Persönlichkeitsbegriff *im engeren Sinne*, d.h. die Gesamtheit aller „typischen" Eigenschaften in Abgrenzung zu dem Begriff Intelligenz als allgemeine „maximale" kognitive Leistungsfähigkeit. Die moderne Psychologie versteht Persönlichkeit nicht mehr als ein Sammelsurium von Charaktertypen oder Charakterzügen (König, 2002), die ein Mensch aufweist und ein anderer nicht unbedingt, sondern als breites Spektrum von Eigenschaftsdimensionen, die allen Menschen innewohnen. Individuen unterscheiden sich lediglich hinsichtlich ihres Ausprägungsgrades auf diesen Dimensionen. Dabei hat sich keine einzelne übergreifende Dimension als gut fundiert herausgestellt, wie es vielleicht einige Forscher neuerdings wieder behaupten (z.B. Rushton & Erwing, 2008), sondern mehrere distinkte, aber breite Persönlichkeitseigenschaften, wie zum Beispiel die Big Five (Goldberg, 1993), bzw. die fünf Persönlichkeitsdomänen des Fünf-Faktoren-Modells (FFM; McCrae & John, 1992): Neurotizismus (versus Emotionale Stabilität), Extraversion (versus Introversion), Offenheit für Erfahrungen, Verträglichkeit (versus Aggressivität) und Gewissenhaftigkeit (versus Rigidität).

Wenn man von „typischen" individuellen Ausprägungen auf einer Persönlichkeitsdimension spricht, impliziert dies eine relative Stabilität. Das grenzt stabile Persönlichkeitseigenschaften (engl. *trait*) als eine Verhaltensdisposition von instabilen Zuständen (engl. *state*) ab. Persönlichkeitseigenschaften prädisponieren (nicht determinieren) mit einer bestimmten Wahrscheinlichkeit zu einem bestimmten Zeitpunkt in einer gegebenen Situation zu einem bestimmten Verhalten. Sie determinieren nicht. Insofern sind Persönlichkeitseigenschaften in einer gegebenen Situation zu einem bestimmten Zeitpunkt nicht direkt beobachtbar (latente Konstrukte). Man benötigt also konsistente (zeit- und situationsübergreifende) Indikatoren zur Erfassung einer Persönlichkeitseigenschaft.

In der Persönlichkeitsforschung allgemein und vor allem in der Untersuchung interindividueller Unterschiede wird oft nur auf eine Methode – meistens Selbstberichte zu verschiedenen situationsübergreifenden Verhaltensaspekten – zur Messung der Persönlichkeit zurückgegriffen, um Aussagen über das Persönlichkeitskonstrukt oder dessen Gültigkeit allgemein zu machen. Obwohl klare Richtlinien und Handlungsanweisungen zur Bestimmung der Konstruktvalidität formuliert wurden (z.B. Campbell & Fiske, 1959), riefen diese im Verhältnis zur großen Anzahl von Zitationen nur wenig Nachhaltigkeit im Umgang mit Persönlichkeitspsychometrie hervor (Fiske

& Campbell, 1992). Auch dem nunmehr Jahrzehnte alten Forschungsfeld von Konsensus und Akkuratheit in der Persönlichkeitsmessung (Funder, 1995; Kenny, 1991) wurde bisher selten Einlass in die Lehrbücher und Anwendung gewährt.

Das beste und aktuellste Beispiel der Folgen mangelnder Nachhaltigkeit ist sicherlich das Postulat eines Generalfaktors der Persönlichkeit. Ähnlich einem Generalfaktor der Intelligenz (Spearman, 1904) – als ein Korrelat aller kognitiven Fähigkeiten – basiert das Postulat eines Generalfaktors der Persönlichkeit auf signifikanten Zusammenhängen (Korrelationen) zwischen den Big Five. Zwar gab es Evidenz für diesen Generalfaktor *konsistent* über verschiedene Datensätze und Fragebögen (Musek, 2007; Rushton & Irwing, 2008; Rushton & Irwing, 2009), jedoch nicht messmethodenübergreifend, d.h. nur ausschließlich innerhalb einer Messmethode (meistens Selbstberichte). Eine Messmethode allein erlaubt jedoch keine Differenzierung dahingehend, ob die Messungen der Big Five auf Substanz im Sinne messmethodenübergreifender Gültigkeit des Konstrukts (*Konvergente Validität*; Campbell & Fiske, 1959) oder auf eine rein artifizielle Ursache im Sinne einer gänzlich methodeninhärenten Natur (*Beurteiler-* oder *Methodenbias*; Paulhus & John, 1998) zurückzuführen sind. Wenngleich die *Konvergente Validität* der Big Five hinlänglich belegt ist, fehlte die Evidenz methodenübergreifender Korrelationen zwischen den Big Five zur Stützung eines Generalfaktors (*Diskriminante Validität*; Biesanz und West, 2004). Neuere Studien (Anusic, Schimmack, Pinkus & Lockwood, 2009; Riemann & Kandler, in Druck), die neben Selbsteinschätzungen auch Berichte guter Bekannter (engl. *multimethod*) zur Messung der fünf Persönlichkeitsdimensionen des FFM (engl. *multitrait*) verwendeten, fanden durchaus Generalfaktoren basierend auf den Korrelationen zwischen den Persönlichkeitsvariablen, jedoch nur innerhalb von Selbst- und Bekanntenberichten. Es fand sich also keine *Konvergente Validität* eines Generalfaktors der Persönlichkeit. Diese Befunde stützen eher die Interpretation eines messmethodenspezifischen Generalfaktors als Messmethodenartefakt und nicht als Substanz im Sinne eines validen messmethodenunabhängigen Persönlichkeitskonstrukts. Da die Persönlichkeitsdimensionen des FFM induktiv analytisch abgeleitet (McCrae & John, 1992), theoretisch verankert (McCrae & Costa, 2008) und interkulturell gut fundiert sind (Yamagata, Suzuki, Ando, Ono, Kijima et al., 2006), möchte ich in der Folge die Betrachtung der Persönlichkeit in erster Linie auf diese Eigenschaftsdimensionen beschränken.

Die Persönlichkeitsforschung wäre keine Wissenschaft, wenn man sich mit Persönlichkeitsbeschreibungen zufrieden geben würde. Vielmehr stellt sich die Frage, welche Rolle genetische und biologische Grundlagen im Gegensatz zur Prägung durch Kultur und Umgebung bei stabilen Persönlichkeitsmerkmalen spielen. Die Designs der Verhaltensgenetik (Zwillingsstudie, Adoptionsstudie) erlauben interindividuelle Persönlichkeitsunterschiede auf Erbanlagen und Umwelteffekte zurückzuführen, was sie mittlerweile nun schon seit 150 Jahren für vielfältige physische, psychische und auch soziale Merkmale tun (Plomin, DeFries, McClearn & McGuffin, 2008). Die verhaltensgenetische Forschung bietet nicht nur die einzigartige Möglichkeit, genetische Faktoren zu identifizieren, ohne auf molekulargenetischer Ebene zu arbeiten, sondern auch

adäquate Umweltforschung zu betreiben, indem sie „wahre“ – um den genetischen Beitrag bereinigte – Umweltfaktoren zu entschlüsseln vermag. Die Persönlichkeitsforschung ist sehr daran interessiert, welche Einflüsse die Erfahrungen im Elternhaus, der Austausch mit Geschwistern oder anderen Verwandten auf die Entwicklung von Persönlichkeitseigenschaften haben. Die Bestimmung solcher „wahren“ Umwelteffekte auf die Persönlichkeit ist nur dann feststellbar, wenn um den Beitrag genetischer Verwandtenähnlichkeit kontrolliert wird. Zwillingsstudien ermöglichen es, den genetischen Anteil zu kontrollieren. Adoptionsstudien erlauben es, Eltern-Kind und Geschwisterbeziehungen zwischen nicht genetisch verwandten Individuen zu untersuchen, um nicht-genetisch-vermittelte Effekte ausfindig zu machen.

Seit der ersten Zwillingsstudie zu Persönlichkeitsmerkmalen von Cattell, Blewett und Beloff (1955) vor über 50 Jahren sind bis zum heutigen Tag weit über 100 weitere genetisch informative Persönlichkeitsstudien durchgeführt worden. So wissen wir heute, dass etwa 50% der beobachtbaren Varianz (*phänotypische Variabilität*) auf genetische (*genotypische Variabilität* = *Erblichkeit*) und 50% auf Umweltfaktoren zurückzuführen ist. Etwa 90% der Studien basieren allerdings lediglich auf einer Methode (meist Selbsteinschätzungen). Welche Probleme diese Einschränkung mit sich bringen kann, wurde schon am Beispiel des Generalfaktor-Postulats erläutert. Man könnte die Korrelation zwischen Selbstberichten eineiiger Zwillinge als untere Grenze *Konvergenter* (oder *Konsensueller*) *Validität* ansehen, wenn genetisch identische Individuen unabhängig voneinander ihre Persönlichkeit einschätzen (Goldsmith & Gottesman, 1977). Der Konsensus von Zwillingen als Maß für die *Konsensuelle Validität* würde um den Anteil nicht erblicher Faktoren auf das Persönlichkeitsmerkmal unterschätzt, welche zwischen den Zwillingen nicht geteilt werden. Der Ansatz ist sicher streitbar und bleibt hinter der *Multitrait-Multimethod* Analyse (Campbell & Fiske, 1959) zurück, jedoch ermöglicht er ferner in Kombination mit dieser zusätzliche Informationen über genetische und umweltbedingte Quellen sowohl von „wahrer“ (im Sinne von konvergent oder konsensuell valide) als auch von methodenspezifischer Varianz. Dieser Kombinationsansatz erlaubt somit Implikationen nicht nur zur Interpretation des betrachteten Persönlichkeitskonstrukts (Annäherung an „wahre“ genetische und Umwelteinflüsse) sondern auch zur Interpretation methodenspezifischer Einflüsse (Bartels, Boomsma, Hudziak, Beijsterveldt & van den Oord, 2007b).

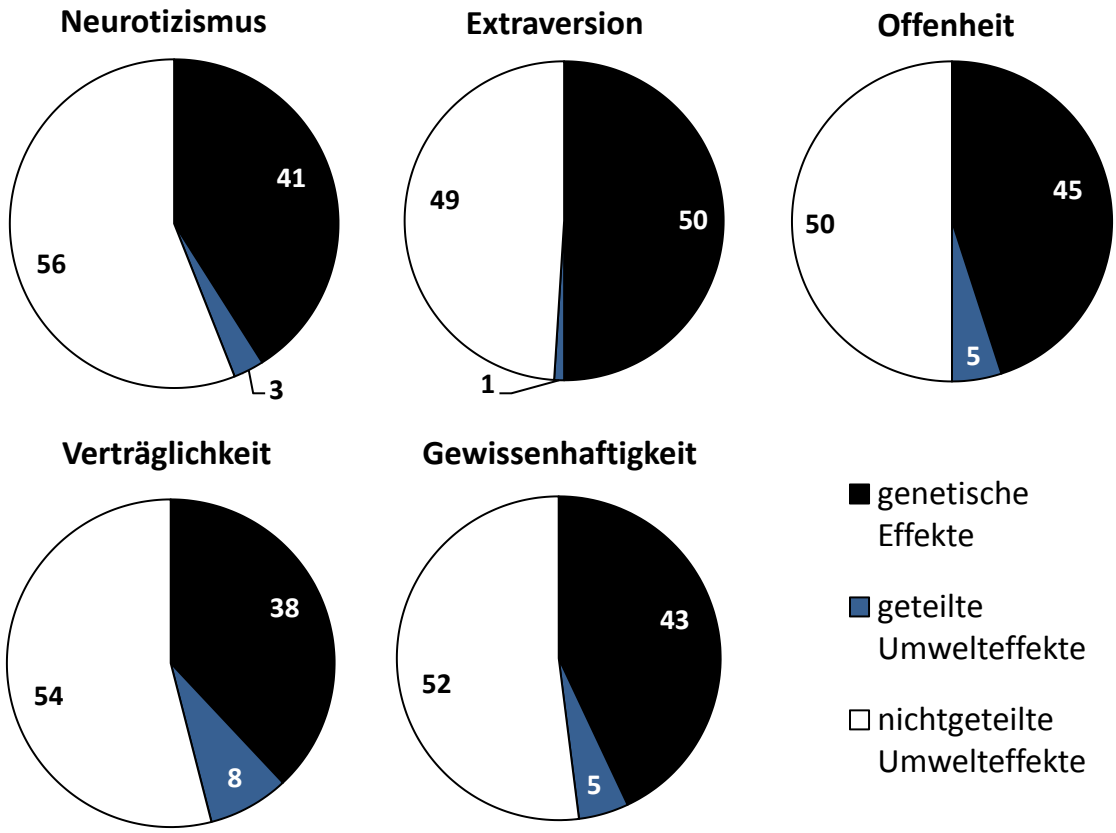
Die vorliegende Arbeit soll unter anderem den Nutzen dieses Kombinationsansatzes veranschaulichen und darstellen, in welcher Weise er das Wissen über die genetische und Umweltbeeinflussung von Persönlichkeitsunterschieden bereichern kann (Punkt 2). Schätzungen genetischer und Umwelteinflüsse sind zunächst populationsebenen-, alters-, zeit- und kulturspezifisch. Eine kulturübergreifende Perspektive (Punkt 2 und 3) und genetisch informative Längsschnittstudien zu verschiedenen Altersstufen erlauben weitere Aufschlüsse. Letzteres ermöglicht Einblick in die Faktoren der Kontinuität und Veränderung der Persönlichkeit über die Lebensspanne (Punkt 3). Ich werde meine eigenen Arbeiten in den aktuellen Stand der Forschung integrieren und das gesamte Befundmuster im Zusammenhang mit dynamischer und statistischer Anlage × Umwelt Interaktion diskutieren (Punkt 4).

2. Zur Validität von Anlage und Umwelt

Die jüngste Metaanalyse von Johnson, Vernon und Feiler (2008) analysierte 145 Studien zu den Big Five und zu diesen grob in Beziehung stehenden Merkmalen (z.B., Temperament, Facetten). Davon beschäftigten sich 110 Studien mit Persönlichkeit im Erwachsenenalter (77 reine Zwillingsstudien, 3 Adoptionsstudien und 30 Kombinationsdesigns). Diese Metaanalyse kam zu dem Ergebnis, dass etwa 40-50% der Varianz in den Big Five auf genetische Einflüsse (additive und nichtadditive genetische Effekte = *Erblichkeit im weiten Sinne*) zurückzuführen sind. Da sich systematische, zwischen den untersuchten Verwandten geteilte Umwelteffekte als vernachlässigbar gering herausgestellt haben, erklären individuelle Umwelteinflüsse (d.h. zwischen Familienmitgliedern nicht geteilte Umweltfaktoren) den größten Teil der nichtgenetischen Persönlichkeitsunterschiede (siehe Abb. 1; siehe auch Bouchard & Loehlin, 2001, und Riemann & Spinath, 2005).

Abbildung 1:

Genetische- und Umwelteffekte auf die Persönlichkeitsdimensionen des FFM in %



Bemerkung. Die Parameterschätzungen wurden aus der Metanalyse von Johnson et al. (2008) abgeleitet und standardisiert.

Die Metanalyse von Johnson et al. (2008) behandelte alle Studien nahezu gleichgewichtig. Jedoch nur eine Handvoll der betrachteten Studien berücksichtigte *unsystematische Messfehler*. Die Psychometrik der Persönlichkeit, wie fast jede Messung in einer Geistes- oder Naturwissenschaft, kann zufälligen Schwankungen aufgrund vielfältiger Ursachen (z.B. Justierung des Messinstruments, Rauschen, Aufmerksamkeit, Wachheit) unterliegen, welche die Zuverlässigkeit der Messungen (Reliabilität) herabsetzen. Solche zufälligen unsystematischen Effekte sind nicht korreliert zwischen verschiedenen Messzeitpunkten und Beurteilern (Steyer & Eid, 2001). Somit können solche Effekte die Merkmalskorrelationen zwischen Zwillingen, Adoptionsgeschwistern oder anderen Familienmitgliedern, welche sich selbst einschätzen oder durch andere beurteilt werden, nicht erhöhen und führen deshalb unweigerlich zu Überschätzungen von spezifischen Umwelteffekten, die zwischen Familienmitgliedern nicht geteilt sind. Daher überrascht es nicht, dass aus reliableren Verfahren zur Messung der Big Five oder aus Studien, welche die Messfehlervarianz kontrollierten, höhere Erblichkeitsschätzungen resultierten (etwa 50-60%; Loehlin, McCrae, Costa & John, 1998). Persönlichkeitsunterschiede aufgrund nichtgeteilter Umwelteffekte (in Abb. 1) sind demnach um den Messfehleranteil überschätzt und die anderen Varianzkomponenten werden anteilig unterschätzt.

Von den in der Metaanalyse betrachteten Studien zur Persönlichkeit basieren 82 lediglich auf einer Messmethode. Messungen der Persönlichkeit im Kindes- und Jugendalter basieren vornehmlich auf Elterneinschätzungen (siehe Punkt 3) und im Erwachsenenalter auf Selbsteinschätzungen. Wie eingangs erwähnt, können Messungen – vor allem Selbstberichte – neben zufälligen Schwankungen auch *systematischen Verzerrungen* (Methodeneffekte) unterliegen, zum Beispiel aufgrund von sozial erwünschten Antworten oder anderen systematischen Fehlern, wie die Tendenz zu einer durchschnittlichen Einschätzung, zur Über- und Unterschätzung oder zum unreflektierten Bejahen von Aussagen (Akquieszenz). Solche Beurteilerfehler (in der Folge als *Beurteilerbias* bezeichnet; Hoyt, 2000) können interindividuell unterschiedlich ausgeprägt sein und somit neben Mittelwertverschiebungen auch zu Verzerrungen der Varianz führen. Wenn man nun in Rechnung stellt, dass auch solche systematischen, methodeninhärenten Effekte die Messungen verzerren können, dann muss man annehmen, dass die Schätzungen sowohl des genetischen als auch des Umweltbeitrags in Selbstberichten aufgrund dieser Methodenartefakte verfälscht sind.

Nur wenige verhaltensgenetische Untersuchungen gingen bisher über die Selbsteinschätzungsmethode zur Bestimmung der Persönlichkeit im Erwachsenenalter hinaus und haben zusätzlich gute Bekannte (vornehmlich Ehe-/Lebenspartner) befragt (z.B. Riemann, Angleitner & Strelau, 1997; **STUDIE I**). Diese Studien berichteten von genetischen und Umwelteffekten auf bekanntenberichtete Persönlichkeitsdimensionen in vergleichbarer Größe zu den Selbstberichtstudien: Etwa 30-50% der Varianz konnte auf genetische Einflüsse und der Restvarianzanteil auf nichtgeteilte Umwelteffekte zurückgeführt werden, da geteilte Umwelteffekte vernachlässigbar gering waren. Eine einzige Zwillingsstudie zur Persönlichkeit im Erwachsenenalter, die videobasierte Beobachtereinschätzungen analysierte (Borkenau, Riemann, Angleitner

& Spinath, 2001), berichtete von Schätzungen geteilter Umwelteffekte (0-30%). Dass sich in Zwillingsstudien auf der Basis von Selbstberichten keine bedeutsamen Effekte geteilter Umwelt finden lassen, kann auf *Kontrasteffekte* zwischen den Selbsteinschätzungen der Zwillinge hindeuten. Da der jeweilige Geschwisterpart eines Zwillingspaars als Abgrenzungsreferenz für die Selbsteinschätzung dienen kann, können stärkere Unterschiede in den Selbstberichten die Folge sein. Andererseits können *Kontrasteffekte* die Parallelität der Ergebnisse aus unabhängigen Bekanntenschätzungen (unterschiedliche Beurteiler für jeden Zwilling eines Paares) zu denen aus Selbstberichten nicht erklären. Da aus Bekanntenberichten keine signifikanten geteilten Umwelteffekte resultieren (Riemann et al., 1997; **STUDIE I**), könnte auch eine andere Erklärung für geteilte Umwelteffekte in Beobachtungsstudien herhalten. Da jedes Zwillingspaar zur selben Zeit, am selben Tag in derselben Laborsituation untersucht wurde (Spinath, Angleitner, Borkenau, Riemann & Wolf, 2002), ist es plausibel, dass situative geteilte Umwelteinflüsse, welchen die Beobachter als auch die Zwillinge zeitgleich unterworfen waren, die Zwillingskorrelationen artifiziell erhöhten und so zur Schätzungen geteilter Umwelteffekte führten.

Dass neben den Selbstberichten aus anderen Methoden zur Persönlichkeitsmessung gleiche Schätzungen für genetische und Umwelteffekte resultierten, ist ein erster Hinweis auf ihre Validität. Jedoch kann nicht ausgeschlossen werden, dass auch gute Bekannte oder gar Beobachter, seien sie noch so gut geschult, systematischen Beurteilerfehlern unterliegen. Auch Bekannte können sozial erwünscht antworten und sich zum Beispiel hinsichtlich ihrer Strenge- versus Milde-Tendenz unterscheiden. Das erzeugt nicht-merkmalsvalide Varianz, was die Schätzungen genetischer und Umwelteffekte verfälscht. Der einfachste Weg zur Kontrolle solcher Effekte erfordert die Bestimmung genetischer und Umwelteffekte auf *konsensuell valide* Varianz (Kenny, 1991). Diese beruht auf der Kovarianz (bzw. Korrelation) zwischen (mindestens) zwei Beurteilern eines Beurteilungsobjektes (Target). Dieser Ansatz erfordert zwei wichtige Voraussetzungen. Die erste Voraussetzung ist die *Unabhängigkeit der Beurteiler* (Campbell & Fiske, 1959). Dabei ist es bei der Untersuchung von Verwandten – wie z.B. bei einer Zwillingsstudie – erforderlich, (mindestens) zwei unabhängige Beurteiler für jeden der beiden Verwandten zu haben. Für die Zwillingsstudie bedeutet das, dass mindestens zwei Beurteiler nur einen der beiden Zwillingsgeschwister einschätzen und zwei weitere den anderen. Beurteiler-inhärente Fehler sollten zwischen den unabhängigen Beurteilern innerhalb eines Targets und zwischen den Zwillingsgeschwistern nicht korrelieren. Die Varianz zwischen den Targets und die Korrelationen innerhalb der Zwillingspaare über die verschiedenen Beurteiler können somit nicht auf *Beurteilerbias* zurückzuführen sein, es sei denn unabhängige Beurteiler nutzen die gleichen „inakkuraten“ Stereotypen für ihr Urteil (Letzring, Wells & Funder, 2006). Verwandte können sich im Mittel nicht nur psychisch sondern auch physisch ähnlicher als zufällig aus einer Population herausgegriffene Personen sein. Aber mit Sicherheit sind sich eineiige Zwillinge (in der Folge auch mit EZ abgekürzt) ähnlicher als zweieiige (in der Folge auch mit ZZ abgekürzt). Wenn also die Urteile auf persönlichkeitsirrelevanten Aspekten (z.B. Brille, geschorene Haare, Tattoos, Kleidung) basieren, können Verwandtenkorrelationen höher sein, bzw. die Zwillings-

korrelationen für EZ höher ausfallen als für ZZ, als es aufgrund des tatsächlichen Merkmals der Fall wäre. Eine persönlichkeitsirrelevante Erblichkeitsschätzung wäre die Folge. Daher ist es neben der *Unabhängigkeit* als eine zweite Voraussetzung wichtig, möglichst *gut informierte Bekannte* als Beurteiler heranzuziehen, denn mit zunehmendem Grad der Bekanntschaft zwischen Target und Beurteiler reduzieren sich Stereotypeneinflüsse (Funder, Kolar & Blackman, 1995; Letzring et al., 2006). Neben den angesprochenen und wichtigsten Voraussetzungen für eine möglichst hohe Akkuratheit der Einschätzungen werden in der Literatur noch weitere diskutiert (siehe hierzu Funder, 1995; Kenny, 1994).

Die Korrelationen zwischen Bekanntenschätzungen als Maß für die *Konsensuelle Validität* (r_{BB}) rangieren um r_{BB} = .50 für die Persönlichkeitsmerkmale des FFM (McCrae, Costa, Martin, Oryol, Rukavishnikov et al., 2004), was einen bedeutsamen geteilten Varianzanteil ausmacht. Die wenigen Studien, die sich nun mit der Frage nach den genetischen und Umweltbeiträgen beschäftigten (Borkenau et al., 2001; Riemann et al., 1997), konnten etwa 60-80% der *konsensuell validen* Varianz in den fünf Persönlichkeitsdimensionen auf genetische Effekte zurückführen, wobei der restliche Varianzanteil nichtgeteilte Umwelteffekte ausmachte. Betrachtet man also „wahre“ Persönlichkeitsvarianz im Sinne *Konsensueller Validität*, so tritt der genetische Beitrag deutlicher in Erscheinung als es Studien mit nur einer Messmethode zeigen können.

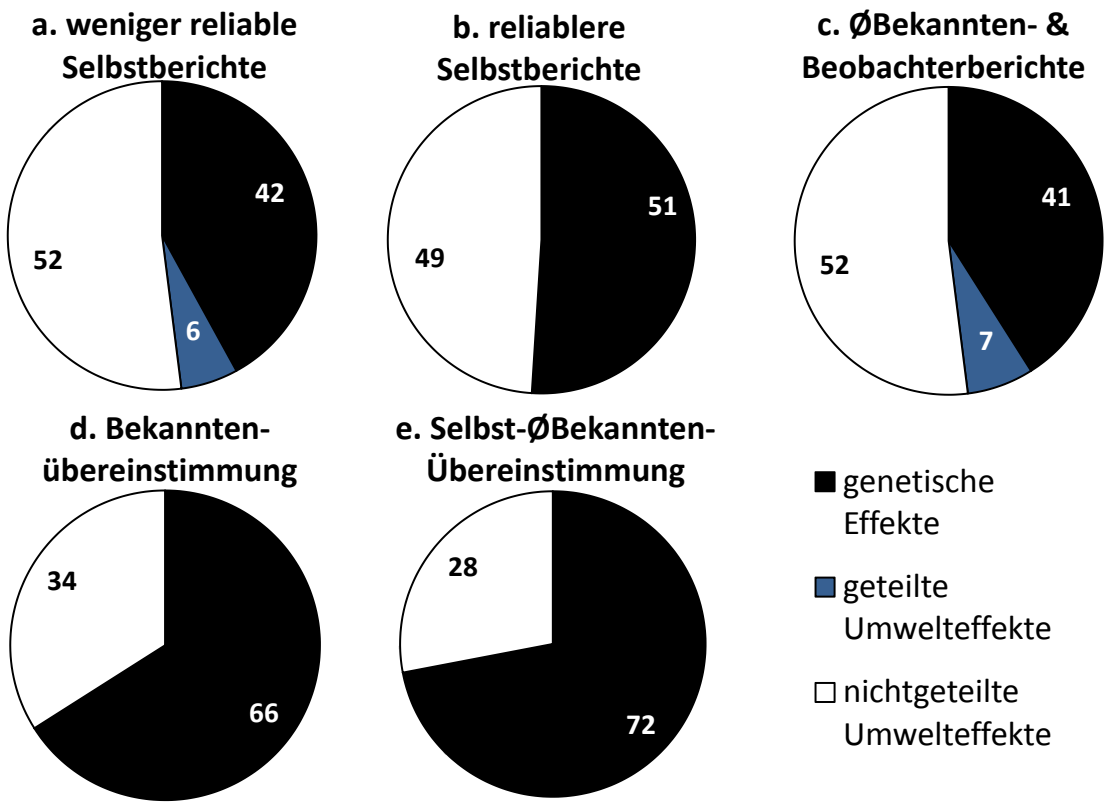
Ein hoher Konsensus zwischen den Bekannten impliziert aber noch keine hohe *Selbst-Bekannten-Übereinstimmung*. Dies ist genau dann der Fall, wenn die Bekannten nicht genug oder keine „guten“ Informationen haben, bzw. keine „guten“ Bekannten urteilen oder keine „guten“ Targets zur Verfügung stehen (siehe Funder, 2001, für eine Ausführung des Begriffes „gut“). In diesem Fall wäre das Bekanntenurteil stark fehlerbehaftet und somit auch der Konsensus. Allerdings könnte ebenso das Selbsturteil stark fehlerbehaftet sein, nämlich genau dann, wenn der Selbsturteiler nicht Willens (soziale Erwünschtheit, falsche Selbstdarstellung) oder unfähig ist (Narzissmus, Selbsttäuschung), akkurate Urteile über sich selbst abzugeben (Hofstee, 1994). In diesem Fall wäre die *Selbst-Bekannten-Übereinstimmung* stark verzerrt. Wir haben also ein Kriterienproblem, denn wir kennen die Gold-Standard-Methode nicht. Dies ist aber insofern unproblematisch, da eine Messmethode als eine Überprüfung der jeweils anderen gelten und die Höhe der *Selbst-Bekannten-Übereinstimmung* für sich genommen interessante Implikationen nach sich ziehen kann (Funder, 2001).

Selbst-Bekannten-Korrelationen (r_{SB}) im Hinblick auf Persönlichkeitseinschätzungen stehen in ihrer Höhe den Korrelationen zwischen Bekannten in nichts nach (r_{SB} = .50; McCrae et al., 2004). Eher können diese Maße der *Konsensuellen Validität* noch höher ausfallen, wenn man Korrelationen zwischen Selbstberichten und gemittelten Bekanntenschätzungen betrachtet (**STUDIE II**), da durch die Mittelung von Bekanntenurteilen schon spezifische *Beurteilerbias* ausgemittelt werden können (Hofstee, 1994). Insofern geht diese Art der Bestimmung der *Konstruktvalidität* über die Korrelation zwischen Bekannten hinaus. Riemann et al. (1997) berichteten von etwa 65-80% genetischer Einflüsse auf den Kovarianzanteil zwischen Selbsteinschät-

zungen und gemittelten Bekanntenberichten, während der übrige Anteil auf nichtgeteilte Umwelteffekte zurückzuführen war.

Abbildung 2:

Genetische- und Umwelteffekte auf Persönlichkeitsdimensionen des FFM in %



Bemerkung. Die Parameterschätzungen beziehen sich nur auf Studien zum FFM und wurden über die fünf Persönlichkeitsdimensionen gemittelt sowie an der Stichprobengröße der ausgewählten Studien gewichtet. Ø = gemittelt.

Abbildung 2 veranschaulicht den Unterschied der Befundmuster zwischen den zahlreichen Selbstberichtstudien und den bisher betrachteten Validierungsansätzen. Abbildung 2a zeigt die Ergebnisse einer ersten Metaanalyse von Loehlin (1992). Er unterteilte die bis dato veröffentlichten genetisch informativen Selbstberichtstudien zu den verschiedenen Persönlichkeitsmerkmalen nach den Big Five. Da diese Studien bis auf eine Ausnahme (Bergeman Chipuer, Plomin, Pederson, McClearn et al., 1993) noch keine Instrumente zur direkten Messung der Big Five (bzw. dem FFM) verwenden konnten, dient Abbildung 2a zur Veranschaulichung von Schätzungen genetischer und Umweltvarianzquellen auf der Basis unzuverlässigerer Messungen der Big Five. Abbildung 2b subsummiert neuere verhaltensgenetische Studien, die reliablere Verfahren (z.B. NEO-FFI oder NEO-PI-R; Borkenau & Ostendorf, 1993; Costa & McCrae, 1992; Ostendorf & Angleitner, 2004) zur direkten Messung der fünf Persönlichkeits-

dimensionen nutzten (Jang, Livesley & Vernon, 1996; **STUDIE I**; Shikishima, Ando, Ono, Toda & Yoshimura, 2006; Riemann et al., 1997) und teilweise sogar um die Messfehlervarianz kontrollierten (Loehlin et al., 1998). Diese Studien umfassen nordamerikanische, mitteleuropäische und ostasiatische Stichproben. Abbildung 2c beschreibt die Befundlage zu gemittelten Bekanntenberichten (Borkenau et al., 2001; **STUDIE I**) und videobasierten Beobachtungen (Borkenau et al., 2001). Die Befundmuster aus Abbildung 2d und 2e stammen aus der leider bisher einzigen Zwillingsstudie von Riemann und Kollegen (1997), die genetische und Umwelteffekte auf latente *konsensuell valide* Persönlichkeitsdimensionen untersuchte. Es gibt jüngere Studien, welche Bekanntenübereinstimmung (Borkenau et al., 2001) und Übereinstimmungen aus Selbst- und gemittelten Bekanntenberichten (**STUDIE II**) zur Validierung genetischer und Umwelteffekte wieder aufgegriffen haben und dieses Muster ziemlich genau replizieren konnten. Die Stichproben überlappen sich jedoch teilweise mit der von Riemann et al. (1997). Daher wartet diese Studie noch auf ihre Replikation über den deutsch-polnischen Sprachraum hinaus. Insgesamt veranschaulicht Abbildung 2 eindrucksvoll, dass Erblichkeitsschätzungen für Persönlichkeitsmerkmale zunehmen, je zuverlässiger und valider diese gemessen werden.

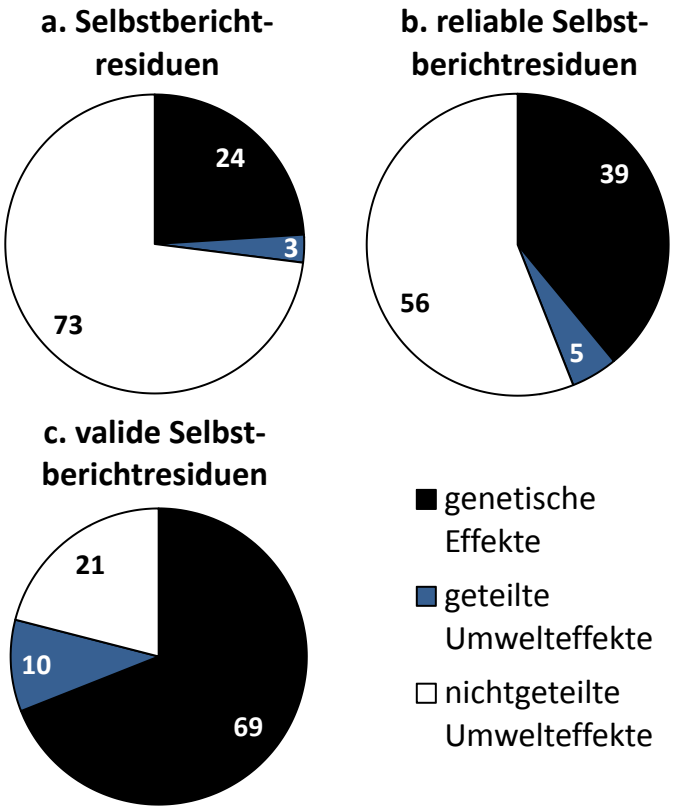
Dieses Muster ist nicht allein auf die fünf Persönlichkeitsdomänen des FFM beschränkt. Es zeigt sich auch für spezifischere Persönlichkeitsmerkmale, wie für die Facetten der jeweiligen Domänen (z.B. Ängstlichkeit, Depressivität, Geselligkeit, Aktivität, Offenheit für Gefühle, Bescheidenheit, Ordnungsliebe, etc.). So konnten Jang, McCrae, Angleitner, Riemann und Livesley (1998) zeigen, dass auch spezifische Varianzanteile der Persönlichkeitsfacetten (bereinigt um die Domänenkomponenten) genetische Einflüsse aufweisen (Abbildung 3a). Darüber hinaus konnte gezeigt werden, dass die Höhe der Erblichkeitsanteile der Facetten denen der Domänen in nichts nachsteht, wenn um Messfehlervarianz korrigiert wird (Abbildung 3b). Eine jüngst durchgeführte Studie (**STUDIE II**) konnte die Befunde von Jang et al. (1998) replizieren und auf die konsensuelle Varianz aus Selbst- und gemittelten Bekanntenberichten erweitern. Wie auf der Ebene der Persönlichkeitsdomänen offenbaren sich höhere Erblichkeitsschätzungen (Abbildung 3c). Erwähnenswert ist auch die Zunahme des Beitrags geteilter Umwelteffekte auf verhaltensnähere Persönlichkeitsfacetten mit der Zunahme validerer Merkmalsvarianz.

Je akkurater also ein Persönlichkeitsmerkmal gemessen wird, desto stärker erscheint es im Licht genetischer Beeinflussung. Dies zeigt sich sogar auf verschiedenen Ebenen einer hierarchischen Persönlichkeitsstruktur (Jang et al., 1998; **STUDIE II**). Daraus resultieren bedeutende Implikationen für die genetische Architektur des FFM. Genetische Faktoren wirken sowohl facettenübergreifend (engl. *top-down*) als auch facettenspezifisch (engl. *bottom-up*). Auch implizieren diese Befunde Interpretationen für nichtgeteilte Umweltvarianz. Diese erscheinen verstärkter im Licht von Messfehlern und methodenspezifischen Effekten, sprich *Beurteilerbias* und *Methodenartefakten*. Trotzdem bleibt ein bedeutender Anteil konsensuell valider Varianz in Persönlichkeitseinschätzungen beeinflusst durch valide Umwelteffekte, die in erster Linie

individuell spezifischer Natur zu sein und sowohl facettenübergreifend (*top-down*) wie auch facettenspezifisch (*bottom-up*) zu wirken scheinen (**STUDIE II**).

Abbildung 3:

Genetische- und Umwelteffekte auf Persönlichkeitsfacettenspezifität des FFM in %



Bemerkung. Die Parameterschätzungen beziehen sich nur auf Studien zum FFM und wurden über die 30 Persönlichkeitsfacetten gemittelt.

Nun heißt aber *Konsensus* zwischen gut informierten Bekannten oder zwischen Selbst- und Bekanntenschätzungen nicht gleich *Akkuratheit*, auch wenn dieser Begriff häufig als Synonym für *Selbst-Bekannten-Übereinstimmung* verwendet wurde (Kenny, 1991; Funder, 1995). Dieser Einwand ist in zweierlei Hinsicht zu sehen: (1) Konsensuell valide Varianz muss nicht ausschließlich „wahre“ Merkmalsvarianz reflektieren und (2) nicht-konsensuell-valide Varianz muss nicht gleich „Junk“-Varianz im Sinne von Methodenartefakten oder *Beurteilerbias* bedeuten (**STUDIE II**).

Ist konsensuell valide Varianz valide? Die meisten Messungen der Persönlichkeit, seien es nun Selbst- oder Bekanntenschätzungen, basieren auf Erfahrungs- oder Adjektivbeschreibungen. Insofern könnte der Konsensus auf die Ähnlichkeit der Messmethode (*Instrumentenbias*) zurückzuführen sein. Auch könnten Korrelationen nur aufgrund ähnlicher impliziter Persönlichkeitstheorien (*IPT-Bias*) der Beurteiler

zustande kommen (Mischel, 1968), da die Beurteiler gelernt haben, welche Merkmale häufiger zusammen auftreten. Diese Verzerrungen können allerdings nicht dazu beitragen, dass die Bekannten- und die Selbst-Bekannten-Korrelationen mit dem Grad der Bekanntschaft zunehmen (Borkenau, 1992; Kenny, Albright, Malloy & Kashy, 1994), dass diese auch zwischen impliziten und expliziten Persönlichkeitsmessungen konvergieren (Hofmann, Geschwender & Schmitt, 2009) und – als ein wesentlicher Beitrag der Verhaltensgenetik – zwischen EZ höher ausfallen als zwischen ZZ (McCrae, Jang, Livesley, Riemann & Angleitner, 2001; **STUDIE II**).

Der bewusstseinsfähige Mensch jedoch, vermag sich nicht nur besser (oder einfach anders) in einem Fragebogen, in einem Interview oder in einer Beobachtungssituation darzustellen, als er es eigentlich ist, sondern auch im alltäglichen Leben (Goffman, 2009). Je mehr der Mensch fähig ist, sich entgegen seiner Natur über vielfältige Situationen und Kontexte zu präsentieren, desto eher kann dieses Bild sogar in einem guten Bekannten repräsentiert sein. Man würde also nicht das zu erfassen gewünschte Persönlichkeitsmerkmal, sondern ein durch das Target gewünschte Merkmal erfassen. In Anlehnung an Paulus & John's *Impression Management* (1998) als bewusst forcierte Selbstdarstellung bezeichne ich diesen Fehler als *IM-Bias*. Dieser *IM-Bias* kann genetisch beeinflusst sein und somit höhere Selbst-Bekannten-Korrelationen zwischen EZ versus ZZ nach sich ziehen.[3] Zu dem Grad, zu dem dieser *IM-Bias* die konsensuell valide Varianz reflektiert, sind die Schätzungen genetischer und Umweltvarianzquellen des eigentlichen (zu betrachten gewünschten) Merkmals verzerrt. Ein *IM-Bias* kann zwar erklären, dass die Informationsquantität (Dauer der Bekanntschaft), nicht aber, dass auch die Informationsqualität (persönlichkeitsrelevante Informationen) die konsensuelle Validität erhöht (Letzring et al., 2006). Auch kann ein *IM-Bias* nicht so einfach erklären, warum der genetische Beitrag für konsensuell valide Varianz größer ausfällt als innerhalb von Selbst-, Bekannten- und Beobachtereinschätzungen (**STUDIE II**; Borkenau et al., 2001). Es sei denn, dieser Bias selbst ist ein stark genetisch beeinflusstes Merkmal. Somit können *Instrumentenbias*, *IPT-Bias* und *IM-Bias* nur schwache Erklärungen für *Konsensus* darstellen. Vielmehr liefern diese Effekte eine Erklärung für nicht-konsensuell-valide Komponenten.

Ist nicht-konsensuell-valide Varianz nicht valide? Aus verschiedenen Beurteilern (z.B. Selbstbeurteiler, Ehe-/Lebenspartner, Arbeitskollege, Vereinssportteamkollege) resultieren automatisch verschiedene *Perspektiven* (z.B. Selbstperspektive, Außenperspektive) und *Kontexte* der Beurteilung (z.B. Familie, Arbeit, Verein). In dem Maße, in dem diese Perspektiven und Kontexte spezifische merkmalsrelevante Informationen zur Verfügung stellen, trägt jeder gute Bekannte valide Informationen (Kraemer, Measelle, Ablow, Essex, Boyce & Kupfer, 2003). Es ist denkbar, dass kontextuelle Umstände zum Ausdruck eines Merkmals beeinflussen. So fanden Letzring

[3] Natürlich ist das nur unter der Annahme gegeben, in der die Natur der Natur Einhalt gebieten kann. Das könnte zum Beispiel dann so sein, wenn ein genetisch beeinflusstes Merkmal, z.B. Soziale Kompetenz, dazu beiträgt, dass sich die Person über verschiedene soziale Situationen und Kontexte hinweg entgegen ihrer veranlagten Unverträglichkeit verträglich verhält.

und Kollegen (2006) mehr Informationsqualität zur Persönlichkeitsbeurteilung in „freien“ versus in „restringierten“ Situationen zum gegenseitigen Kennenlernen. Sogar gut informierte Lebenspartner teilen in der Regel nicht alle Kontexte und können somit nur begrenzte Informationen über die Merkmalsausprägung ihrer Partner haben.

Es ist ebenso denkbar, dass die Persönlichkeitseinschätzungen der Selbstbeurteiler und Bekanntenbeurteiler auf teilweise verschiedenen Informationen basieren. So könnten zum Beispiel Selbstbeurteiler eher interne Aspekte der Persönlichkeit zu ihrer Einschätzung heranziehen (z.B. Gedanken, Gefühle, Selbstkonzept), während Bekanntenbeurteiler sich eher auf das veräußerte Verhalten oder soziale Konsequenzen des Verhaltens beziehen (**STUDIE II**). Anderson (1984) fand höhere Selbst-Andere-Übereinstimmungen, wenn Fremdurteiler einem Interview beisaßen, in dem Gedanken und Gefühle geäußert wurden, im Vergleich zu einem Interview über Hobbies und Aktivitäten. Eine aktuellere Studie konnte zeigen, dass spezifische Selbst- und Bekanntenberichtkomponenten in Einschätzungen des alltäglichen Verhaltens spezifische Verhaltensaspekte, erfasst über Act-Frequencies (Buss & Craik, 1983), vorhersagen können (Vazire & Mehl, 2008). Solche Studien unterstützen die inkrementelle Validität der Spezifität von Selbst- und Bekanntenberichten jenseits von Methodenartefakten.

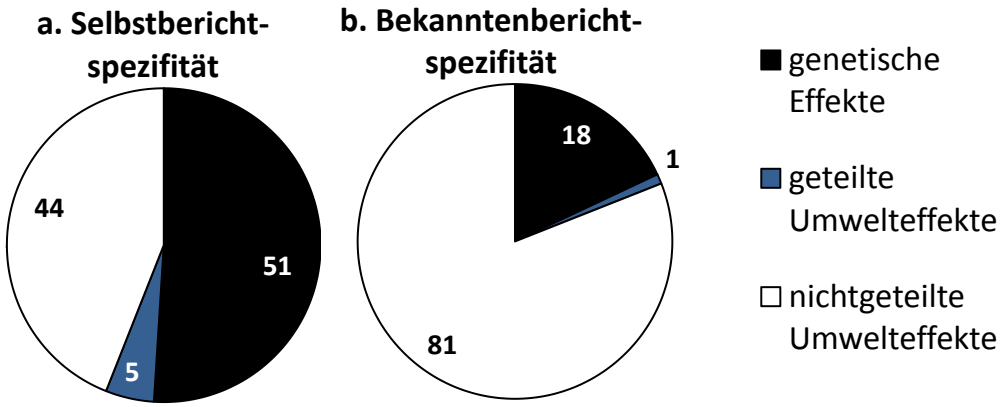
Verhaltensgenetische Studien erlauben darüber hinaus Schätzungen genetischer und Umweltfaktoren auf diese Komponenten und ermöglichen somit Implikationen zur Interpretation dieser. Zwei kürzlich durchgeführte Zwillingsstudien schätzten genetische und Umwelteffekte auf Selbst- und Bekanntenberichtspezifität. Auf der Ebene der Persönlichkeitsdomänen fanden sich signifikante genetische Effekte auf die Selbstberichtspezifität über alle fünf Persönlichkeitsfaktoren und auf die Bekanntenberichtspezifität für Neurotizismus und Verträglichkeit (**STUDIE II**; siehe Abb. 4). Auf einer domänenübergreifenden Ebene, auf der Selbst- und Bekanntenberichtspezifität in zwei Faktoren im Sinne von Digman’s α und β unterteilt wurden (Digman, 1997), fanden sich substantielle genetische Effekte für die selbstberichtspezifischen α und β, jedoch nur moderate genetische Effekte für den bekanntenberichtspezifischen α-Faktor (Riemann & Kandler, in Druck). Darüber hinaus konnten meine Kollegen und ich in einer anderen Studie (**STUDIE I**) zeigen, dass der selbstberichtspezifische genetische α-Faktor neben einem konsensuell validen genetischen Faktor zusätzlich genetische Varianz in den Einschätzungen familiärer Unterstützung erklärte.

Sicherlich kann ein Selbstberichtbias genetisch beeinflusst sein (z.B. Narzissmus; Selbsttäuschung, Impression Management; Vernon, Villani, Vickers, & Harris, 2008). Für einen Bekanntenberichtbias kann diese Erklärung allerdings nicht gelten (siehe hierzu **STUDIE I**; **STUDIE II**; McCrae, Yamagata, Jang, Riemann, Ando et al., 2008; Riemann & Kandler, in Druck). Der Unterschied in der Höhe der genetischen Beeinflussung zwischen Selbst- und Bekanntenberichtspezifität (siehe Abb. 4) kann auf einen genetisch beeinflussten Selbstberichtbias zurückzuführen sein. Die Tatsache jedoch, dass auch nicht zu vernachlässigende Anteile der Bekanntenberichtsspezifität genetisch beeinflusst sind, deutet auf eine Erklärung im Sinne von Merkmalssubstanz

hin (siehe **STUDIE II** für eine detaillierte Diskussion). Diese verhaltensgenetischen Befunde untermauern die schon erwähnten Befunde aus experimentellen Studien (Anderson, 1984; Letzring et al., 2006; Hoffmann et al., 2009) und aus Studien zur Untersuchungen der prädiktiven Validität von Beurteilerperspektiven (Vazire & Mehl, 2008). Zusammengenommen ergibt sich ein Bild, welches die Konstruktvalidität von Beurteilerperspektiven erhärtet.

Abbildung 4:

Genetische- und Umwelteffekte auf Selbst- und Bekanntenberichtspezifität in %



Bemerkung. Die Parameterschätzungen beziehen sich auf die Befunde von **STUDIE II** und wurden über die 5 Persönlichkeitsdomänen gemittelt.

Zusammenfassung. Je akkurater ein Persönlichkeitsmerkmal erfasst wird, desto deutlicher offenbart es seine veranlagte Natur (etwa 2/3). Daneben spielen individuell spezifische Umwelteffekte eine wichtige Rolle (etwa 1/3). Konsensus aus verschiedenen Beurteilerperspektiven vermag die genetische Architektur der hierarchischen Persönlichkeitsstruktur des FFM klarer zu reflektieren, aber ein ausschließlicher Blick auf die konsensuelle Validität liefert wahrscheinlich ein unvollständiges Bild der komplexen Persönlichkeit.

3. Zur Kontinuität von Anlage und Umwelt

Beschäftigt man sich mit den genetischen und umweltbedingten Faktoren der Persönlichkeit, reicht es nicht, die absoluten Anteile und deren Validität zu betrachten. Einen bedeutsamen Anteil von Umweltfaktoren zu identifizieren, heißt nicht gleich Beeinflussbarkeit im Sinne von fortwährenden Lernprozessen und Veränderungen. Genauso wenig sind genetische Faktoren mit Kontinuität gleichzusetzen. Sicherlich verändert sich die DNA nicht, aber genetische Faktoren (Genexpression) können es, indem sie zur Reifung der Persönlichkeit betragen. Es stellt sich also die Frage nach der Höhe des genetischen und umweltbedingten Beitrags zur Kontinuität und zur Veränderung von Persönlichkeitseigenschaften.

Aus phänotypischen Studien ist hinlänglich bekannt, dass der durchschnittliche Mensch im Alter emotional stabiler, verträglicher und gewissenhafter wird (McCrae, Costa, Ostendorf, Angleitner, Hrebickova et al., 2000; Roberts, Walton & Viechtbauer, 2006). Das Bild fällt für Extraversion und Offenheit etwas differenzierter aus. Verschiedene Facetten von Extraversion zeigen gegenläufige Effekte. Nach einer anfänglichen Zunahme von Offenheit stellt sich womöglich eine Abnahme im mittleren Erwachsenenalter ein. Neben normativen Veränderungen werden die Persönlichkeitsunterschiede zwischen Menschen mit dem Alter immer stabiler, was die Zunahme der Kontinuitätskoeffizienten[4] (bei gleichbleibenden Messzeitintervall) von r_P = .30 in den ersten Lebensjahren bis über r_P = .70 bei über Fünfzigjährigen zeigt (Roberts & DelVecchio, 2000). Dass die Kontinuität jedoch mit der Größe des Zeitintervalls abnimmt, ganz unabhängig vom Altersbereich, verweist auf Persönlichkeitsveränderung (Fraley & Roberts, 2005) bis weit in das Erwachsenenalter hinein (Terracciano, Costa & McCrae, 2006). Wie jedes Merkmal in einem Menschen sind demnach auch Persönlichkeitseigenschaften Entwicklungsprozessen unterworfen. Die Frage, welche Rolle dabei die genetische Reifung und soziale Erfahrungen spielen, hat einige Theorien nach sich gezogen.

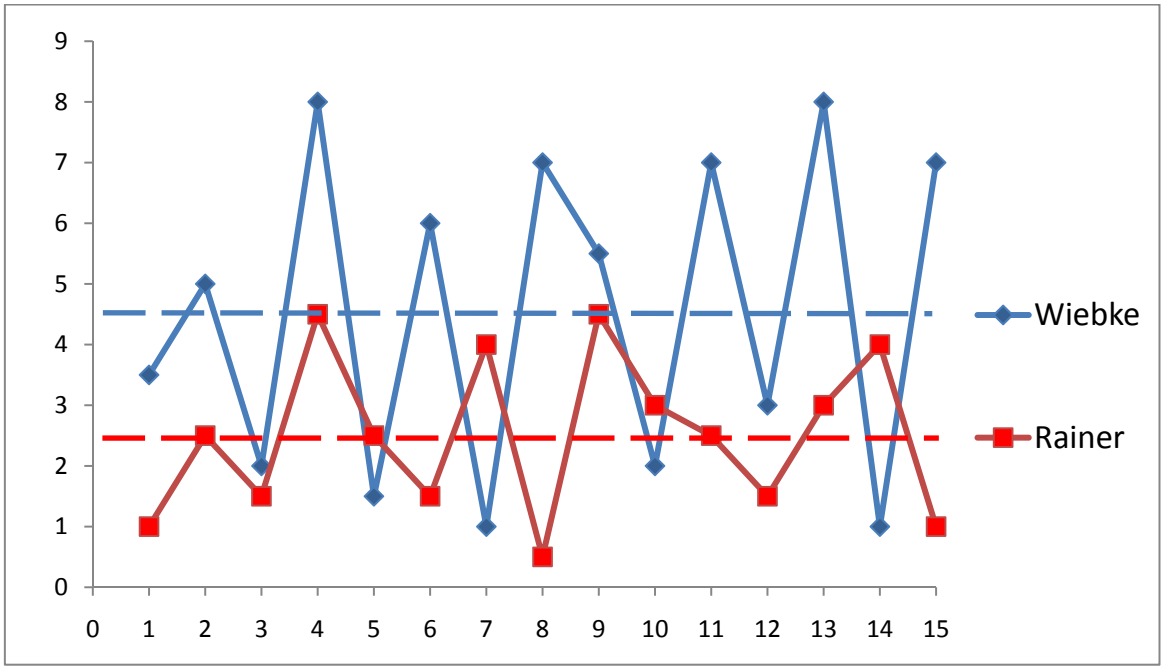
Carey (2002) sieht die Ursache der Persönlichkeitskontinuität in individuellen genetisch verankerten Set-Points, aufgrund derer sich Menschen unterscheiden. Dass die Stabilität über die Zeit nicht r_P = 1.0 beträgt, liege an „kurzzeitigen“ – Stunden, Tage, Wochen, oder sogar Monate andauernden – umweltabhängigen Fluktuationen, die das Individuum (im wahrsten Sinne der Set-Point Theorie) aus der Bahn werfen können. Langfristig regrediere das Individuum immer wieder auf seinen individuellen genetischen Set-Point. So würde also eine gemessene Persönlichkeitseigenschaft eines Individuums (z.B. Wiebke oder Rainer, siehe Abb. 5) zu einem Zeitpunkt in Abhängigkeit der gegebenen Umweltbedingungen entweder über oder unter dem individuellen Set-Point zu finden sein. Carey (2002) begründet seine Theorie mit dem Befund, dass Zwillingskorrelationen zwischen zwei Zeitpunkten nicht signifikant geringer ausfallen als innerhalb eines Messzeitpunktes (McGue, Bacon & Lykken, 1993). Auch wird seine Theorie dadurch gestützt, dass systematische und individuelle Umweltvariablen wie elterliche Unterstützung und Familienstruktur offenbar keine nachhalti-

[4] r_P = phanotypische Korrelation über die Zeit nicht um Messfehler korrigiert.

gen Einflüsse auf die Persönlichkeit im Erwachsenenalter ausüben (**STUDIE I**; Krueger, Markon & Bouchard, 2003).

Abbildung 5:

Set-Point-Theorie



Bemerkung. X-Achse = Messzeitpunkt, Y-Achse = Merkmalsausprägung, gestrichelte Line markiert den individuellen genetischen Set-Point von Rainer und Wiebke.

McCrae und Kollegen (2000) formulierten eine ähnliche Theorie. Persönlichkeitsdimensionen sehen sie selbst als stabile genetische Faktoren (*Basistendenzen*) und führen Abweichung der Kontinuität von r_P = 1.0 ebenfalls auf kurzzeitige Umwelteinflüsse, aber auch auf systematische und unsystematische Messfehler zurück. Dass die Kontrolle von Artefakten größere Erblichkeitsschätzungen (Jang et al., 1998; Loehlin et al., 1998; Riemann et al., 1997) ergab, stützt diese Position. Darüber hinaus führen sie die systematischen mittleren Veränderungen in den Persönlichkeitseigenschaften über das Lebensalter auf genetische Reifung zurück. Dies begründeten sie darin, dass die gleichen Entwicklungstrends sprach- und kulturübergreifend gefunden werden konnten (McCrae et al., 2000).[5]

[5] Eine revidierte Fassung dieser Theorie (McCrae & Costa, 2008) erlaubt, dass sich Umwelteffekte durchaus verankern und zeitlich stabil manifestieren können, jedoch nur vermittelt über biologische Mechanismen, die wiederum stark genetisch beeinflusst sind (z.B. traumatische Erlebnisse oder Drogen beeinflussen Neurotransmitter-Aktivität).

Caspi, Roberts und Shiner (2005) beleuchteten in Anbetracht der bis dato gegebenen Forschung drei Prinzipien der Persönlichkeitsentwicklung. Mit dem *Prinzip der Reifung* beschrieben sie die mittleren Veränderungen in den Persönlichkeitseigenschaften in eine gesellschaftlich funktionalere Richtung (emotional stabiler, verträglicher, gewissenhafter) als eine dem Menschen gegebene Kapazität, als Mitglied der Gesellschaft *produktiv* und *involviert* zu sein. Dabei legten sie sich nicht fest, ob diese Kapazität genetisch oder erfahrungsabhängig ist. Mit dem *Prinzip der kumulativen Kontinuität* fokussierten sie auf die stetige Zunahme der Stabilitätskoeffizienten über die Lebensspanne bei gleichbleibenden Messzeitintervallen. Die Autoren diskutierten das im Zusammenhang mit genetischen Faktoren, Nischenbildungsprozessen, Identitätsausdifferenzierung und normativen Entwicklungsübergängen. Das *Prinzip der Korrespondenz* kombiniert den Aspekt der Abnahme der Stabilitätskoeffizienten bei zunehmenden Messzeitintervallen mit den anderen Prinzipien, indem es Persönlichkeitsentwicklung als ein Koexistieren von Kontinuität und Veränderung beschreibt (Roberst & Caspi, 2003). Dies diskutierten die Autoren als einen dynamischen Person × Umwelt Interaktionismus[6]. Das Individuum wählt sich Umwelten entsprechend seiner Persönlichkeitseigenschaften (*soziale Selektion*). Diese Umwelten wiederum stellen Erfahrungen bereit, welche die Funktionalität und gegebenenfalls die Entwicklung der Persönlichkeit beeinflussen (*sozialer Einfluss*). Aus diesen Prinzipien resultiert die Annahme eines kontinuierlichen Zusammen- und Wechselwirkens zwischen genetischen und Umweltfaktoren als unterschiedliche Prozessebenen in der Persönlichkeitsentwicklung. Genetische und Umweltfaktoren beeinflussen somit sowohl die Stabilität als auch die Veränderung.

Die Überprüfung solcher Theorien erfordert Längsschnittstudien. Obwohl eine große Anzahl von genetisch informativen Studien zur Persönlichkeit mit einer großen Bandbreite hinsichtlich des Alters identifiziert werden konnte (Johnson et al., 2008), sind Längsschnittstudien vergleichsweise rar.[7] Der Mangel an solchen Studien liegt sicher nicht am Mangel adäquater Auswertungsmethoden, wie ein Blick auf die Literatur zeigt (z.B. Boomsma & Molenaar, 1987; Hewitt, Eaves, Neale & Meyer, 1988; McArdle, 1986), vielmehr ist er durch die intensiven Kosten und den oft verkannten Nutzen der quantitativen Verhaltensgenetik als adäquate Umweltforschung begründet (Loehlin, 2009). In Folge dieses Mangels habe ich die Betrachtung genetisch informativer Längsschnittstudien zu den Big Five auf die Persönlichkeitstaxonomien von Eysenck (Eysenck & Eysenck, 1985) und Tellegen (Tellegen & Waller, 2008) erweitert. Die Parallelität dieser Taxonomien sind hinlänglich beschrieben (Bouchard & Loehlin, 2001; Markon, Krueger & Watson, 2005). Trotzdem konnte ich gerade einmal 17 verhaltensgenetische Längsschnittstudien zur Persönlichkeitsentwicklung mit Hilfe von Fachdatenbanken ausfindig machen (siehe Tabelle 1). Da erst in den letzten 5 Jahren

[6] Das meint nicht Interaktion im statistischen Sinne.

[7] Ausgenommen seien hier die Zwillings- und Familienstudien zur Untersuchung von frühen Temperamentsmerkmalen und Problemverhalten im Kindesalter. (z.B. Bartels, van Beijrsterveldt, Derks, Stroet, Polderman et al., 2007b; Goldsmith, Lemery-Chalfant, Schmidt, Arneson & Schmidt, 2007; siehe auch Bartels, 2007).

die Zahl der publizierten Studien nahezu verdoppelt wurde (Studien unter der fettgestrichelten Linie für das Erwachsenenalter in Tabelle 1), konnten die oben dargestellten Theorien nur auf wenigen Studien fundiert werden. Der gegenwärtige Wissensstand erlaubt ein differenzierteres Bild in Bezug auf die Rolle genetischer und Umweltfaktoren bei der Persönlichkeitsstabilität und –entwicklung in Abhängigkeit des Alters.

Tabelle 1:

Genetisch informative Längsschnittstudien zur Persönlichkeit

Autoren (Jahr)	Stichprobe (Design)	Alter (Intervall)	Messinstrument (Messmethode)	Persönlichkeits-merkmale
Eaves & Eysenck (1976)	253 EZ, 188 ZZ	> 18 (2 Jahre)	11-Itemskala (Selbstbericht)	Neurotizismus
Dworkin et al. (1976)*	25 EZ, 17 ZZ	16-28 (12 Jahre)	MMPI, CPI (Selbstbericht)	14 MMPI- und 18 CPI-Skalen
Pogue-Geile & Rose (1985)*	71 EZ, 62 ZZ	20-25 (5 Jahre)	MMPI (Selbstbericht)	6 MMPI-Skalen
Loehlin et al. (1990)	229 AV, 83 BV	3-24 (10 Jahre)	16-PF, MMPI (Elternbericht)	Extraversion Neurotizismus Sozialisation
McGue et al. (1993)	79 EZ, 48 ZZ	17-37 (10 Jahre)	MPQ (Selbstbericht)	Neg. Emotionalität Pos. Emotionalität Zwanghaftigkeit (+ Primärskalen)
Viken et al. (1994)	4922 EZ, 10010 ZZ	18-53 (6 Jahre)	EPI (Selbstbericht)	Extraversion Neurotizismus
Pedersen & Reynolds (1998)	237 EZ (80 getrennt), 437 ZZ (211 getrennt)	> 50 (3 × 3 Jahre)	EPI, NEO-PI (Selbstbericht)	Extraversion Neurotizismus Offenheit
Loehlin & Martin (2001)	2330 EZ, 3465 ZZ	17-92 (8/6 Jahre)	EPQ (Selbstbericht)	Neurotizismus Extraversion Psychotizismus
Gillespie et al. (2004)	253 EZ, 417 ZZ	12-16 (2 × 2 Jahre)	Junior- EPQ (Selbstbericht)	Neurotizismus Extraversion Psychotizismus
Johnson et al. (2005)	384 EZ, 274 ZZ	27-99 (5 Jahre)	MPQ (Selbstbericht)	Neg. Emotionalität Pos. Emotionalität Zwanghaftigkeit (+ Primärskalen)
De Fruyt et al. (2006)	79 EZ, 124 ZZ	5-14 (1.5 Jahre)	HiPIC (Elternbericht)	Neurotizismus Extraversion Imagination Gütmütigkeit Gewissenhaftigkeit

Fortsetzung von Tabelle 1

Autoren (Jahr)	Stichprobe (Design)	Alter (Intervall)	Messinstrument (Messmethode)	Persönlichkeits-merkmale
Read et al. (2006)	149 EZ, 202 ZZ	> 80 (2 × 2 Jahre)	EPI (Selbstbericht)	Neurotizismus Extraversion
Bratko & Butkovic (2007)	75 EZ, 85 ZZ	15-23 (4 Jahre)	EPQ (Selbstbericht)	Neurotizismus Extraversion Psychotizismus
Wray et al. (2007)	4999 Familien	> 17 (22 Jahre)	EPQ, NEO-FFI (Selbstbericht)	Neurotizismus
Blonigen et al. (2008)	411 EZ, 215 ZZ	17-24 (7 Jahre)	MPQ (Selbstbericht)	Pos. Emotionalität Neg. Emotionalität Zwanghaftigkeit (+ Primärskalen)
Bleidorn et al. (2009)	126 EZ, 61 ZZ	18-59 (2 × 5 Jahre)	NEO-PI-R (Selbstbericht)	Neurotizismus Extraversion Offenheit Verträglichkeit Gewissenhaftigkeit (+ Facetten)
Kandler et al. (**STUDIE III**)	696 EZ, 387 ZZ	16-75 (2 × 6½ Jahre)	NEO-FFI (Selbstbericht + gemittelter Bekannten-bericht)	Neurotizismus Extraversion Offenheit Verträglichkeit Gewissenhaftigkeit

Bemerkung. EZ = eineiige Zwillinge; ZZ = zweieiige Zwillinge; BV = biologische Verwandte; AV = Adoptionsverwandte; * Diese Studien sind nur im weiten Sinne mit den im Text zitierten Taxonomien vereinbar.

Welche Rolle spielen genetische Faktoren? Mittlerweile kann kaum mehr daran gezweifelt werden, dass neue genetische Faktoren über das Kindes- und Jugendalter emergieren und bei der Persönlichkeitsreifung eine Rolle spielen, was genetische Korrelationen $r_G < 1.0$ indizieren (De Fruyt, Bartels, Van Leeuwen, De Clercq, Decuyper & Mervielde, 2006; Gillespie, Evans, Wright & Martin, 2004; Loehlin, Horn & Willerman, 1990). Die genetisch beeinflusste Entwicklung erstreckt sich sogar bis in das junge Erwachsenenalter (Blonigen, Carlson, Hicks, Krueger & Iacono, 2008; Bratko & Butkovic, 2007; McGue et al., 1993; Viken, Rose, Kaprio & Koskenvuo, 1994). Im mittleren und späten Erwachsenenalter finden sich nahezu keine neuen genetischen Effekte mehr ($r_G = 1.00$; Johnson, McGue & Krueger, 2005; Pedersen & Reynolds, 1998; Read, Vogler, Pedersen & Johansson, 2006; Viken et al., 1994). Dieses Muster der genetischen Kontinuität für das junge und mittlere Erwachsenenalter konnte kürzlich über verschiedene Messinstrumente (für Neurotizismus; Wray, Birley, Sullivan, Visscher & Martin, 2007) und Beurteilerperspektiven validiert werden (**STUDIE III**). Somit scheint die stetige Zunahme phänotypischer Stabilität vom Kindes- bis zum jungen Erwachsenenalter ($r_P = .35$ - $.60$; Roberts & DelVecchio, 2000) mit genetischen

Reifungsprozessen zusammenzuhängen, während die relativ hohe Stabilität interindividueller Persönlichkeitsunterschiede im mittleren und hohen Erwachsenenalter (r_P > .65; Terracciano et al., 2006) mit der genetischen Stabilität einhergeht.

Normative Veränderungen lassen sich nicht in genetische und Umweltkomponenten zerlegen. Wachstumskurvenmodelle erlauben es jedoch, interindividuelle Unterschiede in intraindividuellen Entwicklungsverläufen in genetische und Umweltvarianz zu dekomponieren (McArdle, 1986). So konnte eine Studie von Bleidorn, Kandler, Riemann, Angleitner & Spinath (2009) zeigen, dass die Varianz intraindividueller Veränderungen in Neurotizismus, Verträglichkeit und Gewissenhaftigkeit stärker auf genetische Effekte zurückzuführen ist als die Veränderung in Extraversion und Offenheit. Da gerade diese Merkmale signifikante mittlere Trends aufweisen, liefert diese Studie einen bemerkenswerten Hinweis auf eine genetische Beteiligung.

Welche Rolle spielen „geteilte" Umweltfaktoren? Querschnittstudien (siehe oben) konnten kaum systematische Umwelteffekte, die von Familienangehörigen (z.B. Zwillinge, Adoptivgeschwister) geteilt werden, auf Persönlichkeitseigenschaften erhärten. Es bleibt aber die Frage, ob diese Faktoren möglicherweise im Kindes- und Jugendalter eine Rolle spielen. Die Befundlage ist hier eher uneinheitlich. De Fruyt et al. (2006) und Gillespie et al. (2004) finden allenfalls schwache Hinweise auf geteilte Umwelteffekte für bestimmte Merkmale in ihren jungen Stichproben. Hingegen fand eine multimodale Querschnittstudie aus mehreren Messmethoden, multiplen Beurteilerperspektiven und Verwandtschaftsbeziehungen, welche persönlichkeitsverwandte Merkmale im Jugendalter untersuchte, deutlichere Hinweise auf geteilte Umwelteffekte (Loehlin, Neiderhiser & Reiss, 2003).

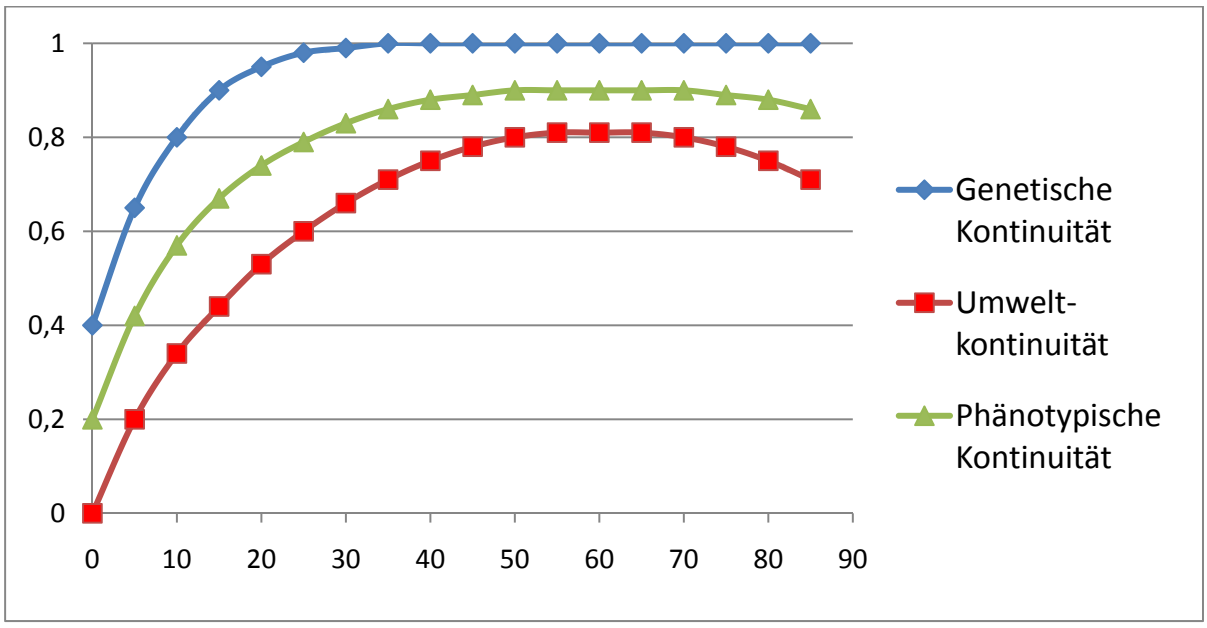
Diese uneinheitliche Befundlage kann mit zwei Problemen verbunden sein. Erstens ist es fraglich, breite Persönlichkeitseigenschaften wie die Big Five schon im Kindesalter anzunehmen. Deswegen wurden im Kindesalter häufig persönlichkeitsverwandte Temperamentsmerkmale und Verhaltenscharakteristika untersucht. Diese können jedoch einer anderen Beeinflussung durch genetische und Umweltfaktoren unterliegen. Zweitens ist es schwierig, Persönlichkeitseigenschaften im Kindesalter zu messen. Die häufigsten Messungen basieren auf Elternberichten. Dadurch können Schätzungen geteilter Umwelteffekte durch *Assimilations*- und *Kontrasteffekte* verzerrt sein (Saudino, 2003; Spinath & Angleitner, 1998). Erste Versuche, beide Probleme zu überwinden, stecken noch in den Kinderschuhen (z.B. Bleidorn & Ostendorf, 2009).

Welche Rolle spielen „nichtgeteilte" Umweltfaktoren? Will man Effekte individueller (von Verwandten nichtgeteilter) Umwelten auf die Stabilität und Veränderung untersuchen, ist es wichtig, den Messfehleranteil zu kontrollieren. Messfehler sind unkorreliert über die Zeit und führen daher zu einer Erniedrigung und bei Nichtberücksichtigung zu einer Unterschätzung von phänotypischen Kontinuitätskoeffizienten. Da der Messfehler mit spezifischen Umwelteffekten konfundiert ist, führt die Nichtberücksichtigung zur Überschätzung von messzeitpunktspezifischen Umwelteffekten und zur Unterschätzung der Kontinuität nichtgeteilter Umwelteffekte. Struk-

turgleichungsmodellierungen zur Auswertung von genetisch informativen Längsschnittdaten erlauben zumindest die Kontrolle des Messfehlers in Bezug auf die Kontinuität[8], ohne auf verschiedene Merkmalsindikatoren oder Messmethoden zurückgreifen zu müssen (Evans, Gillespie & Martin, 2002; Lemery & Goldsmith, 1999; Neale & McArdle, 2000).

Abbildung 6:

Phänotypische, genetische und Umweltkontinuität in Abhängigkeit des Alters



Bemerkung. X-Achse = Alter; Y-Achse = 1-Jahres-Stabilitäten; die Abbildung zeigt den hypothetischen Verlauf der Kontinuitäten (korrigiert um den Messfehler) in Anlehnung an die gegenwärtige Befundlage.

Gillespie et al. (2006) fanden 2-Jahres-Stabilitäten von spezifischen Umweltfaktoren der Eysenck'schen Persönlichkeitsdimensionen (Eysenck & Eysenck, 1985) im Teenager-Alter zwischen r_E* = .16 (Psychotizismus) und r_E* = .53 (Neurotizismus). Erwartungsgemäß fielen die 4-Jahres-Stabilitäten geringer aus: r_E* = .06 (Psychotizismus) bis r_E* = .36 (Neurotizismus). Aus der Studie von Blonigen et al. (2008), die den Übergang vom Jugendalter in das junge Erwachsenenalter betrachteten, lassen sich Umweltstabilitäten von r_E* = .12 (Positive Emotionalität) bis r_E* = .26 (Zwanghaftigkeit) über 7 Jahre für die Sekundärfaktoren des MPQ (Tellegen & Waller, 2008) ableiten. Für das mittlere bis hohe Erwachsenenalter berichteten Johnson et al. (2005) schon Umweltkontinuitäten zwischen r_E* = .64 (Zwanghaftigkeit) und r_E* = .73 (Positive Emotionalität) für die MPQ-Faktoren und zwischen r_E* = .53 und r_E* = .70 für die

[8] Korrigierte phänotypische Kontinuitätskoeffizienten: r_P* = r_P/(1 – Messfehler); korrigierte Umweltkontinuitätskoeffizienten: r_E* = r_E/(1 – Messfehler)

Primärskalen. Im hohen Erwachsenenalter scheinen die Umweltkorrelationen allerdings wieder abzunehmen (Pedersen & Reynolds, 1998; Read et al., 2006). Diese Befunde deuten darauf hin, dass die rasant zunehmende Stabilität der Persönlichkeit bis ins Erwachsenenalter auch mit einer Zunahme stabiler Umweltfaktoren zusammenhängt, während die weitere Zunahme über das Erwachsenenalter und eine Abnahme der Kontinuitätskoeffizienten im hohen Alter lediglich mit zunehmender Variabilität individueller Umweltfaktoren (z.B. Krankheit, Tod des Partners) einhergeht (siehe Abb. 6).

Zwar können Längsschnittstudien Effekte des Messfehlers kontrollieren. Die Kontrolle eines systematischen *Beurteilerbias* erfordert allerdings dennoch mehrere Beurteilerperspektiven. Soweit mir bekannt ist, gab es bisher erst eine Studie, welche die Befundlage über Selbst- und Bekanntenberichte validierte (**STUDIE III**). Meine Kollegen und ich **(STUDIE III)** berichteten konsensuell valide umweltbedingte 13-Jahresstabilitäten zwischen r_E* = .25 (Neurotizismus) und r_E* = .28 (Extraversion) im jungen Erwachsenenalter, während die Umweltkontinuitäten im mittleren Erwachsenenalter zwischen r_E* = .41 (Gewissenhaftigkeit) und r_E* = .67 (Extraversion) rangierten. Die fortwährende (jedoch negativ beschleunigte) Zunahme von phänotypischen Kontinuitätskoeffizienten jenseits von 30 (Terracciano et al., 2006) kann also auf kumulierende Umweltkontinuität zurückgeführt werden (siehe Abb. 6). Wie es schon Viken et al. (1994) für Extraversion und Neurotizismus beschrieben, kumulieren Umwelteffekte über die Lebensspanne (Zunahme der Variabilität aufgrund spezifischer Umwelteffekte). Da genetische Faktoren konstant bleiben, ist eine tendenzielle Abnahme der Höhe der Erblichkeitsschätzungen über das Erwachsenenalter die Folge.

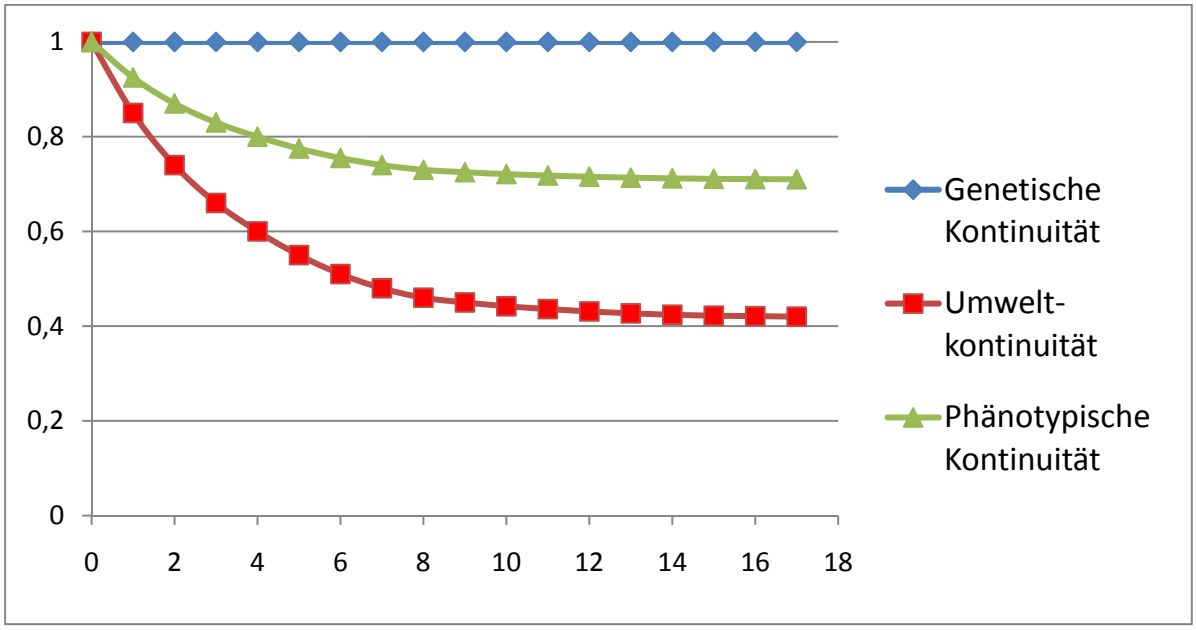
Während genetische Kontinuitätskoeffizienten spätestens ab dem mittleren Erwachsenenalter unabhängig von der Größe des Messzeitintervalls konstant bleiben, nehmen Kontinuitäten von Umweltfaktoren mit zunehmenden Messzeitintervall wieder ab (siehe Abb. 7; **STUDIE III**). Persönlichkeitsveränderung und -entwicklung (nach Vollendung der genetischen Reife) jenseits von 30 scheinen also umweltbedingt zu sein. Wray und Kollegen (2007) untersuchten die 22-Jahres-Kontinuität von Neurotizismus als Komposition aus zwei Messinstrumenten und bestimmten eine Umweltkontinuität von r_E* = .42. Da sich bisher keine bemerkenswerten Unterschiede zwischen den Persönlichkeitsdimensionen hinsichtlich des Musters ihrer Stabilität zeigten, ließ ich die Umweltkontinuitäten in Abbildung 7 entsprechend dieser Langzeitstudie r_E* = .42 annähern. Der Annäherungspunkt der phänotypischen Kontinuität ist für r_P* = .71 gewählt. Wenn man die Kontrolle des Messfehlers in Rechnung stellt, entspricht das ganz gut den Studien von Roberts & DelVecchio (2000) sowie Fraley & Roberts (2005).

Jenseits der Betrachtung von Veränderungen individueller Unterschiede berichteten Bleidorn et al. (2009) signifikante Umwelteffekte auf die Varianz systematischer intraindividueller Veränderungen in Extraversion und Offenheit. Neyer & Asendorpf (2001) fanden eine Zunahme emotionaler Stabilität einhergehend mit dem Beginn der ersten Partnerschaft zwischen dem 18. und 30. Lebensjahr. Normative Ent-

wicklungstrends können als auch mit normativen Veränderungen der Umweltfaktoren einhergehen.

Abbildung 7:

Phänotypische, genetische und Umweltkontinuitäten in Abhängigkeit der Länge des Messzeitintervalls



Bemerkung. X-Achse = Messzeitintervall; Y-Achse = Stabilität. die Abbildung zeigt den hypothetischen Verlauf der Kontinuitäten (korrigiert um den Messfehler) in Anlehnung an die gegenwärtige Befundlage. Der angenommene erste Messzeitpunkt liegt im mittleren Erwachsenenalter nach Vollendung der genetischen Reife.

Zusammenfassung. Der aktuelle Stand der Forschung erlaubt ein differenzierteres Bild genetischer und Umweltfaktoren der Persönlichkeitsentwicklung über die Lebensspanne (Abb. 6 und 7). Die genetische Kontinuität nimmt negativ beschleunigt über das Kindes- und Jugendalter zu bis diese spätestens ab dem mittleren Erwachsenenalter konstant bleibt, was als genetischer Reifungsprozess interpretiert werden kann. Die Kontinuität von Umweltfaktoren nimmt bis in das hohe Erwachsenenalter zu, erreicht jedoch nie absolute Stabilität und nimmt im späten Erwachsenenalter sogar wieder ab. Die genetische Kontinuität verläuft somit nach dem jungen Erwachsenenalter unabhängig vom Messzeitintervall, während die Umweltkontinuität mit Zunahme des Messzeitintervalls wieder abnehmen kann. Somit bilden Umwelteffekte die primäre Quelle der Persönlichkeitsveränderung im Erwachsenenalter. Umwelteffekte kumulieren über die Lebensspanne, was eine Abnahme von Erblichkeitsschätzungen mit dem Alter nach sich zieht.

4. Zur Interaktion von Anlage und Umwelt

Bisher habe ich mich auf die Betrachtung von absoluten Beiträgen genetischer und Umweltfaktoren per se an der phänotypischen Varianz beschränkt. Das gleiche gilt selbstverständlich auch für die Korrelationen zwischen Messmethoden eines Phänotyps und die Korrelationen über verschiedene Messzeitpunkte innerhalb eines Phänotyps. Diese Per-Se-Betrachtung impliziert keine Aussagen über dahinterliegende Prozesse und Mechanismen von genetischen und Umweltfaktoren, bzw. das Zusammen- und Wechselwirken dieser.

Der Wert, die Molekulargenetik und direkte Maße der Umwelt in verhaltensgenetische Studien zu integrieren, wurde von Verhaltensgenetikern schon lange erkannt (Plomin, DeFries & Loehlin, 1977).[9] Einem Review von 43 Studien zu gemessenen spezifischen Umweltvariablen folgte aber große Ernüchterung. Nicht einmal 2% der Varianz spezifischer Umweltfaktoren in Persönlichkeitseigenschaften konnte durch Varianz in objektiven Umweltmaßen (z.B. Familienkonstellation, differentielles Erziehungsverhalten, Geschwister- und Peerinteraktion) erklärt werden (Turkheimer & Waldron, 2000). Interessanterweise stellte sich die Suche nach effektiven Genen (z.B. DNA-Markern und SNP's) mit Hilfe von mittlerweile sogar Genom-weiten Assoziationsstudien als noch ernüchternder heraus. Nicht ein menschliches Gen konnte bisher robust genug mit Persönlichkeitseigenschaften in Verbindung gebracht werden (Benjamin, Ebstein & Belmaker, 2002; Flint & Willis-Owen, 2009; Plomin et al., 2008). Das wirft eine berechtigte Frage auf: *Warum können effektive genetische und Umweltvarianz aus quantitativen genetischen Studien nicht oder nicht robust genug mit Genvarianten und objektiven Umweltmaßen in Verbindung gebracht werden?*

Kumulation sehr kleiner Effekte multipler Gene und multipler Umwelten. Es könnte sein, dass der Effekt eines einzelnen Gens wie auch der einer einzelnen Umwelterfahrung verschwindend gering sein können. Nur die Kumulation über viele Gene und Umwelten macht große Effekte. Für die komplexen psychischen Krankheitsbilder Schizophrenie und Bipolare Störung konnte jüngst eine bahnbrechende Studie Evidenz für eine polygene Komponente mit tausenden Allelen, die für sich genommen einen verschwindend kleinen Effekt aufweisen, als substantiellen molekulargenetischen Risikofaktor erbringen (Purcell, Wray, Stone, Visscher, O'Donovan et al., 2009). Persönlichkeitseigenschaften gelten als komplexere menschliche Merkmale als Schizophrenie und Bipolare Störungen. Insofern erscheint die Annahme der Kumulation kleiner Geneffekte als hochgradig plausibel. Auf Seiten der Umwelt erhöhe die Aggregation verschiedener objektiver Umweltmaße den Anteil effektiver Umweltvarianz auf 13% (Plomin, Asbury & Dunn, 2001). Dazu ist das Auffinden spezifischer objektiver Umweltfaktoren zusätzlich dadurch erschwert, dass diese über die gesamte Lebensspanne betrachtet zeitinstabil sind (siehe Punkt 3; **STUDIE III**) und so verschiedene Umweltfaktoren in verschiedenen Altersabschnitten eine Rolle spielen können. Das macht eine lebensabschnittsspezifische Untersuchung notwendig.

[9] Umgekehrt wurde der Nutzen verhaltensgenetischer Designs für die Molekulargenetik und für Sozialisationsstudien seltener erwähnt (Diewald, 2010).

Genetische und Umwelteffekte auf verschiedenen Ebenen der Merkmalsabstraktion. Persönlichkeitseigenschaften stellen selbst eine Kompositionen typischer menschlicher Eigenschaften dar. Auf verschiedenen Ebenen der Abstraktion sind wahrscheinlich unterschiedliche genetische und Umweltfaktoren wirksam **(STUDIE II)**, was eine ebenenspezifische (*bottom-up*) und ebenenübergreifende (*top-down*) Betrachtung erforderlich macht.

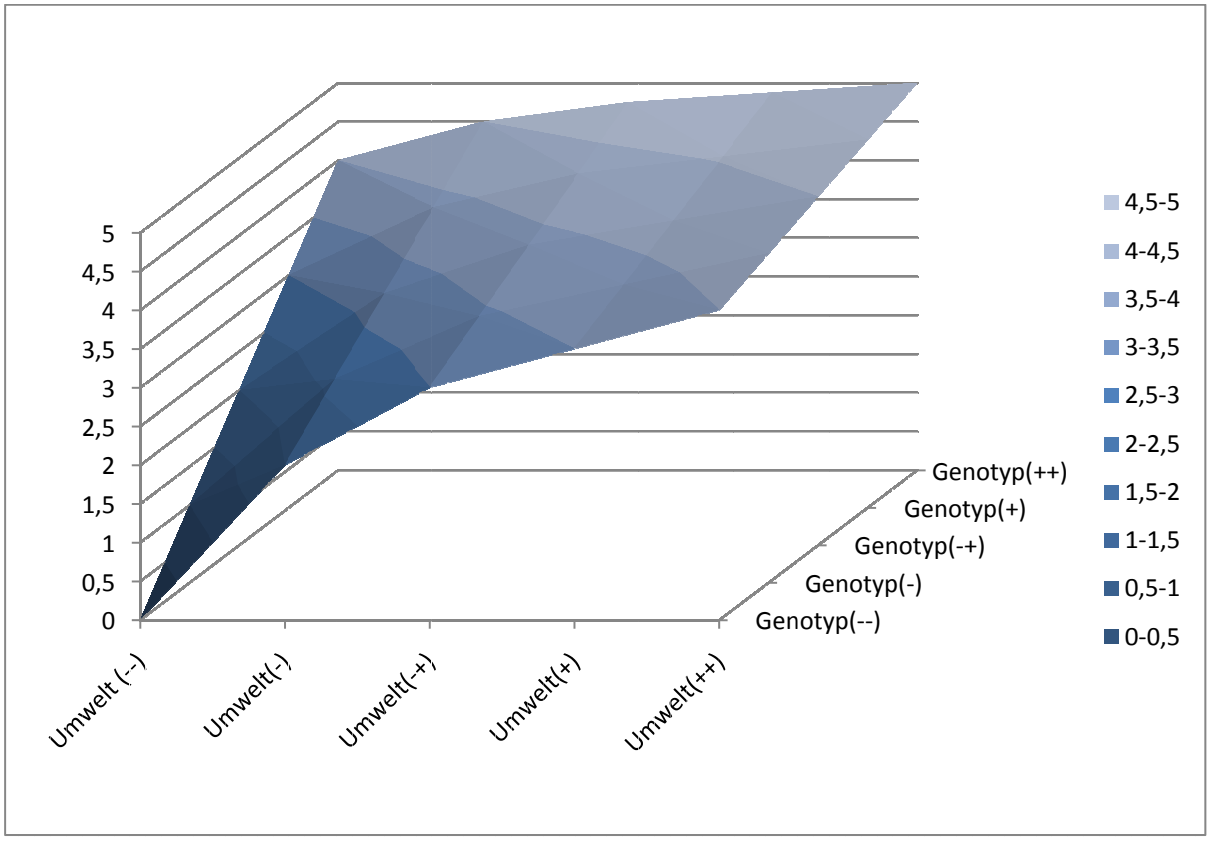
Anlage × Umwelt Interaktion. Die Bestimmung additiver Komponenten (Punkt 2 und 3) resultiert aus einer sehr restriktiven Annahme für Anlage und Umweltfaktoren eines Persönlichkeitsmerkmals. Wechselwirkungen zwischen genetischen und Umweltfaktoren sind plausibel. Diese besagen, dass der genetische Effekt von Umwelteinflüssen abhängen kann (z.B. schützt eine lebenslange phenylalanin-arme Diät vor den problematischen Auswirkungen eines PKU-Gens). Umgekehrt kann der Effekt der Umwelt vom genetischen Einfluss abhängen (z.B. genetische Komponente von Resilienz als Schutz vor Auswirkungen kritischer Lebensereignisse). Verschiedene Verhaltensgenetiker haben diesen Aspekt schon lange in Betracht gezogen. Sowohl in zahlreichen quantitativ verhaltensgenetisch als auch in molekulargenetisch informativen Studien wurden Effekte von *Anlage × Umwelt Interaktion* mit komplexen Merkmalen untersucht und identifiziert (z.B. Caspi, Sudgen, Moffitt, Taylor, Craig et al., 2003; Jang, Dick, Wolf, Livesley & Paris, 2005). Abbildung 8 zeigt eine hypothetische *Anlage × Umwelt Interaktion* für ein Merkmal mit den Ausprägungen 0 bis 5. Genetische Unterschiede treten deutlicher in den extrem negativen Umweltbedingungen (--) in Erscheinung (phänotypische Ausprägung liegen hier zwischen 0 und 4). Die genetische Variabilität ist in den extrem positiven Umweltbedingungen (++) nur halb so groß (phänotypische Ausprägungen liegen zwischen 4 und 6). Unter der Bedingung extrem negativer Umweltfaktoren (--) kann also eine positive genetische Komponente (+;++) vor negativen phänotypischen Auswirkungen bewahren (3 und 4). Solche Effekte sind für Persönlichkeitsmerkmale überaus plausibel und erfordern eine Ausrichtung der genetisch informativen Studien und Modelle zur Berücksichtigung dieser (South & Krueger, 2008). Da eine Interaktion zwischen additiven genetischen und spezifischen Umwelteffekten bei Nichtberücksichtigung in verhaltensgenetischen Modellen mit der spezifischen Umweltkomponente konfundiert ist (Purcell, 2002), führt die Anwesenheit solcher Effekte zu einer Überschätzung der spezifischen Umweltkomponente. Interaktionen zwischen additiv genetischen und geteilten Umweltfaktoren führen jedoch entgegen der früheren Annahmen zu einer Überschätzung von additiver genetischer Varianz.

Gen × Gen Interaktion und *Umwelt × Umwelt Interaktion*. Neben *Anlage × Umwelt Interaktion* ist es natürlich auch denkbar, dass *multiple* genetische und Umweltfaktoren selbst nicht additiv wirken. Da das Befundmuster zu nichtadditiven genetischen Effekten (Dominanzabweichung und Epistase) im Zusammenhang mit Persönlichkeitseigenschaften relativ uneinheitlich über verschiedene Studien ist (Jang et al., 1996; **STUDIE I**; Keller, Coventry, Heath, & Martin, 2005) und die gleichen Korrelationsmuster für getrennt und gemeinsam aufgewachsene Zwillinge zu finden sind (Riemann & Spinath, 2005), erscheint die Annahme von deutlich nichtadditiven gene-

tischen Effekten auf Persönlichkeitsunterschiede (Penke, Denissen & Miller, 2007) als nicht plausibel. Plausibel hingegen erscheint die Annahme nichtlinearer Umwelteffekte (Scarr, 1993). Leider lässt sich diese Annahme nicht so ohne Weiteres überprüfen. Dafür ist die Betrachtung objektiver Umweltmaße erforderlich. Abbildung 8 veranschaulicht einen möglichen nichtlinearen Umwelteffekt. Im positiven Umweltspektrum (+;++) sind phänotypische Auswirkungen (3 – 5) weniger dramatisch als im negativen Spektrum (0 – 4,5). In der Tat finden sich Effekte physischer Disziplinierungsmaßnahmen auf die jugendliche Aggressivität, wobei die Größe des Effekts von der Disziplinierungsstrenge, dem kulturellen Kontext und dem Beziehungshintergrund zwischen Eltern und Kind abhängt (Deater-Deckard & Dodge, 1997). Auch gibt es Hinweise auf Interaktionen zwischen spezifischen und geteilten Umweltfaktoren. Eine engere Eltern-Kind-Verbundenheit puffert den Langzeiteffekt stressvoller Lebensereignisse auf die adoleszente Depressivität (Ge, Natsuaki, Neiderhiser & Reiss, 2009). Solche Effekte erscheinen bei Nichtberücksichtigung als Effekte nichtgeteilter Umwelten (Purcell, 2002). Nach diesen Überlegungen ist es also nicht mehr verwunderlich, dass die Suche nach Haupteffekten spezifischer Umwelten (Turkheimer & Waldron, 2000) relativ erfolglos war.

Abbildung 8:

Anlage × Umwelt Interaktion in Gegenwart nichtlinearer Umwelteffekte



Bemerkung. Phanotypische Merkmalsausprägung = 0 – 5; 5 Genotypen = merkmalsrelevante genetische Ausprägungen (--,-,-+,+,++) mit additivem Effekt; 5 Umweltausprägungen (--,-,-+,+,++) mit nichtadditivem Effekt.

Dynamische Anlage × Umwelt Interaktion. Der Mensch ist ein bewusst und selbständig handelndes Individuum in der Gesellschaft und gestaltet seine Entwicklung selbst mit. Der Mensch sucht sich seine Umwelten aus, welche zu seiner Persönlichkeit passen. Er vermag Umwelten zu beeinflussen und sogar zu verändern.[10] Da die Persönlichkeit genetisch beeinflusst ist, führt dieser Mechanismus unweigerlich dazu, dass sich bestimmte genetische Ausprägungen (Genotypen) in bestimmten Umwelten häufiger finden und bestimmte Phänotypen bestimmte Erfahrungen häufiger machen. Diese *dynamische Anlage × Umwelt Interaktion* ist in der Verhaltensgenetik besser bekannt als *Anlage × Umwelt Korrelation* (r_{GE}; Plomin et al., 1977). Es lassen sich drei Typen solcher Mechanismen unterscheiden. Individuen suchen sich bestimmte Umwelten, die zu ihren genetisch beeinflussten Merkmalen passen und eher merkmalsförderlich sind (aktive r_{GE}). Unabhängig von diesem aktiven Prozess, kann die Umwelt in unterschiedlicher Weise auf unterschiedliche Genotypen reagieren (reaktive r_{GE}). Biologische Eltern können Familienumwelten in Folge aktiver r_{GE} bereitstellen, welche demnach auch mit den Genotypen ihrer Kinder korreliert sind (passive r_{GE}).

Eine Möglichkeit der Aufdeckung solcher Mechanismen besteht in der Hinzuziehung von objektiven Umweltmaßen in verhaltensgenetischen Persönlichkeitsstudien. Nahezu jedes Umweltmaß, basiert es nun auf subjektiven Einschätzungen (Selbstberichte) oder objektiven Messungen (Bekanntenbericht, Beobachtung), weist eine genetische Komponente auf. Eine Metaanalyse aus 55 Studien ergab eine mittlere gewichtete Erblichkeitsschätzung von 27% (9% - 39%) für Lebensereignisse, Erziehungsverhalten, Familienumwelten, soziale Unterstützung, Peerinteraktion und Beziehungsqualität (Kendler & Baker, 2006). Ein Großteil dieser zunächst wenig plausiblen genetischen Varianz in verschiedenen Umweltmaßen kann durch die genetische Komponente der Persönlichkeitseigenschaften aufgeklärt werden (Chipuer, Plomin, Pedersen, McClearn & Nesselroade, 1993; **STUDIE I**; Krueger et al., 2003; Saudino, Pedersen, Lichtenstein, McClearn & Plomin, 1997). Der Zusammenhang zwischen Persönlichkeits- und Umweltmerkmalen findet sich nahezu immer als ausschließlich genetisch vermittelt. Wenn r_{GE}-Prozesse in der Persönlichkeitsentwicklung eine Rolle spielen und bei der Varianzzerlegung nicht berücksichtigt werden, wird die additive genetische Komponente überschätzt (Purcell, 2002). *Anlage × Umwelt Korrelationen* im klassischen Sinne implizieren aber bidirektionale Effekte (Plomin et al., 1977): Bestimmte Persönlichkeitsgenotypen finden sich in bestimmten Umwelten häufiger (aufwärtsgerichtete Pfeile in Abb. 9). Diese Umwelten wiederum stellen Erfahrungen bereit, welche die Persönlichkeit beeinflussen (abwärtsgerichtete Pfeile in Abb. 9).

Leider ist mir bisher keine verhaltensgenetische Längsschnittstudie bekannt, welche den bidirektionalen Mechanismus der r_{GE} genauer untersucht hätte (siehe Abb. 9). Phänotypische Längsschnittstudien finden eher Hinweise auf eine unidirektionale Erklärung. Persönlichkeitseigenschaften führen zu Umwelterfahrungen, aber

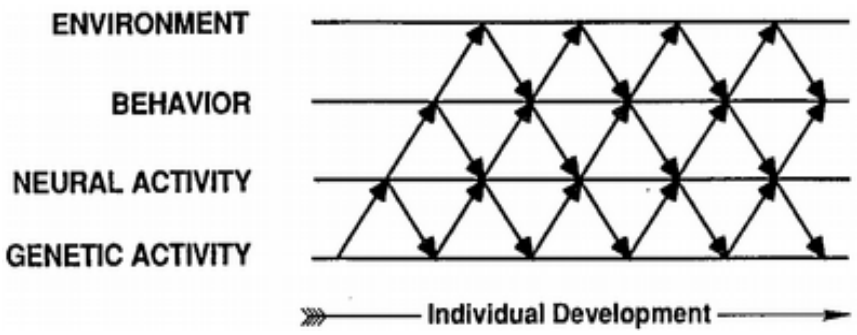
[10] Damit möchte ich nicht auf eine Debatte in Bezug auf Freiheit versus Determinismus fokussieren. Die Verhaltensgenetik vermag es nicht, eine Antwort auf diese Frage zu geben (Riemann, 2008).

Umwelterfahrungen wiederrum beeinflussen die Persönlichkeit nicht: Extraversion beeinflusst zu eher positiven und Neurotizismus zu eher negativen Erfahrungen (Headey & Wearing, 1989; Magnus, Diener, Fujita, & Pavot, 1993). Eine neuere Studie berichtete zumindest einen kleinen Effekt extrem negativer Erfahrungen auf Neurotizismus (Löckenhoff, Terracciano, Patriciu, Eaton & Costa, 2009). Solange keine genetisch informative Längsschnittstudie r_{GE} in Bezug auf interindividuelle Persönlichkeitsunterschiede untersucht, bleiben die bidirektionalen Einflüsse zwischen Genotyp und Umwelt (vermittelt über die Persönlichkeit) reine Spekulation.

Man stelle sich nun noch vor, dass auf jeder Ebene in Abbildung 9 vielfache genetische und Umwelteffekte statistisch miteinander interagieren können. Neurophysiologische Prozesse, zum Beispiel, können sowohl durch die Umwelt (z.B. traumatische Erlebnisse, Drogen, Sonnenlicht) als auch durch genetische Komponenten (morphologische Konstitution) in nichtadditiver Verknüpfung (Anlage × Umwelt Interaktionen) beeinflusst sein. Diese Vorstellung vermag einen Forscher zur Verzweiflung zu bringen. Wo wir an die Grenzen unserer Vorstellungskraft stoßen, beginnt die Natur der Persönlichkeit erst ihre Geheimnisse zu offenbaren.

Abbildung 9:

Dynamische Anlage × Umwelt Interaktion im Entwicklungsprozess



Bemerkung. Vereinfachtes Schema bidirektionaler Entwicklungseinflüsse (nach Gottlieb, 1991).

Zusammenfassung. Minimale Effekte multipler genetischer Faktoren, dynamische und statistische Anlage × Umwelt Interaktionen sind konfundiert mit Schätzungen additiver genetischer Varianz und liefern somit eine Erklärungen für die Schwierigkeit des Auffindens von robusten Assoziationen zwischen Genvarianten und Persönlichkeitsvariablen. Minimale Effekte multipler Umwelterfahrungen, statistische Anlage × Umwelt Interaktionen und nichtlineare Umwelteffekte sind konfundiert mit Schätzungen spezifischer Umwelteffekte und können somit die schwachen Haupteffekte von objektiven Umweltmaßen auf die Persönlichkeitseigenschaften erklären und ergänzen.

5. Zusammenfassung und Schlussfolgerungen

Zusammenfassung. Die dargestellten Befunde im Zusammenhang mit meinen eigenen Studien lassen sich in 14 Thesen zusammenfassen:

1. Familienumweltmerkmale haben keinen direkten umweltvermittelten Effekt auf die Persönlichkeitseigenschaften im Erwachsenenalter (**STUDIE I**; Krueger et al., 2003).
2. Studien mit nur einer Messmethode unterschätzen den genetischen Beitrag auf breite Persönlichkeitsdomänen wie die Big Five (**STUDIE II**; Riemann et al., 1997).
3. Dies gilt auch für spezifischere Persönlichkeitsmerkmale wie die Persönlichkeitsfacetten (**STUDIE II**).
4. Je akkurater die hierarchische Struktur des FFM erfasst wird, desto deutlicher offenbart es seine zugrunde liegende genetische Struktur (Jang et al., 1998; **STUDIE II**).
5. Die hierarchische genetische Struktur des FFM wird durch eine hierarchische Struktur aus spezifischen Umwelteffekten gespiegelt (**STUDIE II**).
6. Es finden sich beurteilerspezifische genetische Faktoren für Selbst- und Bekanntenein schätzungen, deutlicher jedoch für Selbstberichte. Diese können als valide interne und externe Beurteilerperspektiven auf die Persönlichkeit interpretiert werden (**STUDIE I; STUDIE II**).
7. Der Beitrag genetischer Faktoren ist für die Selbstberichtspezifität größer, was auch die Interpretation eines genetisch beeinflussten Selbstberichtbias rechtfertigt (**STUDIE II**; Riemann & Kandler, in Druck).
8. Diese spezifische Selbstberichtperspektive erweist sich als stabil über das junge und mittlere Erwachsenenalter, was in erster Linie mit genetischen Faktoren einhergeht (**STUDIE III**).
9. Genetische Faktoren sind die primäre Quelle stabiler Persönlichkeitsunterschiede (spätestens) im (mittleren) Erwachsenenalter (Johnson et al., 2005; **STUDIE III**; Read et al., 2006).
10. Umweltfaktoren tragen bei gleichbleibenden Messzeitintervallen und stabilen genetischen Komponenten zu einer fortwährenden Zunahme phänotypischer Stabilität im mittleren Erwachsenenalter bei (**STUDIE III**; Viken et al., 1994).
11. Umweltfaktoren führen zu einer Abnahme phänotypischer Kontinuität bei zunehmenden Messzeitintervallen und bilden somit die primäre Quelle phänotypischer Veränderung im Erwachsenenalter (**STUDIE III**).
12. Umwelteffekte kumulieren über die Lebensspanne. Eine Abnahme von Erblichkeitsschätzungen ist die Folge (**STUDIE III**; Viken et al., 1994).
13. Genetische Komponenten in selbst- und bekanntenberichteten Persönlichkeitsvariablen erklären genetische Einflüsse auf Familienumweltvariablen. Das kann als genetisch beeinflusste, über Persönlichkeit vermittelte Wahrnehmung von Familienumwelten oder als Anlage × Umwelt Korrelationen interpretiert werden (**STUDIE I**; Krueger et al., 2003).

14. Kumulierende Umwelteffekte bei genetischen Faktoren als einzige Quelle der Langzeitstabilität im jungen Erwachsenenalter sind mit dem *Prinzip der Korrespondenz* vereinbar, wenn die primäre Quelle der Selektion genetischen Ursprungs ist (Caspi et al., 2005; **STUDIE III**; Scarr & McCartney, 1983).

Schlussfolgerungen. Genetische Faktoren bilden die primäre Quelle *valider* und *stabiler* interindividueller Persönlichkeitsunterschiede (**STUDIE II; STUDIE III**). Durch die Annahme von und Befunden zu *dynamischen* und *statistischen Anlage × Umwelt Interaktionen* sowie *nichtlinearen Umwelteffekten* kann eine Brücke zwischen den Befunden aus quantitativer Verhaltensgenetik und den unbefriedigenden Befunden aus molekulargenetischer und Umweltforschung geschlagen werden. Umwelten (physisch, sozial, kulturell) stellen Möglichkeiten bereit und begrenzen sie (McCrae & Costa, 2008). Sie fungieren als Kanalisation der Stabilisierung und Veränderung. Die gerichteten genetischen Pfade sind es, welche meistens die Zusammenhänge mit Umweltvariablen erklären **(STUDIE I)**. Der entscheidende Impuls als treibende Kraft der Persönlichkeitsentwicklung liegt also sehr wahrscheinlich im Genotyp (**STUDIE III**; Scarr & McCartney, 1983). Ich stimme mit Sandra Scarr (1993) überein, dass nur eine solche Entwicklungstheorie mit einer Darwinistischen Evolutionstheorie vereinbar ist. Dabei ist es wichtig zu erwähnen, dass diese Theorie nichts darüber aussagt, ob der Mensch nun Fahrzeug oder Fahrer ist. Die Verantwortung des Menschen liegt so oder so im „Fahren".

Zum Wert der Verhaltensgenetik. Die quantitative Verhaltensgenetik war in der Vergangenheit immer wieder zahlreicher Kritik ausgesetzt, die natürlich richtig aber auch teilweise unfair war. Die Einschränkung auf die Betrachtung absoluter Varianzkomponenten impliziere keine Entwicklungsprozesse, keine Kausalität und keine Relevanz für den Einzelfall. Diesen Kritikpunkten stimme ich natürlich zu, wenn die Kritik auch auf sehr hohem Niveau angesiedelt ist und nicht so stehen gelassen werden sollte. Die eingeschränkte Betrachtung einer verhaltensgenetischen Studie resultiert nicht aus den Möglichkeiten der formalen Genetik, sondern aus den Begrenzungen eines Studiendesigns (Hood, Halpern, Greenberg & Lerner, 2010). Ich hoffe, mit meinen Arbeiten und denen meiner Kollegen in ausreichender Weise gezeigt zu haben, was schon einfache Zwillingsstudien für das Bild der Persönlichkeitsentwicklung zu leisten im Stande sind. Wie ich es schon eingangs erwähnt habe, liegt ein großer Nutzen der quantitativen Verhaltensgenetik in der adäquaten Umweltforschung, indem sie erlaubt, spezifische Umweltfaktoren unter Kontrolle des genetischen Beitrags zu identifizieren. Durch die Berücksichtigung des Prinzips der Unkonfundiertheit (Nachtigall, Steyer & Wüthrich-Martone, 2001), d.h. der Kontrolle des genetischen Beitrags, sind verhaltensgenetische Studien näher an der Kausalität und daher von großem Nutzen für die Sozialisationsforschung. Was die Kausalität im Weiteren und die Relevanz für den Einzelfall betrifft, bleibt ganz klar die Unterscheidung zwischen Determinismus und Probabilismus zu machen. Sofern eine betrachtete Person Teil der betrachteten Population in einer bestimmten Zeit ist, können wir nur Wahrscheinlichkeitsaussagen über diese Person treffen. Das ist kein Mangel der Verhal-

tensgenetik, das ist ein methodologisches Problem der gesamten Sozialwissenschaften (Turkheimer, 2000). Ich hoffe mit dieser Synopse auch gezeigt zu haben, welchen Wert die Verhaltensgenetik nicht nur für die Differentielle Psychologie hat, sondern wie sie auch die Sozialisationsforschung und die Entwicklungspsychologie bereichern kann. Kritische Einwände sind stets willkommen und bringen die Forschung voran, doch Kritik allein löst keine Probleme.

Literaturverzeichnis

Anderson, S. M. (1984). Self-knowledge and social inference: II. The diagnosticity of cognitive/affective and behavioural data. *Journal of Personality and Social Psychology*, *69*, 518-529.

Anusic, I., Schimmack, U., Pinkus, R. T. & Lockwood, P. (2009). The nature ans structure of correlations among Big Five ratings: The Halo-Alpha-Beta model. *Journal of Personality and Social Psychology*, *97*, 1142-1156.

Asendorpf, J. B. (2009). *Persönlichkeitspsychologie.* Heidelberg: Springer.

Bartels, M. (2007). An update on longitudinal twin studies and family studies. *Twin Research and Human Genetics*, *10*, 1-2.

Bartels, M., van Beijsterveldt, T. C. E. M., Derks, E. M., Stroet, T. M., Polderman, T. J. C., Hudziak, J. J. & Boomsma, D. I. (2007a). Young Netherlands Twin Register (Y-NTR): A longitudinal multiple informant study of problem behavior. *Twin Research and Human Genetics*, *10*, 3-11.

Bartels, M., Boomsma, D. I., van Beijsterveldt, T. C. E. M., Hudziak, J. J. & van den Oord, E. J. G. (2007b). Twins and the study of rater (dis)agreement. *Psychological Methods*, *12*, 451-466.

Benjamin, J., Ebstein, R. P. & Belmaker, R. H. (2002). *Molecular Genetics and Human Personality.* Washington: American Psychiatric Publishing.

Bergeman, C. S., Chipuer, H. M., Plomin, R., Pederson, N. L., McClearn, G. E., Nesselroade, J. R., Costa, P. T. & McCrae, R. R. (1993). Genetic and environmental effects on Openness to Experience, Agreeableness, and Conscientiousness: An adoption/twin study. *Journal of Personality*, *61*, 159-179.

Biesanz, J. C., & West, S. G. (2004). Toward understanding assessments of the Big Five: Multitrait-multimethod analyses of convergent and discriminant validity across measurement occasion and type of observer. *Journal of Personality*, *72*, 845-876.

Blackman, M. C. & Funder, D. C. (1998). The effect of information on consensus and accuracy in personality judgment. *Journal of Experimental Social Psychology*, *34*, 164-181.

Bleidorn, W., Kandler, C., Riemann, R., Angleitner, A., & Spinath, F. (2009). Patterns and sources of adult personality development: Growth curve analyses of the NEO-PI-R scales in a longitudinal twin study. *Journal of Personality and Social Psychology*, *97*, 142-155.

Bleidorn, W. & Ostendorf, F. (2009). Ein Big Five-Inventar für Kinder und Jugendliche: Die deutsche Version des Hierarchical Personality Inventory for Children (HiPIC). *Diagnostica*, *55*, 160-173.

Blonigen, D.M., Carlson, M.D., Hicks, B.H., Krueger, R.F., & Iacono, W.G. (2008). Stability and change in personality traits from late adolescence to early adulthood: A longitudinal twin study. *Journal of Personality*, *76*, 229-266.

Boomsma, D. I. & Molenaar, P. C. M. (1987). The genetic analysis of repeated measures. I. Simplex models. *Behavior Genetics*, *17*, 111-123.

Borkenau, P. (1992). Implicit personality theory and the five-factor model. *Journal of Personality*, *60*, 295-327.

Borkenau, P. & Ostendorf, F. (1993). *NEO-Fünf-Faktoren-Inventar (NEO-FFI)*. Göttingen: Hogrefe.

Borkenau, P., Riemann, R., Angleitner, A. & Spinath, F. (2001). Genetic and environmental influences on observed personality: Evidence from the German Observational Study of Adult Twins. *Journal of Personality and Social Psychology*, *80*, 655-668.

Bouchard, T. J., Jr., & Loehlin, J. C. (2001). Genes, evolution, and personality. *Behavior Genetics*, *31*, 243-273.

Bratko, D. & Butkovic, A. (2007). Stability and genetic and environmental effects from adolescence to young adulthood: Results of Croatian longitudinal twin study of personality. *Twin Research and Human Genetics*, *10*, 151-157.

Buss, D. M., & Craik, K. H. (1983). The act frequency approach to personality. *Psychological Review, 90,* 105–126.

Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81-105.

Carey, G. (2002). *Human Genetics for the Social Sciences*. Thousand Oaks, CA: Sage.

Caspi, A., Sudgen, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., McClay, J., Mill, J., Martin, J., Braithwaite, A. & Poulton, R. (2003). Influence of life stress on depression: Moderation by a polymorphism in the 5HTT gene. *Science*, *301*, 386-389.

Caspi, A., Roberts, B.W., & Shiner, R.L. (2005). Personality development: Stability and change. *Annual Review of Psychology*, *56*, 453-484.

Cattell, R. B., Blewett, D. B. & Beloff, J. R. (1955). The inheritance of personality. American *Journal of Human Genetics*, *7*, 122-146.

Chipuer, H. M., Plomin, R., Pedersen, N. L., McClearn, G. E., & Nesselroade, J. R. (1993). Genetic influence on family environment: The role of personality. *Developmental Psychology*, *29*, 110-118.

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual.* Odessa, FL: Psychological Assessment Resources.

Deater-Deckard, K. & Dodge, K. A. (1997). Externalizing behavior problems and discipline revisited: Nonlinear effects and variation by culture, context, and gender. Psychological Inquiry, 8, 161-175.

De Fruyt, F., Bartels, M., Van Leeuwen, K. G., De Clercq, B., Decuyper, M. & Mervielde, I. (2006). Five Types of personality continuity in childhood and adolescence. *Journal of Personality and Social Psychology*, *91*, 538-552.

Diewald, M. (2010). Zur Bedeutung genetischer Variation für die soziologische Ungleichheitsforschung. *Zeitschrift für Soziologie*, *39*, 4-21.

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 73,* 1246–1256.

Dworkin, R. H., Burke, B. W., Maher, B. A. & Gottesman, I. I. (1976). A longitudinal study of the genetics of personality. *Journal of Personality and Social Psychology*, *34*, 510-518.

Eaves, L. & Eysenck, H. (1976). Genetic and environmental components of inconcistency and unrepeatability in twins' responses to a neuroticism questionnaire. *Behavior Genetics*, *6*, 145-160.

Evans, D. M., Gillespie, N. A. & Martin, N. G. (2002). Biometrical genetics. *Biological Psychology*, *61*, 33-51.

Eysenck, H. J. & Eysenck, M. W. (1985). *Personality and Individual Differences.* New York: Plenum.

Fiske, D. W. & Campbell, D. T. (1992). Citations do not solve problems. *Psychological Bulletin*, *112*, 393-395.

Flint, J. & Willis-Owen, S. (2009). The genetics of personality. In Speicher, M. R., Antonarakis, S. E. & Motulsky, A. E. (Hrsg.), *Vogel and Moulsky's Human Genetics*. Berlin: Springer.

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, *102*, 652-670.

Funder, D. C. (2001). Accuracy in personality judgment. Research and theory concerning on obvious questions. In Roberts, B. W. & Hogan, R. (Eds.), *Personality Psychology in the Workplace. Decade of Behavior.* Washington: APA.

Funder, D. C., Kolar, D. C. & Blackman, M. C. (1995). Agreement among judges of personality: Interpersonal relations, similarity, and acquaintanceship. *Journal of Personality and Social Psychology*, *69*, 656-672.

Fraley, C. & Roberts, B.W. (2005). Patterns of continuity: A dynamic model for conceptualizing the stability of individual differences in psychological constructs across the life course. *Psychological Review*, *112*, 60-74.

Ge, X., Natsuaki, M. N., Neiderhiser, J. M. & Reiss, D. (2009). The longitudinal effects of stressful life events on adolescent despression are buffered by parent-child closeness. *Development and Psychopathology*, *21*, 621-635.

Gillespie, N. A., Evans, D. E., Wright, M. M. & Martin, N. G. (2004). Genetic simplex modeling of Eysenck's dimensions of personality in a sample of young Australian *twins. Twin Reseach*, *7*, 637-648.

Goffman, E. (2003). *Wir spielen alle Theater. Die Selbstdarstellung im Alltag.* München: Piper Verlag.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, *48*, 26-34.

Goldsmith, H. H. & Gottesman, I. I. (1977). An extension of construct validity for personality scales using twin-based criteria. *Journal of Research in Personality*, *11*, 381-397.

Goldsmith, H. H., Lemery-Chalfant, K., Schmidt, N. L., Arneson, C. L. & Schmidt, C. K. (2007). Longitudinal study of affect, temperament, and childhood psychopathology. Twin *Research and Human Genetics*, *10*, 118-126.

Gottlieb, G. (1991). Experimental canalization of behavioral development: Theory. *Developmental Psychology*, *27*, 187-203.

Headey, B. W., & Wearing, A. J. (1989). Personality, life events, and subjective well-being: Toward a dynamic equilibrium model. *Journal of Personality and Social Psychology*, *57*, 731-739.

Hewitt, J.K., Eaves, L.J., Neale, M.C., &, Meyer, J.M. (1988). Resolving causes of developmental continuity or "tracking". I. Longitudinal twin studies during growth. *Behavior Genetics*, *18*, 133-151.

Hofmann, W., Gschwender, T., & Schmitt, M. (2009). The road to the unconscious self not taken: Discrepancies between self- and observer-inferences about implicit dispositions from nonverbal behavioural cues. *European Journal of Personality*, *23*, 343-366.

Hofstee, W. K. B. (1994). Who should own the definition of personality? *European Journal of Personality*, *8*, 149-162.

Hood, K. E.., Halpern, C. T., Greenberg, G. & Lerner, R. M. (2010). *Handbook of Developmental Systems, Behavior, and Genetics.* Malden: Blackwell.

Hoyt, W. T. (2000). Rater bias in psychological research: When is it a problem and what can we do about it? *Psychological Methods*, *5*, 64-86.

Jang, K., Dick, D. M., Wolf, H., Livesley, W. J. & Paris, J. (2005). Psychosocial adversity and emotional instability: An application of gene-environment interaction *models. European Journal of Personality*, *19*, 359-372.

Jang, K. L., Livesley, W. J., & Vernon, P. A. (1996). Heritability of Big Five dimensions and their facets: A twin study. *Journal of Personality*, *64*, 577-591.

Jang, K. L., McCrae, R. R., Angleitner, A., Riemann, R., & Livesley, W. J. (1998). Heritability of facet-level traits in cross-cultural twin study: Support for a hierarchical model of personality. *Journal of Personality and Social Psychology*, *74*, 1556-1565.

Johnson, W., McGue, M., & Krueger, R. F. (2005). Personality stability in late adulthood: A behavioral genetic analysis. *Journal of Personality*, *73*, 523-551.

Johnson, A. M., Vernon, P. A. & Feiler, A. R. (2008). Behavioral genetic studies of personality: An introduction and review of the results of 50+ years of research. In Boyle, G. J., Matthews, G., & Saklofske, D. H. (Eds.), *The SAGE Handbook of Personality Theory and Assessment: Volume 1 Personality Theories and Models*. London: SAGE.

Keller, M. C., Coventry, W. L., Heath, A. C., & Martin, N. G. (2005). Widespread evidence for non-additive genetic variation in Cloninger's and Eysenck's personality dimensions using a twin plus sibling design. *Behavior Genetics*, *35*, 707-721.

Kendler, K. S. & Baker, J. H. (2007). Genetic influences on measures of the environment: a systematic review. *Psychological Medicine*, *37*, 615-626.

Kenny, D. A. (1991). A general model of consensus and accuracy in interpersonal perception. *Psychological Review*, *98*, 155-163.

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis.* New York: Guilford.

Kenny, D. A., Albright, L., Malloy, T. E. & Kashy, D. A. (1994). Consensus and interpersonal perception: Acquaintance and the Big Five. *Psychological Bulletin*, *116*, 245-258.

König, K. (2002). *Kleine psychoanalytische Charakterkunde* (7. Auflage). Göttingen: Vandenhoeck & Rubrecht.

Kraemer, H. C., Measelle, J. R., Ablow, J. C., Essex, M. J., Boyce, W. T., & Kupfer, D. J. (2003). A new approach to integrating data from multiple informants in psychiatric assessment and research: Mixing and matching contexts and perspectives. *American Journal of Psychiatry*, *160*, 1566-1577.

Krueger, R. F., Markon, K. E., & Bouchard, T. J., Jr. (2003). The extended genotype: The heritability of personality accounts for the heritability of recalled family environments in twins reared apart. *Journal of Personality*, *71(5)*, 809-834.

Lemery, K. S. & Goldsmith, H. H. (1999). Genetically informative designs for the study of behavioural development. *International Journal of Behavioral Development*, *23*, 293-317.

Letzring, T. D., Wells, S. M. & Funder, D. C. (2006). Information quantity and quality affect the realistic accuracy of personality judgment. *Journal of Personality and Social Psychology*, *9*, 111-123.

Loehlin, J. C. (1992). *Genes and Environment in Personality Development.* Newbury Park: SAGE.

Loehlin, J. C. (2009). Environment and the behavior genetics of personality: Let me count the ways. Personality and Individual Differences, doi: 10.1016/j.paid.2009.10.035.

Loehlin, J. C., Horn, J. M. & Willerman, L. (1990). Heredity, environment, and personality change: Evidence from the Texas adoption project. *Journal of Personality*, *58*, 221-243.

Loehlin, J. C. & Martin, N. G. (2001). Age changes in personality traits and their heiritabilities during the adult years: Evidence from Australian twin registry samples. *Personality and Individual Differences*, *30*, 1147-1160.

Loehlin, J. C., McCrae, R. R., Costa, P. T., & John, O. P. (1998). Heritabilities of common and measure-specific components of the Big Five personality factors. *Journal of Research in Personality*, *32*, 431-453.

Loehlin, J. C., Neiderhiser, J. M. & Reiss, D. (2003). The behavior genetics of personality and the NEAD study. *Journal of Research in Personality*, *37*, 373-387.

Löckenhoff, C. E., Terracciano, A., Patriciu, N. S., Eaton, W. W., & Costa, P. T. Jr. (2009). Self-reported extremely adverse life events and longitudinal changes in five-factor model personality traits in an urban sample. *Journal of Traumatic Stress*, *22*, 53-59.

Magnus, K., Diener, E., Fujita, F., & Pavot, W. (1993). Extraversion and Neuroticism as predictors of objective life events: A longitudinal analysis. *Journal of Personality and Social Psychology*, *65*, 1046-1053.

Markon, K. E., Krueger, R. & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology*, *88*, 139-157.

McArdle, J. J. (1986). Latent variable growth within behavior genetic models*. Behavior Genetics*, *16*, 163-200.

McCrae, R.R. & Costa, P.T. (2008). The five-factor theory of personality. In John, O.P., Robins, R.W., & Pervin, L.A. (Eds.), *Handbook of Personality: Theory and Research* (3rd ed.). New York: Guilford.

McCrae, R. R., Costa, P. T., Martin, T. A., Oryol, V. E., Rukavishnikov, A. A., Senin, I. G., Hrebickova, M. & Urbanek, T. (2004). Consensual validation of personality traits across cultures. *Journal of Research in Personality*, *38*, 179-201.

McCrae, R.R., Costa, P.T., Ostendorf, F., Angleitner, A., Hrebickova, M., Avia, M.D., Sanz, J., Sanchez-Bernardos, M. L., Kusdil, M. E., Woodfield, R., Saunders, P. R. & Smith, P. B. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology*, *78*, 173-186.

McCrae, R. R., Jang, K. L., Livesley, W. J., Riemann, R., & Angleitner, A. (2001). Sources of structure: Genetic, environmental, and artifactual influences on the covariation of personality traits. *Journal of Personality*, *69*, 511-535.

McCrae, R. R. & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*, 175-215.

McCrae, R. R., Yamagata, S., Jang, K. L., Riemann, R., Ando, J., Ono, Y. et al. (2008). Substance and artifact in the higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 95,* 442-455.

McGue, M., Bacon, S., & Lykken, D.T. (1993). Personality stability and change in early adulthood: A behavioral genetic analysis. *Developmental Psychology*, *29*, 96-109.

Mischel, W. (1968). *Personality and Assessment.* New York: Wiley.

Musek, J. (2007). A general factor of personality: Evidence for the Big One in the five-factor model. *Journal of Research in Personality*, *41*, 1213-1233.

Nachtigall, C., Steyer, R. & Wüthrich-Martone, O. (2001). Causal effects in empirical research. In M. May & U. Ostermeier (Eds.), *Interdisciplinary perspectives on causality. Bern Studies in the History and Philosophy of Science*. Norderstedt: Libri Books on Demand.

Neale, M. C. & McArdle, J. (2000). Structured latent growth curves for twin data. *Twin Research*, *3*, 165-177.

Neyer, F. J. & Asendorpf, J. B. (2001). Personality-relationship transaction in young adulthood. *Journal of Personality and Social Psychology*, *81*, 1190-1204.

Ostendorf, F., & Angleitner, A. (2004). *NEO-Persönlichkeitsinventar, revidierte Form, NEO-PI-R nach Costa und McCrae*. Göttingen, Germany: Hogrefe.

Paulhus, D. L. & John, O. P. (1998). Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality*, *66*, 1025-1060.

Pedersen, N.L. & Reynolds, C.A. (1998). Stability and change in adult personality: Genetic and environmental components. *European Journal of Personality*, *12*, 365-386.

Penke, L., Denisson, J. J. A. & Miller, G. F. (2007). The evolutionary genetics of personality. *European Journal of Personality*, *21*, 549-587.

Plomin, R., Asbury, K. & Dunn, J. (2001). Why are children in the same family so different? Nonshared environment a decade later. *Canadian Journal of Psychiatry*, *46*, 225-233.

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype- environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin*, *84*, 309-322.

Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2008). *Behavioral Genetics* (5th edition). New York: Worth Publishers.

Pogue-Geile, M. F. & Rose, R. J. (1985). Developmental genetic studies of adult personality. *Developmental Psychology*, *21*, 547-557.

Purcell, S. (2002). Variance components models for gene–environment interaction in twin analysis. *Twin Research and Human Genetics*, *5*, 554-571.

Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O'Donovan, M. C., Sullivan, P. F. & Sklar, P. (2009). Common polygenetic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, *460*, 748-752.

Read, S., Vogler, G. P., Pedersen, N. L. & Johansson, B. (2006). Stability and change in genetic and environmental components of personality in old age. *Personality and Individual Differences*, *40*, 1637-1647.

Riemann, R. (2008). Können wir unser Verhalten steuern? Verhaltensgenetische Aspekte. In R. Kögerler , F. Gruber & M. Dürnberger (Hrsg.) *Homo animal materiale. Die materielle Bestimmtheit des Menschen.* Linz: Wagner.

Riemann, R., Angleitner, A. & Stelau, J. (1997). Genetic and environmental influences on personality: A study of twins reared together using the self- and peer report NEO-FFI scales. *Journal of Personality*, *65*, 449-475.

Riemann, R. & Kandler, C. (in Druck). Construct validation using multitrait-multimethod-twin data: The case of a general factor of personality. *European Journal of Personality.*

Riemann, R. & Spinath, F. M. (2005). Genetik der Persönlichkeit. In J. Hennig & P. Netter (Hrsg.), *Biopsychologische Grundlagen der Persönlichkeit* (S. 539-628). Heidelberg: Spektrum-Verlag.

Roberts, B.W. & Caspi, A. (2003). The cumulative continuity model of personality development: Striking a balance between continuity and change in personality traits across the life course. In U.M. Staudinger & U. Lindenberger (Eds.) *Understanding Human Development: Dialoges with Lifespan Psychology* (pp. 183-214). Dorrecht, Netherlands: Kluwer Academic Publishers.

Roberts, B. W. & DelVecchio, W.F. (2000). The rank-order consistency of personality from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, *126*, 3-25.

Roberts, B. W., Walton, K. E. & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life-course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *132*, 1-25.

Rushton, J. P. & Irwing, P. (2008). A general factor of personality (GFP) from two meta-analyses of the Big Five: Digman (1997) and Mount, Barrick, Scullen, and Rounds (2005). *Personality and Individual Differences*, *45*, 679-683.

Rushton, J. P. & Irwing, P. (2009). A general factor of personality in the Comrey Personality Scales, the Minnesota Multiphasic Personality Inventory-2, and the Multicultural Personality Questionnaire. *Personality and Individual Differences*, *46*, 437-442.

Saudino, K. J. (2003). Parent ratings from twin temperament: Lesson from twin studies. *Infant Behavior and Development*, *26*, 100-107.

Saudino, K. J., Pedersen, N. L., Lichtenstein, P., McClearn, G. E. & Plomin, R. (1997). Can personality explain genetic influences on life events? *Journal of Personality and Social Psychology*, *72*, 196-206.

Scarr, S. (1993). Biological and cultural diversity: The legacy of Darwin for development. *Child Development, 64*, 1333-1353.

Scarr, S. & McCartney, K. (1983). How people make their own environments: A theory of genotype → environment effects. *Child Development*, *54*, 424-435.

Shikishima, C., Ando, J., Ono, Y., Toda, T. & Yoshimura, K. (2006). Registry of adolescent and young adult twins in the Tokyo area. *Twin Research & Human Genetics*, *9*, 811-816.

South, S. C., & Krueger, R. F. (2008). An interactionist perspective on genetic and environmental contributions to personality. *Social and Personality Psychology Compass*, *2*, 929-948.

Spearman, C. (1904). General intelligence: Objectively determined and measured. *American Journal of Psychology*, *15*, 201-293.

Spinath, F. M. & Angleitner, A. (1998). Contrast effects in Buss and Plomin's EAS questionnaire: A behavioral-genetic study on early developing personality traits assessed through parental ratings. *Personality and Individual Differences*, *25*, 947-963.

Spinath, F. M., Angleitner, A., Borkenau, P., Riemann, R. & Wolf, H. (2002). German Observational Study of Adult Twins (GOSAT): A multimodal investigation of personality, temperament and cognitive ability. *Twin Research and Human Genetics*, *5*, 372-375.

Steyer, R. & Eid, M. (2001). *Messen und Testen* (2. korrigierte Auflage). Berlin: Springer.

Tellegen, A. & Waller, N. G. (2008). Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire. In Boyle, G. J., Matthews, G., & Saklofske, D. H. (Eds.), *The SAGE Handbook of Personality Theory and Assessment: Volume 2 Personality Measuring and Testing*. London: SAGE.

Terracciano, A., Costa, P. T., Jr., & McCrae, R. R. (2006). Personality plasticity after age 30. *Personality and Social Psychology Bulletin*, *32*, 999-1009.

Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions of Psychological Science*, *9*, 160-164.

Turkheimer, E., & Waldron, M. (2000). Nonshared environment: A theoretical, methodological, and quantitative review. *Psychological Bulletin*, *126*, 78-108.

Vernon, P. A., Villani, V. C., Vickers, L. C., & Harris, J. A. (2008). A behavioural genetic investigation of the dark triad and the big five. *Personality and Individual Differences*, *44*, 445-452.

Viken, R.J., Rose, R.J., Kaprio, J., & Koskenvuo, M. (1994). A developmental genetic analysis of adult personality: Extraversion and Neuroticism from 18 to 59 years of age. *Journal of Personality and Social Psychology*, *66*, 722-730.

Wray, N. R., Birley, A. J., Sullivan, P. F., Visscher, P. M. & Martin, N. G. (2007). Genetic and phenotypic stability of measures of Neuroticism over 22 years. Twin Research and Human Genetics, 10, 695-702.

Yamagata, S., Suzuki, A., Ando, J., Ono, Y., Kijima, N., Yoshimura, K., et al. (2006). Is the genetic structure of human personality universal? A cross-cultural twin study from North America, Europe, and Asia. *Journal of Personality and Social Psychology*, *90*, 987-998.

ANHANG I

Behavior Genetics

(2009)

(http://www.springerlink.com/content/84253p70768t452x/)

Genetic and Environmental Mediation between Measures of Personality and Family Environment in Twins Reared Together

Christian Kandler

Rainer Riemann

Nicole Kämpfe

Abstract. In this study we analyzed the etiology of the relationship between personality traits and retrospectively recalled family environment. The data of 226 identical and 168 fraternal twin pairs reared together from the Jena Twin Study of Social Attitudes (JeTSSA) were available. Personality traits were measured using the self- and peer report versions of the German NEO-Personality Inventory-Revised. A German version of Blocks Environmental Questionnaire was applied to measure two broad dimensions of the family environment retrospectively: Support and Organization. We could replicate earlier findings that retrospective reports of these family environment dimensions were in part genetically influenced. 66% of the genetic variance in Support and 24% in Organization could be accounted for by heritable variance in self-rated personality. That was replicated by using peer reports of personality, 41% explained genetic variance in Support and 17% in Organization. Environmental mediations were negligible. This indicates that the relationship between personality and retrospectively recalled family environment is largely genetically mediated.

In recent years the interplay between genes and environment has increasingly aroused the interest of behavioral genetic research. In addition to focusing on environmental and genetic contributions to phenotypic variance, researchers are studying the role of genes in the perception and interpretation, selection and creation of environments as well as the genetic control of exposure to various environments (Rowe, 1981, 1983; Kendler, 2001; Rutter & Silberg, 2002; Rutter et al., 2006). In addition, differential reactions of the social environment (e.g., parental behavior) to genetically affected personality and temperament traits have been demonstrated (Lytton, 1977; O'Connor et al., 1995, 1998). The goal of the present paper is to examine the relationship between personality traits of the Five-Factor Model (FFM; McCrae & Costa, 1985) and retrospective accounts of the childhood family environment by using measures of environment and personality in a study of twins reared together. While the etiology of personality has been studied extensively (Bouchard & Loehlin, 2001; Yamagata et al., 2006), only a few studies relate to the gene and environment interplay between personality and family environment.

Behavior genetic research on family environment mostly used questionnaires like the Family Environment Scale (FES; Moos & Moos, 1986) and the Block Environmental Questionnaire (BEQ; Hur & Bouchard, 1995). Factor analyses of the FES and BEQ scales yielded two robust factors underlying the family environment measures. In a pioneering study using the FES, Rowe (1983) labeled these two factors Acceptance-Rejection and Restrictiveness-Permissiveness, whereas a recent study (Herndon et al., 2005) specified the dimensions as Support and Structure. Hur and Bouchard (1995) used separate factor analyses of the scales in FES and BEQ and yielded two robust factors which they named Support and Organization. Using FES and BEQ as well as two other environmental questionnaires for a combined factor analysis, Krueger et al. (2003) also found two factors of the family environment which they labeled Cohesion and Status. Although the two factors reported were given different labels, they are highly similar in meaning across BEQ and FES. One factor reflects perceived parental acceptance versus rejection (Rowe, 1983) and cohesion versus conflict in the family environment (Krueger et al., 2003). The second factor, Organization, resembles the restrictiveness versus permissiveness dimension (Rowe, 1983) of family structure as well as the familial intellectual-cultural pursuits (Herndon et al., 2005) and parental socioeconomic status (Krueger et al., 2003). Over all studies, the two factor solution accounted for more variance in the scales of the BEQ than in those of the FES. Behavioral genetic studies on the etiology of retrospective family environment measures consistently showed a moderate genetic influence (Plomin et al., 1994, 2001). Hur and Bouchard (1995) studied retrospective reports on the rearing family environment in a sample of twins reared apart. Latent structural equation modeling of the FES and BEQ factors yielded a heritability estimate of 44% on Support and 28% on Organization.

At first glance, genetic influences on measures of the environment are counterintuitive. In the literature three explanations of this effect have been suggested. First, genetic effects on retrospective assessments of the childhood rearing environment may be interpreted as a genetically influenced memory bias of retrospective information. That means genetic factors act, for example, on the selective recall of experiences. Some individuals may recall more positive, others more negative experiences. However, adult (Plomin et al., 1989) and adolescent twins' (Herndon et al., 2005) accounts of their *current* family environment were not less heritable than retrospective reports, rendering the memory bias explanation less plausible. Second, genes may influence the way of processing experiences (Kendler, 2001). In other words, genetic factors act on the individual's interpretation or perception of

the environment. In this view, reports on the family environment (e.g., parental treatment) by a single observer (e.g., one child) reflect characteristics of the environment and individuals completing these (Vernon et al., 1997). Finally, studies using observational measures of family environment (O'Connor et al., 1995) also showed genetic influences. This has been interpreted as genetic control of exposure to the environment (Kendler, 2001). The individual genotype may influence the probability of exposure to certain events by evoking reactions, selecting and seeking out settings, changing and creating situations. These processes are known as genotype–environment correlation. Three kinds have been suggested: the passive, active, and reactive type (Plomin et al., 1977; Scarr & McCartney, 1983). For example, biological parents provide a rearing environment that is related to the child's genotype (a passive kind), child's genotype may also create and actively shapes the family environments, like family cohesion (an active type), or child's genotype receives responses from family members, e.g. parental support (a reactive kind).

In sum, there are two plausible explanations of genetic effects on self-reports of the environments: Perceptions of the environment or exposure to certain environments may be influenced by genetically predisposed characteristics of individuals. Genetically influenced personality traits are promising characteristics of individuals to explain the genetic influence on environmental measures, because personality affects how people create, interpret, or perceive their environments or evoke reactions from other people (Plomin et al., 2001).

There are some studies examining the role of genetic influences for explaining the phenotypic correlation between personality and measures of the family environment. Chipuer et al. (1993) found that genetic effects on Neuroticism and Extraversion significantly overlap with genetic influences on reports of current family environments. Jang et al. (2000) report moderate genetic correlations between indices of personality pathology and family environment factors derived from the Family Environment Scale (FES). Krueger et al. (2003) used a multivariate approach to examine phenotypic correlations between higher-order factors of the Multidimensional Personality Questionnaire (MPQ; Tellegen, 2000) and the two dimensions Support and Organization (labeled as Cohesion and Status) derived from retrospective measures of family environment. They found a negative relationship between Negative Emotionality and Support and a positive correlation between Constraint and Support. Positive Emotionality was uncorrelated with Support but was the only personality factor correlating positively with Organization. All significant phenotypic correlations were mediated genetically but not environmentally. In contrast to the general finding, Vernon et al. (1997) found no genetic effects on factors of the family environment. Small correlations between measures of personality and family environments were entirely attributable to correlated non-shared environmental influences.

However, these studies of personality–environment correlations relied on measures of personality and family environment provided by the same informants (self-reports). Genetic effects may influence not only the way we assess our environment in childhood but also the way we perceive or describe ourselves. Therefore, a genetically influenced self-report specific bias has to be taken into account as a source of genetic variance and genetic correlation (McCrae et al., 2008). We collected peer reports of personality to disentangle this explanation from the others, because rater bias is not shared among different informants. Note, that we used the word "bias" for a systematic disagreement between self- and peer raters that may be due to either (1) systematic response errors on reports from the same rater (social desirable responding, acquiescence, leniency) or (2) true self-rater unique

perception of his- or herself *as well as* his or her own environment (specificity of viewpoint). Obviously, correlations mediated by effects of the specific environment may likewise reflect biased self-reports, but environmental correlations between peer reports of personality and self-reports on the rearing environment reflect an overlap of true environmental influences. The latter correlations indicate environmental influences originating from the rearing family that affect one sibling's personality specifically (Plomin & Daniels, 1987).

The current study estimates genetic and environmental mediations between personality and retrospective accounts of the rearing environment in a sample of twins reared together. Consistent with previous research, we expect that the variation in FFM personality traits is due to genetic and non-shared environmental effects (Jang et al., 1996; Riemann et al., 1997), whereas differential accounts of the rearing environments are affected by genetic, shared and non-shared environmental effects (Hur & Bouchard, 1995; Herndon et al., 2005). Furthermore, consistent with earlier research (Chipuer et al., 1993; Krueger et al., 2003), we expect significant phenotypic correlations between personality traits and family environment, which are genetically mediated. Extending previous research, we collected self- as well as peer reports of twins' personality to disentangle a bias account from substance explanations of observed correlations between personality domains and recalled family environments. Rater bias is absent in correlations between peer rated personality and retrospective self-reports of family environments.

Methods

Participants

The sample consisted of 226 monozygotic and 168 dizygotic twin pairs (including 67 opposite-sex pairs) from the Jena Twin Registry (Stößel et al., 2006). Sampling was based in part on registers of multiple births and data from registration offices (41%), and in part on a volunteer sample approached by media calls and twin clubs (59%). Twins were offered a personality profile and a compensation of 12 € for participation. About 20% of twin pairs from the Jena study (Stößel et al., 2006) are also registered in the Bielefeld Twin Registry (Spinath et al., 2002) and 13% also participated in the Berlin Twin Study (Busjahn, 2006).

Zygosity was diagnosed by a self- report questionnaire (Oniszczenko et al., 1993) that assesses the frequency of confusing the twins by different relatives, teachers and peers across the life span as well as physical similarity criteria (concordance with genetic finger printing data is 93.2%; Becker et al., 1993). Our sample replicates earlier experiences with volunteer samples. 79% of MZ twins and 80% of same-sex DZ twins were females. The mean age of the twin participants was M = 34.30 years (SD = 13.63). The sample was heterogeneous with regard to education and occupational status and there were no marked differences between MZ and DZ twins. For 78% of the participating twins at least one peer report was available (MZ: N = 178; DZ: N = 136).

Measures

Personality. We administered the self- and peer report version of the German Neuroticism Extraversion Openness-Personality Inventory-Revised (NEO-PI-R; Ostendorf & Angleitner, 2004). The NEO-PI-R is a 241-item inventory designed for measuring personality on five domain scales (Neuroticism, Extraversion, Openness to Experiences, Agreeableness, and

Conscientiousness) and 30 facet scales. The German version (Ostendorf & Angleitner, 2004) is highly similar to the American version (Costa & McCrae, 1992). Detailed characteristics of the scales and their constituent items as well as evidence on the reliability and validity are presented in the manual (Ostendorf & Angleitner, 2004). In the present sample, Cronbach's α for the five self-rated domains ranged from 0.85 to 0.91 (highest for Neuroticism, lowest for Agreeableness) and within peer reports from 0.87 to 0.92 (highest for Conscientiousness, lowest for Openness). The Spearman-Brown corrected agreement among peers for Neuroticism, Extraversion, Openness, Agreeableness and Conscientiousness in that order were 0.62, 0.70, 0.65, 0.56 and 0.65, and the correlations between self- and averaged peer reports were 0.50, 0.62, 0.58, 0.44 and 0.53, respectively. Only the self- and averaged peer reported raw scores of the five higher-order domains were included in the subsequent analyses.

Family Environment. We used the German version of the Block Environmental Questionnaire (BEQ, Riemann & Wagner, 2000) derived from the American version (Hur and Bouchard, 1995) to measure twins' childhood family environment retrospectively. The questionnaire consists of six scales: (1) Acceptance/Rejection by mother and (2) father, (3) Family Cohesion, (4) Intellectual Orientation of mother and (5) father, and (6) Family Organization. All primary scales reported by twins revealed acceptable up to good levels of internal consistency (Cronbach's α): Maternal Acceptance/Rejection: α = 0.88; Paternal Acceptance/Rejection: α = 0.90; Family Cohesion: α = 0.90; Maternal Intellectual Orientation: α = 0.69; Paternal Intellectual Orientation: α = 0.68; Family Organization: α = 0.61. The twins' parents (with respect to complete twin pairs, 82% mothers and 69% fathers) independently rated on the following four scales: Family Cohesion and Family Organization as well as their own Acceptance/Rejection and Intellectual Orientation. The correlations between twins' and mothers' reports for Acceptance/Rejection, Family Cohesion, Intellectual Orientation and Family Organization were 0.27, 0.34, 0.35 and 0.38 (separately calculated for each twin and then averaged across twins), and the correlations between twins' and fathers' reports were 0.37, 0.29, 0.32 and 0.41, respectively.[11] Twins' raw scores of the six primary scales were used in the following analyses.

Statistical Analyses

As the existence of age and gender effects can bias the estimates of twin similarity, self- and averaged peer reported raw scores of the five personality domains and the six BEQ primary scales reported by twins were adjusted for sex and linear age effects using a regression procedure. Standardized residuals from these regressions were used in the subsequent analyses.

The sex and age effects corrected primary scales of BEQ were factor analyzed by an exploratory principle components factor analysis with varimax rotation. We expected a two factor solution suggesting the existence of two independent family dimensions consistent with Hur and Bouchard (1995). The solution is presented in the results. Using a regression method based on the combined sample of MZ and DZ twins we estimated individual factor scores.

[11] In a sample of twins reared together, naturally only one biological mother and one biological father were available to assess rearing environments of both twins and reports of them are genetically correlated with reports of their offspring because of their genetic kinship. Thus, parental reports of twins reared together are not useful and informative to address a differentiation between the genetic perception and the genotype-environment correlation hypothesis, and hence we did not use parental reports in our present analyses.

MZ and DZ variance-covariance-matrices within and between phenotypes were estimated for structural equation analyses using a listwise deletion procedure for handling missing values. The matrices were analyzed by fitting univariate and multivariate genetically informative models described below via maximum likelihood using the statistical software package Mx (Neale et al., 2003). The overall model fit was evaluated by using the Root Mean Square Error of Approximation (RMSEA) in conjunction with the Akaike's Information Criterion (AIC). Values of RMSEA < 0.05 indicate a good fit; values of RMSEA between 0.05 and 0.08 are interpretable as acceptable fit (Steyer et al., 2005). The most negative value of AIC indicates the best fit model (χ^2- 2 *df*). We compared nested models using the likelihood ratio-test (LR- test or χ^2- Difference Test) whereby we tested the hypothesis: Models with more parameters fit the data significantly better than reduced models. For comparisons among unnested models we used the AIC indices for descriptive comparisons (Steyer et al., 2005).

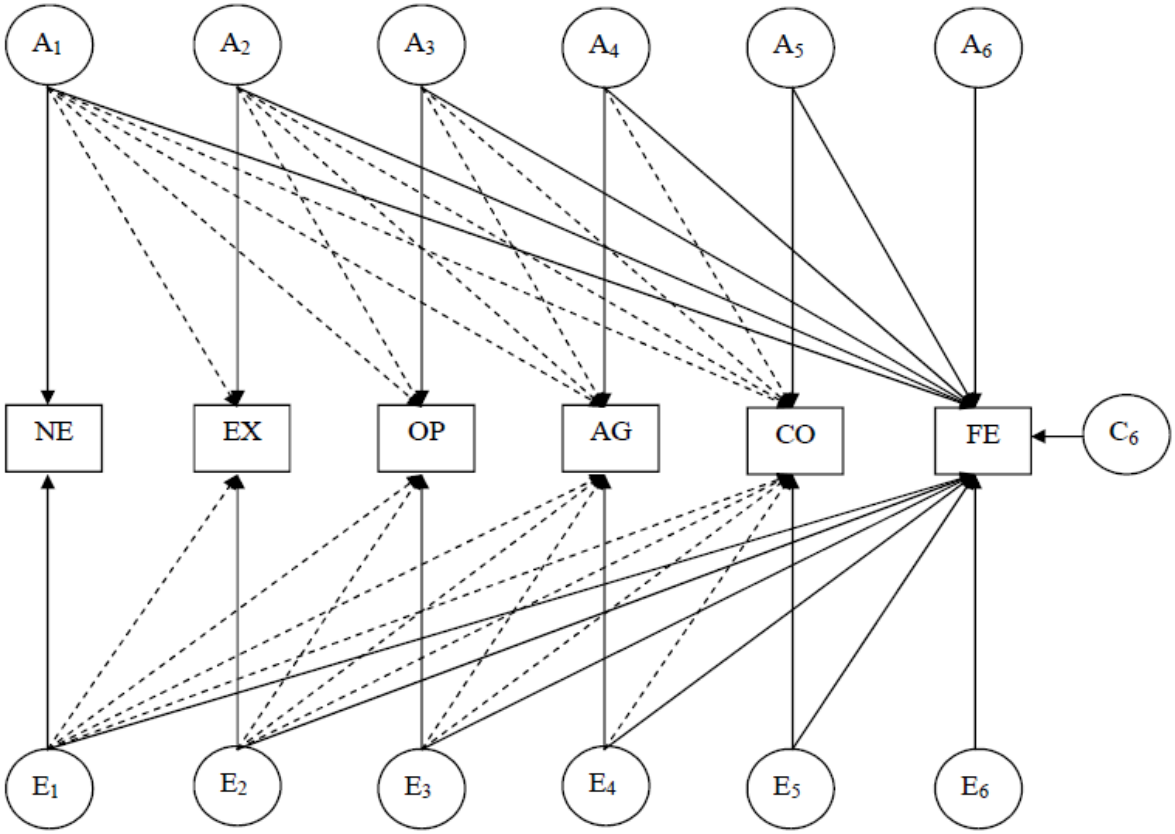
An univariate behavioral genetic structural equation model for twins reared together was applied assessing the effects of additive genetic (a^2), shared environmental (c^2) and non-shared environmental effects (e^2) within twin pairs, and standard procedures to test reduced models (Neale & Maes, 2004). If no model accomplished an overall acceptable fit at least, we checked for possible significant differences in variance between MZ and DZ twin samples or tested alternative models allowing for non-additive genetic effects (Neale & Maes, 2004).

The overall magnitude of genetic and environmental mediations between the five NEO-PI-R domains and each of the two BEQ-factors was estimated by fitting a six-variable structural equation model known as "Cholesky decomposition" (Figure 1) to the MZ and DZ phenotypic variance-covariance matrices via maximum likelihood. For model identification, the variances of the latent variables were fixed to one in order to obtain estimates for all paths in the model. Paths between the latent variables A_1 to A_6 and manifest variables (rectangles in Figure 1) reflect genetic influences, and paths between E_1 to E_6 and the observed variables represent non-shared environmental influences. Note, that our models allowed for genetic and environmental paths among personality domains (marked with dotted lines in Figure 1). But these coefficients (Jang et al., 2006) were not the focus in this study. We compared reduced models against the initial model to test four hypotheses: (1) Phenotypic correlations are mediated by both genes and environment, (2) observed relationships are due to genetic influences or (3) environmental influences and (4) there is no significant mediation. The procedure of testing the overall magnitude of genetic and environmental influences on phenotypic relationships is not affected by the order of manifest personality variables in the model.

In addition to assessing path coefficients between personality and BEQ variables, we converted these coefficients into genetic and environmental correlations. The latent correlations were calculated for the best fitting models and their statistical significance was determined using a maximum likelihood 95% confidence interval. Bivariate genetic and non-shared environmental correlations indicate the extent to which an effect on a personality variable (disregarding the effects of other personality variables) also affects the factors Support and/or Organization.

Figure 1:

Initial Six-Variables Cholesky Model of the Relationship between Personality Traits and One Family Environment Factor (FE)

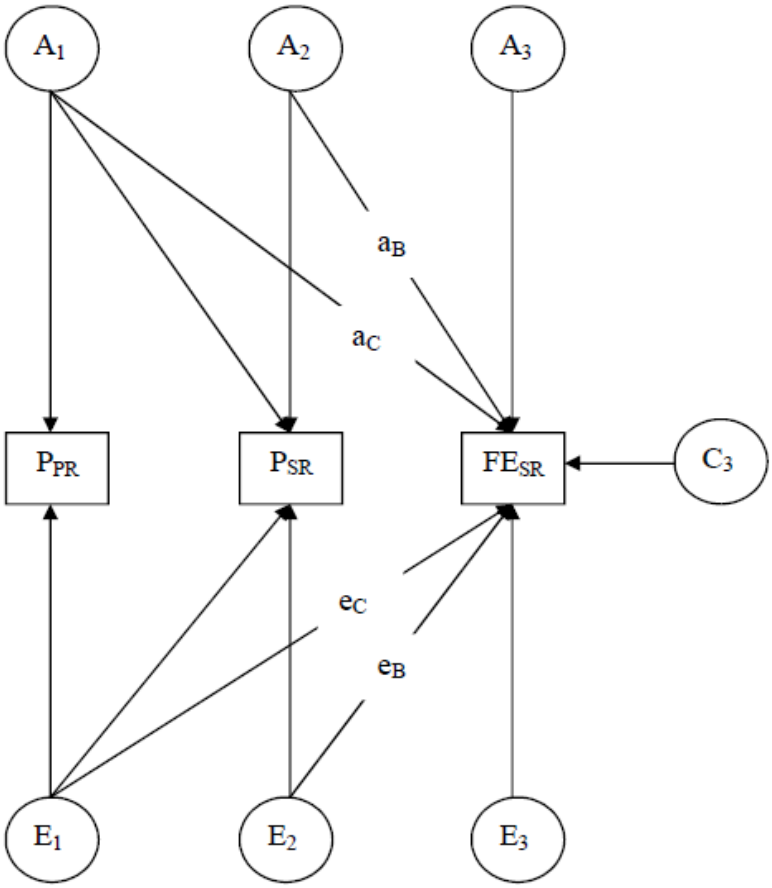


Note. For simplicity, the model is shown only for one twin. A = additive genetic variance, C = shared environmental effects, E = non-shared environmental effects, NE = Neuroticism, EX = Extraversion, OP = Openness, AG = Agreeableness, CO = Conscientiousness. Further description in the text.

Furthermore, the magnitude of genetic and non-genetic biases was estimated by fitting a three-variable Cholesky Model including averaged peer reports, self-reports of personality traits and family environment (Figure 2). Using this Cholesky decomposition, the order of variables in the model is relevant. If a self-report variable of personality explains additional (genetic or environmental) variance - reflected by the paths a_B and e_B - to that explained by the averaged peer reports - reflected by the paths a_C and e_C, this will indicate the upper limit of self-report bias. We tested the significance of biases by comparing different model modifications: (1) the full model (a_B and $e_B \neq 0$), (2) non-genetic bias model ($a_B = 0$), (3) genetic bias model ($e_B = 0$), and (4) unbiased self-reports model (a_B and $e_B = 0$). Furthermore, we tested for significant genetic and non-shared environmental correlations: (a) no genetic correlation ($a_C = 0$), (b) no specific environmental correlation ($e_C = 0$) and (c) no correlation (a_C and $e_C = 0$).

Figure 2:

Initial Three-Variables Cholesky Model Detecting Genetic and Environmental Biases



Note. For simplicity, the model is shown only for one twin. A = additive genetic variance, C = shared environmental effects, E = non-shared environmental effects, P = Personality, FE = family environment factor, PR = averaged peer reports, SR = self-reports. Further description in the text.

Results

Factor Analysis of Environmental Measures

Principle components analyses of the environmental primary scales with varimax rotation were conducted for the combined sample and the Twin A and Twin B subsamples in which one member of each twin pair was randomly selected (Table I). Screeplot suggested a two factor solution. Eigenvalues dropped off markedly after two largest values 2.46 and 1.88. Highly similar results were obtained in subsamples that were built according to zygosity, gender or age. The two principle components could easily be interpreted as Support (parental acceptance, family cohesion) and Organization (parental intellectual and cultural orientation, familial structure) consistent with Hur & Bouchard (1995).

Table I:

Varimax Rotaded Factor Loadings for the Six Primary Scales of the BEQ for the Combined Sample of Twins and Subsamples

	BEQ Factor					
	All (*N* = 757)		Twin A (*N* = 380)		Twin B (*N* = 377)	
Scale	Support	Organization	Support	Organization	Support	Organization
M. Acceptance/Rejection	**0.89**	0.09	**0.88**	0.10	**0.89**	0.07
P. Acceptance/Rejection	**0.82**	0.19	**0.85**	0.18	**0.80**	0.20
Family Cohesion	**0.91**	0.20	**0.92**	0.17	**0.91**	0.22
M. Intellectual Orientation	0.32	**0.73**	0.29	**0.73**	0.35	**0.72**
P. Intellectual Orientation	0.26	**0.77**	0.22	**0.78**	0.28	**0.78**
Family Organization	-0.06	**0.82**	-0.04	**0.79**	-0.07	**0.84**
% Variance accounted for	41	31	41	31	41	32

Note. M. = Maternal; P. = Paternal; Factor loadings greater than 0.50 are shown in bold; statistics are based on values corrected for sex and linear age effects.

Univariate Analyses

For all scales and factor scores the observed MZ Intra-Class Correlations were greater than the DZ correlations (see column 2 and 3 of Table II). Significant greater MZ than DZ correlations suggest the existence of genetic influences. Little difference in the MZ and DZ correlation implicates the existence of shared environmental influences. Overall high MZ correlations indicate small non-shared environmental effects.

The best fitting univariate models are shown in Table II. For all self- and peer reported personality domains the ACE model gave no improvement in fit over the AE model using a likelihood ratio test ($\Delta\chi^2$ (1) = 3.84, $p < 0.05$) and for all dimensions the AE model gave an improvement in fit over the CE model by AIC (the most negative value) and a significantly better fit than an E model. Only for self-reported Neuroticism an alternative model (AIE, Neale & Maes, 2004) allowing for non-additive (epistasis) genetic effects fitted the data significantly better than the standard models. The full ACE model provided a good fit (RMSEA < 0.05) for Support and an acceptable fit (RMSEA < 0.08) for Organization; it fitted the data significantly better than the reduced models.

Multivariate Analyses

Phenotypic correlations are shown in the second column of Table III. All NEO-PI-R domains but Openness were significantly correlated with Support ($p < 0.05$). Neuroticism showed the highest (negative) correlation with Support, but was unrelated with Organization. Self- and peer reports of Openness did not correlate with Support but showed the highest association of all personality traits with Organization. All correlations between personality traits and BEQ-factors supported a positive correspondence between self- and peer reports of personality, with two exceptions. The small correlation between Organization and self-reported Extraversion as well as Conscientiousness differed significantly from zero whereas the corresponding correlations between Organization and peer reports did not. Overall we observed only small correlations between personality and environmental measures, which tended to be even smaller for peer reports.

Table II:

Univariate Modelfitting: Best Fitting Models and Percentage of Phenotypic Variance Explained by Additive Genetic (a^2), Non-Additive Genetic (i^2), Shared Environmental (c^2) and Specific Environmental Effects (e^2)

	ICC		Best fit	Fit Statistic			Effects in %		
Variables	MZ	DZ	model	χ^2 (*df*)	RMSEA	AIC	a^2	c^2/i^2	e^2
NEO-PI-R self-reports of 226 MZ pairs and 168 DZ pairs									
Neuroticism	0.59	0.14	AE	10.30(4)	0.09	2.30	57.9		42.1
			AIE	3.79(3)	0.03	-2.21	25.7	35.7	38.6
Extraversion	0.57	0.27	AE	3.99(4)	0.03	-4.01	56.8		43.2
Openness	0.59	0.28	AE	6.13(4)	0.05	-1.87	58.1		41.9
Agreeableness	0.49	0.26	AE	6.18(4)	0.05	-1.82	50.8		49.2
Conscientiousness	0.56	0.33	AE	6.06(4)	0.05	-1.94	59.9		40.1
Averaged NEO-PI-R peer reports of 178 MZ pairs and 136 DZ pairs									
Neuroticism	0.29	0.07	AE	5.04(4)	0.03	-2.96	28.3		71.7
Extraversion	0.49	0.26	AE	4.09(4)	0.02	-3.91	50.5		49.5
Openness	0.45	0.13	AE	2.39(4)	0.00	-5.61	42.3		57.7
Agreeableness	0.36	0.27	AE	5.01(4)	0.04	-2.99	39.0		61.0
Conscientiousness	0.30	0.11	AE	8.67(4)	0.07	0.67	31.4		68.6
BEQ-factor scores of 210 MZ pairs and 157 DZ pairs									
Support	0.77	0.54	ACE	1.09(3)	0.00	-4.91	47.5	30.1	22.4
Organization	0.64	0.49	ACE	5.37(3)	0.06	-0.63	30.0	33.3	36.7

Note. ICC = Intra-Class Correlation; MZ = monozygotic twins; DZ = dizygotic twins; A = additive genetic influence; C = shared environmental influence; E = specific environmental influence; I = effects of epistasis; the reported models are the best fitting models in *LR*- testing using a critical $\Delta\chi^2(1) = 3.84$ ($p < 0.05$) and smallest AIC.

Univariate model fitting analyses did not reveal any significant effects of shared environment on personality variables, nor were non-additive genetic effects on the BEQ-factors found. Furthermore, Cholesky models including non-additive genetic factors on Neuroticism did not lead to a change in the amount of overlapped genetic variance between N and the BEQ-factors. Thus, we did not consider shared environmental influences and non-additive genetic effects in the result presentation of the relationship between personality and family environment measures. We aimed to identify the model based on the six-variable Cholesky Model with the most negative AIC value that also maintained a RMSEA < 0.08 and did not fit significantly more poorly than a model with more parameters ($\Delta\chi^2(5) = 11.07$, $p < 0.05$).

Additive genetic and non-shared environmental correlations (Table III) were derived from the best fitting model (smallest AIC) presented in Table IV. Nearly all significant observed correlations between self- as well as averaged peer rated personality traits and the BEQ-factors reflected significant genetic correlations (95% confidence interval does not include 0). Several environmental correlations were statistically significant, but were unsystematic with regard to the phenotypic correlations and mostly smaller than the corresponding genetic correlations. In most cases, genetic correlations were lower in magnitude in the analysis of peer rated personality measures compared to self-reports, especially for the personality variables Neuroticism, Agreeableness and Conscientiousness in relationship with

Support. Modeling self-reports exclusively, higher genetic correlations and small but significant non-shared environmental correlations indicated the upper limit of biases which inflates the bivariate phenotypic correlation of self-reports.

Table III:

Phenotypic, Genetic and Environmental Correlations between NEO-PI-R Domain Self Reports as well as Averaged Peer Reports and the Family Environment Factors Support and Organization

	averaged	Genetic Correlations			Environmental Correlations		
NEO-PI-R scales	$r_{P,F}$	r_G	LCL	UCL	r_E	LCL	UCL
				Support			
Self-reports							
Neuroticism	**-0.33**	**-0.57**	-0.79	-0.39	**-0.14**	-0.27	-0.01
Extraversion	**0.25**	**0.39**	0.22	0.59	**0.16**	0.03	0.28
Openness	-0.05	-0.11	-0.30	0.06	0.02	-0.11	0.15
Agreeableness	**0.19**	**0.34**	0.16	0.55	0.10	-0.03	0.22
Conscientiousness	**0.29**	**0.50**	0.33	0.72	0.04	-0.09	0.18
Averaged peer reports							
Neuroticism	**-0.20**	**-0.42**	-0.71	-0.20	0.00		
Extraversion	**0.19**	**0.29**	0.11	0.49	0.00		
Openness	-0.08	**-0.25**	-0.46	-0.06	0.00		
Agreeableness	**0.10**	0.17	-0.02	0.38	0.00		
Conscientiousness	**0.13**	**0.29**	0.07	0.55	0.00		
				Organization			
Self-reports							
Neuroticism	0.03	0.07	-0.13	0.31	0.02	-0.11	0.15
Extraversion	**0.14**	**0.20**	0.00	0.44	**0.14**	0.02	0.27
Openness	**0.23**	**0.38**	0.19	0.68	**0.16**	0.03	0.28
Agreeableness	0.05	0.01	-0.22	0.23	0.11	-0.02	0.23
Conscientiousness	**0.13**	0.18	-0.01	0.42	0.11	-0.02	0.24
Averaged peer reports							
Neuroticism	0.02	0.04	-0.22	0.33	0.00		
Extraversion	0.04	0.10	-0.10	0.34	0.00		
Openness	**0.14**	**0.30**	0.10	0.60	0.00		
Agreeableness	0.02	0.04	-0.19	0.26	0.00		
Conscientiousness	0.07	0.19	-0.05	0.49	0.00		

Note. Statistics based on self-report data of 367 twin pairs (MZ: *N* = 210; DZ: *N* = 157) and on averaged peer report data of 300 twin pairs (MZ: *N* = 169; DZ: *N* = 131) and derived from the best fitting models (smallest AIC) presented in Table IV.; phenotypic correlations ($r_{P,F}$) are averaged Pearson's correlations of subsamples in which one member of each twin pair was randomly selected; significant correlations based on 95% confidence intervals via maximum likelihood are shown in bold; LCL = lower confidence limit; UCL = upper confidence limit.

Table IV presents model fitting statistics and estimates of the absolute magnitude of additive genetic and/or non-shared environmental mediation between all five personality variables and the BEQ-factors. All initial models (model 1) showed at least an acceptable fit (RMSEA < 0.08). Using the LR-test ($\Delta\chi^2(5) = 11.07$, $p < 0.05$) the reduced model allowing only for genetic mediation (model 2) did not fit the data significantly more poorly than the initial full model except for the relationship between self-rated personality and Organization. With respect to the smallest AIC, in the relationship between self-rated personality and Support the full model also showed an increase in fit over the genetic mediation model. Above all relationships, other reduced model modifications (model 3 and 4) did not lead to an improvement in fit over the initial or the genetic mediation model.

Table IV:

Multivariate Modelfitting: Six-Variable Cholesky Decomposition Models and Percentage of Phenotypic Variance Components in BEQ-Factors

						BEQ-factor variance components in %				
						Explained by personality		Specific BEQ-factor variance		
		Fit Statistic								
NEO-PI-R	Model	χ^2	df	RMSEA	AIC	a_C^2	e_C^2	a_R^2	c_R^2	e_R^2
					Support					
Self-reports	**1**	**131.90**	**113**	**0.03**	**-94.10**	**22.3**	**1.0**	**20.9**	**34.7**	**21.1**
	2	**142.11**	**118**	**0.03**	**-93.89**	**27.3**		**14.0**	**37.2**	**21.5**
	3	199.06	118	0.03	-36.94		6.2	27.4	41.5	24.9
	4	276.48	123	0.08	30.48			27.3	45.2	27.5
Averaged	1	146.53	113	0.05	-79.48	17.6	0.4	31.7	25.4	24.9
peer reports	**2**	**148.67**	**118**	**0.04**	**-87.33**	**20.4**		**29.8**	**25.0**	**24.8**
	3	174.62	118	0.06	-61.38		2.1	43.1	28.1	26.7
	4	201.65	123	0.07	-44.35			38.6	30.0	30.4
					Organization					
Self-reports	**1**	**138.14**	**113**	**0.03**	**-87.86**	**8.8**	**2.2**	**27.3**	**27.5**	**34.2**
	2	152.87	118	0.04	-83.13	14.4		17.3	32.2	36.1
	3	161.28	118	0.04	-74.72		5.5	23.8	34.9	35.8
	4	208.64	123	0.06	-37.36			30.7	30.9	38.4
Averaged	1	174.00	113	0.06	-52.00	10.2	1.0	28.8	25.9	34.1
peer reports	**2**	**179.32**	**118**	**0.06**	**-56.68**	**6.8**		**33.5**	**24.8**	**34.9**
	3	185.08	118	0.06	-50.92		0.6	35.7	29.1	34.6
	4	193.86	123	0.06	-52.14			32.5	29.1	38.4

Note. Statistics based on self-report data of 367 twin pairs (MZ: $N = 210$; DZ: $N = 157$) and on averaged peer report data of 300 twin pairs (MZ: $N = 169$; DZ: $N = 131$); best fitting models based on LR-testing ($\Delta\chi^2(5) = 11.07$, $p < 0.05$) and AIC (most smallest) are shown in bold; Model 1: additive genetic and non-shared environmental mediation; Model 2: additive genetic mediation; Model 3: environmental mediation; Model 4: no mediation; a_C^2 and e_C^2 = common additive genetic and non-shared environmental effects between NEO-PI-R scales and BEQ-factors; a_R^2, c_R^2 and e_R^2 = residual additive genetic, shared and non-shared environmental variance components in BEQ-factors.

Within self-reports, we pointed out from the best fitting model ($\Delta\chi^2(5) = 11.07$, $p < 0.05$), that 66% of the additive genetic effects on Support and 24% of the additive genetic variance in Organization could accounted for by heritable variance in personality variables ($a_C^2 / [a_C^2 + a_R^2]$) whereas the non-shared environmental effects on BEQ-factors explained by the corresponding effects on NEO-PI-R scales ($e_C^2 / [e_C^2 + e_R^2]$) were negligible or very small (0% and 6%). That could be replicated in the analyses of the relationship between the averaged peer reports of personality and the BEQ-factors. Indeed, the proportion of additive genetic variance in the BEQ-factors accounted for by additive genetic variance in averaged peer rated personality scales was smaller, 41% of Support and 17% of Organization, and environmental mediations were generally negligible.

Table V:

Multivariate Modelfitting: Best Fit Three-Variable Cholesky Decomposition Models and Percentage of Phenotypic Variance Components in BEQ-Factors

						BEQ-factor variance components in %						
						Explained by personality		Explained by biases		Specific BEQ-factor variance		
NEO-PI-R scales	Best Fit Model	Fit statistic χ^2	*df*	RMSEA	AIC	a_C^2	e_C^2	a_B^2	e_B^2	a_R^2	c_R^2	e_R^2
					Support							
NE	3b	39.42	31	0.04	-22.56	12.6		7.3		27.2	32.5	20.4
EX	2b	28.38	31	0.02	-33.62	6.6			1.1	44.4	27.3	20.6
OP	4b	25.69	32	0.00	-38.31	1.5				53.1	24.4	21.0
AG	3b	17.69	31	0.00	-44.31	2.1		5.0		46.1	25.9	20.9
CO	3b	27.46	31	0.01	-34.54	5.1		14.6		33.9	25.5	20.9
					Organization							
NE	4c	40.03	33	0.03	-25.97					38.8	25.4	35.8
EX	2b	31.94	31	0.03	-30.06	1.0			1.1	38.2	24.4	35.3
OP	2b	34.19	31	0.02	-27.81	5.2			1.4	37.8	20.6	35.0
AG	4c	35.77	33	0.02	-30.23					40.2	24.3	35.5
CO	4b	40.08	32	0.03	-23.92	2.5				40.3	21.5	35.7

Note. NE = Neuroticism; EX = Extraversion; OP = Openness; AG = Agreeableness; CO = Conscientiousness; statistics based on complete data of 300 twin pairs (MZ: $N = 169$; DZ: $N = 131$); best fitting models based on LR-testing ($\Delta\chi^2(1) = 3.84$, $p < 0.05$) and AIC (most smallest); a_C^2 and e_C^2 = common additive genetic and non-shared environmental effects between NEO-PI-R scales and BEQ-factors; a_B^2 and e_B^2 = genetic and non-genetic biases; a_R^2, c_R^2 and e_R^2 = residual additive genetic, shared and non-shared environmental variance components in BEQ-factors; Model 1: genetic and non-genetic bias model; Model 2: Non-genetic bias model; Model 3: genetic bias model; Model 4: unbiased self-reports model; Model a: $a_C^2 = 0$; Model b: $e_C^2 = 0$; Model c: $a_C^2 = e_C^2 = 0$.

The comparison of personality self- with peer reports indicates the effect of genetic and non-genetic biases on self-reports. To estimate the significance of biases, we identified the model with the most negative AIC value that also maintained a RMSEA < 0.08 and did not fit the data significantly more poorly than a model with more parameters ($\Delta\chi^2(1) = 3.84$, $p < 0.05$). The best fitting models are shown in Table V. We found that a genetic bias explains a significant portion of the common additive genetic influences ($a_B^2 / [a_B^2 + a_C^2]$) on the cor-

relations between self-reported Neuroticism (37%), Agreeableness (70%) as well as Conscientiousness (74%) with Support. Additionally, the trivariate model fitting analysis revealed evidence for a small non-genetic bias, which completely accounted for the non-shared environmental relationship (e_B^2/ [e_B^2 + e_C^2]) in self-reports. Furthermore, the models demonstrated the evidence of unbiased overlaps of genetic effects (a_C^2/ [a_B^2 + a_C^2 + a_R^2]) between personality variables and family environmental factors ranging from 3% (Openness) to 27% (Neuroticism) for Support and from 3% (Extraversion) to 12% (Openness) for Organization.

Discussion

Estimating the role of genetic influence on the variance in self- and peer rated personality traits to explain genetic variance in retrospectively recalled family environments as well as examining the role of biases in self-report measures was the major aim of our study. The results support our hypotheses. First, variation in self- and peer reported personality is due to genetic and non-shared environmental effects, whereas accounts of the rearing environments are also affected by shared environmental influences. Second, genetic effects on personality and family environment overlap whereas environmental effects do not, indicating that phenotypic relationships between these measures are genetically mediated. Third, phenotypic correlations between the BEQ-factors and self-reports of personality tended to be higher than the corresponding correlations with peer-reports. The significant differences are due to a self-rater bias, which is particularly influenced by genetic factors.

More specifically, our findings of genetic influences on BEQ-factors Support (48%) and Organization (30%) were consistent with the previous American study (Hur & Bouchard, 1995). In contrast to Hur and Bouchard (1995) who studied twins reared apart, our sample consists of twins reared together who rated identical family environments. These “objectively” shared family environments are reflected in shared environmental effects (30% on Support and of 33% on Organization) and these true shared environmental influences can be interpreted as “effectively” shared experiences (Plomin & Daniels, 1987). The remaining variance was explained by non-shared environmental effects (22% on Support and 37% on Organization) reflecting “effectively” non-shared experiences within twins confounded with measurement error. It is notable that designs of reared- together twins, who assessed the same parental environment (in our German study), and designs of reared-apart twins, who reported different parental environments (Hur & Bouchard, 1995), do not differ in the amount of genetic effects on retrospectively rated family environments. Therefore, the magnitude of genetic influence on retrospective accounts of family environments seems to be invariant across culture (US and Germany) and neither depends on whether twins report the “objectively” shared or non-shared family environments.

According to our results, genetic influences on self- and peer reported personality traits are good candidates to explain a substantial proportion of genetic effects on retrospective accounts of family environment whereas environmental mediations were negligible. Negligible environmental correlations indicate that differential environmental influences originating from the rearing family do not affect differences in adults’ personality. Genes contribute to the correlation of personality traits with Support as well as to the correlation of Extraversion, Openness and Conscientiousness with Organization. These results are consistent with two interpretations. First, genetically predisposed personality affects the perception or interpretation of family environments. Second, the parental environment is asso-

ciated with or reacts to genetically influenced differences in offspring's behavior – perhaps through childhood temperament – which are genetically correlated with the personality in adulthood. For example, ratings of parental support may be causally determined by the individual's genetically predisposed emotional stability. Otherwise, genetically influenced emotional instable offspring may show behaviors and emotional reactions that decrease parental acceptance and support which in turn increase Neuroticism (reactive genotype–environment correlation). With regard to the relationships of Extraversion and Conscientiousness with familial support, a genetically predisposed extraverted individual evokes more familial support and parents react on scrupulous children with more acceptance. The assumption that more extraverted and conscientious individuals *perceive* their family more supportive seems to be less plausible. For the relationship of Openness with Organization, an additional explanation is suitable, namely passive genotype–environment correlation. Familial structure as well as intellectual and cultural orientation correlate with parental openness to experiences which in turn is genetically related to the offspring's Openness. Please note, that our study design cannot disentangle in a testable way between the genetic perception hypothesis and the genotype–environment correlation hypothesis, because environments weren't measured directly.

The joint analyses of self- and peer reports revealed effects of a genetically influenced self-report bias, which increases genetic and as a consequence phenotypic correlations among self-report measures. Small up to moderate portions of genetic relationships between the three personality traits Neuroticism, Agreeableness and Conscientiousness with Support were explained by such a genetic bias. This result is consistent with our hypothesis that genes not only affect how people perceive or interpret their environment but also how they perceive or rate themselves. For example, a positive outlook or a positive self-presentation spuriously inflates the negative correlation between judgments of familial support and self-reports on N as well as the positive correlation of Support with A and C. Similarly, the significant environmental mediation of self-rated Openness and Extraversion with Organization indicates a non-genetic bias. In other words, external (e.g. positive experiences) or internal (e.g. positive mood) situational influences or non-genetic response styles which lead to higher scores in Openness to Experiences and Extraversion also lead to more positive reports in retrospectively recalled family structure and cultural orientation.

The relative high proportion of genetic overlap of Support with N, A, and C due to a genetic bias is not surprisingly. Perhaps, these dimensions are more socially stated than E and O. Common genetic variance in N, A and C (Jang et al., 2006; McCrae et al., 2008) might reflect genetic influences on individually differential response styles (Digman, 1997). Regarding the measures of support, Bergeman et al. (1990) found stronger genetic influence on the assessments of the quality than of the frequency of social support. It is plausible that judgments of the quality of experiences are more influenced by response biases than reports on factual aspects. Since the factor Organization reflects rather factual information about the family than Support, the higher heritability of Support reinforces the conjecture of a genetic bias. As already mentioned, we used the word "bias" for a systematic disagreement among self- and peer reports of personality that may be due to either (1) systematic response errors or (2) true self-rater unique perception. Both, true specificity of rater's viewpoint and specific response behavior, may be shared by twin pairs dependent on their genetic resemblance reflecting a genetic bias on self-reports.

However, a moderate proportion of the genetic variance in Support and a substantial part of the genetic influence on Organization remained unexplained by personality variables. These portions may be accounted for by other psychological characteristics in individuals. It is plausible, for example, that cognitive abilities, which are strongly genetically influenced and correlated between family members, may be good candidates to explain genetic effects on the focused environmental measures particularly on intellectual and cultural orientation in a family. In addition, physiological attributes may also be correlated with parental behaviors and familial influences. For example, mothers of more attractive infants are more affectionate and supportive with their offspring than are mothers of less attractive infants (Langlois et al., 1995).

In sum, the results of our study are consistent with four main conclusions. First, genetic influences lead to individual differences in personality, which in turn influence differences in the description of recalled rearing environments. Second, this can be shown for self- and peer reports of personality and is interpretable as a genetically influenced perception process or genotype–environment correlations (Hur & Bouchard, 1995; Kendler, 2001; Krueger et al., 2003; Herndon et al., 2005). Third, genetic factors affect the impression (specificity of viewpoint) or description (response bias) of both environments and behavioral characteristics, reflected by a genetic self-report bias (McCrae et al., 2008). Fourth, as negligible environmental mediations have been demonstrated, parental support and familial organization seemingly have but modest influence on individual differences in adult personality.

References

Becker, A., Busjahn, A., Faulhaber, H.- D., Bähring, S., Robertson, J., Schuster, H. et al. (1997). Twin zygosity: Automated determination with microsatellites. *Journal of Reproductive Medicine*, *42*, 260-266.

Bergeman, C. S., Plomin, R., Pederson, N. L., McClearn, G. E., & Nesselroade, J. R. (1990). Genetic and environmental influences on social support: The Swedish Adoption Twin Study of Aging (SATSA). *Journal of Gerontology*, *45(3*), 101-106.

Bouchard, T. J., & Loehlin, J. C. (2001). Genes, evolution, and personality. *Behavior Genetics*, *31(3)*, 243-273.

Busjahn, A. (2006). HealthTwiSt: The Berlin Twin Registry. *Twin Research and Human Genetics*, *9*, 778-782.

Chipuer, H. M., Plomin, R., Pederson, N. L., McClearn, G. E., & Nesselroade, J. R. (1993). Genetic influence on family environment: The role of personality. *Developmental Psychology*, *29*, 110-118.

Costa, P.T., Jr., & McCrae, R.R. (1992). *Reviced NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory: Professional manual.* Odessa, FL: Psychological Assessment Resources.

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, *73 (6)*, 1246-1256.

Herndon, R. W., McGue, M., Krueger, R. F., & Iacono, W. (2005). Genetic and environmental influences on adolescents' perceptions of current family environment. *Behavior Genetics*, *35 (4)*, 373-380.

Hur, Y.-M., & Bouchard, T. J., Jr. (1995). Genetic influences on perceptions of childhood family environment: A reared apart twin study. *Child Development*, *66*, 330-345.

Jang, K. L., Livesley, W. J., & Vernon, P. A. (1996). Heritability of Big Five dimensions and their facets: A twin study. *Journal of Personality*, *64*, 577-591.

Jang, K. L., Vernon, P. A., & Livesley, W. J. (2000). Personality disorder traits, family environment, and alcohol misuse: A multivariate behavioural genetic analysis. *Addiction*, *95*, 873-888.

Jang, K. L., Livesley, W. J., Ando, J., Yamagata, S., Suzuki, A., Angleitner, A., et al. (2006). Behavioral genetics of the higher-order factors of the Big Five. *Personality and Individual Differences*, *41*, 261-272.

Kendler, K. S. (2001). Twin studies of psychiatric illness. An Update. *Arch Gen Psychiatry*, *58*, 1005-1014.

Krueger, R. F., Markon, K. E., & Bouchard, T. J., Jr. (2003). The extended genotype: The heritability of personality accounts for the heritability of recalled family environments in twins reared apart. *Journal of Personality*, *71(5)*, 809-834.

Langlois, J. H., Ritter, J. M., Casey, R. J., & Sawin, D. B. (1995). Infant attractiveness predicts maternal behaviours and attitudes. *Developmental Psychology*, *31 (3)*, 464-472.

Lytton, H. (1977). Do parents create or respond to differences in twins? *Developmental Psychology*, *13*, 456-459.

McCrae, R. R., & Costa, P. T., Jr. (1985). Updating Norman's "adequate taxonomy": Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology*, *49*, 710-721.

McCrae, R. R., Jang, K. L., Ando, J., Ono, Y., Yamagata, S., Riemann, R., et al. (2008). Substance and artifact in the higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, *95 (2)*, 442-455.

Moos, R. H., & Moos, B. S. (1986). *Family Environment Scale manual.* Palo Alto, CA: Consulting Psychologists Press.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. M. (2003). *Mx: Statistical Modeling.* VCU Box 900126, Richmond, VA 23298: Department of Psychiatry. 6th edition.

Neale, M. C., & Maes, H. H. M. (2004). *Methodology for genetic studies of twins and families.* Dordrecht: Kluwer Academic Publishers B.V.

O'Connor, T. G., Deater-Deckard, K., Fulker, D., Rutter, M., & Plomin, R. (1998). Genotype–environment correlations in late childhood and early adolescence: Antisocial behavioural problems and coercive parenting. *Developmental Psychology*, *34(5)*, 970-981.

O'Connor, T. G., Hetherington, E. M., Reiss, D., & Plomin, R. (1995). A twin- sibling study of observed parent-adolescent interactions. *Child Development*, *66*, 812-829.

Oniszczenko, W., Angleitner, A., Strelau, J., & Angert, T. (1993). *The Questionnaire of Twins' Physical Resemblance*. Unpublished questionnaire, University of Warsaw, Warsaw, Poland, and University of Bielefeld, Bielefeld, Germany.

Ostendorf, F., & Angleitner, A. (2004). *NEO-PI-R: NEO- Persönlichkeitsinventar nach Costa und McCrae, revidierte Fassung [NEO- Personality Inventory revisited by Costa and McCrae].* Göttingen: Hogrefe.

Plomin, R., & Daniels, D. (1987). Why are children from the same family so different from each other? *The Behavioral and Brain Science*, *10*, 1-16.

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype- environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin*, *84*, 309-322.

Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2001). *Behavioral Genetics* (4th edition). New York: Worth Publishers.

I

Plomin, R., McClearn, G. E., Pederson, N. L., Nesselroade, J. R., & Bergeman, C. S. (1989). Genetic influence on adults' ratings of their current family environment. *Journal of Marriage and the Family*, *51*, 791-803.

Plomin, R., Reiss, D., Hetherington, E. M., & Howe, G. W. (1994). Nature and nurture: Genetic contributions to measures of family environment. *Developmental Psychology*, *30*, 32-43.

Riemann, R., Angleitner, A., & Stelau, J. (1997). Genetic and environmental influences on personality: A study of twins reared together using the self- and peer report NEO-FFI scales. *Journal of Personality, 65(3)*, 449-475.

Riemann, R. & Wagner, R. (2000). *Deutsche Version des Block Environmental Questionnaire nach Hur & Bouchard [German Version of BEQ].* Unpublished Questionnaire.

Rowe, D. C. (1981). Environmental and genetic influences on dimensions of perceived parenting: A twin study. *Developmental Psychology*, *17*, 203-208.

Rowe, D. C. (1983). A biometrical analysis of perceptions of family environment: A study of twin and singleton sibling kinships. *Child Development*, *54*, 416-423.

Rutter, M., Moffitt, T. E., & Caspi, A. (2006). Gene–environment interplay and psychopathology: multiple varieties but real effects. *Journal of Child Psychology and Psychiatry*, *47(3/4)*, 226-261.

Rutter, M., & Silberg, J. (2002). Gene–environment interplay in relation to emotional and behavioural disturbance. *Annu. Rev. Psychol.*, *53*, 463-490.

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype–environment effects. *Child Development*, *54*, 424-435.

Spinath, F. M., Angleitner, A., Borkenau, P., Riemann, R., & Wolf, H. (2002). German observational study of adult twins (GOSAT): A multimodal investigation of personality, temperament and cognitive ability. *Twin Research*, *5*, 372-375.

Stößel, K., Kämpfe, N., & Riemann, R. (2006). The Jena Twin Registry and the Jena Twin Study of Social Attitudes (JeTSSA). *Twin Research*, *9 (6)*, 783-786.

Tellegen, A. (2000). *Manual of the Multidimensional Personality Questionnaire*. University of Minnesota Press: Minneapolis, MN.

Vernon, P. A., Jang, K. L., Harris, J. A., & McCarthy, J. M. (1997). Environmental predictors of personality differences: A Twin and Sibling Study. *Journal of Personality and Social Psychology*, *72 (1)*, 177-183.

Yamagata, S., Suzuki, A., Ando, J., Ono, Y., Kijima, N., Yoshimura, K., et al. (2006). Is the genetic structure of human personality universal? A cross-cultural twin study from North America, Europe, and Asia. *Journal of Personality and Social Psychology*, *90*, 987-998.

ANHANG II

Journal of Personality

(2010)

II

Sources of Variance in Personality Facets: A Multiple-Rater Twin Study of Self-Peer, Peer-Peer, and Self-Self (Dis-)Agreement

Christian Kandler

Rainer Riemann

Frank M. Spinath

Alois Angleitner

Abstract. This study considered validity of the personality structure based on the five-factor model using both self- and peer reports on twins' NEO-PI-R facets. Separating common from specific genetic variance in self- and peer reports, this study examined genetic substance of different trait-levels and rater-specific perspectives relating to personality judgments. Data of 919 twin pairs were analyzed using a multiple-rater twin model to disentangle genetic and environmental effects on domain-level trait, facet-specific trait, and rater-specific variance. About two third of both the domain-level trait variance and the facet-specific trait variance were attributable to genetic factors. This suggests that the more personality is measured accurately the better these measures reflect the genetic structure. Specific variance in self- and peer reports also showed modest up to substantial genetic influence. This may indicate genetically influenced self-rater biases but also substance components specific for self- and peer raters' perspectives on traits actually measured.

II

Introduction

Few behavioral genetic studies have focused on personality traits measured with two or more indicators like facet scales (Jang, Livesley, Angleitner, Riemann, & Vernon, 2002) or multiple raters such as self- and peer judges (Riemann, Angleitner, & Strelau, 1997). No study has combined the advantages of both so far, although a major advantage of multiple indicators and multiple raters lies in the disentanglement of consensually valid, method, and unsystematic residual variance. Variance can be considered 'valid' to the degree to which different raters' assessments converge (Campbell & Fiske, 1959). Previous research on the accuracy of personality judgment has already shown that quantity and quality of personality-relevant information yield an increase of self-other correlation supporting the validity of common variance of self-rater and well-informed rater judgments on personality (e.g., Blackman & Funder, 1998; Letzring, Wells, & Funder, 2006). However, further studies have shown that rater-specific components may provide incremental validity (e.g., Kraemer, Measelle, Ablow, Essex, Boyce, et al., 2003; Vazire & Mehl, 2008).

Multiple-rater twin studies offer additional insight into the sources of both common (also referred to as self-peer agreement or trait variance) and method variance (also referred to as self-peer disagreement or rater-specific variance). They have already illustrated the usefulness of genetically informative data to improve the interpretation of common and unique variance in parental, teachers', and children's reports on psychopathological traits in childhood (e.g., Bartels, Boomsma, Hudziak, Beijsterveldt, & Oord, 2007; Simonoff, Pickles, Hewitt, Silberg, Rutter, et al., 1995). The identification and the size of genetic influences on both common variance and specific variance in self-reports and well- informed peer reports on adult personality can have important implications for how to interpret trait and method variance and thus for the research on accuracy in personality judgments.

This study examined the sources of common and specific trait variance in personality facet scores reflecting different levels of trait generality based on the Five-Factor Model of personality (FFM; McCrae & John, 1992). That is, this study extended previous studies on the hierarchical structure of personality (Jang, McCrae, Angleitner, Riemann, & Livesley, 1998) since it analyzed the sources of common variance of self- and peer reports. Furthermore, genetic and non-genetic sources of specific variance in self- and peer judgments were estimated in order to consider genetic substance for these variance components.

Self-Rater Twin Studies of the FFM of Personality

The FFM (McCrae & Costa, 1987; McCrae & John, 1992) describes the personality structure on five broad traits: Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. Although the five domains are assumed to have a genetic basis (McCrae, Costa, Hrebícková, Ostendorf, Angleitner et al., 2000; McCrae & Costa, 2008), several twin studies in different cultures (e.g., Jang, Livesley, & Vernon, 1996) found only about half of the variance in personality domains to be genetically influenced. The remaining variance is due to environmental influences not shared by family members and measurement error. Environmental effects shared by family members do not significantly contribute to individual differences on the domain level (see Bouchard & Loehlin, 2001, for a review). Jang et al. (1998) focused on the proposed hierarchical structure of personality domains (Costa & McCrae, 1992) and found that variance in residual scores of nearly all facets (controlled for domain variance) showed genetic effects (on average 26%). This suggested that also more specific facet-level traits are also discrete personality constructs with a genetic basis.

Although most of the facet-specific trait variance was due to non-shared environmental effects, it must be noted that in behavioral genetic studies of single self-report measures estimates of non-shared environmental effects are confounded with random measurement error. In terms of classical test theory, random error variance is defined to be neither consistent across indicators of the same construct nor consistent over different methods of measurement, and thus does not correlate within twin pairs. Therefore, random error variance cannot to be affected by genetic factors. A consequence for mono-indicator and (or) mono-rater studies is that estimates of non-shared environmental variance are overestimated. Consequently, correcting for unreliability, heritability of facet residual scores increased from 26% to 47% (Jang et al., 1998).

Loehlin, McCrae, Costa, and John (1998) analyzed three self-report measures (adjectives, rating scales, and questionnaires) on the domain-level and found that after controlling for measurement-specific effects, on average 55% of the variance in *latent* phenotypes was due to genetic effects. The remaining variance was attributable to non-shared environmental influences. With regard to the facet-level scales as indicators for domain factors, Yamagata, Suzuki, Ando, Ono, Kijima et al. (2006) utilized factor analyses on twin data and found that across different cultures the phenotypic five-factor structure of personality is the result of a genetic five-factor structure rather than of an environmental factor structure. That is, phenotypic correlations among facet scores primarily reflect genetic correlations providing "evidence that the FFM reflects underlying genetic structure" (p. 996).

Though the application of multiple indicators provides estimates corrected for random error these studies have been limited to their reliance on self-report measures only. It is well known that mono-rater measures of personality constructs are susceptible to random and systematic error components (Campbell & Fiske, 1959). While random error variance reduces the reliability of measurement, systematic error variance increases reliability. The domain scales of the FFM are highly reliable (about $\alpha = .90$) indicating small random error. Convergent validity estimated via self-peer correlations is substantially lower (ranging between .40 and .60). Because validity is lower than reliability, rater biases (e.g., self-enhancement or self-deception) or insufficient self-perception may play a substantial role in personality judgments (Paulhus & John, 1998). Thus, self-reports might be distorted because of the subjects' unwillingness or incapability to impart necessary and valid data.[12]

Multiple-Rater Twin Studies of the FFM of Personality

Surprisingly few behavioral genetic studies on adult personality have utilized data beyond self reports (see Borkenau, Riemann, Angleitner, & Spinath, 2000, for a review of observational studies). One observational study of adult twins[13] (Borkenau, Riemann, An-

[12] The focus on the NEO-PI-R structure should not arouse the impression that we wanted to ignore or belittle other hierarchical models of personality. The parallels between different models appear to overbalance the varieties (see Bouchard & Loehlin, 2001; Markon, Krueger, & Watson, 2005).

[13] In contrast to other studies, Borkenau et al. (2001) found evidence of environmental effects shared by family members (even if not consistently significant) suggesting that contrast effects reduce correlations between fraternal twins in self-report studies. However, the analysis of reports by independent well-informed peer raters for each twin sibling in the same study and in another (Riemann et al., 1997) yielded no indication of contrast effects. Since each twin pair arrived at and was tested in the laboratory to the same time on the same day, it is also plausible that situational effects (Kraemer et al., 2003) shared by the observers as well as twins could have caused artificially higher MZ and DZ correlations leading to estimates of small shared environmental effects.

gleitner, & Spinath, 2001) estimated genetic effects (46%) on video-based ratings of personality traits to be of similar magnitude compared to findings from self-report studies. Similarly, the sources of variance in personality scale scores obtained from reports of acquaintances are comparable to self-reports (Angleitner, Riemann, & Strelau, 1995): genetic and non-shared environmental effects explain about half of the total variance.

When personality traits were modeled as latent variables across self- and peer reports, phenotypic variance was more attributable to genetic effects (66% to 79%; Riemann et al., 1997). The remaining variance of common variance in personality judgments was due to non-shared environmental effects. A factor analysis of facet scores among twins' self- and spouses' reports (McCrae, Jang, Livesley, Riemann, & Angleitner, 2001) showed that the five personality domains appear to be "almost entirely the result of genetic influences" (p. 530). Since self-peer covariance reflects convergent validity (Campbell & Fiske, 1959) and thus variance in personality measures which is free from rater-specific variance as well as random error, these studies showed that the more personality factors are measured accurately the more these reflect genetic factors.

Previous twin studies on data from multiple raters (e.g., Riemann et al., 1997) have been limited to the examination of genetic influences on the domain level of personality structure. Self-report studies have found clear evidence of genetic variance in most facet residuals corrected for the domain variance (Jang et al., 1998; Jang et al., 2002). To our knowledge, this finding has not been validated on data from both self- and peer raters.

Rater-Specific Perspectives in Personality Judgments

Multiple-rater twin studies on adult personality have usually relied on the assumptions that *not only* random error *but also* rater-specific variance are not influenced by genetic factors. This unproved assumption seems questionable, because self-enhancement and self-deception or other rater biases (e.g., response styles) might be affected by genetic factors. Other heritable individual traits, not the traits actually measured, may bias judgments on one's own personality. For example, Narcissism and Machiavellianism are heritable (Vernon, Villani, Vickers, & Harris, 2008) and may systematically distort self-ratings resulting in more socially desired self-descriptions or self-deception. As a consequence, specific variance in self-reports may be influenced by genetic factors. Based on these considerations, it is not surprising that a recent study, relying partially on the same data set that was analyzed here, found genetic effects on self-report-specific variance in latent higher-order factors of the Big Five (McCrae, Yamagata, Jang, Riemann, Ando, et al., 2008).[14]

Rater biases should not correlate between independent informants who rate only one sibling of a twin pair, and should thus not contribute to twin correlations unless different peer raters share inaccurate stereotypes (Funder, 1995; Letzring et al., 2006). These stereotypes may be present when others base their judgments on a limited set of cues (Blackman & Funder, 1998) such as appearance. For example, a target's tattoos or shorn hair may lead to a judgment that the person is less agreeable. It is well known that identical twins look more alike than fraternal twins. Therefore, such stereotypes shared by different raters may artificially produce estimates of genetic effects on peer-rater-specific variance. Howev-

[14] The previous study (McCrae et al., 2008) focused on the higher-order structure of the Big Five. Our study focused on domain-level and facet-specific traits. In addition, we allowed for estimates of genetic effects on self- *as well as* peer-rater-specific variance, but both controlled for random error.

er, effects due to stereotypes should decrease with increasing availability of relevant cues to make good judgments (Funder, 1995; Kenny, 1991). This should be warranted for ratings on target's personality by acquaintances who know the target person very well (Funder, Kolar, & Blackman, 1995). Consequently, in line with the bias explanation rater-specific variance in independent well-informed acquaintances' reports on twins' personality should not show genetic effects.

Otherwise, different informants capture different perspectives and may thus have privileged access to different trait-relevant information (Kraemer et al., 2003). Recent research provided evidence that both the self- and other's perspective possess unique insight into the target's behavior (Vazire & Mehl, 2008). Self- and peer raters might also form their personality assessments on partially different sets of behavioral cues, for example, when self-raters rather base their assessments on internal aspects of personality (e.g., motives, feelings, self-concept) and when observers rather base their ratings on expressed behavior or social consequences of behavior (e.g., social reputation). These different aspects of personality might be genetically influenced. In line with this explanation, self- *as well as* peer-rater-specific variance may contain substance of the traits actually measured and may thus show specific genetic variance. A number of behavioral genetic studies on psychopathology support this position (e.g., Bartels et al., 2007; Simonoff et al., 1995; van der Valk, van den Oord, Verhulst, & Boomsma, 2003) showing genetic components of unique variance in mothers' and fathers' reports on traits of their children.

Aims of the Present Study

The present study combined a multiple-indicator (Jang et al., 2002) and a multiple-rater twin model (Riemann et al., 1997) which was extended to a hierarchical model of personality constructs based on the FFM. On the one hand, this model disentangled genetic and environmental effects on common as well as specific trait variance in personality facets. As a consequence, the model allowed us to provide evidence of the hypothesis that the more personality traits are measured accurately the more these reflect genetic factors *not only* on broad personality dimensions (Riemann et al., 1997) *but also* on facet-specific traits, since systematic and random error can be controlled (Jang et al., 1998). In other words, our study examined genetic substance of self-peer agreement in personality facet judgments. If convergent valid variance reflects more accurate variance in heritable personality traits, heritability estimates should be much larger than proposed in mono-rater studies.

On the other hand, the model allowed for the separation of genetic and non-genetic method factors to examine the sources of rater-specific variance. Thus, our model shares the basic idea to use behavioral genetic data for the study of rater-specific variance with previous multiple-rater twin models on psychopathology (e.g., Bartels et al., 2007). But it should be noted that these models are based on assessments of both twin siblings by the same observer (e.g., mother and father rate their twin children) while our model requires independent measures for each twin sibling (i.e., different peer raters for each twin sibling). Our study extended previous multiple-rater twin studies of personality (e.g., McCrae et al., 2001, 2008; Riemann et al., 1997) to the estimation of genetic and non-genetic factors on self- *as well as* peer-rater-specific variance in order to provide validity for different rater-specific perspectives relating to personality. Specific variance in *both* self- *and* peer raters should show significant levels of heritability if these components reflect substance.

II

Method

Participants

Data from two German twin studies was combined for the present investigation: the sample (N = 844) from the third wave of the Bielefeld Longitudinal Study of Adult Twins (BiLSAT; Spinath, Angleitner, Borkenau, Riemann, & Wolf, 2002) and the sample (N = 869) of the Jena Twin Study of Social Attitudes (JeTSSA; Stößel, Kämpfe, & Riemann, 2006). Data from participants in the JeTSSA, who already took part in the BiLSAT (N = 98), were removed, unless more complete data were available in the JeTSSA. In contrast to a previous analysis, using a combined sample of BiLSAT and JeTSSA (McCrae et al., 2008), we included data from unmatched twin pairs (UM), when data were available for one twin sibling without zygosity diagnosis (N = 81), because in an unmatched twin group it is not relevant whether the individual is a part of MZ or DZ twins. The resulting sample consisted of 1,615 individuals from 919 twin pairs (433 MZ, 263 DZ, and 223 UM). Participants ranged in age from 17 to 82 years (M = 36.3, SD = 13.1), 1,243 (77%) of them were females. Although the sample was not representative with regard to zygosity and sex, it was heterogeneous with regard to education and occupational status.

Twins were instructed to ask acquaintances, who knew them but not their twin sibling very well, to provide the peer judgments. That is, we used different peer reports for each sibling of a twin pair. Most peers were friends and spouses. In 92.6% of cases at least one peer report was available (MZ: 92.5%, DZ: 92.2%, UM: 93.7%), and for 86.1% of participants two independent peers provided complete data (MZ: 85.9%, DZ: 86.5%, UM: 86.1%).

Measures

Twins completed the self-report version of the German NEO Personality Inventory Revised (NEO-PI-R; Costa & McCrae, 1992; Ostendorf & Angleitner, 2004) and peers completed the third person version (Form R). The NEO-PI-R consists of 240 items designed for measuring personality on 30 facet scales and five domain scales. Each domain scales encompasses six facet scales (see Table 1). Detailed characteristics of the facets and their psychometric quality are available (Costa & McCrae, 1992; Ostendorf & Angleitner, 2004).

In the current study, internal consistencies (Cronbach's alpha) of facet scales ranged from α = .48 (Openness to Values) to α = .83 (Anxiety) for self-reports (averaged α = .69), from α = .49 (Openness to Values) to α = .84 (Self-Discipline) for single peer reports (averaged α = .71), and from α = .52 (Openness to Values) to α = .86 (Self-Discipline) for mean peer reports (averaged α = .77). Both peer-peer correlations and correlations between self-reports and randomly selected peer reports across facet scales are on average r = .38, ranging between r = .21 (Straightforwardness) and r = .54 (Assertiveness). Correlations between self- and mean peer reports across facets come up to the average of r = .44, ranging between r = .27 (Straightforwardness) and r = .59 (Assertiveness). Because different observers are unlikely to share the same biases, the aggregation of peer judgments reduces both random and systematic error (Hofstee, 1994) and was thus used as reference of personality supplied by well-informed observers. We used the raw scores of self- and averaged peer reports on the 30 facet scales for the subsequent analyses.

As the existence of age and sex effects can bias the twin correlation, self- and mean peer reports of the facet scales were adjusted for sex and linear as well as quadratic age ef-

fects using a regression procedure. Regression residuals were used in subsequent analyses. We did not use residuals with unit variance, because observed variances in self- and averaged peer judgments can differ due to the fact that aggregation typically reduces random error and rater-specific method variance. Self-peer correlations within individuals as well as mono- and hetero-method correlations within twins were estimated for each personality variable and subsample (e.g., MZ, DZ, and UM) using an Expectation Maximization (EM) algorithm (Little & Rubin, 2002) for handling missing values (Table 1).

Table 1:

The EM-Estimated Self-Self, Peer-Peer, and Self-Peer Correlations

Personality facets	Hetero-method correlations					Mono-method twin correlations				Hetero-method twin correlations			
	$Self_X$ - $Peer_X$			$Self_Y$ - $Peer_Y$		$Self_X$ - $Self_Y$		$Peer_X$ - $Peer_Y$		$Self_X$ - $Peer_Y$		$Self_Y$ - $Peer_X$	
	MZ	DZ	UM	MZ	DZ	MZ	DZ	MZ	DZ	MZ	DZ	MZ	DZ
Anxiety	.36	.44	.56	.44	.59	.46	.20	.35	.13	.27	.14	.21	.08
Angry Hostility	.40	.44	.45	.49	.54	.45	.17	.41	.20	.31	.16	.26	.13
Depression	.45	.49	.57	.46	.56	.42	.17	.28	.13	.25	.16	.23	.01
Self-Consciousness	.34	.40	.42	.31	.46	.49	.24	.36	.12	.26	.14	.28	.04
Impulsivity	.30	.37	.44	.41	.43	.46	.14	.35	.19	.31	.11	.22	.04
Vulnerability	.40	.46	.41	.45	.50	.42	.24	.28	.12	.21	.13	.16	.05
Warmth	.51	.41	.35	.53	.52	.46	.20	.38	.21	.37	.13	.37	.16
Gregariousness	.56	.53	.52	.60	.61	.50	.30	.52	.32	.43	.37	.45	.21
Assertiveness	.61	.54	.57	.62	.66	.49	.22	.38	.12	.40	.14	.37	.09
Activity	.45	.51	.51	.46	.57	.44	.14	.40	.15	.29	.12	.25	.18
Excitement-Seeking	.51	.57	.55	.57	.63	.44	.27	.50	.38	.34	.30	.38	.27
Positive Emotions	.52	.43	.47	.50	.51	.44	.22	.34	.19	.35	.21	.32	.09
O. to Fantasy	.25	.36	.31	.42	.44	.40	.16	.36	.08	.20	.08	.21	.13
O. to Aestethics	.56	.59	.50	.55	.56	.57	.27	.50	.15	.45	.29	.46	.18
O. to Feelings	.35	.45	.38	.41	.49	.41	.22	.35	.18	.22	.14	.23	.06
O. to Actions	.47	.48	.53	.50	.47	.46	.19	.39	.14	.32	.20	.39	.18
O. to Ideas	.52	.55	.46	.51	.46	.49	.16	.48	.20	.39	.22	.43	.19
O. to Values	.35	.39	.27	.39	.38	.47	.32	.31	.22	.21	.15	.34	.17
Trust	.35	.31	.32	.44	.32	.50	.24	.31	.18	.26	.10	.23	.20
Straightforwardness	.26	.32	.19	.23	.37	.34	.22	.34	.22	.16	.11	.24	.18
Altruism	.35	.32	.37	.40	.49	.47	.16	.36	.16	.32	.14	.26	.14
Compliance	.44	.51	.45	.45	.50	.37	.18	.46	.26	.31	.12	.32	.24
Modesty	.31	.43	.42	.36	.51	.43	.23	.32	.15	.25	.10	.24	.08
Tender-Mindedness	.37	.39	.34	.39	.38	.48	.27	.35	.26	.29	.31	.29	.25
Competence	.34	.42	.33	.37	.50	.45	.24	.34	.15	.29	.14	.22	.18
Order	.49	.53	.57	.55	.55	.44	.12	.39	.11	.41	.15	.30	.00
Dutifulness	.44	.44	.40	.32	.43	.39	.27	.36	.15	.28	.21	.29	.11
Achievement Striving	.45	.49	.46	.44	.50	.45	.32	.34	.07	.32	.20	.36	.22
Self-Discipline	.51	.47	.45	.40	.57	.48	.26	.40	.10	.33	.13	.31	.03
Deliberation	.38	.41	.46	.42	.51	.41	.13	.35	.17	.24	.19	.25	.09

Note. *X* and *Y* = randomly assigned twin and co-twin sibling samples, MZ = monozygotic twins; DZ = dizygotic twins; UM = unmatched twin group.

Structural Equation Modeling

As already mentioned, we combined a multiple-indicator and a multiple-rater twin model (MIMRT). Such a model consists of at least two indicators for each combination of a method (self- or peer reports) and two targets (each sibling of a twin pair). The minimal model is presented in Figure 1. Data from multiple informants in a study of twins does not only allow the analysis of correlations among self- (index *S*) and among peer reports (index *P*) but also the analysis of self-peer within-individual (providing convergent validity) and cross-twin correlations (providing genetic substance and convergent validity). Specifically, the variance of the observed variables *X* (phenotype of one twin sibling) and *Y* (phenotype of the other) can be decomposed into a trait component (*T*), a method component (*M*), and a random error component (ε). In our application, the trait factor (*T*), which results from self-peer correlations, reflected convergent valid variance of a personality construct (Campbell & Fiske, 1959), also called universe variance in generalizability theory (Cronbach, Gleser, Nanda, & Rajaratnam, 1972), or simply target variance (Hoyt, 2000). The method factors (M_S and M_P) reflected specific variance for self- and mean peer reports.

As the design consisted of twins reared together, each trait and each method factor could in turn be decomposed into an additive genetic component (*A*), an environmental component shared by twins (*C*), and an environmental component not shared by twins (*E*). The additive genetic component (*A*) reflects effects of combined genes that can be simply added together. The shared environmental component (*C*) reflects common environmental effects for twins, whereas the non-shared component (*E*) reflects individually specific factors. This twin model assumes the absence of non-additive genetic effects, assortative mating of twins' parents, gene-environment interaction and gene-environment correlation. Shared (*C*) and non-shared (*E*) environmental effects were assumed to be equal across MZ and DZ twins (see Neale & Maes, 2004, for more details).

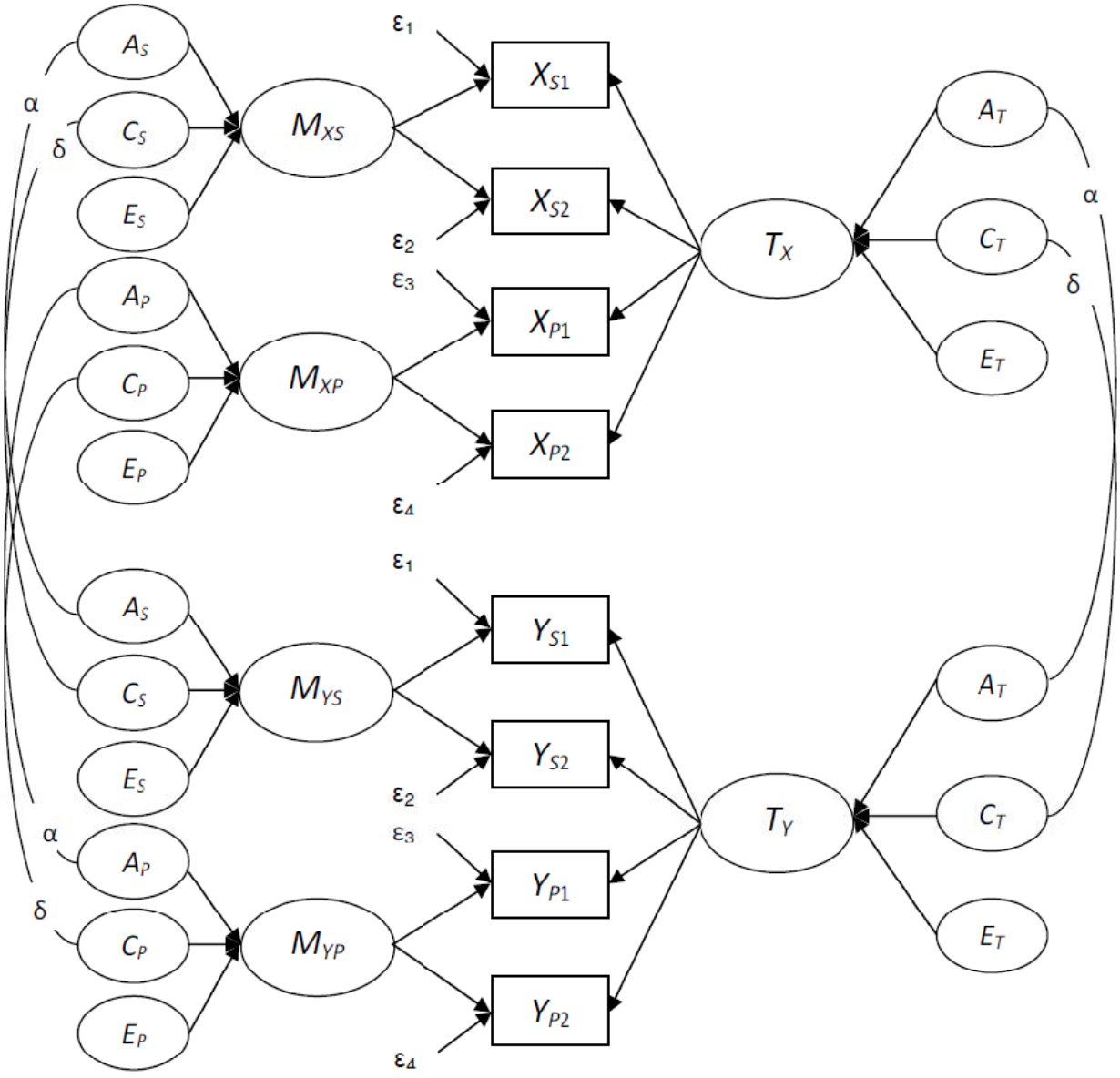
Hetero-method correlations within individuals (*X* or *Y*, see Table 1) indicate (convergent valid) trait components (*T*, see Figure 1). Larger MZ than DZ twin correlations across self- and peer reports would provide evidence for genetic influences on that trait (A_T). No or small differences in hetero-method twin correlations between MZ and DZ twins would indicate effects of shared environmental factors (C_T). Non-shared environmental influences on trait variance (E_T) would be inferred from smaller hetero-method MZ twin correlations compared to hetero-method correlations within individuals.

If mono-method (within-informant: self-self and peer-peer) twin correlations are larger than the corresponding hetero-method (across-informant: self-peer) twin correlations (see Table 1), the difference can be attributed to rater-specific effects (M_S and M_P, see Figure 1). Moreover, if mono-method twin correlations, corrected for hetero-method twin correlations, are larger for MZ than for DZ twins, there should be genetic influence on rater-specific factors (A_S, A_P). No or small differences would suggest the influence of C_S and C_P factors reflecting non-genetic rater-specific effects shared within self- and within peer raters. These possibly indicate method-specific effects within the self-report or peer report method. In an analogous manner, the E_S and E_P factors (non-genetic method effects not shared by different raters within the same method) indicate rater-specific influences independent of the self- or peer report method and would be inferred from smaller hetero-method MZ twin correlations compared to mono-method MZ twin correlation. Finally, the model (Figure 1) separates random measurement error (ε) from systematic variance components. In sum, the MIMRT model decomposes reliably measured variance into convergent valid genetic and

environmental trait components as well as self- and peer- report-specific variance components due to genetic and non-genetic influences.

Figure 1:

Multiple-Indicator Multiple-Rater Twin (MIMRT) Model of Self- (*S*) and Peer Reports (*P*)



Note. For simplicity, only two indicators (1 and 2) are shown for every trait-rater-twin combination. *T* = convergent valid trait factor; *M* = method factor, reflecting rater-specific variance; ε = measurement error; *X* = observed variable of one twin sibling; *Y* = observed variable of co-twin; α = 1.0 for MZ twins and 0.5 for DZ twins; δ = 1.0 for all twins; *A* = additive genetic factors; *C* = shared environmental factors; *E* = non-shared environmental factors.

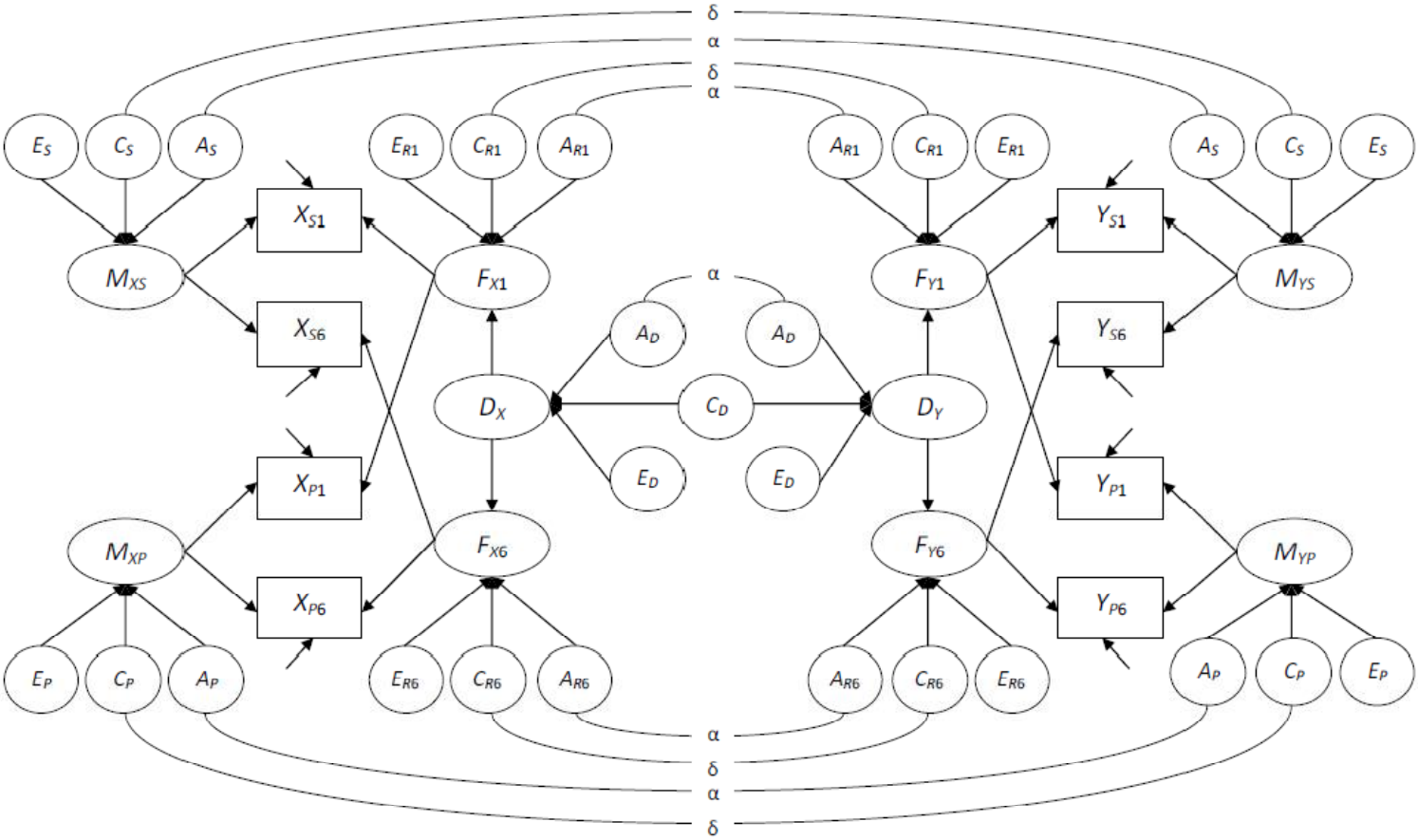
In the present study, the minimal MIMRT model (Figure 1) was extended to a hierarchical model of personality constructs based on the NEO-PI-R structure to examine genetic and environmental effects on domain-level and facet-specific trait variance (Figure 2) named as hierarchical trait multiple-rater twin (HTMRT) model. This model permitted the estimation of the sources of individual differences in latent personality domains (*D*) based on the inter

correlations of the corresponding six facet scales across multiple raters (*F*). In addition, the model allowed us to examine the sources of variance in facet trait residuals (*R*), adjusted for domain, rater-specific, and random error variance. That is, the model provided more accurate estimates of genetic (A_D) and environmental factors (C_D, E_D) on latent personality domain traits as well as estimates of genetic (A_R) and environmental influences (C_R, E_R) on facet-specific traits. Finally, the model allowed the separation of genetic (A_S, A_P) and non-genetic method factors to examine the sources of rater-specific variance which is common among twin sibling self-raters (C_S) and among independent peer raters (C_P) as well as not shared by different raters (E_S, E_P).

MZ, DZ, and UM twins' variance-covariance-matrices of corrected facet scores were estimated using an EM algorithm (Little & Rubin, 2002) and analyzed by fitting five HTMRT models (Figure 2), one for each personality domain. As twin siblings in the unmatched twin group were not available, we examined only the variance-covariance-matrices within individuals for that group and hence fitted an appropriately reduced model (one half of the complete twin model presented in Figure 2) to the matrices of this group. The additional modeling of the UM group increased power to detect small phenotypic variance components.

Figure 2:

Hierarchical Trait Multiple-Rater Twin (HTMRT) Model of Self- (*S*) and Peer Reports (*P*)



Note. For simplicity only two facet scores (1 and 6) are shown for every domain-method-twin combination. *D* = convergent valid domain factor; *F* = convergent valid facet factor; *R* = convergent valid facet trait residuals; *M* = method factor, reflecting rater-specific variance; empty arrows = measurement errors; *X* = observed variable of one twin sibling; *Y* = observed variable of co-twin; α = 1.0 for MZ twins and 0.5 for DZ twins; δ = 1.0 for all twins; *A* = additive genetic factors; *C* = shared environmental factors; *E* = non-shared environmental factors.

For model identification we used the path coefficient approach (Neale & Maes, 2004): variances of all exogenous (unexplained) latent variables were fixed to one in order to obtain estimates for all paths in the model. All phenotypic first-level paths from the latent variables *F* and *M* to the manifest variables (rectangles in Figure 2) were fixed to be equal within each facet and method. As personality facets can be differently indicative for personality domains (Yamagata et al., 2006) we tested a less constrained model with unequal second-level loadings on *D* and compared it against the more restrictive model with equal loadings. All genetic and environmental paths were equalized across twins (X and Y) and groups (MZ, DZ, and UM). That is, we decomposed each of the phenotypic variance components (except random error ε) into genetic and non-genetic components.

The structural equation models were fitted to the variance-covariance-matrices via maximum likelihood using the statistical software package AMOS 17.0 (Arbuckle, 2007). We evaluated the overall model fit by the Root Mean Square Error of Approximation (RMSEA; Steiger, 1990) in conjunction with its 90% confidence interval. Lower limit values of RMSEA < 0.05 indicate a good fit (Browne & Cudeck, 1993). We compared more parsimonious models against less constrained models using the χ^2- difference test and the parsimony comparative fit index (PCFI), which is a parsimony adjustment to the CFI (Bentler, 1993). A higher PCFI indicates a better fit (values ranged from 0 to 1). To evaluate the significance of all parameters we used the critical ratio (C.R.) for each path coefficient.

Results

For all personality dimensions the model with unequal *D*-factor loadings showed a good overall fit: Neuroticism: $\chi^2(634) = 1543.31$, RMSEA = .040, 90%CI[.037-.042], PCFI = .942; Extraversion: $\chi^2(634) = 1795.01$, RMSEA = .045, 90%CI[.042-.047], PCFI = .894; Openness: $\chi^2(634) = 1435.32$, RMSEA = .037, 90%CI[.035-.040], PCFI = .905; Agreeableness: $\chi^2(634) = 1487.44$, RMSEA = .038, 90%CI[.036-.041], PCFI = .897; Conscientiousness: $\chi^2(634) = 1330.05$, RMSEA = .035, 90%CI[.032-.037], PCFI = .956. All upper limit RMSEA 90% confidence intervals were lower than .05. The model fitted the data significantly better than the restricted model (equal loadings) indicated by a higher PCFI and a significant χ^2- difference ($\Delta\chi^2(5) > 11.07$, $p \leq .05$): Neuroticism: $\Delta\chi^2(5) = 900.47$, PCFI = .866; Extraversion: $\Delta\chi^2(5) = 241.04$, PCFI = .874; Openness: $\Delta\chi^2(5) = 470.60$, PCFI = .842; Agreeableness: $\Delta\chi^2(5) = 210.56$, PCFI = .873; Conscientiousness: $\Delta\chi^2(5) = 304.60$, PCFI = .903.

Phenotypic parameters derived from the best fitting model are presented in Table 2. As we wanted to separate unsystematic measurement error (ε) from systematic (reliable) individual differences that are due to personality facet variance (facet consistency) and the respective rater-specific variance (method specificity), reliability coefficients (*Rel*) were computed as the amount of variance of an observed variable which is explained by the facet trait and method factors. Across all facet variables, the averaged reliability for self-reports was *Rel* = .56 and *Rel* = .69 for mean peer reports. The difference indicated that mean peer reports were more reliable than self-reports. The reliabilities implied in the model were smaller than the Cronbach's Alphas (on average α = .69 for self-reports and α = .77 for mean peer reports) indicating the presence of facet-specific method influences which were confounded with random error (1 – *Rel*). It is important to note that method factors (rater-specific variance) were computed as *domain-specific*. That is, measurement error and *facet-specific* (indicator-specific) method effects (Eid, Lischetzke, Nussbeck, & Trierweiler, 2003) could not

be separated in our model. Therefore, our model implies reliabilities of manifest variables which could be smaller than the internal consistencies of the observed variables that include *facet-specific* method effects.

II

Table 2:

Hierarchical Trait Multiple-Rater Twin Models: Reliabilities, Consistency, Method, and Facet Specificity Coefficients

Manifest	Self-report			Mean peer report			Facet trait score	
variables	*Rel*	Con_F	Spe_M	*Rel*	Con_F	Spe_M	Con_D	Spe_F
				NEUROTICISM				
Anxiety	.64	.40	.24	.82	.62	.20	.78	.22
Angry Hostility	.69	.40	.29	.72	.53	.19	.43	.57
Depression	.69	.46	.23	.82	.65	.17	.95	.05
Self-Consciousness	.60	.30	.30	.71	.47	.24	.61	.39
Impulsivity	.48	.18	.30	.56	.30	.26	.02	.98
Vulnerability	.71	.39	.32	.77	.54	.23	.86	.14
				EXTRAVERSION				
Warmth	.63	.43	.20	.69	.57	.12	.80	.20
Gregariousness	.60	.47	.13	.75	.67	.08	.51	.49
Assertiveness	.66	.54	.12	.71	.64	.07	.25	.75
Activity	.60	.42	.18	.66	.55	.11	.28	.72
Excitement-Seeking	.59	.44	.15	.72	.62	.09	.17	.83
Positive Emotions	.56	.43	.13	.73	.63	.09	.90	.10
				OPENNESS TO EXPERIENCE				
O. to Fantasy	.39	.27	.12	.57	.48	.09	.57	.43
O. to Aestethics	.62	.51	.11	.71	.65	.06	.73	.27
O. to Feelings	.52	.35	.17	.64	.53	.11	.89	.11
O. to Actions	.53	.37	.16	.65	.55	.10	.21	.79
O. to Ideas	.56	.44	.12	.61	.55	.06	.26	.74
O. to Values	.45	.20	.25	.54	.35	.19	.16	.84
				AGREEABLENESS				
Trust	.41	.31	.10	.65	.46	.19	.56	.44
Straightforwardness	.35	.20	.15	.58	.30	.28	.28	.72
Altruism	.54	.39	.15	.69	.46	.23	.88	.12
Compliance	.54	.43	.11	.67	.51	.16	.59	.41
Modesty	.42	.31	.11	.63	.44	.19	.22	.78
Tender-Mindedness	.49	.31	.18	.62	.36	.26	.48	.52
				CONSCIENTIOUSNESS				
Competence	.59	.35	.24	.78	.46	.32	.53	.47
Order	.61	.44	.17	.74	.53	.21	.62	.38
Dutifulness	.62	.36	.26	.79	.46	.33	.71	.29
Achievement Striving	.61	.42	.17	.76	.54	.22	.60	.40
Self-Discipline	.62	.46	.16	.82	.61	.31	.95	.05
Deliberation	.49	.34	.15	.69	.47	.22	.32	.68
AVERAGE	**.56**	**.38**	**.18**	**.69**	**.51**	**.18**	**.54**	**.46**

Note. Rel = reliability; Con_F = facet consistency; Spe_M = method specificity; Con_D = domain consistency; Spe_F = facet specificity; see text for explanation of these coefficients.

The facet consistency coefficients (Con_F) were estimated as the degree of reliable variance of an observed facet scale, which is explained by the respective facet trait factor. Across all facets we found a larger facet consistency for mean peer reports (on average Con_F = .51) than for self-reports (on average Con_F = .38) indicating that averaged peer reports are more accurate. That obviously followed from the greater reliability of aggregated peer reports. Method specificity coefficients (Spe_M) were computed as the amount of reliable variance of an observed variable due to the respective method factor. Although we found no differences in the averaged degree of method specificity between self- and mean peer reports (on average Spe_M = .18 for both), self-report specificity was larger for Neuroticism, Extraversion, and Openness, whereas the model proposed a higher degree of mean peer report specificity for Agreeableness and Conscientiousness.

As the second-order factor *D* (see Figure 2) reflects common variance of facet traits defining a domain trait, we separated domain-level trait variance from facet-specific trait variance. Similar to the variance decomposition of the reliable variance, facet trait variance (facet consistency) was decomposed into common trait (domain consistency) and facet-specific variance components (facet specificity). For a better exemplification, domain consistency (Con_D) and facet specificity (Spe_F) were computed in such way that they added up to one (see Table 2). Facet traits differed remarkably in the proportion of common variance with their corresponding domain traits. For example, Depression, Anxiety, and Vulnerability were clear indicators of Neuroticism, whereas Impulsivity was not. Similarly, Warmth and Positive Emotions were clear indicators for Extraversion, whereas Excitement Seeking was not. However, only Impulsivity was not significantly consistent (C.R. < 1.96; $p > .05$) indicating complete independence from Neuroticism.

Genetic and non-genetic influences on the domain trait and rater-specific method factors are presented in Table 3. Latent domain trait (*D*) variance was influenced by additive genetic (a^2) and non-shared environmental effects (e^2), whereas shared environmental influences (c^2) were non-significant. Except for Neuroticism, additive genetic influences on valid personality differences were larger than environmental influences.

Self-report specificity consistently showed genetic influences across domains and no significant non-genetic influences shared by twins, except for Openness. Thus, correlations between self-report-specific factors (M_S) depended on self-raters' genetic resemblance. For Extraversion, Openness, and Conscientiousness, the peer report method factors (M_P) were not significantly correlated between averaged peer ratings on each sibling of twin pairs. The peer-report-specific factors of Neuroticism and Agreeableness were significantly correlated as a function of twins' genetic resemblance, indicating a peer viewpoint-specific genetic component. However, the largest part of method variance in averaged peer reports was explained by non-genetic influences.

The final aim of our study was to examine the sources of facet-specific traits estimated as facet trait residuals (*R*). Phenotypic (facet specificity), genetic, and environmental parameters are shown in Table 4. For two facets, Depression and Self-Discipline, the phenotypic residual components were not significant. That is, variance in Depression was completely accounted for by the variance in Neuroticism and variance in Self-Discipline was completely accounted for by the variance in Conscientiousness. As a consequence genetic and environmental variance components could not be significant either. It is important to note that the statistical significance of additive genetic, shared and non-shared environmental components depended on the amount of facet specificity (Spe_F).

II

For 24 of 28 facets the larger amount of facet specificity was explained by genetic influences. Facet specificity in Positive Emotions, Altruism, and Tender-Mindedness was primarily explained by shared environmental factors. Non-shared environmental effects primarily explained the facet specificity in Vulnerability. We found significant non-shared environmental effects for 23 of 28 facets, although the magnitude was small across facets.[15]

Table 3:

Standardized Components of Domain-Specific Trait and Rater-Specific Variance: Additive Genetic (a^2), Shared (c^2), and Non-Shared Environmental (e^2) Effects

Latent variables	Variance components		
	a^2	c^2	e^2
	NEUROTICISM		
D	.47*	.00	.53*
M_S	.54*	.00	.46*
M_P	.31*	.00	.69*
	EXTRAVERSION		
D	.62*	.00	.38*
M_S	.51*	.01	.48*
M_P	.08	.06	.86*
	OPENNESS TO EXPERIENCES		
D	.68*	.00	.32*
M_S	.35*	.20*	.45*
M_P	.09	.01	.90*
	AGREEABLENESS		
D	.67*	.00	.33*
M_S	.66*	.00	.34*
M_P	.33*	.00	.67*
	CONSCIENTIOUSNESS		
D	.72*	.00	.28*
M_S	.50*	.04	.46*
M_P	.09	.00	.91*
	AVERAGE		
D	**.63**	**.00**	**.37**
$\boldsymbol{M_S}$	**.51**	**.05**	**.44**
$\boldsymbol{M_P}$	**.18**	**.01**	**.81**

Note. *D* = domain-specific convergent valid variance; M_S = self-report-specific variance; M_P = mean-peer-report-specific variance; see text for explanations of these coefficients. * Variance components are significant with $p \leq .05$ (C.R. > 1.96).

[15] We also ran the analysis with only one (randomly selected) peer report for each twin sibling. The model implied reliability of single peer reports (averaged *Rel* = .60) was lower than of mean peer reports and on a similar level as self-reports. This is attributable to the higher level of random error in single-rater variance. As a consequence, the facet trait consistency for single peer reports (averaged Con_F = .38) was smaller than for mean peer reports and even on the same level as for self-reports because of the slightly higher level of method specificity in single peer report variance (averaged Spe_M = .22). Because of their lower psychometric quality, (single) peer report specific variance showed a lower level of heritability but still significant for Neuroticism (18%) and Agreeableness (26%). As a consequence of the lower self-peer covariance, this analysis showed on average slightly lower levels of genetic influences on common domain variance (60%) and slightly higher levels of genetic influences on self-report-specific variance (54%). Altogether, this supports the usefulness of aggregated peer reports as a compact and better reference measure of the observable personality (Hofstee, 1994).

Table 4:

Standardized Components of Facet-Specific Trait Variance: Additive Genetic (a^2), Shared (c^2), and Non-Shared Environmental (e^2) Effects

Personality Facets	Facet specificity	Variance components (%)		
		a^2	c^2	e^2
	NEUROTICISM			
Anxiety	.22*	.16 (74)*	.00 (0)	.06 (26)*
Angry Hostility	.57*	.48 (84)*	.00 (0)	.09 (16)*
Depression	.05	.05 (100)	.00 (0)	.00 (0)
Self-Consciousness	.39*	.28 (72)*	.04 (9)	.07 (19)*
Impulsivity	.98*	.90 (92)*	.00 (0)	.08 (8)*
Vulnerability	.14*	.02 (12)	.04 (29)	.08 (59)*
	EXTRAVERSION			
Warmth	.20*	.20 (100)*	.00 (0)	.00 (0)
Gregariousness	.49*	.27 (56)*	.13 (26)*	.09 (18)*
Assertiveness	.75*	.49 (65)*	.00 (0)	.26 (35)*
Activity	.72*	.49 (68)*	.00 (0)	.23 (32)*
Excitement-Seeking	.83*	.56 (67)*	.08 (10)	.19 (23)*
Positive Emotions	.10*	.00 (0)	.06 (64)	.04 (36)
	OPENNESS TO EXPERIENCES			
O. to Fantasy	.43*	.37 (86)*	.00 (0)	.06 (14)*
O. to Aestethics	.27*	.22 (80)*	.05 (20)	.00 (0)
O. to Feelings	.11*	.10 (96)*	.01 (4)	.00 (0)
O. to Actions	.79*	.60 (76)*	.00 (0)	.19 (24)*
O. to Ideas	.74*	.64 (86)*	.00 (0)	.10 (14)*
O. to Values	.84*	.57 (68)*	.16 (19)*	.11 (13)*
	AGREEABLENESS			
Trust	.44*	.36 (81)*	.00 (0)	.08 (19)*
Straightforwardness	.72*	.57 (79)*	.00 (0)	.15 (21)*
Altruism	.12*	.01 (11)	.11 (89)*	.00 (0)
Compliance	.41*	.35 (86)*	.00 (0)	.06 (14)*
Modesty	.78*	.54 (69)*	.00 (0)	.24 (31)*
Tender-Mindedness	.52*	.00 (0)	.46 (89)*	.06 (11)*
	CONSCIENTIOUSNESS			
Competence	.47*	.23 (49)*	.10 (22)*	.14 (29)*
Order	.38*	.22 (57)*	.00 (0)	.16 (43)*
Dutifulness	.29*	.17 (59)*	.00 (0)	.12 (41)*
Achievement Striving	.40*	.19 (47)*	.08 (21)	.13 (32)*
Self-Discipline	.05	.02 (30)	.00 (0)	.03 (70)
Deliberation	.68*	.46 (67)*	.00 (0)	.22 (33)*
Averaged estimates for facet specificity				
$Spe_F = Con_F - Con_D$	**.46**	**.32 (69)**	**.04 (10)**	**.10 (21)**

Note. Con_F = facet consistency; Con_D = domain consistency; Spe_F = facet specificity; * Variance components are significant with $p \leq .05$ (C.R. > 1.96).

Discussion

The results of our study support three major conclusions: First, mono-rater studies of single personality measures tend to underestimate genetic effects *not only* on broad personality dimensions *but also* on facet-specific traits, since rater-specific effects and random error cannot be controlled. The more the FFM-structure is measured accurately the better these measures reflect the proposed genetic structure. Second, after correcting for all genetic trait components, rater-specific components, sex and age effects, as well as error of measurement, there remains convergent valid variance in personality domain traits and facet-specific traits that is primarily attributable to non-shared environmental influences. This suggests that personality traits, global domains as well as specific facets, are never independent from their external influences. Third, across all personality dimensions about half of the self-report-specific variance is genetically influenced, whereas low (not statistically significant) up to moderate genetic effects contribute to peer rater-specific variance. This may indicate specific heritable personal characteristics distorting self-reports on personality but also specific substance components in self- *as well as* in mean peer reports on traits actually measured.

Genotypic Personality Structure

As we analyzed not only self- and averaged peer reports but also multiple indicators (facets) of global personality dimensions (domains), we were able to examine the heritability of domain-level as well as facet-specific traits. On average, about one half of phenotypic variance in facet traits was explained by the proposed corresponding personality domain (McCrae & Costa, 1992). Not surprisingly, domain trait variance components showed on average stronger genetic influence than variance in single self-reports (Jang et al., 1996), mean peer reports (Kandler, Riemann, & Kämpfe, 2009), or multiple self-report measures (Loehlin et al., 1998), since mono-rater studies cannot separate systematic and random error variance from trait variance. Mono-rater studies overestimate non-shared environmental effects and underestimate genetic factors on personality domains (Riemann et al., 1997).

Consequently and as already hypothesized by Jang et al. (1998), genetic effects on more accurately estimated facet-residual traits surpassed the amount of non-shared environmental influences, since it increased on average from 47% (of reliable facet-specific variance; Jang et al.) up to 69% (of convergent valid facet-specific variance). That is, there were unique genetic influences accounting for about a third of the variance in facet-specific traits, but there were also common genetic effects across personality levels of abstraction accounting for another third of variance. This is consistent with a theory of a hierarchical structure of personality traits (McCrae & Costa, 2008). Furthermore, a similar genetic model to the one recently proposed for cognitive abilities (Plomin, DeFries, McClearn, & McGuffin, 2008) also appears to characterize the structure of each of the five broad personality dimensions and their corresponding facets. This model describes a combination of a bottom-up and a top-down model of genetic effects. It encompasses genetic factors specifically associated with almost each personality facet (bottom-up) as well as genetic factors associated with personality domains that are not associated with lower level traits when domain variance is controlled (top-down).

In sum, our study extends and validates the former mono-method study on the hierarchical structure of the NEO-PI-R dimensions (Jang et al., 1998). Applied to the search for genes in the personality realm, this calls for facet-specific molecular genetic efforts, because facet-level traits include both the facet-specific (bottom up) and the corresponding domain-

specific (top down) genetic components. Moreover, the use of self- and peer reports might also be fruitful for genotype-phenotype-association studies since it is convergent valid variance in particular that reflects underlying genetic variance. Concentrating on such variance components might increase the power to detect and replicate small effects of quantitative-trait-loci on facet-specific *and* domain-level traits. If we concentrate only on one method, we cannot decide whether the association is with the method or with the trait.

Environmental Personality Structure

Though on average most of the convergent valid variance in personality facets is influenced by genetic factors, we found no significant genetic influence on facet-specificity of Altruism, Vulnerability, Positive Emotions, and Tender-Mindedness. Furthermore, shared and (primarily) non-shared environmental effects play a significant role beyond method specificity and error, accounting for the remaining third of trait variance in personality facets. Some findings of shared environmental effects (e.g., Altruism) show stronger commonality with previous self-report studies (e.g., Jang et al., 1998), other findings do not (e.g., Tender-Mindedness). However, regarding the disentanglement of shared and non-shared environmental effects, it is important to note that "objectively shared" environmental effects by reared-together twins (e.g., parental environment) might be "individually non-shared" (Dunn & Plomin, 1990). That is, the most of the environmental variance might be non-shared because most experiences have unsystematic effects on all of the individuals (e.g., siblings) who have these experiences.

Different explanations for environmental effects on personality are conceivable. First, estimates of environmental effects may reflect an accumulation of small effects due to many experiences (Plomin & Daniels, 1987): e.g., in the workplace, at clubs, and in the family. Referring to the genetic set-point theory (Carey, 2003), in which genetic factors act as individual set-points to which individuals regress on a long-term basis, environmental effects might also be interpreted as short-time situational effects. Another explanation is that non-shared environmental variance may reflect effects of interactions between personality relevant genotype and environments (Plomin, DeFries, & McClearn, 1977), since interaction between additive genetic and non-shared environmental effects is confounded with non-shared environmental variance (Purcell, 2002) when this interaction is not controlled for. Referring to the Five-Factor-Theory (McCrae & Costa, 2008), such effects may be interpreted as *Characteristic Adaptations* of personality genotypes. That is, genotypes interact with opportunities and incentives of the specific social environment, likewise personalities of others, social roles, developmental tasks, etc. (McAdams & Pals, 2006). Finally, a further explanation is supposable. Perhaps, it does not reflect actual or direct environmental effects at all but humans' capability of conscious will to overcome temperamental characteristics and to represent specific facets. Without conscious will no cooperative society (i.e., tender-minded and altruistic) and no helpful psychotherapy would be possible (Wilber, 2000). Consciously designed facets (represented to the world) should be well observable by peers. If specific effects of several contexts are generally small, short-term effective, and interact with genetic factors, or not direct at all this might explain the difficulty to detect manifest specific environments affecting personality (Turkheimer & Waldron, 2000).

The readers may suppose that non-shared environmental trait variance might also reflect artificial influences, because self-peer covariance could reflect common method variance due to the choice of the assessment instrument or if biases of averaged ratings of informed acquaintances are not fully independent of self-reports. For example, individuals

have learned the true associations between personality traits, known as implicit personality theory (Borkenau, 1992; McCrae et al., 2001). However, if such a learned semantic scheme or instrument-based biases are shared between persons providing self-reports and peer judgments, these biases should also be shared by twin self-raters and among all peer raters across twins. Therefore, such effects should result in estimates of systematic and thus shared environmental effects on trait variance. Our findings do not support this explanation.

Genetic Influence on Rater-Specific Perspectives

Though self-peer agreement provides a clearer reflection of the genetic architecture, self-rater-specific variance is also substantially influenced by genetic factors, whereas specific variance in averaged peer reports shows only moderate or insignificant genetic influences. This difference may indicate genetically influenced rater biases on self-reports (McCrae et al., 2008; Kandler et al., 2009). That is, genetic factors may affect response styles or different cognitive processes distorting self-reports (e.g., self-enhancement or self-deception). However, genetic effects on peer-rater-specific variance are not explainable by genetically influenced biases. This variance component should reflect real behavioral differences which were observed by well-informed peers but in a peer viewpoint-specific manner. Our results suggest that observers may have privileged access to information cues relevant for judging a target person's personality. This interpretation is in line with studies showing that peer reports on personality provide significant incremental variance over self-ratings in predicting personality-relevant criterions (Fiedler, Oltmanns, Turkheimer, 2004; Mount, Barrick, & Strauss, 1994). A recent study found evidence that self-raters do not pay as much attention to or make as much use of available behavioral information cues as observers (Hofmann, Gschwender, & Schmitt, 2009). It appears that other-ratings predict observable personality relevant behavior better than self-ratings.

In an analogous manner, specific genetic variance in self-reports may also reflect substance components of the focused personality traits that are not readily accessible to the peer rater. Self-raters probably pay more attention to other cues for self-diagnosis (e.g., motives or mental states). According to this consideration, phenotypic validity studies controlling for response style variance measured by validity scales (e.g., Piedmont, McCrae, Riemann, & Angleitner, 2000) do not support the assumption that self-reports are significantly distorted resulting in self-peer disagreement. In line with a substance interpretation, self-ratings of personality outperformed informant ratings at predicting self-report ratings of emotion (Spain, Eaton, & Funder, 2000) and perceptions of family environment (Kandler et al., 2009). Such studies strengthen the position that self-reports capture valid information that is not accessible by peers.

It is important to note that peers, even if well-informed (e.g., spouses, colleagues, and teammates), may see the targets only in limited contexts (e.g., home, clubs, and job environment). As mentioned above, people may show different behavior depending on their social environments, or in other words, personality genotypes may interact with specific contexts. Consequently, reports of selective peers with different social roles, even if aggregated, can only reflect a limited proportion of perceivable target personality. This may explain the difference between self- and peer-rater-specific components in the amount of genetic influences beyond genetically influenced self-rater biases.

The interpretation of different valid perspectives is in line with Kendler's (2001) two pathways in which genes may be effective: *Within*- and *Outside-the-Skin*. *Within-the-Skin*,

genetic expression takes place through internal personality traits (e.g., motives and emotions). *Outside-the-Skin*, genetic factors may also affect personality-relevant behavior and its social consequences. The *Within-the-Skin* pathway may be more accessible for self-raters whereas the *Outside-the-Skin* pathway could be to some extent more perceivable by peers. Thus, it might be argued that expressions of personality genotypes have to be differentiated into internal (experienced intentions and emotions) and external (expressed behavior and its social consequences) effects. As a consequence, an exclusive focus on variance shared by self- and peer reports or a complete reliance on self- or peer reports may provide an incomplete reflection of personality. From this follows that it would not be more fruitful to look for the specific genes involved in personality by solely focusing on common variance in self- and peer reports.

Though it is well established that personality structure is substantially heritable, much remains to be learned about the genetics of the processes involved in judgments on and perceptions of personality traits. Future research on personality traits should aim at finding testable ways to resolve whether rater-specific variance in self-reports is essential, artificial, or both. Other methods, e.g., implicit measures (Hofmann et al., 2009), real-life measures of act frequencies (Vazire & Mehl, 2008), or measures of endophenotypes (Ebstein, 2006; Plomin et al., 2008) might be useful to rule out rater biases (e.g., response styles, self-enhancement, and self-deception). As peers see targets only in limited contexts, collecting assessments from each rater across contexts and from different raters within each context may additionally rule out contextual effects on personality judgments (Kraemer et al., 2003). There should be no doubt that the combination of multi-method and genetically informative designs delivers additional insight into whether traits, which are considered to have a genetic basis, are measured accurately. By now, one thing appears to be sure: self-peer agreement provides a clearer reflection of the genetic architecture of personality traits, but an exclusive focus on self-peer agreement provides an incomplete picture of the complex personality relevant behavior.

References

Angleitner, A., Riemann, R., & Strelau, J. (1995). *A study of twins using the self-report and peer report NEO-FFI scales.* Paper presented at the 7th meeting of the International Society for the Study of Individual Differences, July 15-19, Warsaw, Poland.

Arbuckle, J. L. (2007). *AMOS Users' Guide 16.0*. Chicago: SPSS.

Bartels, M., Boomsma, D. I., van Beijsterveldt, T. C. E. M., Hudziak, J. J., & van den Oord, E. J. G. (2007). Twins and the study of rater (dis)agreement. *Psychological Methods*, *12*, 451-466.

Bentler, P. M. (1990). Comparative fit indices in structural models. *Psychological Bulletin*, *107*, 238-246.

Blackman, M. C., & Funder, D. C. (1998). The effect of information on consensus and accuracy in personality judgment. *Journal of Experimental Social Psychology*, *34*, 164-181.

Borkenau, P. (1992). Implicit personality theory and the Five-Factor Model. *Journal of Personality*, *60*, 295-327.

Borkenau, P., Riemann, R., Angleitner, A., & Spinath, F. (2001). Genetic and environmental influences on observed personality: Evidence from the German Observational Study of Adult Twins. *Journal of Personality and Social Psychology*, *80*, 655-668

Borkenau, P., Riemann, R., Spinath, F.M. & Angleitner, A. (2000). Behavior Genetics of Personality: The Case of Observational Studies. In S. Hampson (Ed.), *Advances in personality psychology, Volume* 1 (pp. 107-137). Hove, England: Psychology Press.

Bouchard, T. J., Jr., & Loehlin, J. (2001). Genes, evolution, and personality. *Behavior Genetics*, *31*, 243-273.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 136-162). Newbury Park: Sage.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81-105.

Carey, G. (2003). *Human Genetics for the Social Sciences*. Thousand Oaks, CA: Sage.

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual.* Odessa, FL: Psychological Assessment Resources.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles.* New York: Wiley.

Dunn, J. & Plomin, R. (1990). *Separate Lives. Why Siblings Are so Different?* New York: Basic Books.

Ebstein, R. (2006). The molecular genetic architecture of human personality: Beyond self-report questionnaires. *Molecular Psychiatry*, *11*, 563-568.

Eid, M., Lischetzke, T., Nussbeck, F. W., & Trierweiler, L. I. (2003). Separating trait effects from trait-specific method effects in multitrait-multimethod models: A multiple-indicator CT-C(*M*-1) model. *Psychological Methods*, *8*, 38-60.

Fiedler, E. R., Oltmanns, T. F., & Turkheimer, E. (2004). Traits associated with personality disorders and adjustment to military life: Predictive validity of self- and peer reports. *Military Medicine*, *169*, 207-211.

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, *102*, 652-670.

Funder, D. C., Kolar, D. C., & Blackman, M. C. (1995). Agreement among judges of personality: Interpersonal relations, similarity, and acquaintanceship. *Journal of Personality and Social Psychology*, *69*, 656-672.

Hofmann, W., Gschwender, T., & Schmitt, M. (2009). The road to the unconscious self not taken: Discrepancies between self- and observer-inferences about implicit dispositions from nonverbal behavioural cues. *European Journal of Personality*, *23*, 343-366.

Hofstee, W. K. B. (1994). Who should own the definition of personality? *European Journal of Personality*, *8*, 149-162.

Hoyt, W. T. (2000). Rater bias in psychological research: When is it a problem and what can we do about it? *Psychological Methods*, *5*, 64-86.

Jang, K. L., Livesley, W. J., Angleitner, A., Riemann, R., & Vernon, P. A. (2002). Genetic and environmental influences on the covariance of facets defining the domains of the five-factor model of personality. *Personality and Individual Differences*, *33*, 83-101.

Jang, K. L., Livesley, W. J., & Vernon, P. A. (1996). Heritability of Big Five dimensions and their facets: A twin study. *Journal of Personality*, *64*, 577-591.

Jang, K. L., McCrae, R. R., Angleitner, A., Riemann, R., & Livesley, W. J. (1998). Heritability of facet-level traits in cross-cultural twin study: Support for a hierarchical model of personality. *Journal of Personality and Social Psychology*, *74*, 1556-1565.

Kandler, C., Riemann, R., & Kämpfe, N. (2009). Genetic and environmental mediation between measures of personality and family environment in twins reared together. *Behavior Genetics*, *39*, 24-35.

Kendler, K. S. (2001). Twin studies of psychiatric illness. An update. *Archives of General Psychiatry*, *58*, 1005-1014.

Kenny, D. A. (1991). A general model of consensus and accuracy in interpersonal perception. *Psychological Review*, *98*, 155-163.

Kraemer, H. C., Measelle, J. R., Ablow, J. C., Essex, M. J., Boyce, W. T., & Kupfer, D. J. (2003). A new approach to integrating data from multiple informants in psychiatric assessment and research: Mixing and matching contexts and perspectives. *American Journal of Psychiatry*, *160*, 1566-1577.

Letzring, T. D., Wells, S. M., & Funder, D. C. (2006). Information quantity and quality affect the realistic accuracy of personality judgment. *Journal of Personality and Social Psychology*, *9*, 111-123.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical Analysis with Missing Data* (2nd ed.). New Jersey: Wiley.

Loehlin, J. C., McCrae, R. R., Costa, P. T., & John, O. P. (1998). Heritabilities of common and measure-specific components of the Big Five personality factors. *Journal of Research in Personality*, *32*, 431-453.

Markon, K. E., Krueger, R., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology*, *88*, 139-157.

McAdams, D. P. & Pals, J. (2006). A new big five. Fundamental principles for an integrative science of personality. *American Psychologist*, *61*, 204-217.

McCrae, R. R., & Costa, P. T., Jr. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*, 81-90.

McCrae, R.R., & Costa, P.T. (2008). The five-factor theory of personality. In John, O.P., Robins, R.W., & Pervin, L.A. (Eds.), *Handbook of Personality: Theory and Research* (3rd ed.). New York: Guilford.

McCrae, R. R., Costa, P. T., Jr., Hrebícková, M., Ostendorf, F., Angleitner, A., Avia, M. D., et al. (2000). Nature over Nurture: Temperament, Personality, and Life Span Development. *Journal of Personality and Social Psychology*, *78*, 173-186.

McCrae, R. R., Jang, K. L., Livesley, W. J., Riemann, R., & Angleitner, A. (2001). Sources of Structure: Genetic, Environmental, and Artifactual Influences on the Covariation of Personality Traits. *Journal of Personality*, *69*, 511-535.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*, 175-215.

McCrae, R. R., Yamagata, S., Jang, K. L., Riemann, R., Ando, J., Ono, Y., et al. (2008). Substance and artifact in the higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, *95*, 442-455.

Mount, M. K., Barrick, M. R., & Strauss, J. P. (1994). Validity of observer ratings of the Big Five personality factors. *Journal of Applied Psychology*, *79*, 272-280.

Neale, M. C., & Maes, H. H. M. (2004). *Methodology for genetic studies of twins and families*. Dordrecht: Kluwer Academic Publishers B.V.

Ostendorf, F., & Angleitner, A. (2004). *NEO-Persönlichkeitsinventar, revidierte Form, NEO-PI-R nach Costa und McCrae* [Revised NEO Personality Inventory, NEO-PI-R of Costa and McCrae]. Göttingen, Germany: Hogrefe.

Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality*, *66*, 1025-1060.

Piedmont, R. L., McCrae, R. R., Riemann, R., & Angleitner, A. (2000). On the invalidity of validity scales: Evidence from self-reports and observer ratings in volunteer samples. *Journal of Personality and Social Psychology, 78*, 582-593.

Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from each other? *Behavioral and Brain Sciences*, *10*, 1-16.

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin*, *84*, 309-322.

Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2008). *Behavioral Genetics* (5th ed.). New York: Woth Publishers.

Purcell, S. (2002). Variance components models for gene–environment interaction in twin analysis. *Twin Research and Human Genetics*, *5*, 554-571.

Riemann, R., Angleitner, A., & Stelau, J. (1997). Genetic and environmental influences on personality: A study of twins reared together using the self- and peer report NEO-FFI scales. *Journal of Personality*, *65*, 449-475.

Simonoff, E., Pickles, A., Hewitt, J. K., Silberg, J., Rutter, M., Loeber, R., et al. (1995). Multiple raters of disruptive child behaviour: Using a genetic strategy to examine shared views and bias. *Behavior Genetics*, *25*, 311-326.

Spain, J. S., Eaton, L. G., & Funder, D. C. (2000). Perspectives on personality: The relative accuracy of self versus others for the prediction of emotion and behaviour. *Journal of Personality*, *68*, 837-867.

Spinath, F. M., Angleitner, A., Borkenau, P., Riemann, R., & Wolf, H. (2002). German Observational Study of Adult Twins (GOSAT): A multimodal investigation of personality, temperament and cognitive ability. *Twin Research and Human Genetics*, *5*, 372-375.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, *25*, 173-180.

Stößel, K., Kämpfe, N., & Riemann, R. (2006). The Jena Twin Registry and the Jena Twin Study of Social Attitudes (JeTSSA). *Twin Research and Human Genetics*, *9*, 783-786.

Turkheimer, E., & Waldron, M. (2000). Nonshared environment: A theoretical, methodological, and quantitative review. *Psychological Bulletin*, *126*, 78-108.

van der Valk, J. C., van den Oord, E. J. C. G., Verhulst, F. C., & Boomsma, D. I. (2003). Using shared and unique parental views to study the etiology of 7-year old twins' internalizing and externalizing problems. *Behavior Genetics*, *33*, 409-420.

Vazire, S., & Mehl, M. R. (2008). Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology*, *95*, 1202-1216.

Vernon, P. A., Villani, V. C., Vickers, L. C., & Harris, J. A. (2008). A behavioural genetic investigation of the dark triad and the big five. *Personality and Individual Differences*, *44*, 445-452.

Wilber, K. (2000). *Intergral Psychology: Consciousness, Spirit, Psychology, Therapy.* Bosten: Shambhala Publications.

Yamagata, S., Suzuki, A., Ando, J., Ono, Y., Kijima, N., Yoshimura, K., et al. (2006). Is the genetic structure of human personality universal? A cross-cultural twin study from North America, Europe, and Asia. *Journal of Personality and Social Psychology*, *90*, 987-998.

ANHANG III

Journal of Personality and Social Psychology

(2010)

III

Sources of Cumulative Continuity in Personality: A Longitudinal Multiple-Rater Twin Study

Christian Kandler

Wiebke Bleidorn

Rainer Riemann

Frank M. Spinath

Wolfgang Thiel

Alois Angleitner

Abstract. This study analyzed the etiology of rank-order stability and change in personality over a time period of 13 years in order to explain cumulative continuity with age. NEO- five factor inventory self- and peer report data from 696 monozygotic and 387 dizygotic twin pairs reared together were analyzed using a combination of multiple-rater twin, latent state-trait, and autoregressive simplex models. Correcting for measurement error, this model disentangled genetic and environmental effects on long- and short-term convergent valid stability, on occasional influences, and on self- and peer report-specific stability. Genetic factors represented the main sources that contributed to phenotypic long-term stability of personality in young and middle adulthood, whereas change was predominantly attributable to environmental factors. Phenotypic continuity increased as a function of cumulative environmental effects which became manifest in stable trait variance and decreasing occasion-specific effects with age. Our findings suggest a complex interplay between genetic and environmental factors resulting in the typical patterns of continuity in personality across young and middle adulthood.

III

Introduction

The meta-analysis by Roberts and DelVecchio (2000) profoundly established that rank-order stability in personality increases across the life course until it reaches its peak in later adulthood after age 50 (Caspi & Roberts, 2001; Fraley & Roberts, 2005). This increasing continuity in personality proved to be a robust finding across self- and other reports, independent of gender and the specific trait considered (e.g., Costa & McCrae, 1988; Terracciano, Costa, & McCrae, 2006). Three prominent theories provide rather conflicting etiological explanations for the cumulative rank-order continuity with age: the genetic set-point model (Carey, 2002), the genetic maturation hypothesis (McCrae, Costa, Ostendorf, Angleitner, Hrebickova, et al., 2000), and a model proposing transactions between genetic and environmental factors (Caspi, Roberts, & Shiner, 2005). We analyzed personality assessments of twins on global personality traits of the five-factor model (McCrae & John, 1992) across three waves of measurement over a period of 13 years in order to test the adequacy of predictions from each of these etiological theories. Since this is the first longitudinal twin study which included self- *and* peer reports, we were able to generalize our findings across multiple raters demonstrating convergent validity (Campbell & Fiske, 1959).

The Genetic Set-Point Hypothesis

A number of behavioral genetic studies have led to the conclusion that genetic factors primarily contribute to stability in personality traits (e.g., McGue, Bacon, & Lykken, 1993; Viken, Rose, Kaprio, & Koskenvuo, 1994). Carey (2002) interpreted this finding in terms of a set-point model, in which environmental fluctuations are assumed to affect short-term changes (a few days, weeks, or even several months) in personality, whereas genetic factors determine individual set-points to which individuals will regress on a long-term basis. These assumptions are in accordance with the intriguing finding that parental environments seem to have no long-term influence on personality in adulthood (e.g., Kandler, Riemann, & Kämpfe, 2009; Krueger, Markon, & Bouchard, 2003).

According to this theory, variance due to individual genetic set-points does not change across time. The cumulative stability of personality across the life-span should thus result exclusively from decreasing effects of environmental fluctuations with age. To the degree that more of the important life transitions occur in young adulthood (e.g., vocational training, finding a job, starting a family) this seems to be a plausible assumption. However, there is powerful evidence that personality stability decreases as the time interval between measurement occasions increases (Fraley & Roberts, 2005) reflecting long-term rank-order change that is not compatible with the genetic set-point hypothesis. If phenotypic scores get closer and closer to the genetic set-point as a function of decreasing short-term environmental influences, then the correlation among scores of initial intervals of time will necessarily increase across a series of later intervals, even though the retest interval is also increasing. This prediction is not in line with the results and continuity functions presented by Fraley and Roberts. Despite these conflicting findings, the genetic set-point model is still a quite appealing and parsimonious model of personality development. Even though the entire model seemed to be too restrictive, the specific assumption of decreasing effects of situational fluctuations might be a relevant mechanism of increasing stability and thus for personality development which is worthwhile to study in more detail.

The Genetic Maturation Hypothesis

The Five-Factor Theory (FFT) provides an alternative explanation of personality development proposing that both rank-order continuity and change in traits, considered as *basic tendencies* (abstract psychological potentials), are exclusively mediated by genetic factors (McCrae et al., 2000). According to this hypothesis, significant environmental effects on traits should merely result from short-term contextual influences and systematic as well as random measurement error.

A number of biometric studies have provided some support for this genetic maturation hypothesis since they found stability to be primarily influenced by genetic factors and have even obtained evidence for the appearance of new genetic factors during young adulthood (e.g., Bratko & Butkovic, 2006; Viken et al., 1994). Actually, when measurement error is controlled for, genetic factors seem to contribute largely to individual stability and growth in Emotional Stability, Agreeableness, and Conscientiousness over a time period of ten years in middle adulthood (Bleidorn, Kandler, Riemann, Angleitner, & Spinath, 2009). However, most findings have contradicted the hypothesis of an exclusive genetic maturation of personality because personality continuity has been found to be also attributable to environmental factors (e.g., Blonigen, Carlson, Hicks, Krueger, & Iacono, 2008; Johnson, McGue, & Krueger, 2005). Differential individual growth in Extraversion and Openness were largely due to environmental factors (Bleidorn et al., 2009).

At this point, it should be noted that previous longitudinal behavioral genetic studies on personality development have exclusively relied on self-reports. As a consequence, these studies could not provide a critical test of the genetic maturation hypothesis, since it was not possible to control for nonrandom bias (method) and random (measurement error) effects, while a multimethod longitudinal behavioral genetic study would address this issue. In a cross sectional study, Riemann, Angleitner, and Strelau (1997) found that the employment of self- and peer reports lead to higher estimates of heritability in personality traits by subtraction of error and method variance. Recently, this finding could be replicated and extended on personality facets (Kandler, Riemann, Spinath, & Angleitner, in press). However, these estimates were still different from unity. Therefore, McCrae and colleagues (2000) acknowledged that the small remaining variance might include true environmental influences including biological sources, such as prenatal infections or different metabolisms, which could distinguish the development of genetically identical individuals. If very early environmental effects contribute to stable differences in personality, they should not change across the life course. Based on the assumptions of this *weaker* genetic maturation hypothesis, cumulative phenotypic continuity should result from cumulative genetic continuity, while environmental effects contribute to stability, situational fluctuations, and systematic as well as random error in personality measures.

The Gene-Environment Transaction Hypothesis

The two theories described above provide elegant and parsimonious explanations for the increasing rank-order stability of personality over the life-span. However, the complete picture of findings appears to be incompatible with both the genetic set-point and the genetic maturation hypothesis. First, large sample longitudinal twin studies (Pederson & Reynolds, 1998, 2002; Viken et al., 1994) have found that phenotypic variance increases with age as a function of increasing non-shared environmental effects. Furthermore, there is evidence for phenotypic stability to increase as a result of increasing environmental stability (Viken et al.,

1994). In view of these findings, McCrae and Costa (2008) revised important tenets of the FFT. They postulated that personality development is determined by *biological* maturation. That is, genetic factors still play a crucial role but the environment can also affect personality traits through biological bases (such as drugs, disease, etc.), and can thus affect personality change.

Caspi et al. (2005) proposed an integrative theory of personality development. They postulated that continuity and change result from transactions between genetic and environmental factors contributing to estimates of both genetic and environmental effects on phenotypic stability and change. The increasing continuity with age is considered as a process of developing and maintaining an identity (Roberts & Caspi, 2003). From this perspective, personality development may be best explained as a result of two mutually supportive life-course dynamics (Caspi et al., 2005): First, people select environments that are correlated with their personality traits (*social selection*); second, experiences in these contexts affect personality functioning (*social influence*) resulting in cumulating effects over the life course. Based on the gene-environment transaction hypothesis, phenotypic rank-order stability should increase with age as a result of an accumulation of environmental influences on trait variance leading to an increase of environmental continuity.

Aims of the Present Study

Only a handful of behavioral genetic studies of personality have estimated continuity and change longitudinally at more than two waves of measurement (e.g., Bleidorn et al., 2009; Pederson & Reynolds, 1998, 2002). Another restriction of previous behavioral genetic studies concerns the sole reliance on self-reports. The present longitudinal study analyzed personality scales of the Five-Factor Model (FFM, McCrae & John, 1992) assessed by self- and peer raters, and spanned a time period of 13 years with three waves of assessment. As a consequence, we were able to answer questions about short- and long-term personality stability and change which was necessary to test the adequacy of the predictions derived from the three conceptions introduced above.

In order to provide a critical test of the three conflicting hypotheses, we combined a multiple-informant twin model (Riemann et al., 1997; Riemann & Kandler, in press), a latent state-trait model (Steyer, Schmitt, & Eid, 1999), and a genetic simplex model (Boomsma & Molenaar, 1987). The availability of self- and peer reports of twins' personality allowed us to decompose convergent valid (i.e., shared by self- and peer reports) variance into (a) long-term stable genetic and environmental sources (about 13 years), (b) "short-term" stable sources (about 6.5 years), and (c) occasion-specific genetic and environmental variance reflecting sources of change and situational fluctuations. We used structural equation modeling to test the three conflicting hypothesis were tested against each other. The most complex model that allowed for genetic and environmental stability and change would reflect the gene-environment transaction hypothesis. The absence of environmental stability and change would argue for the genetic maturation hypothesis, whereas additional stability due to environmental factors would speak for the weaker position of the genetic maturation hypothesis. Finally, the most restricted model that only allowed for genetic stability (over a period of 13 years) and decreasing short-term environmental effects (< 6.5 years) would provide evidence for the genetic set-point hypothesis. For testing parameter equivalence between different age groups, the complete sample was subdivided into two age subsam-

ples representing young and middle adulthood. Because of the cumulative principle of continuity, we hypothesized stability to be larger in the older subsample.[16]

Method

Participants and Procedure

Our study utilized data from the first, third, and fifth wave (in the following referred to as time 1, time 2, and time 3) of the Bielefeld Longitudinal Study of Adult Twins (BiLSAT; Spinath, Wolf, Angleitner, Borkenau, & Riemann, 2005; Bleidorn et al., 2009). At these approximately equidistant measurement occasions both self- and peer reports of personality were gathered between 1993 and 2008. The intervals between time 1 and 2 averaged 6.35 years (*SD* = 1.22); between time 2 and 3 the mean interval was 6.30 years (*SD* = 0.47). The complete time interval spanned almost 13 years.

Participants were excluded from analyses if they were younger than 16 at time 1, since below this age problems in understanding some items of the personality measure were reported (Borkenau & Ostendorf, 1993). Because of mortality that led to a higher degree of dropout, participants older than 75 at time 3 (older than 62 years at time 1) were also excluded. The resulting sample consisted of 696 monozygotic (MZ; 154 male and 542 female) and 387 dizygotic (DZ; 60 male, 213 female and 114 opposite sex) twin pairs at time 1 who provided self-reports of personality. The number of participating twin pairs at each measurement occasion including dropout rates is displayed in Table 1.

Table 1:

Zygosity × Age Group Subsamples With Valid Values at Each Measurement Occasion

Age groups	Time	Age median	Age mean (*SD*)	*N* of pairs		Dropout in %	
				MZ	DZ	MZ	DZ
Young adulthood	1	23	22.7 (3.9)	382	205	-	-
	2	29	28.8 (4.1)	156	95	59	54
	3	35	35.2 (4.3)	84	44	78	79
Middle adulthood	1	39	41.2 (9.1)	314	182	-	-
	2	46.5	48.1 (8.8)	140	88	55	52
	3	54	55.0 (8.8)	103	66	67	64
Total	1	28	31.2 (11.5)	696	387	-	-
	2	35	38.0 (11.8)	296	183	57	53
	3	44	46.2 (12.2)	187	110	73	72

Note. MZ = monozygotic twins; DZ = dizygotic twins; the dropout rates are referred to time 1.

[16] At this point, it should be noted that this study was not aimed to test the three theories in their entireties, but was aimed to compare and test specific predictions regarding rank-order stability and change. It should also be noted that environmental factors may contain non-genetic psychological, sociological, biological, and historical factors. That is, this study was aimed to test the genetic maturation hypotheses with regard to rank-order stability, but was not able to disentangle the *biological* maturation hypothesis (McCrae & Costa, 2008) versus the hypothesis of gene-environment transactions (Caspi et al., 2005) as a primary basis of personality development. It is in line with both hypotheses that predict environmental effects to accumulate across the life course, either more directly (Caspi et al., 2005) or mediated by biological processes (McCrae & Costa, 2008).

Missing values were not completely at random for age using the MCAR-test (χ^2 = 37.65; df = 5; p = .00; Little, 1988). T-tests showed that dropout was larger for younger people (see also Table 1). As a consequence, we used a median split of age regarding available data at all points of time to subdivide the complete sample into two equally large age groups across measurement occasions: young (16-29 at time 1) and middle adulthood (30-62 at time 1). For 99.2% of the participants at the first, 98.2% at the second, and 96.4% at the last assessment at least one peer report was available, received from peers who knew one twin but (preferably) not the co-twin very well.

Measure

We administered the self- and peer report versions of the German Neuroticism-Extraversion-Openness-Five-Factor-Inventory (NEO-FFI; Borkenau & Ostendorf, 1993; Costa & McCrae, 1989). The NEO-FFI is a 60-item inventory designed for measuring personality on five dimensions: Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. At the second measurement occasion, the NEO-FFI scales were computed from the NEO-Personality Inventory-Revised (NEO-PI-R; Costa & McCrae, 1992; Ostendorf & Angleitner, 2004), in which the NEO-FFI items are included. Cronbach's α for the five scales are presented in Table 2. Differences in internal consistencies between subsamples of twins and co-twins (regarding the dependence of twin siblings in a combined sample) and between age groups were not significant. The reliabilities for self-reported Agreeableness were slightly lower than for the corresponding peer reports. Openness derived from the NEO-PI-R (at time 2) yielded somewhat higher internal consistencies compared to the NEO-FFI assessments.

Table 2:

Internal Consistency (Cronbach's α): Self- and Peer Reports at Each Measurement Occasion

Scales	Time 1			Time 2			Time 3		
	Self	Peer1	Peer2	Self	Peer1	Peer2	Self	Peer1	Peer2
Neuroticism	.85	.85	.85	.87	.87	.87	.88	.88	.86
Extraversion	.80	.80	.80	.82	.81	.79	.82	.79	.79
Openness	.63	.64	.62	.71	.70	.69	.61	.61	.60
Agreeableness	.69	.78	.78	.69	.77	.79	.71	.79	.80
Conscientiousness	.82	.84	.85	.82	.86	.85	.79	.85	.86

Note. Statistics based on the complete sample ($N_{time\ 1}$ = 2086; $N_{time\ 2}$ = 796; $N_{time\ 3}$ = 564).

Correlations between peers ranged between .38 (for Agreeableness at time 3) and .54 (for Neuroticism at time 3) with an average of .44. We did not found noticeable differences in the degree of agreement among assessment waves, although peer raters were not necessarily the same across measurement occasions. As averaging peer reports reduces measurement error and rater bias (Hofstee, 1994), averaged peer reported scale scores were used in all subsequent analyses. The correlations between self-reports and mean peer reports ranged between .43 (for Agreeableness at time 3) and .62 (for Extraversion at time 1) with an average of .53. There were no significant differences among points of time or between age subsamples. Consistently, the lowest degree of self-peer agreement was found for Agreeableness, the largest for Extraversion.

Analyses

The existence of age and sex differences can increase variance biasing twin covariance. Thus, self- and averaged peer reports were adjusted for linear and quadratic age effects as well as sex differences within each measurement occasion and age subsample using a regression procedure. This correction did not affect the age differences across measurement occasions but adjusted for age effects at a given point in time. Therefore, each measurement occasion represented the respective mean age of subsamples at this given point in time (see Table 1).

Phenotypic differential stability in self- and averaged peer reports was examined via Pearson's product-moment correlation on the basis of complete data using a pairwise deletion procedure for handling missing values. We estimated stability coefficients for the young and the middle adult subsamples as well as for the short-term (between time 1 and 2 and between time 2 and 3) and full-term intervals (13 years) since we expected to find higher stabilities in the older subsample and within short-term intervals (Fraley & Roberts, 2005; Roberts & DelVecchio, 2000).

As we wanted to analyze all available data in biometric analyses, we tested whether the missing values were completely at random using the MCAR-test (Little, 1988) for each personality variable and each of the four twin data sets (young MZ and DZ as well as middle aged MZ and DZ). MCAR-tests were not significant ($p < .05$). Thus, dropout was completely at random with reasonable certainty. Randomization of missing values was the precondition to receive unbiased results due to missing values and to analyze all available data via raw maximum likelihood modeling to detect genetic and environmental influences (Derks, Dolan, & Boomsma, 2006). This procedure is implemented in the statistical software package Mx (Neale, Boker, Xie, & Maes, 2003) and used for all biometric analyses.

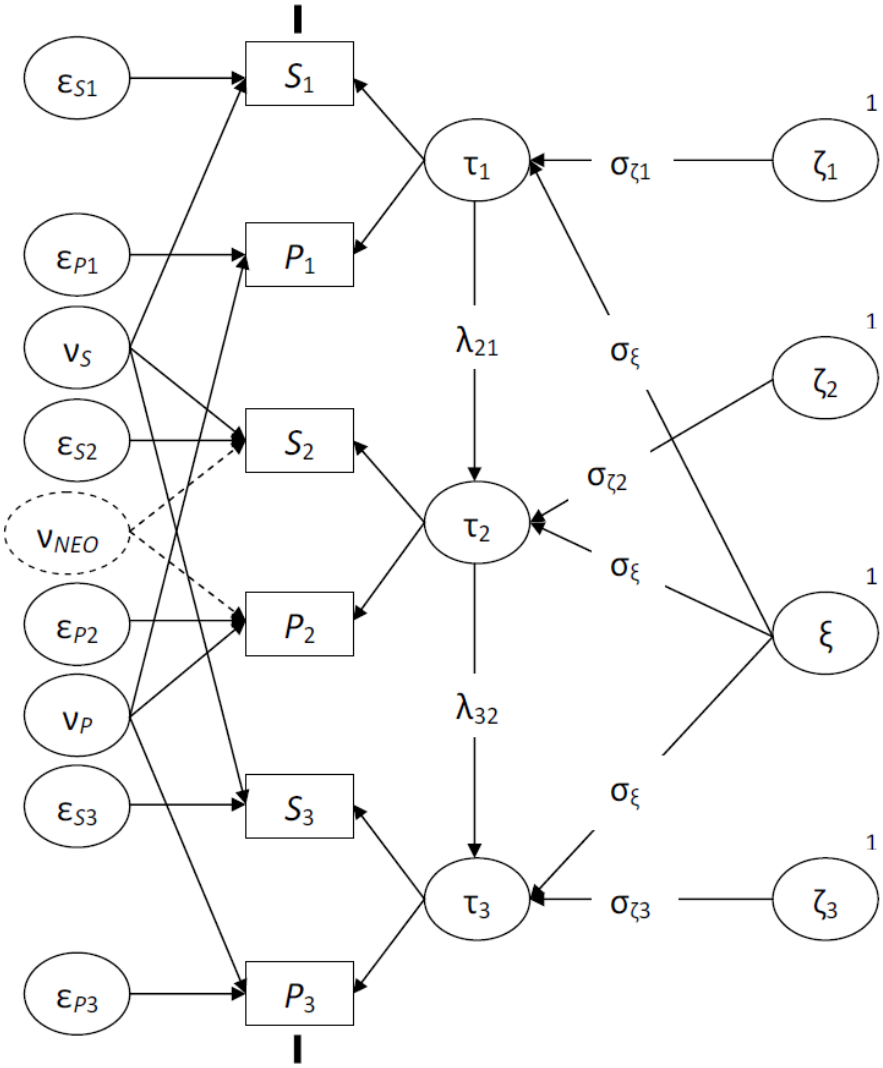
By combining a multiple-informant twin model (Riemann et al., 1997), a latent state-trait model (Steyer, Schmitt, & Eid, 1999), and a genetic simplex model (Boomsma & Molenaar, 1987), we were able to rule out some drawbacks that would arise with the single use of each of these models apart. For example, the latent state-trait model can be employed to examine the accumulation of trait stability but is static and cannot assess the decrease in stability that arises when the time interval between measurement occasions increases. In contrast, the autoregressive simplex model can be used for the latter analysis, but is not suitable to determine a single stable trait component that does not change with time. The combined model takes into consideration that rank-order stability increases with age and decreases with longer intervals. Furthermore, the extension of that model for twins reared together offers the possibility to disentangle genetic and environmental effects on several latent variables (see Hewitt, Eaves, Neale, & Meyer, 1988, for a description of such models). In the following, we describe the model in terms of the usual notation for structural equation modeling.

On the phenotypic level, the model (see Figure 1) allowed us to decompose self- (*S*) and mean peer reports (*P*) at each point of time (rectangles in the figure) into a valid true score (τ), a method (ν) and a residual component (ε). The true score parameters can be considered valid to the degree to which self- and averaged peer raters' assessments converge. In other words, common variance in self- and peer reports reflects convergent validity (Campbell & Fiske, 1959). That is, true score parameters were corrected for self- and peer report specific factors and random error.

On the convergent valid structure level (see the right side of Figure 1), each latent true score variable (τ) was further decomposed into a stable trait (ξ), in the following termed as set-point, and an occasion-specific residual component (ζ). To disentangle short-term stability or – as the other side of the same coin – long-term rank-order change from the set-point component (ξ), we included regressions (λ_{21} and λ_{32}) between neighboring true score variables. In other words, the set-point factor (ξ) was modeled to explain variance due to long-term rank-order stability. The regressions (λ_{21} and λ_{32}) were modeled to consider the fact that rank-order stability may increase with age and decrease as the time interval increases, reflecting long-term rank-order change. That is, covariance between true score 1 (τ_1) and true score 2 (τ_2) may be smaller than the covariance between true score 2 (τ_2) and true score 3 (τ_3). However, covariance between true score 1 and true score 3 may be even smaller than the covariance between neighboring true scores, since the product $\lambda_{21} \times \lambda_{32}$ ($0 \leq \lambda \leq 1$) is always smaller than λ_{21} and λ_{32}. The true score residual component (ζ) was modeled to explain variance specific to each measurement occasion.

Figure 1:

Full Phenotypic Structural Equation Model



Note. *S* = self-report; *P* = peer report; indices 1, 2, and 3 = points of time; τ = convergent valid true score variable; ξ = set-point variable; ζ = true score residual; λ = linear regression of the true score variable on the previous true score variable; ν_S = stable self-report specific factors; ν_P = stable peer report specific factors; ε = measurement error; ν_{NEO} = method error with regard to the fact that NEO-FFI scales were computed from the NEO-PI-R at time 2.

The modeling of self- and mean peer reports allowed us to estimate method factors. In our application these factors reflected method-specific stability (ν_S and ν_P; on the left side of Figure 1). Non-stable factors specific to self- and peer reports were confounded with measurement error (ε) in our model. In addition, we modeled a systematic method factor accounting for differences between the administered instruments, acknowledging the fact that NEO-FFI scales were computed from the NEO-PI-R at the second measurement occasion (ν_{NEO}).

For identification of this *phenotypic model*, it is possible to fix *second-level exogenous* latent variable variances (variances in ξ, ζ_1, ζ_2, and ζ_3; marked with '1' on these latent variables in Figure 1) and fix paths (marked with unlabeled arrows in Figure 1) from each *first-level exogenous* (ν_S, ν_P, ε_{S1}, ε_{S2}, ε_{S3}, ε_{P1}, ε_{P2}, and ε_{P3}) and *endogenous* (τ_1, τ_2, and τ_3) latent variables in order to estimate variance components of all *exogenous* variables (ξ, ζ_1, ζ_2, ζ_3, ν_S, ν_P, ε_{S1}, ε_{S2}, ε_{S3}, ε_{P1}, ε_{P2}, and ε_{P3}) and regressions (λ_{21} and λ_{32}). The squares of standardized loadings reflect the respective latent variable variances (σ_ξ^2, $\sigma_{\zeta1}^2$, $\sigma_{\zeta2}^2$, and $\sigma_{\zeta3}^2$). However, it is also possible to fix second-level paths (aside from regressions) in order to estimate second-level variable variances. The results should be the same. The ν_{NEO} factor, which reflects non-genetic instrument-specific biases, is not identified in the *phenotypic model*. However, it is identified in a twin model or a multigroup model by equalizing this parameter across twins and (or) groups.

The next step was the extension of the *phenotypic model* to a *biometric model* (see Figure 2). We disentangled genetic (*G*) and environmental (*E*) effects on every latent variable and regressions, except on random error (ε) and the inventory method factor (ν_{NEO}). Random error is defined to be neither consistent over different methods and occasions of measurements, nor correlated within twin pairs. Effects of the inventories should be equal across MZ and DZ twins, self- and peer raters, young and middle adulthood. Thus, both components cannot be affected by genetic factors. However, we estimated the amount of genetic and environmental effects on self- and peer report method factors which might reflect stable substantial rater-specific components of personality (e.g., self-concept, social consequences of behavior) or (and) artificial rater biases (e.g., leniency, self-enhancement).

In our application, all *exogenous* latent variable (*G* and *E*) variances were fixed to one in order to obtain estimates for all paths (σ_G and σ_E) *on the biometric level* (marked with dotted lined arrows in Figure 2). Biometric variance components were computed from the squares of these freely estimated path coefficients (σ_G^2 and σ_E^2). Phenotypic components (σ_P^2) were computed from the sum of corresponding biometric components (see Neale & Maes, 2004):

[1] $\sigma_P^2 = \sigma_G^2 + \sigma_E^2$.

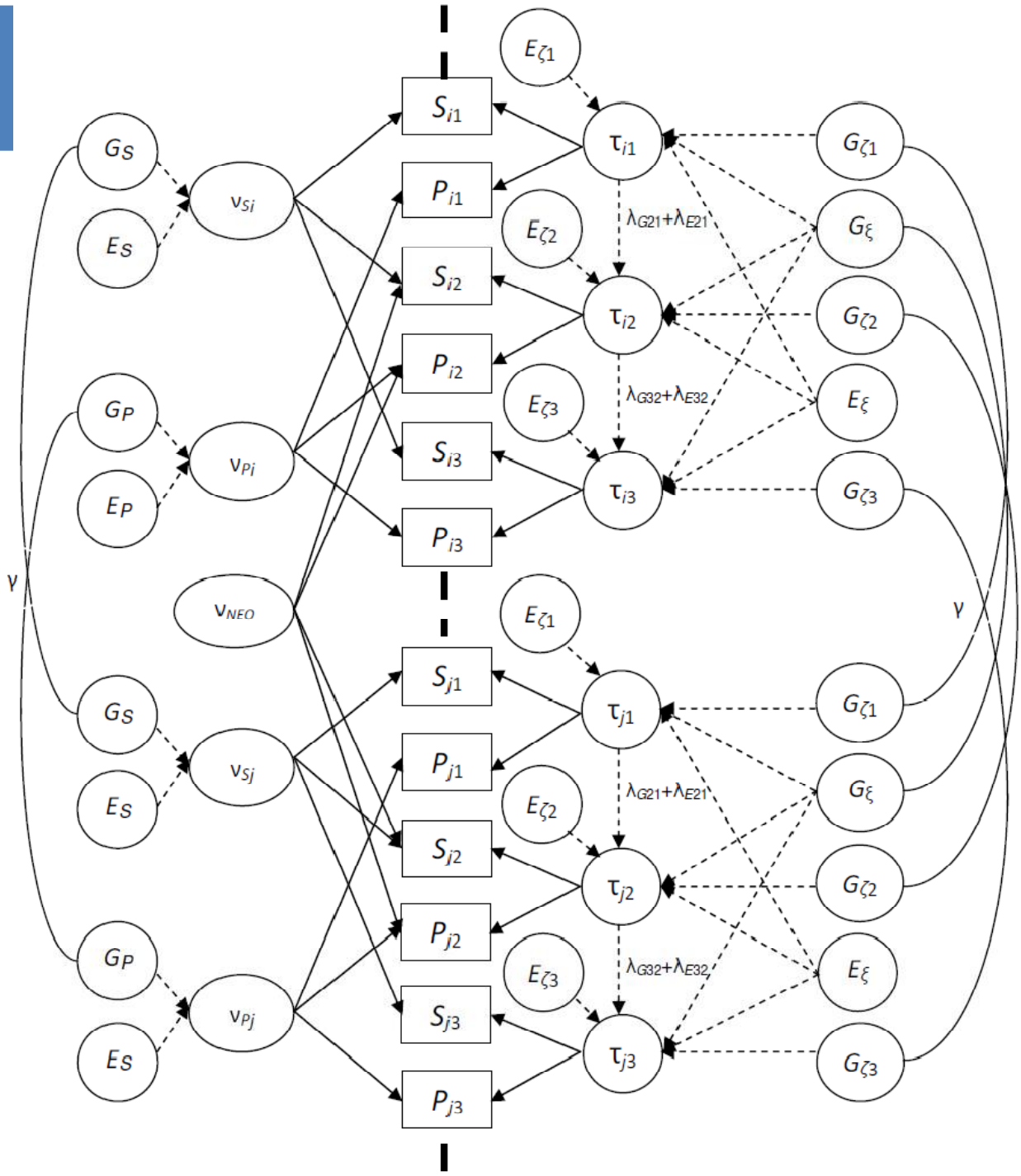
As already mentioned, such decompositions were conducted with all phenotypic parameters (e.g., $\lambda_P = \lambda_G + \lambda_E$) except with random error and the inventory method factor. According to quantitative genetic theory, genetically identical (MZ) twins share 100% and fraternal (DZ) twins share on average 50% of their segregating genes. Cross-twin (cross *i* and *j*, see Figure 2) covariance for MZ twins (σ_{MZ}, genetic correlation: $\gamma = 1$) is equivalent to the genetic variance:

[2] $\sigma_{MZ} = \sigma_G^2$.

And cross-twin covariance for fraternal (DZ) twins (σ_{DZ}, genetic correlation: $\gamma = .50$) is equivalent to a half of the genetic variance:

[3] $\sigma_{DZ} = ½\ \sigma_G^2$.

Figure 2:
Full Biometric Structural Equation Model



Note. *G* = genetic factors; *E* = environmental factors; *S* = self-report; *P* = peer report; indices *i* and *j* = twin and co-twin; indices 1, 2, and 3 = points of time; τ = convergent valid true score variable; ξ = set-point variable; ζ = true score residual; λ = linear regression of the true score variable on the previous true score variable; v_S = stable self-report specific factors; v_P = stable peer report specific factors; v_{NEO} = method error with regard to the fact that NEO-FFI scales were computed from the NEO-PI-R at time 2; γ = 1.0 for MZ twins and 0.5 for DZ twins; for a better readability labels of path coefficients and measurement residual variables reflecting random error are not shown.

From this it follows that environmental effects (σ_E^2) are implicated to the degree that MZ twins differ from one another. This genetically informative model may further be extended to non-additive genetic influences or environmental effects shared by twin siblings (see Pederson & Reynolds, 1998, 2002, for a description of biometric common factor/simplex models). Because of inconsistent findings about non-additive genetic effects across different studies and different methods of assessment and considering the lack of power to detect non-additive genetic effects in the classical twin design (Kandler et al., 2009; Keller, Coventry, Heath, & Martin, 2005; Riemann et al., 1997), genetic effects were assumed to be additive. Environmental effects reflected sibling-specific influences of experiences referred to as non-shared environment, since environmental effects on personality shared by siblings can be assumed to be negligible (Bouchard & Loehlin, 2001; Yamagata, Suzuki, Ando, Ono, Kijima et al., 2006).

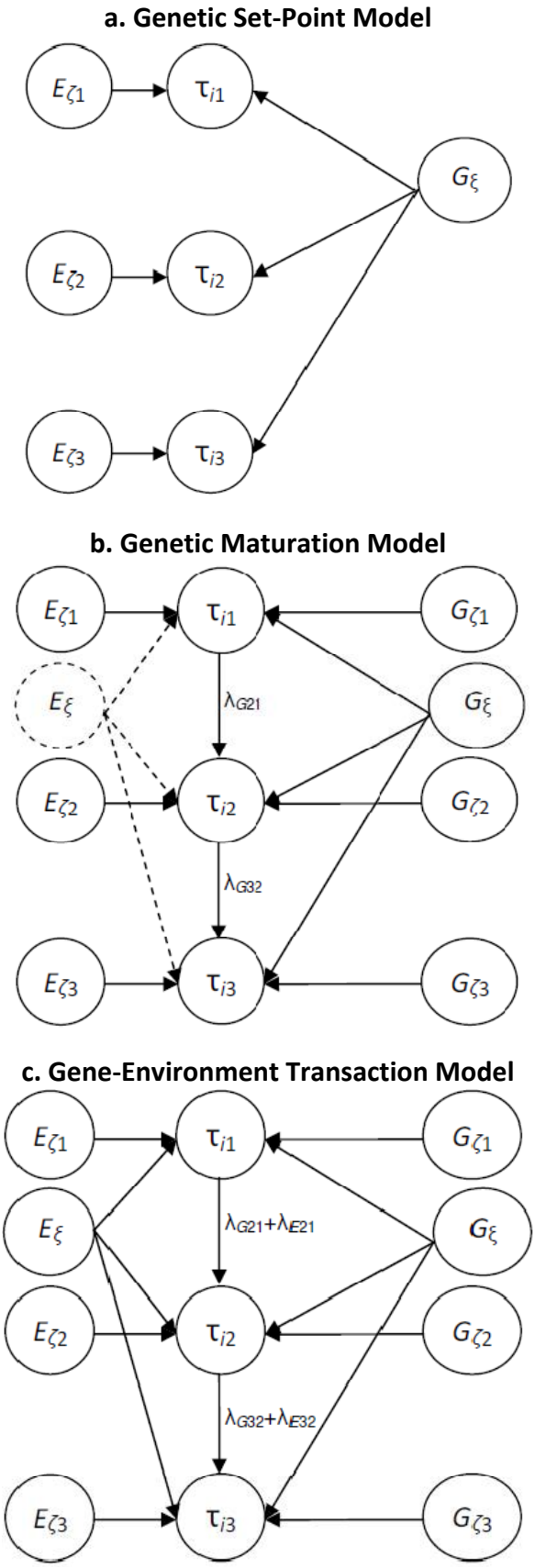
The complete structure model on the right side of Figure 2 reflects the gene-environment transaction model. More restricted models which reflect the two other conceptions (Figures 3a and 3b) are nested in that model, and in turn the genetic set-point model (Figure 3a) is nested in all other models (Figures 3b and 3c). Thus, we were able to compare the models via -2 log-likelihood (-2*LL*) ratio tests (Neale et al., 2003). Furthermore, model modifications were tested that reflect alteration of the three conceptual models. For example, we could differentiate between a strong and a weak position of the genetic maturation hypothesis. Compared to the strong position model, the weak model allows for environmental or non-genetic biological set-point effects (McCrae et al., 2000; marked with dotted lines in Figure 3b).

Prior to model comparisons, we tested for the significance of the inventory method factor. Comparing models, we began with the most parsimonious model, the genetic set-point model (Figure 3a: $E_\xi = G_{\zeta 1} = G_{\zeta 2} = G_{\zeta 3} = \lambda_{G21} = \lambda_{E21} = \lambda_{G32} = \lambda_{E32} = 0$), as the baseline model and compared it with more complex models. First, we compared the baseline model with a model allowing for a genetic set-point in the presence of genetic change (Figure 3b: $E_\xi = \lambda_{E21} = \lambda_{E32} = 0$). This model reflects the strong position of the genetic maturation hypothesis. Then we tested the significance of an additional environmental set-point variable ($\lambda_{G21} = \lambda_{E21} = \lambda_{G32} = \lambda_{E32} = 0$) as well as environmental change ($E_\xi = \lambda_{G21} = \lambda_{G32} = 0$) in presence of a genetic set-point. Proceeding with this bottom-up strategy, more complex models were compared with nested models. For example, the strong genetic maturation model could be compared with the weak genetic maturation model that allowed for an environmental set-point (Figure 3b: $\lambda_{E21} = \lambda_{E32} = 0$). All reduced models were nested in the full model reflecting the gene-environment transaction model (Figure 3c).

We utilized a four-group twin model (young and middle aged MZ and DZ twins) to examine differences in convergent valid parameter estimates between the two age subsamples. For all model comparisons, self- and peer report specific as well as random effects were freely estimated. After we had identified the best fitted structural equation model, we tested for the equivalence of set-point variance components between the young and the middle aged twins. Finally, we computed phenotypic, genetic, and environmental continuity coefficients, which were corrected for measurement error and method-specific effects. These coefficients were computed for each age group and for the short-term (between true score 1 and 2 as well as between true score 2 and 3) and long-term intervals (between true score 1 and 3).

III

Figure 3:



Note. These models only represent portions of the full model presented in Figure 2. *G* = genetic factors; *E* = environmental factors; indices 1, 2, and 3 = time 1, 2, and 3; τ = convergent valid true score variable; ξ = set-point variable; ζ = true score residual; λ = linear regression of the true score variable on the previous true score variable; further description in the text.

Results

Uncorrected phenotypic rank-order stability coefficients for self-reports and mean peer reports show apparent differences between the young and middle aged subsamples (Table 3). Across all personality variables, stability increased with time and age and decreased as the time interval increased. Stability coefficients were consistently smaller for mean peer reports.[17] This refers to factors affecting stability but self-report-specifically. We did not find differences between twin sibling *i* and *j* subsamples.

Table 3:

Phenotypic Rank-Order Stabilities of Self-Reports and Averaged Peer Reports

Scales	Time 1 to 2 (6.35 years)		Time 2 to 3 (6.30 years)		Time 1 to 3 (12.61 years)	
	Self	Peer	Self	Peer	Self	Peer
			Young adulthood			
N =	414	413	184	184	256	256
Neuroticism	.65	.44	.67	.59	.60	.51
Extraversion	.70	.55	.81	.63	.65	.50
Openness	.65	.54	.72	.56	.55	.42
Agreeableness	.60	.48	.66	.49	.47	.34
Conscientiousness	.69	.54	.64	.58	.53	.41
MEAN	**.66**	**.51**	**.70**	**.57**	**.56**	**.44**
			Middle adulthood			
N =	380	377	237	229	322	316
Neuroticism	.62	.60	.75	.69	.62	.55
Extraversion	.75	.63	.77	.64	.73	.51
Openness	.67	.61	.66	.62	.62	.55
Agreeableness	.66	.51	.73	.60	.59	.47
Conscientiousness	.67	.54	.74	.57	.63	.53
MEAN	**.68**	**.58**	**.73**	**.63**	**.64**	**.52**

Note. Statistics are based on pairwise deletion. All correlations were significant ($p < .05$).

Multiple group structural equation modeling is summarized in Table 4. Starting with the genetic set-point model (baseline model 0: $G_\xi + E_\zeta$), model fitting analyses did not reveal significant effects of genetic change (model 1: $G_\xi + E_\zeta + G_\zeta + \lambda_G$) in the complete sample, but significant genetic change was found for Agreeableness and Conscientiousness in the younger subsample. That is, genetic maturation affecting rank-order change was only significant for Agreeableness and Conscientiousness in young adulthood. The inclusion of an environmental set-point variable (model 2: $G_\xi + E_\zeta + E_\xi$) led to an increase in fit over the baseline model ($G_\xi + E_\zeta$) across all personality variables and subsamples, except for Agreeableness in the young subsample. The model allowing for environmental change in addition to a genetic set-point (model 3: $G_\xi + E_\zeta + \lambda_E$) fitted the data significantly better than the baseline model ($G_\xi + E_\zeta$) across all personality variables and subsamples. The model allowing for both envi-

[17] Stability coefficients are presumably lower for averaged peer reports, because targets could have been rated by different raters at the three measurement occasions. This specificity might also account for the difference in stability of method effects between self-reports and peer ratings.

ronmental set-point and environmental change in the presence of a genetic set-point (model 4: $G_\xi + E_\zeta + E_\xi + \lambda_E$) improved fit over that model allowing only for an additional environmental set-point (model 2: $G_\xi + E_\zeta + E_\xi$), consistently for all personality variables. However, it did not lead to an improvement in fit over that model allowing only for environmental change (model 3: $G_\xi + E_\zeta + \lambda_E$), except for Neuroticism, Extraversion, and Agreeableness in the older subsample. In sum, model fitting analyses provided no direct support for a genetic set-point model, because short- as well as long-term (convergent valid) environmental change was significant for all personality traits and age groups. Depending on age (young adults vs. middle adulthood) and the specific trait, different models that reflected compromises of the genetic maturation and the gene-environment transaction model were most suitable to describe the data. That is, a model that allow for a genetic set-point, genetic maturation (for Agreeableness and Conscientiousness), and environmental change ($G_\xi + G_\zeta + \lambda_G + E_\zeta + \lambda_E$) provided the best fit for the data of the young adulthood subsample, whereas a model allowing for a genetic set-point, an environmental set-point (for Neurotcism, Extraversion, and Agreeableness), and environmental change ($G_\xi + E_\xi + E_\zeta + \lambda_E$) were most suitable to describe the data of the middle adulthood sample.

Table 4:

Multiple Group Model Fit Statistics

		Variables				
Model	Fit statistic	N	E	O	A	C
Baseline (BL):	-2*LL*(*df*)	45692.06	42710.40	40335.64	41697.39	42982.37
(0) G_ξ+E_ζ		(7025)	(7025)	(7024)	(7024)	(7022)
	Complete sample					
(1) BL+G_ζ+λ_G	-2*LL*(*df*)	45687.53	42702.35	40327.19	41685.68	42966.42
		(7015)	(7015)	(7014)	(7014)	(7012)
vs. (0) BL	Δ-2*LL*(10)	4.53	8.05	8.45	11.71	15.95
(2) BL+E_ξ	-2*LL*(*df*)	45648.96	42585.40	40299.78	41671.25	42960.80
		(7023)	(7023)	(7022)	(7022)	(7020)
vs. (0) BL	Δ-2*LL*(2)	43.10***	125.00***	35.86***	26.14***	21.57***
(3) BL+λ_E	-2*LL*(*df*)	45623.86	42582.76	40291.20	41667.38	42952.79
		(7021)	(7021)	(7020)	(7020)	(7018)
vs. (0) BL	Δ-2*LL*(4)	68.20***	127.64***	44.44***	30.01***	29.58***
(4) BL+E_ξ+λ_E	-2*LL*(*df*)	45620.69	42572.29	40289.11	41660.81	42951.39
		(7019)	(7019)	(7018)	(7018)	(7016)
vs. (2)BL+E_ξ	Δ-2*LL*(4)	28.27***	13.11***	10.67**	10.44**	9.41*
vs. (3)BL+λ_E	Δ-2*LL*(2)	3.17	10.47***	2.09	6.57**	1.40
	Young adulthood group					
(1) BL+G_ζ+λ_G	-2*LL*(*df*)	45689.98	42703.90	40333.91	41686.32	42972.13
		(7020)	(7020)	(7019)	(7019)	(7017)
vs. (0) BL	Δ-2*LL*(5)	2.08	6.50	1.73	11.07*	10.24*
(2) BL+E_ξ	-2*LL*(*df*)	45677.97	42676.54	40328.93	41694.95	42978.90
		(7024)	(7024)	(7023)	(7023)	(7021)
vs. (0) BL	Δ-2*LL*(1)	14.09***	33.86***	6.71***	2.44	3.47*
(3) BL+λ_E	-2*LL*(*df*)	45665.54	42666.90	40322.91	41689.85	42972.44
		(7023)	(7023)	(7022)	(7022)	(7020)
vs. (0) BL	Δ-2*LL*(2)	26.52***	43.50***	12.73***	8.39**	9.93***

- Continuance of Table 4 -

(4) BL+E_ξ+λ_E	-2*LL*(*df*)	45664.84 (7022)	42665.12 (7022)	40322.91 (7021)	41689.85 (7021)	42972.44 (7019)
vs. (2)BL+E_ξ	Δ-2*LL*(2)	13.13***	11.42***	6.02**	5.10*	6.46**
vs. (3)BL+λ_E	Δ-2*LL*(1)	0.70	1.78	0.00	0.00	0.00
	Middle adulthood group					
(1) BL+G_ζ+λ_G	-2*LL*(*df*)	45689.61 (7020)	42708.88 (7020)	40328.95 (7019)	41696.76 (7019)	42976.70 (7017)
vs. (0) BL	Δ-2*LL*(5)	2.45	1.52	6.69	0.63	5.67
(2) BL+E_ξ	-2*LL*(*df*)	45664.27 (7024)	42621.44 (7024)	40307.48 (7023)	41674.07 (7023)	42966.76 (7021)
vs. (0) BL	Δ-2*LL*(1)	27.79***	89.00***	28.16***	23.32***	15.61***
(3) BL+λ_E	-2*LL*(*df*)	45651.99 (7023)	42627.79 (7023)	40304.30 (7022)	41674.85 (7022)	42963.30 (7020)
vs. (0) BL	Δ-2*LL*(2)	40.07***	82.61***	31.34***	22.54***	19.07***
(4) BL+E_ξ+λ_E	-2*LL*(*df*)	45649.27 (7022)	42619.49 (7022)	40.302.67 (7021)	41668.57 (7021)	42962.07 (7019)
vs. (2)BL+E_ξ	Δ-2*LL*(2)	14.90***	57.05***	4.81*	5.50*	4.69*
vs. (3)BL+λ_E	Δ-2*LL*(1)	2.72*	8.30***	1.63	6.28**	1.23
Best fitting model	-2*LL*(*df*)	**45621.50 (7020)**	**42574.31 (7020)**	**40291.20 (7020)**	**41648.97 (7014)**	**42941.82 (7013)**
vs. BL	Δ-2*LL*(Δ*df*)	**70.56*** (5)**	**136.09*** (5)**	**44.44*** (4)**	**48.42*** (10)**	**40.55*** (9)**

Note. G_ξ = genetic set-point; E_ζ = occasion-specific environmental influences; $G_\zeta + \lambda_G$ = genetic maturation; E_ξ = environmental set-point; λ_E = long-term environmental change; -2*LL* = -2 log-likelihood; Δ = ratio; * p <.10; ** p <.05; *** p <.01.

Across all personality variables, the exclusion of the inventory method factor did not lead to a decline in fit, -2*LL* differences ranged between 0.00 and 0.78 ($\Delta\text{-}2LL_{crit,p<.10}(1)$ = 2.71). Thus, the choice of inventories (NEO-FFI vs. NEO-PI-R) did not affect our results. Testing the equivalence of the degree of genetic set-point variance between age groups did not lead to a decline in fit, too (Δ-2*LL* ranged between 0.00 and 1.80; $\Delta\text{-}2LL_{crit,p<.10}(1)$ = 2.71).

The view to the phenotypic and biometric parameters derived from the best fitting model offers more specific information. Table 5 represents all latent variable variance components derived from the best fitting models (last rows of Table 4) aside from random error (ε). Set-point (ξ) variance in the young subsample was exclusively influenced genetically. In Neuroticism, Extraversion, and Agreeableness, there was an increase in set-point (ξ) variance across age samples as a result of significant environmental set-point variance in the older subsample. Variance due to short- and long-term change (τ – ξ) was smaller in middle adulthood (except time 3 for Conscientiousness) and solely affected by the environment. The younger subsample, on the other hand, showed a larger degree of rank-order change and occasion-specific effects. The decrease of occasion-specific effects and the exclusive genetic set-point variance in young adulthood beyond environmental influences was in line with the genetic set-point as well as the strong genetic maturation hypotheses. The increase, however, of environmental set-point variance across age supported the gene-environment transaction hypothesis.

III

Table 5:

Best Fitting Models: Phenotypic, Genetic, and Environmental Variance Components

Parameters	Age group	Latent variables	Personality variables				
			N	E	O	A	C
Phenotypic variance components	Young	ξ	17.16	14.38	10.83	7.07	12.28
		τ_1- ξ	10.58	9.25	4.35	9.45	13.57
		τ_2- ξ	10.48	7.57	5.10	8.90	7.71
		τ_3- ξ	16.25	8.93	5.45	5.57	4.30
		ν_S	22.65	13.52	8.78	7.24	13.37
		ν_P	0.22	0.00	0.00	1.98	1.68
	Middle	ξ	21.39	18.59	10.83	9.72	12.28
		τ_1- ξ	7.22	5.25	3.45	3.03	4.94
		τ_2- ξ	7.08	3.51	3.83	0.88	5.14
		τ_3- ξ	8.34	3.42	2.66	2.50	5.93
		ν_S	18.15	12.72	7.71	6.62	11.13
		ν_P	5.47	0.00	0.22	4.17	2.52
Set-point components (%)	Young	G_ξ	17.16(100)	14.38(100)	10.83(100)	7.07(100)	12.28(100)
		E_ξ	-	-	-	-	-
	Middle	G_ξ	17.16(80)	14.38(77)	10.83(100)	7.07(73)	12.28(100)
		E_ξ	4.23(20)	4.21(23)	-	2.65(27)	-
Trait change and occasional specificity (%)	Young	$G_{\tau 1-\xi}$	-	-	-	3.99(42)	5.23(39)
		$G_{\tau 2-\xi}$	-	-	-	4.93(55)	1.46(19)
		$G_{\tau 3-\xi}$	-	-	-	2.85(51)	1.79(42)
		$E_{\tau 1-\xi}$	10.58(100)	9.25(100)	4.35(100)	5.46(58)	8.34(61)
		$E_{\tau 2-\xi}$	10.48(100)	7.57(100)	5.10(100)	3.97(45)	6.25(81)
		$E_{\tau 3-\xi}$	16.25(100)	8.93(100)	5.45(100)	2.72(49)	2.51(58)
	Middle	$G_{\tau 1-\xi}$	-	-	-	-	-
		$G_{\tau 2-\xi}$	-	-	-	-	-
		$G_{\tau 3-\xi}$	-	-	-	-	-
		$E_{\tau 1-\xi}$	7.22(100)	5.25(100)	3.45(100)	3.03(100)	4.94(100)
		$E_{\tau 2-\xi}$	7.08(100)	3.51(100)	3.83(100)	0.88(100)	5.14(100)
		$E_{\tau 3-\xi}$	8.34(100)	3.42(100)	2.66(100)	2.50(100)	5.93(100)
Self-report specificity (%)	Young	G_{vs}	15.62(69)	8.54(63)	6.59(75)	3.77(52)	9.15(68)
		E_{vs}	6.03(31)	4.98(37)	2.19(25)	3.47(48)	4.22(32)
	Middle	G_{vs}	12.34(68)	9.18(72)	3.17(41)	3.92(59)	7.79(70)
		E_{vs}	5.81(32)	3.54(28)	4.54(59)	2.70(41)	3.34(30)
Peer report specificity (%)	Young	G_{vp}	0	0	0	0.46(23)	0.82(49)
		E_{vp}	0.22(100)	0	0	1.52(77)	0.86(51)
	Middle	G_{vp}	0	0	0	1.83(44)	0
		E_{vp}	5.47(100)	0	0.22(100)	2.34(56)	2.52(100)

Note. ξ = set point; τ - ξ = trait change + occasion-specific effects; indices 1 to 3 = points of time; ν_S = self-report specific component; ν_P = peer report specific component; *G* = genetic factor; *E* = environmental factor; dashed lines substitute parameters that were fixed in the best fitted model.

Table 6:

Best Fitting Models: Latent Regressions and Convergent Valid Continuity Coefficients

Parameters	Mean age interval	Latent parameters	Variables N	E	O	A	C
Standardized regression parameters	23-29	λ_{G21}	-	-	-	.30	.16
	29-35	λ_{G32}	-	-	-	.29	.07
	23-29	λ_{E21}	.37	.50	.37	.43	.49
	29-35	λ_{E32}	.73	.80	.67	.67	.65
	41-48	λ_{G21}	-	-	-	-	-
	48-55	λ_{G32}	-	-	-	-	-
	41-48	λ_{E21}	.21	.33	.79	.17	.59
	48-55	λ_{E32}	.59	.34	.82	.32	.72
Phenotypic continuity	23-29	$r_{1\text{-}2}$	.76	.81	.81	.77	.80
	29-35	$r_{2\text{-}3}$	.88	.93	.89	.88	.87
	23--35	$r_{1\text{--}3}$	.67	.72	.78	.64	.68
	41-48	$r_{1\text{-}2}$	.83	.93	.95	.90	.88
	48-55	$r_{2\text{-}3}$	.97	.96	.96	.98	.93
	41--55	$r_{1\text{--}3}$	.78	.88	.88	.80	.82
Genetic continuity	23-29	$r_{G1\text{-}2}$	1.00	1.00	1.00	.91	.95
	29-35	$r_{G2\text{-}3}$	1.00	1.00	1.00	.94	.95
	23--35	$r_{G1\text{--}3}$	1.00	1.00	1.00	.79	.82
	41-48	$r_{G1\text{-}2}$	1.00	1.00	1.00	1.00	1.00
	48-55	$r_{G2\text{-}3}$	1.00	1.00	1.00	1.00	1.00
	41--55	$r_{G1\text{--}3}$	1.00	1.00	1.00	1.00	1.00
Environmental continuity	23-29	$r_{E1\text{-}2}$	.37	.50	.37	.43	.49
	29-35	$r_{E2\text{-}3}$	.73	.80	.67	.67	.65
	23--35	$r_{E1\text{--}3}$	.25	.28	.28	.26	.27
	41-48	$r_{E1\text{-}2}$	.58	.82	.79	.76	.59
	48-55	$r_{E2\text{-}3}$	.94	.89	.82	.95	.72
	41--55	$r_{E1\text{--}3}$	.47	.67	.44	.54	.41

Note. λ_{21} = regression coefficients of true score 2 on true score 1; λ_{32} = regression coefficients of true score 3 on true score 2; *r* = latent correlations between different points of time 1, 2, and 3 corrected for specific effects of self- and peer reports and measurement error; *G* = genetic component; *E* = environmental component; dashed lines substitute parameters that were fixed in the best fitted model; number of hyphens between mean ages and points of time reflect number of time intervals.

Somewhat surprisingly, a very large proportion of self-report specific (v_S) variance was found to be stable across time and age subsamples. This stable component showed substantial genetic influences consistently across personality variables and age subsamples indicating that self-reports include large portions of self-report-specific components (not shared with peer reports) which were long-term stable and basically attributable to genetic factors.

Adding up all genetic and environmental variance components for each method at each measurement occasion corrected for instable method effects and random error, we calculated heritability estimates of 59-79% (median = 68%) for self-reports and 51-81% (me-

dian = 67%) for peer reports in the young sample as well as 57-71% (median = 64%) for self-reports and 49-78% (median = 60%) for peer reports in the middle-aged sample. The same pattern was found for true score variance corrected for stable method effects (v_S and v_P): heritability decreased from 68% (51-85%) in the young sample to 65% (55-80%) in the middle-aged sample. This indicates, first, a larger heritability for personality reports when corrected for both instable and stable method effects as well as random error, and second, slight decreases of heritability for both self- and peer reports across age, a finding which has already been reported for (uncorrected) self-reports by Viken et al. (1994). Larger heritability coefficients for true score variance components provided support for the genetic maturation hypothesis that postulated strong rater-specific components due to non-genetic method effects. The decrease of heritability, however, was exclusively in line with the gene-environment transaction hypothesis.

Beyond the variance components, we computed latent phenotypic correlations between true scores as well as correlations due to genetic and environmental factors (Table 6) for each mean age interval.[18] These coefficients reflected continuity corrected for measurement error and method-specific effects. Based on the best fitting model some parameters were zero (e.g., E_ξ). This simplified the formulas whereby, for example, standardized regression coefficients (e.g., λ_{E21} and λ_{E32}) accorded with continuity coefficients (e.g., $r_{E1\text{-}2}$ and $r_{E2\text{-}3}$). Phenotypic continuity increased with time and age and was lower in long-term intervals that reflect rank-order change. Genetic factors influenced long-term stability, whereas genetic effects on rank-order change played only a role in young adulthood (for Agreeableness and Conscientiousness). Environmental factors primarily affected short-term stability and rank-order change in personality. Environmental continuity cumulated with time and age while this increase was consistently larger in young adulthood. Generally, long- and short-term environmental continuity was found to be larger in middle than in young adulthood. The increase of continuity due to environmental factors provided strong support for the gene-environment transaction hypothesis.

Discussion

The primary aim of the present study was to examine the sources of personality rank-order continuity and change. In general, our analyses yielded most support for the gene-environment transaction hypothesis (Caspi et al., 2005), because environmental factors were the primary source of change within and across young and middle adulthood. However, our findings also support aspects of the genetic maturation hypothesis (McCrae et al., 2000) and the genetic set-point hypothesis (Carey, 2002), since set-point variance in young adulthood was exclusively influenced by genetic factors. Occasion-specific effects appear to decrease with age which was a specific deduction from the genetic set-point hypothesis (Carey, 2002). Moreover, different models fitted the data depending on the age group data and traits which were analyzed. Considering the whole pattern of findings, we thus propose a combined model that integrates relevant assumptions from each of the three theoretical approaches. The resulting conception is in fact very similar to an earlier developmental conception by Scarr and McCartney (1983) proposing that experiences are directed by genotypes.

[18] Genetic continuity refers to enduring effects shared by different methods as a function of twins' resemblance. Environmental continuity refers to enduring effects that are found in both self-reports and peer ratings but are not shared by twins.

Cumulative Continuity in Personality

The increasing phenotypic continuity of personality can be interpreted as a process of developing, committing to, and maintaining an identity (Roberts & Caspi, 2003). Caspi et al. (2005) proposed two mutually supportive life-course dynamics: *social selection* and *social influence*. They assume that individuals select environments that are correlated with their personality traits which in turn provide experiences that affect personality.

In line with previous behavioral genetic studies on personality development (e.g., McGue et al., 1993; Viken et al., 1994), we found personality stability to be primarily influenced by genetic factors. In particular, we found 13-year continuity of convergent valid true scores in young adulthood to be exclusively influenced by genetic factors. Thereby, the amount of genetic variance did not differ between young and middle aged adult twins. The exclusive genetic influence on long-term stability in personality traits of young adults supports the assumption of genetic *set-points* (Carey, 2002) or *basic tendencies* (McCrae et al., 2000). Regarding the theory of *social selection* (Caspi et al., 2005), it might be emphasized that young adults select environments that are correlated with their genotypic (not environmental) personality set-points. Genotypes could affect emerging personality phenotypes of young individuals possibly both directly and through prompting new experiences (Scarr & McCartney, 1983).

Though long-term stability in young adulthood is not affected by environmental factors in our study, environmental continuity seems to increase (negatively accelerated) with age. In middle adulthood, environmental factors play a significant role in 13-year continuity of Neuroticism, Extraversion, and Agreeableness. The increase of environmental continuity might be the result of a negatively accelerated accumulation of individual-specific *social influences* across the life course, as mentioned by Caspi et al. (2005). However, the process linking environment to personality development has not been studied here. Thus it remains open whether environmental influences are of social nature (Caspi et al., 2005) or mediated through biological changes (McCrae & Costa, 2008). The negatively accelerated accumulation might be the result of an increase of personality stabilization (committing to and maintaining an identity) *and* a decrease of occasion-specific contextual effects (minor effects of or fewer life transitions). Both mechanisms may lead to the well established cumulative phenotypic continuity of personality (Roberts & DelVecchio, 2000).

The combination of the two mutually supportive life-course dynamics (Caspi et al., 2005) is also known as genotype–environment correlation (Plomin, DeFries, & Loehlin, 1977; Scarr & McCartney, 1983). The personality genotype influences the probability of exposure to certain events (*social selection*) by evoking reactions and seeking out settings as well as modifying and creating situations. In turn, the selected social contexts allocate experiences that affect individuals' development of personality (*social influence*) accumulating across the life-span. Thereby, *social selection* directed by personality genotypes should be correlated among relatives because of genetic relatedness (e.g., attending university and majoring in the same field of study), but correlated contexts do not necessarily allocate the same *social influences* on personality (e.g., different study conditions and different fellow students). Maybe, the chance (e.g., allocation to different colleges) plays a minor role in the short but a major role in the long run separating twins' development (Dunn & Plomin, 1990). As a consequence, heritability of personality should rather decrease across adulthood because of the accumulation of specific environmental influences, which is in line with findings on self-reports by Viken et al. (1994) and was also revealed in our multiple-informant study. Scarr

and McCartney (1983) mentioned that the impetus for certain experiences comes from the genotypes, whereas phenotypes are elaborated and maintained by environments.

Rank-Order Change

The design of our study allowed us to differentiate between genetic and environmental long-term change and occasion-specific effects. Generally, there was no evidence for genetic change, except for Agreeableness and Conscientiousness in young adulthood. A previous study (Bleidorn et al., 2009), utilizing monomethod data partially from the same twin sample of BiLSAT, found that variance in individual-level change of Agreeableness, Conscientiousness, and Neuroticism was primarily attributable to genetic factors. Our study did not find evidence for genetic change in Neuroticism. One explanation of this divergence might be the sole reliance on self-reports in the previous study. On the other hand, it should be noted that the structural equation models in the present study focused exclusively on the relative ordering of individuals and were sensitive to detect the relative change of individuals' ranks. The models were not sensitive to detect genetic and environmental variance in systematic intraindividual-level growth or decline over time as it was the focus in the previous study. Our analyses were sensitive to detect increase and decrease of relative change over time. Genetic change in Agreeableness and Conscientiousness decreased with age in young adulthood and was not significant in the middle adulthood, indicating decreasing effects of genetic maturation on rank-order change across young and middle adulthood.

In contrast, long-term change in personality was consistently found to be predominantly attributable to environmental factors in both young and middle adulthood. Across personality variables, change seems to decrease with age. The higher degree of long-term change in young adulthood may be attributed to more or larger effects of major life transitions in this period of life. In young adulthood, individuals usually have to decide which life goals (e.g., career, family) they primarily want to pursue and how to shape their life course (e.g., vocational training, starting a family). The therewith associated transitions may also contribute to a higher degree of personality trait change. In spite of a higher degree of continuity in later adulthood, long-term stability was still found to be lower than short-term stability indicating that personality is not fixed in that period of life and change way beyond young adulthood, too.

Occasion-Specific Effects

According to the definition of true score residuals termed as "state residuals" in the latent state-trait theory (Steyer et al., 1999), occasion-specific effects contain influences of the situation in which the individual's phenotype is measured and (or) effects due to the interaction between person and situation. Unfortunately, in our analyses it was not possible to control for interaction effects. Purcell (2002) showed that interaction between additive genetic factors and specific environments acts like effects of the specific environment in the classical design of twins reared together when interaction is present but not estimated. Thus, environmental occasion-specific effects may be also due to influences of genotype–environment interaction (Plomin, DeFries, & Loehlin, 1977): genetic effects depend on the environments or, the other way round, environmental effects depend on the genotypes.

Referring to the FFT (McCrae & Costa, 2008), occasion-specific environmental influences may thus be interpreted in terms of *characteristic adaptations* of personality genotypes (*basic tendencies*) that respond to the opportunities and incentives of social contexts.

As already mentioned for long-term change, the higher degree of occasion-specific effects in young adulthood may be attributed to larger effects of or easily to more major life transitions in this period of life. Likewise, the revealed occasion-specific genetic effects on Agreeableness and Conscientiousness in young adulthood may be interpreted as *characteristic adaptations* of personality genotypes which are shared between twins reared together, because interaction between additive genetic factors and shared environments acts like additive genetic effects when interaction is present but not estimated (Purcell, 2002).

Self-Report Specific Stability

A large proportion of method-specific variance in self-reports was stable whereas the corresponding variance in averaged peer reports was not. In addition, non-shared environmental effects on long-term rank-order stability appear to be first and foremost self-report specific. These results were certainly not the major focus of our study but deserve discussion (Kandler et al., in press). It may indicate that there are stable individual differences in self-report response styles (McCrae & Costa, 2008). In this regard it should be noted that the self-report method factors across all personality variables and both age subsamples was substantially influenced by genetic factors. Thus, this component may also reflect personality characteristics that are not readily accessible to the peer raters (e.g., motives, self-concept).

This interpretation is in line with Kendler's (2001) two pathways in which genes may be effective: *within-the-skin* and *outside-the-skin*. Genes may affect personality through these two pathways. *Within-the-skin*, genetic expression takes place in internal personality features (e.g. motives and emotions). *Outside-the-skin*, genes can also affect personality-relevant behavior and its social consequences. The *outside-the-skin* pathway is more readily perceivable by peers. Thus, it might be argued that expressions of personality genotypes have to be differentiated into internal (intentions and self-concept) and external (expressed behavior and its social consequences) effects. A recent study found evidence that self- and other raters appear to pay attention to different information cues when judging personality (Hofmann, Gschwender, & Schmitt, 2009). Moreover, observers seem to focus rather on states, whereas self-raters primarily focus on their stable attributes, even when they are instructed to focus on their states. This is in line with our results and the idea that stable method-specific variance in self-reports may reflect valid information on personality not accessible by peers. These results call for future studies addressing this issue. Much remains to be learned about the primary processes of introspective and external personality judgments and perceptions.

Conclusions

In sum, the results of our study led us to formulate main conclusions. First, genetic factors affect rank-order stability in personality directly and possibly through experiences resulting from genotype–environment correlations (Scarr & McCartney, 1983). Second, genetic factors remain stable across adulthood, whereas environmental influences trigger both an increase in phenotypic continuity with age and a decrease in phenotypic continuity with increasing time intervals between assessments. Third, self-report specific variance is largely stable and genetically influenced which might reflect internal effects on personality phenotypes that are less perceivable to other persons. Furthermore, we can conclude that much remains to be learned about the primary processes involved in *social selection* and *social influence* as well as the mechanisms underlying gene- environment transaction (Caspi et al., 2005) and genotype-environment correlation affecting personality development. The pre-

sent work may just serve as a basis in order to continue with research into the processes and mechanisms of personality development over the life span.

References

Bleidorn, W., Kandler, C., Riemann, R., Angleitner, A., & Spinath, F. (2009). Patterns and sources of adult personality development: Growth curve analyses of the NEO-PI-R scales in a longitudinal twin study. *Journal of Personality and Social Psychology*, *97*, 142-155.

Blonigen, D.M., Carlson, M.D., Hicks, B.H., Krueger, R.F., & Iacono, W.G. (2008). Stability and change in personality traits from late adolescence to early adulthood: A longitudinal twin study. *Journal of Personality*, *76*, 229-266.

Boomsma, D.I. & Molenaar, P.C.M. (1987). The genetic analysis of repeated measures. I. simplex models. *Behavior Genetics*, *17*, 111-123.

Borkenau, P. & Ostendorf, F. (1993). *NEO-Fünf-Faktoren-Inventar (NEO-FFI)* [NEO Five-Factor Inventory]. Göttingen: Hogrefe.

Bouchard, T.J. & Loehlin, J.C. (2001). Genes, evolution, and personality. *Behavior Genetics*, *31*, 243-273.

Bratko, D. & Butkovic, A. (2007). Stability and genetic and environmental effects from adolescence to young adulthood: Results of Croatian longitudinal twin study of personality. *Twin Research and Human Genetics*, *10*, 151-157.

Campbell, D.T. & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81-105.

Carey, G. (2002). *Human Genetics for the Social Sciences*. Thousand Oaks, CA: Sage.

Caspi, A. & Roberts, B.W. (2001). Personality development across the life course: The argument for change and continuity. *Psychological Inquiry*, *12*, 49-66.

Caspi, A., Roberts, B.W., & Shiner, R.L. (2005). Personality development: Stability and change. *Annual Review of Psychology*, *56*, 453-484.

Costa, P.T. & McCrae, R.R. (1988). Personality in adulthood: A six year longitudinal study of self-reports and spouse ratings of the NEO personality inventory. *Journal of Personality and Social Psychology*, *54*, 853-863.

Costa, P.T. & McCrae, R.R. (1989). *NEO Personality Inventory: Manual Form S and Form R*. Odessa, FL: Psychological Assessment Resources.

Costa, P.T. & McCrae, R.R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory: Professional manual.* Odessa, FL: Psychological Assessment Resources.

Derks, E.M., Dolan, C.V., & Boomsma, D.I. (2006). Statistical power to detect genetic and environmental influences in the presence of data missing at random. *Twin Research and Human Genetics*, *10*, 159-167.

Dunn, J. & Plomin, R. (1990). *Separate Lives. Why Siblings Are so Different?* New York: Basic Books.

Fraley, C. & Roberts, B.W. (2005). Patterns of continuity: A dynamic model for conceptualizing the stability of individual differences in psychological constructs across the life course. *Psychological Review*, *112*, 60-74.

Hewitt, J.K., Eaves, L.J., Neale, M.C., &, Meyer, J.M. (1988). Resolving causes of developmental continuity or “tracking”. I. Longitudinal twin studies during growth. *Behavior Genetics*, *18*, 133-151.

Hofmann, W., Gschwender, T., & Schmitt, M. (2009). The road to the unconscious self not taken: Discrepancies between self- and observer-inferences about implicit dispositions from nonverbal behavioural cues. *European Journal of Personality*, *23*, 343-366.

Hofstee, W.K.B. (1994). Who should own the definition of personality? *European Journal of Personality*, *8*, 149-162.

Johnson, W., McGue, M., & Krueger, R. F. (2005). Personality stability in late adulthood: A behavioral genetic analysis. *Journal of Personality*, *73*, 523-551.

Kandler, C., Riemann, R., & Kämpfe, N. (2009). Genetic and environmental mediation between measures of personality and family environment in twins reared together. *Behavior Genetics*, *39*, 24-35.

Kandler, C., Riemann, R., Spinath, F. M., & Angleitner, A. (in press). Sources of variance in personality facets: A multiple-rater twin study of self-peer, peer-peer, and self-self (dis-) agreement. *Journal of Personality.*

Keller, M. C., Coventry, W. L., Heath, A. C., & Martin, N. G. (2005). Widespread evidence for non-additive genetic variation in Cloninger's and Eysenck's personality dimensions using a twin plus sibling design. *Behavior Genetics*, *35*, 707-721.

Kendler, K.S. (2001). Twin studies of psychiatric illness. An update. *Archives of General Psychiatry*, *58*, 1005-1014.

Krueger, R.F., Markon, K.E., & Bouchard, T.J.Jr. (2003). The extended genotype: The heritability of personality accounts for the heritability of recalled family environments in twins reared apart. *Journal of Personality*, *71*, 809-834.

Little, R.J.A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of American Statistical Association*, *83*, 1198-1202.

McCrae, R.R. & Costa, P.T. (2008). The five-factor theory of personality. In John, O.P., Robins, R.W., & Pervin, L.A. (Eds.), *Handbook of Personality: Theory and Research* (3rd ed.). New York: Guilford.

McCrae, R.R., Costa, P.T., Ostendorf, F., Angleitner, A., Hrebickova, M., Avia, M.D., et al. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology*, *78*, 173-186.

McCrae, R.R. & John, O.P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*, 175-215.

McGue, M., Bacon, S., & Lykken, D.T. (1993). Personality stability and change in early adulthood: A behavioral genetic analysis. *Developmental Psychology*, *29*, 96-109.

Neale, M.C. & Maes, H.H.M. (2004). *Methodology for Genetic Studies of Twins and Families*. Dordrecht: Kluwer Academic Publishers B.V.

Neale, M.C., Boker, S.M., Xie, G., & Maes, H.H.M. (2003). *Mx: statistical modeling*. VCU Box 900126, Richmond, VA 23298, 6th ed. Department of Psychiatry

Ostendorf, F. & Angleitner, A. (2004). *NEO-PI-R: NEO-Persönlichkeitsinventar nach Costa und McCrae, revidierte Fassung* [NEO-Personality Inventory revised by Costa and McCrae]. Göttingen: Hogrefe.

Pederson, N.L. & Reynolds, C.A. (1998). Stability and change in adult personality: Genetic and environmental components. *European Journal of Personality*, *12*, 365-386.

Pederson, N.L. & Reynolds, C.A. (2002). Stability and change in adult personality: Genetic and environmental components [Erratum]. *European Journal of Personality*, *16*, 77-78.

Plomin, R., DeFries, J.C., & Loehlin, J.C. (1977). Genotype – environment intercation and correlation in the analysis of human behavior. *Psychological Bulletin*, *84*, 309-322.

Purcell, S. (2002). Variance components models for gene-environment interaction in twin analysis. *Twin Research*, *5*, 554-571.

Riemann, R., Angleitner, A., & Strelau, J. (1997). Genetic and environmental influences on personality: A study of twins reared together using self- and peer report NEO-FFI scales. *Journal of Personality*, *65*, 449-475.

Riemann, R., & Kandler, C. (in press). Construct validation using multitrait-multimethod twin data: The case of a general factor of personality. *European Journal of Personality.*

Roberts, B.W. & Caspi, A. (2003). The cumulative continuity model of personality development: Striking a balance between continuity and change in personality traits across the life course. In U.M. Staudinger & U. Lindenberger (Eds.) *Understanding Human Development: Dialoges with Lifespan Psychology* (pp. 183-214). Dorrecht, Netherlands: Kluwer Academic Publishers.

Roberts, B.W. & DelVecchio, W.F. (2000). The rank-order consistency of personality from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, *126*, 3-25.

Scarr, S. & McCartney, K. (1983). How people make their own environments: A theory of genotype → environment effects. *Child Development*, *54*, 424-435.

Spinath, F. M., Wolf, H., Angleitner, A., Borkenau, P., & Riemann, R. (2005). Multimodale Untersuchung von Persönlichkeit und kognitiven Fähigkeiten. Ergebnisse der deutschen Zwillingsstudie BiLSAT und GOSAT. *Zeitschrift für Soziologie in der Erziehung*, *25*, 146-161.

Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state-trait theory and research in personality and individual differences. *European Journal of Personality*, *13*, 389-408.

Terracciano, A., Costa, P. T., Jr., & McCrae, R. R. (2006). Personality plasticity after age 30. *Personality and Social Psychology Bulletin*, *32*, 999-1009.

Viken, R.J., Rose, R.J., Kaprio, J., & Koskenvuo, M. (1994). A developmental genetic analysis of adult personality: Extraversion and Neuroticism from 18 to 59 years of age. *Journal of Personality and Social Psychology*, *66*, 722-730.

Yamagata, S., Suzuki, A., Ando, J., Ono, Y., Kijima, N., Yoshimura, K., et al. (2006). Is the genetic structure of human personality universal? A cross-cultural twin study from North America, Europe, and Asia. *Journal of Personality and Social Psychology*, *90*, 987-998.

ANHANG IV
CURRICULUM VITAE

Christian Kandler, geb. Döbrich (amtl.), geb. Eichhorn (real), wurde am 20.01.1982 in Neuhaus am Rennweg geboren. Als leiblicher Sohn von Martina Diez und Frank Koch ist er Kind einer multifaktoriellen Patch-Work-Familie (Einzelkind ersten Grades, eine Stiefmutter, ein Stiefvater, vier Halbschwestern, zwei Halbbrüder, eine Stiefschwester, vier Stiefbrüder). Seit 2008 ist er verheiratet mit Anja.

Nach dem Erreichen der Hochschulreife 2000 absolvierte er einen zweijährigen Wehrdienst. Er wurde zum Beobachtungsoffizier ausgebildet. Nach Abschluss der aktiven Dienstzeit besuchte er verschiedene Wehrübungen (2003 Beförderung zum Leutnant der Reserve). Er studierte von 2002 bis 2007 an der Friedrich-Schiller-Universität Jena Psychologie. In dieser Zeit war er zwei Jahre lang Mitglied im Fachschaftrat (2004-2005) und bei Prof. Dr. Riemann im Projekt JeTSSA als Hilfswissenschaftler tätig (2005-2006). Seit 2007 arbeitet er in der Arbeitsgruppe von Prof. Dr. Riemann an der Universität Bielefeld als Wissenschaftlicher Angestellter mit dem Forschungsschwerpunkt Verhaltensgenetik.

Seit 2009 ist er zusätzlich als Gutachter in der lokalen Ethikkommission an der Fakultät für Psychologie (Universität Bielefeld) tätig. Von 2009 bis 2010 führte er eine fakultätsübergreifenden Erfassung von Einstellungen der Studierenden gegenüber englischsprachiger Lehre durch. Er ist Mitglied im Bundeswehrverband (seit 2000), im Thüringer Turn- und Sportverein (seit 2006) und in der Deutschen Gesellschaft für Psychologie (Fachgruppe für Differentielle Psychologie, Persönlichkeitspsychologie und Psychologische Diagnostik; seit 2008).