



NEPS WORKING PAPERS

Michael Bayer, Florian Wohlkinger, Jan-David

Freund, Hartmut Ditton und Sabine Weinert

TEMPERAMENT BEI KLEINKINDERN – THEORETISCHER HINTERGRUND, OPERATIONALISIERUNG IM NATION- ALEN BILDUNGSPANEL (NEPS) UND EMPIRISCHE BEFUNDE AUS DEM FORSCHUNGSPROJEKT VIVA

NEPS Working Paper No. 58
Bamberg, Mai 2015

Working Papers of the German National Educational Panel Study (NEPS)

at the Leibniz Institute for Educational Trajectories (LifBi) at the University of Bamberg

The NEPS Working Papers publish articles, expertises, and findings related to the German National Educational Panel Study (NEPS).

The NEPS Working Papers are edited by a board of researchers representing the wide range of disciplines covered by NEPS. The series started in 2011.

Papers appear in this series as work in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character.

Any opinions expressed in this series are those of the author(s) and not those of the NEPS Consortium.

The NEPS Working Papers are available at

<https://www.neps-data.de/projektübersicht/publikationen/nepsworkingpapers>

Editorial Board:

Jutta Allmendinger, WZB Berlin

Cordula Artelt, University of Bamberg

Jürgen Baumert, MPIB Berlin

Hans-Peter Blossfeld, EUI Florence

Wilfried Bos, University of Dortmund

Claus H. Carstensen, University of Bamberg

Henriette Engelhardt-Wölfler, University of Bamberg

Frank Kalter, University of Mannheim

Corinna Kleinert, IAB Nürnberg

Eckhard Klieme, DIPF Frankfurt

Cornelia Kristen, University of Bamberg

Wolfgang Ludwig-Mayerhofer, University of Siegen

Thomas Martens, DIPF Frankfurt

Manfred Prenzel, TU Munich

Susanne Rässler, University of Bamberg

Marc Rittberger, DIPF Frankfurt

Hans-Günther Roßbach, LifBi

Hildegard Schaeper, DZHW Hannover

Thorsten Schneider, University of Leipzig

Heike Solga, WZB Berlin

Petra Stanat, IQB Berlin

Volker Stocké, University of Kassel

Olaf Struck, University of Bamberg

Ulrich Trautwein, University of Tübingen

Jutta von Maurice, LifBi

Sabine Weinert, University of Bamberg

Contact: German National Educational Panel Study (NEPS) – Leibniz Institute for Educational Trajectories – Wilhelmsplatz 3 – 96047 Bamberg – Germany – contact@lifbi.de

Temperament bei Kleinkindern – Theoretischer Hintergrund, Operationalisierung im Nationalen Bildungspanel (NEPS) und empirische Befunde aus dem Forschungsprojekt ViVA

Michael Bayer

Leibniz-Institut für Bildungsverläufe, Bamberg

Florian Wohlkinger

Ludwig-Maximilians-Universität, München

Jan-David Freund

Otto-Friedrich-Universität, Bamberg

Hartmut Ditton

Ludwig-Maximilians-Universität, München

Sabine Weinert

Otto-Friedrich-Universität, Bamberg

E-Mail-Adresse des Erstautors:

michael.bayer@lifbi.de

Bibliographische Angaben:

Bayer, M., Wohlkinger, F., Freund, J.-D., Ditton, H. & Weinert, S. (2015). Temperament bei Kleinkindern – Theoretischer Hintergrund, Operationalisierung im Nationalen Bildungspanel (NEPS) und empirische Befunde aus dem Forschungsprojekt ViVA (NEPS Working Paper No. 58). Bamberg: Leibniz-Institut für Bildungsverläufe, Nationales Bildungspanel.

Temperament bei Kleinkindern – Theoretischer Hintergrund, Operationalisierung im Nationalen Bildungspanel (NEPS) und empirische Befunde aus dem Forschungsprojekt ViVA

Abstract

Das Temperament von Kindern in den ersten Lebensjahren stellt eine wichtige Forschungsperspektive sowohl in Hinblick auf deren Persönlichkeitsentwicklung als auch ihre sozial-emotionale Entwicklung dar. Gleichzeitig jedoch sind die vorhandenen und in der Temperamentsforschung gebräuchlichen Messinstrumente meist sehr umfangreich und damit für eine thematisch breit angelegte Studie wie das Nationale Bildungspanel (NEPS) eher ungeeignet. Im Arbeitspapier wird die Entwicklung sowie die Operationalisierung und Validierung eines Kurzinstrumentes zur Messung von Temperamentsdimensionen mittels Fremdeinschätzung durch einen Elternteil vorgestellt. Hierfür wurde auf vorhandene Instrumente zurückgegriffen, die im Rahmen einer Teststudie zu einem im NEPS einsetzbaren Kurzinstrument umgestaltet wurden. Das anschließend in den Erhebungen des Nationalen Bildungspanels eingesetzte Instrument umfasst zwei zentrale Temperamentsdimensionen und wurde im Rahmen des Forschungsprojekts ViVA mit dem ursprünglichen, deutlich umfangreicheren Instrument verglichen. Damit können erste Ergebnisse zur Güte des Kurzinstrumentes nachvollzogen werden. Das Arbeitspapier dokumentiert sämtliche Entscheidungen, die im Prozess der Anpassung und vor allem der Kürzung bzw. der Streichung von Items getroffen wurden.

Schlagworte

Temperament, IBQ-R, Persönlichkeit

1. Theoretischer Hintergrund und Forschungsstand

1.1 Temperament als Persönlichkeitsaspekt

Die Messung von Temperament als ein begrifflich und konzeptionell distinktes Phänomen basiert auf einer langen Forschungstradition (für einen Überblick Rothbart & Bates, 1998). Bereits in der Antike wurden Menschen mittels eines auf vier Grundtypen basierenden Temperamentkonzepts typologisch beschrieben. Die sich in solchen Typologien ausdrückende Idee, dass das Temperament eine Art emotionaler Grundgestimmtheit des menschlichen Charakters darstellt, findet sich nach wie vor in Teilbereichen der Temperamentsforschung wieder. Der Temperamentsbegriff ist zudem in vielfältiger Weise in die Alltagssprache diffundiert, was eine begriffliche Vereinheitlichung deutlich erschwert. Stellte „Charakter“ in den ersten Jahrzehnten des 20. Jahrhunderts eine Art Oberbegriff auch für den Begriff des Temperaments dar (vgl. etwa Webb, 1915; Baumgarten, 1933), so lässt sich im weiteren zeitlichen Verlauf eine Verschiebung hin zu einer Zuordnung zum Rahmenkonzept der Persönlichkeit feststellen (so etwa Allport, 1961). Die grundsätzliche Frage, was denn unter Temperament zu verstehen ist, ist zwar keineswegs endgültig geklärt; nichtsdestotrotz lässt sich ein Konsens dahingehend feststellen, dass Temperament bzw. Temperamentsentwicklung als ein Teilprozess der allgemeineren Persönlichkeitsentwicklung begriffen wird. Ob und inwieweit es sich hierbei um ein vor allem emotional oder aber kognitiv verankertes Phänomen handelt, ist hingegen weder geklärt noch dürfte dies abschließend zu klären sein. Jenseits der Forschung im Kontext von eher persönlichkeitspsychologischen Ansätzen wird Temperament und Temperamentsentwicklung auch im Bereich der komparativen Psychologie (Elliot & Thrash, 2002) und im Feld der Verhaltensgenetik (Nigg & Goldsmith, 1998) untersucht.

Insbesondere die Arbeiten von Rothbart (1989), aber auch von Bates (1980) sowie Buss und Plomin (1984), welche Versuche darstellen zu einer operationalen Definition von Temperament zu kommen, hatten gleichzeitig auch das Anliegen, die Distanz zwischen den beiden getrennten Bereichen „Temperament“ und „Persönlichkeit“ zu überbrücken. In der Definition von Rothbart und Bates – „we define temperament as constitutionally based on individual differences in emotional, motor, and attentional reactivity and self-regulation“ (Rothbart & Bates, 1998, S.109) – wird ein entsprechend breites Verständnis von Temperament angelegt, welches ermöglicht, die Frage nach der „Natur“ des Phänomens sowie die Frage nach den zentralen konstitutiven Merkmalen weitgehend offen zu halten. Damit sind es nicht nur emotionale Grundgestimmtheiten und motorische Ausdrücke, welche im Zentrum einer operationalen Definition stehen sondern mit dem Begriff der Selbst-Regulation findet eine dezidiert kognitive Facette Eingang in das so definierte Temperamentskonzept.

Auf dieses multidimensionale Verständnis von Temperament setzen die Überlegungen auf, welche für die Implementierung eines Messinstruments des kindlichen Temperaments im Rahmen der Erhebungen des Nationalen Bildungspanels bei Kleinkindern vorgenommen wurden. Es geht hierbei nicht um eine Klärung der Frage nach dem Erzeugungsgrund für individuelle Differenzen; vielmehr soll mit der Messung des Temperaments die Möglichkeit geboten werden, (1) Temperamentsentwicklung durch wiederholte Messungen nachzuzeichnen und (2) darüber hinaus Temperament als eine Art Vorläuferkonstrukt von

Persönlichkeitseigenschaften sowie sozial-emotionalen Kompetenzen im späteren Leben darzustellen.

Vor allem der letzte Punkt, Temperament in den Rahmen allgemeinerer Persönlichkeits(entwicklung) einzuordnen, stellt eine theoretisch-konzeptionelle Vorentscheidung dar, bedeutet jedoch keineswegs eine Zurückweisung andersartiger Auffassungen. Die nach wie vor virulente Frage, ob und inwieweit das je individuelle Temperament und dessen Entwicklung biologisch-genetische Ursachen hat (vgl. Eysenck, 1967; Gray, 1982; Zuckerman, 1991) steht außerhalb des Fokus des hier vorzustellenden Konzepts.

Die Temperamentsentwicklung in ihren emotionalen wie auch kognitiven Dimensionen im Kontext einer Bildungsstudie zu untersuchen stellt keinen Selbstzweck dar, sondern schließt an Forschungen an, welche sich mit den (ursächlichen) Wirkungen von Temperament auf spätere Entwicklungen und insbesondere auf schulische und berufliche Outcomes beschäftigen (vgl. v.a. Lawson & Ruff, 2004; Karreman, de Haas, van Tuijl, van Aken & Dekovic, 2010; Laake & Bridgett, 2014). In welcher Weise und auf welchen Wegen das kindliche Temperament jedoch einerseits die spätere Persönlichkeit prägt und andererseits Wirkungen etwa auf schulische Leistungs- und Kompetenzentwicklung hat, kann nur im Rahmen längsschnittlicher Analysedesigns untersucht werden. Hierfür stellen die Erhebungen des Nationalen Bildungspanels einen entsprechenden Rahmen zur Verfügung.

Die Vielfalt an vorhandenen Instrumenten zu Bestimmung des Temperaments auch und vor allem bei (Klein-)Kindern und die Notwendigkeiten und Restriktionen eines Large-Scale-Surveys machten umfangreiche Vorarbeiten notwendig, um ein valides, reliables und ökonomisches Instrument zu entwickeln, welches an vorhandene Instrumente anknüpft und damit eine gewisse Vergleichbarkeit gewährleistet.

1.2 Temperament als Entwicklungsaspekt

Das Temperament von Kleinkindern stellt im Rahmen des Nationalen Bildungspanels wie bereits angedeutet keinen Selbstzweck dar, und es geht auch nicht um die Vermessung einer Alterskohorte hinsichtlich deren Temperamentsausprägungen. Vielmehr soll die frühe und sich wiederholende Messung des kindlichen Temperaments es ermöglichen, dessen Einfluss auf spätere, bildungsrelevante Ereignisse, Entscheidungen und Entwicklungen zu bearbeiten. Dieses Ziel wird nicht im Sinne einer Entwicklungsmodellierung verfolgt; vielmehr stellt Temperament ein wiederholt querschnittlich erhobenes Konstrukt dar, mit dem zwar Veränderungen auch auf individueller Ebene verfolgt werden können, dies jedoch in den Grenzen der Anwendung identischer Instrumente. So wechseln die konkreten Items zur Messung des kindlichen Temperaments mit zunehmendem Alter der Kinder, so dass auf Itemebene keine direkte Vergleichbarkeit gegeben ist.

Da Temperament im Nationalen Bildungspanel als Vorläuferkonstrukt von Persönlichkeit und als ein Aspekt der sozial-emotionalen Entwicklungen verstanden wird, stellt die *Temperamentsentwicklung* einen zentralen Aspekt des konzeptionellen Designs dar.

1.3 Temperamentsmessung durch Fremdeinschätzungen

Neben theoretisch-konzeptionellen Fragen, gibt es hinsichtlich der „richtigen“ Erhebungsmethodik der Messung (früh-)kindlichen Temperaments unterschiedliche Positionen (für eine tabellarische Übersicht vgl. Rothbart & Bates, 1998, S.120). Insbesondere Eltern-Einschätzungen (Parent Reports) stellen ein häufig verwendetes, aber auch durchaus kontrovers diskutiertes Verfahren der Temperamentsmessung dar. Während Kagan (1994) der Nutzung von elterlichen Temperamenteinschätzungen eher kritisch gegenübersteht und diese für weitgehend wertlos hält, führen Rothbart und Bates (1998) an, dass diese Einschätzungen auf der entsprechenden Interpretation der oftmals geringen bis moderaten Korrelationen zwischen unterschiedlichen Fremdeinschätzungen desselben Kindes durch unterschiedliche Rater basiert.

Rothbart und Bates argumentieren an dieser Stelle vor allem methodisch bzw. statistisch; aus einer eher konzeptionellen Perspektive ist darüber hinaus anzumerken, dass wenn man Temperament nicht (ausschließlich) als ein biologisch-genetisch grundiertes Phänomen ansieht, die Einschätzung primärer Sozialisatoren zumindest eine wirklichkeitsverändernde Kraft besitzen kann. Dies empirisch zu untersuchen weist jedoch weit über den Rahmen einer Instrumentendarstellung hinaus und kann an dieser Stelle kaum weiterverfolgt werden.

2. Die Messung von Temperament

2.1 Temperament und Temperamentsdimensionen – Konturen eines Konstrukts

Ausgangspunkt des Auswahlprozesses eines für die Erhebung im NEPS (Blossfeld, Roßbach & von Maurice, 2011) geeigneten Konstrukts stellte der von Mary Rothbart (1981) konzipierte „Infant Behavior Questionnaire“ (IBQ) dar, den sie zusammen mit Gartstein zum „Infant Behavior Questionnaire - Revised“ (IBQ-R) (vgl. Gartstein & Rothbart, 2003) weiterentwickelte. Der IBQ kam in einer Vielzahl von meist in Laboratorien durchgeführten Studien zum Einsatz, die das Instrument mit einer großen Validierungsbreite versahen (vgl. etwa Bridges, Palmer, Morales, Hurtado, & Tsai, 1993; Crockenberg & Acredolo, 1983; Goldsmith & Rothbart, 1991; Reznick, Gibbons, Johnson, & McDonough, 1989; Rothbart, 1986).

Rothbart (2004) berichtet, dass Temperament bzw. die einzelnen Facetten des Temperaments interessante und wichtige Zusammenhänge etwa zur Entwicklung des Gewissens oder zur Herausbildung von Scham- bzw. Schuld-Gefühlen zeigen. Unter anderem konnten immer wieder Zusammenhänge zwischen dem Temperament und dem Geschlecht des Kindes nachgewiesen werden (vgl. Kohnstamm, 1989).

Der originale IBQ (vgl. Rothbart, 1981) umfasst insgesamt 87 Items, mit denen sechs Facetten von Temperament (*activity level, smiling and laughter, fear, distress to limitations, soothability, duration of orienting*) gemessen werden. Die revidierte Fassung des Instruments, der IBQ-R, erfasst neben den bereits in der ursprünglichen Version abgebildeten sechs Facetten ergänzend acht weitere (*approach, vocal reactivity, high pleasure, perceptual sensitivity, sadness, falling reactivity, low pleasure, cuddliness*), so dass insgesamt 14 Subskalen mittels 184 Items gemessen werden.

Die *Very Short Form* der revidierten Variante des IBQ (IBQ-R-VsF) (Helbig, Putnam, Gartstein & Rothbart, 2009; vgl. auch Putnam, Garstein, Helbig & Rothbart, 2014) konzentriert sich dann auf die diesen 14 Dimensionen wiederum zu Grunde liegenden drei latenten Dimensionen (zweiter Ordnung) *Surgency/Extraversion*, *Negative Affectivity* und *Orienting/Regulatory Capacity*, die mittels 37 Items erfasst werden können. Hierfür existierte zum Zeitpunkt der Vorbereitung der Erhebung von Temperament bei Kleinkindern im Rahmen des Nationalen Bildungspanels keine deutschsprachige Version, so dass hinsichtlich der Formulierungen zunächst auf eine durch Kristen, Eisenbeis, Thoermer und Sodian (2007) adaptierte deutschsprachige Variante des IBQ-R zurückgegriffen wurde.

Die 37 Items der Kurzversion des *IBQ-R-VsF* wurden so ausgewählt, dass sie gleichzeitig noch die ursprünglichen 14 Temperamentsdimensionen abzubilden imstande sind. Diese Abbildung der „first-order“ Dimensionen des Temperaments ließ sich aufgrund des weiteren Kürzungsbedarfs nicht aufrechterhalten, so dass die im NEPS implementierte Variante einer revidierten und gekürzten Version des *IBQ-R-VsF* lediglich die Hauptkomponenten abbilden soll(te).

2.2 Auswahl, Implementierung und Analyse einer gekürzten Variante des IBQ-R-VsF

Den Ausgangspunkt für eine weitere Itemreduzierung stellte die von Helbig et al. (2009) entwickelte *Very Short Form* des *IBQ-R* dar. Aus den 37 Items sollten für eine erste Pilotierung 15 Items (je fünf Items pro Hauptkomponente) ausgewählt werden. Vonderlin, Ropeter und Pauen (2012) haben die von Kristen et al. (2007) adaptierte deutschsprachige Version des *IBQ-R* in einer Studie an N = 119 Müttern mit Kindern im Alter zwischen sieben und neun Monaten eingesetzt. Die Autoren stellten uns die entsprechenden Kennwertanalysen für die deutschsprachigen Items in Vorbereitung einer im *NEPS* einzusetzenden Variante freundlicherweise zur Verfügung. Für die weitere Reduktion wurden insbesondere die Trennschärfekoeffizienten analysiert. Die von Vonderlin, Ropeter und Pauen (2012) vorgenommenen sprachlichen Anpassungen der Kristen-Version wurden für die zu testende Version übernommen.

2.2.1 Pilotierung einer gekürzten Variante der deutschsprachigen Version des IBQ-R-VsF

In einer Pilotstudie für das Nationale Bildungspanel wurden insgesamt N = 466 Mütter mit einem CAPI-Instrument (*Computer Assisted Personal Interviews*) befragt, welches neben den Fragen zum Temperament eine Fülle weiterer inhaltlicher Fragstellungen umfasste. Die durchschnittliche Dauer der Interviews, inklusive einigen mit den Kindern durchgeführten Spielen, betrug 106 Minuten. Die Stichprobe umfasste zwei Altersgruppen. In der ersten befanden sich Kinder, welche im März/April 2011 geboren wurden, die zweite Gruppe umfasste Kinder, die im Mai/Juni 2011 zur Welt kamen. Die Interviews wurden alle zwischen dem 21.10.2011 und dem 21.01.2012 durchgeführt. Für die Analysen wurden die Kinder zu zwei Altersgruppen zusammengefasst (5-7 Monate und 8-9 Monate), da die Entwicklungsunterschiede in diesem Altersbereich durchaus beträchtlich sein können.

Im Rahmen der Pilotstudie wurde eine Testversion eingesetzt, welche die drei Hauptkomponenten, die mit jeweils fünf Items abgefragt wurden, umfasste. Die Items und ihre dimensionale Zugehörigkeit sind in Tabelle 1 aufgeführt. In den Klammerausdrücken findet sich der wörtliche Stimulus des CAPI.

Tabelle 1 Übersicht über die eingesetzten Items in der Pilotstudie

Temperaments- dimension	Nr.	Items
Surgency/ Extraversion	1	Wenn Ihr Baby an- oder ausgezogen wurde, wie oft hat es sich gewunden und/oder versucht wegzurollen? <i>[Wenn <Name des Zielkindes> während der letzten 7 Tage an- oder ausgezogen wurde, wie oft hat sie/er versucht sich zu wehren oder wegzurollen?]</i>
	2	Wie oft hat Ihr Baby sich neuen Objekten schnell angenähert? <i>[Wie oft während der letzten 7 Tage hat <Name des Zielkindes> sich schnell auf neue Gegenstände zubewegt?]</i>
	3	Wenn Ihr Baby ins Badewasser gesetzt wurde, wie oft hat es gelacht? <i>[Wie oft während der letzten 7 Tage hat <Name des Zielkindes> gelacht, wenn sie ins Badewasser gesetzt wurde?]</i>
	4	Während des „Kuckuck“ Spielens, wie oft hat Ihr Baby gelacht? <i>[Wie oft während der letzten 7 Tage hat <Name des Zielkindes> während des „Kuckuck“ Spielens gelacht?]</i>
	5	Wie oft hat Ihr Baby gelächelt oder gelacht, wenn ihr/ihm ein Spielzeug gegeben wurde? <i>[Wie oft während der letzten 7 Tage hat <Name des Zielkindes> gelächelt oder gelacht, wenn ihr ein Spielzeug gegeben wurde?]</i>
Negative Affectivity	6	Wenn Ihr Baby müde war, wie oft hat es Anzeichen von Stress und Unbehagen gezeigt? <i>[Wie oft während der letzten 7 Tage hat <Name des Zielkindes> Anzeichen von Stress und Unwohlsein gezeigt, wenn sie müde war?]</i>
	7	Wie oft erschien Ihr Baby wütend (Weinen und Quengeln), wenn Sie es im Bettchen gelassen haben? <i>[Wie oft während der letzten 7 Tage schien <Name des Zielkindes> wütend zu werden und reagierte z.B. mit Weinen oder Quengeln, wenn Sie sie im Bettchen gelassen haben?]</i>
	8	Am Ende eines aufregenden Tages, wie oft ist Ihr Baby weinerlich geworden? <i>[Wie oft während der letzten 7 Tage ist <Name des Zielkindes> am Ende eines aufregenden Tages weinerlich geworden?]</i>
	9	Wenn Sie mit etwas Anderem beschäftigt waren und Ihr Baby nicht in der Lage war, Ihre Aufmerksamkeit zu erlangen, wie oft hat es geweint? <i>[Wie oft während der letzten 7 Tage hat <Name des Zielkindes> geweint, wenn Sie mit etwas anderem beschäftigt waren und sie nicht in der Lage war, Ihre Aufmerksamkeit zu erlangen?]</i>
	10	Wenn Ihr Baby etwas haben wollte, wie oft war es aufgebracht, wenn es nicht das haben konnte, was es wollte? <i>[Wie oft während der letzten 7 Tage ist <Name des Zielkindes> wütend geworden, wenn sie nicht das haben konnte, was sie wollte?]</i>
Orienting / Regulatory Capacity	11	Wie oft hat Ihr Baby es genossen wenn es, wie beim Wiegen und Schaukeln, sanft und rhythmisch bewegt wurde? <i>[Wie oft während der letzten 7 Tage schien <Name des Zielkindes> es zu genießen, wenn sie sanft und rhythmisch gewiegt und geschaukelt wurde?]</i>
	12	Wenn Ihr Baby gehalten wurde, wie oft hat es sich scheinbar amüsiert? <i>[Wie oft während der letzten 7 Tage schien <Name des Zielkindes> sich zu vergnügen, wenn sie auf dem Arm oder Schoß gehalten wurde?]</i>
	13	Wenn Sie Ihrem Baby vorgesungen oder mit ihm gesprochen haben, wie oft hat es sich sofort beruhigt? <i>[Wie oft in den letzten !!14 Tagen!! hat <Name des Zielkindes> sich sofort beruhigt, wenn Sie ihr vorgesungen oder mit ihr gesprochen haben?]</i>
	14	Wenn Ihr Baby gewiegt oder umarmt wurde, wie oft hat es dies anscheinend genossen? <i>[Wie oft während der letzten 7 Tage schien <Name des Zielkindes> es zu genießen, wenn sie geschaukelt oder umarmt wurde?]</i>
	15	Wenn Sie Ihr Baby sanft am Körper geklopft oder darüber gestrichen haben, wie oft hat es sich sofort beruhigt? <i>[Wie oft in den letzten !!14 Tagen!! hat <Name des Zielkindes> sich sofort beruhigt, wenn Sie sie sanft gestreichelt haben?]</i>

Entsprechend dem Originalinstrument wurde zunächst im Stimulus ein Zeitbezug von – im Regelfall – sieben Tagen verankert.¹ Die befragten Mütter wurden gebeten, die Auftrittshäufigkeit des im Item beschriebenen Verhaltens ihres Kindes einzuschätzen.

2.2.2 Empirische Ergebnisse der Pilotierung

Die deskriptiven Ergebnisse sowie die Analyse der Skalenkonsistenzen und Trennschärfen sind im Anhang in den Tabellen A1 und A2 aufgeführt. Insbesondere bei der Komponente *Surgency/Extraversion* ergaben sich in der Pilotstudie einige Probleme. Die interne Skalenkonsistenz war vor allem bei der Gruppe der Mütter mit Kindern im Alter von 8 bis 9 Monaten sehr niedrig. Würde man in dieser Komponente diejenigen beiden Items mit der geringsten Trennschärfe streichen, blieben nur noch „Lach-Items“ zur Ermittlung von *Surgency/Extraversion* übrig. Aus diesem Grund wurde auf die Erhebung dieser Temperamentsdimension, zumindest im Altersbereich, der durch den IBQ abgedeckt wird, verzichtet. Ferner wurde das Item „Wenn Ihr Baby gewiegt oder umarmt wurde, wie oft hat es dies anscheinend genossen?“ ebenfalls gestrichen, da die Befragten größere Schwierigkeiten hinsichtlich der Beantwortung dieses Items signalisierten. Darüber hinaus wurde die 7er-Skala, die in der Pilotstudie noch voll gelabelt war, auf eine Endlabelung „nie“ und „immer“ umgestellt.

Das überarbeitete Instrument enthält dementsprechend neun Items, mit denen die beiden Dimensionen *Negative Affectivity* sowie *Orienting/Regulatory Capacity* erfasst werden (vgl. Tabelle 2). Über die Streichung von insgesamt sechs Items hinaus wurde eine weitere Anpassung auf Basis der Ergebnisse der Pilotstudie und den Rückmeldungen aus dem Feld vorgenommen. So wurde der zeitliche Bezug („die letzten 7 Tage“ oder „in den letzten 14 Tagen“) gestrichen und abweichend vom Ursprungsinstrument danach gefragt, wie sich das Kind „üblicherweise“ in solchen Situationen verhält. Damit wurde auf Irritationen der befragten Mütter reagiert, die gehäuft darauf verwiesen, dass der anvisierte Zeitraum aus unterschiedlichen Gründen überhaupt nicht typisch sei.

Tabelle 2 Übersicht über die Items der Hauptstudie (in der Reihenfolge der Abfrage)

Temperamentsdimension	Nr.	Items
Negative Affectivity	6	Wenn „Name des Kindes“ müde ist, wie oft zeigt er/sie Anzeichen von Stress und Unwohlsein?
Orienting / Regulatory Capacity	11	Wenn „Name des Kindes“ rhythmisch gewiegt und geschaukelt wird, wie oft genießt er/sie es?
Negative Affectivity	7	Wenn es Zeit fürs Bett ist, und „Name des Kindes“ nicht schlafen will, wie oft quengelt er/sie dann?
Orienting / Regulatory Capacity	14	Wenn „Name des Kindes“ geschaukelt oder umarmt wird, wie oft scheint er/sie es zu genießen?
Negative Affectivity	9	Wenn Sie mit etwas anderem beschäftigt sind und „Name des Kindes“ nicht in der Lage ist, Ihre Aufmerksamkeit zu erlangen, wie oft weint er/sie dann?
Orienting / Regulatory Capacity	12	Wenn „Name des Kindes“ auf dem Arm oder Schoß gehalten wird, wie oft scheint er/sie sich zu vergnügen?

¹ Nach Durchführen einiger Pretestinterviews wurde bei zwei Items dieser Zeitbezug auf 14 Tage erweitert, da die Mütter angaben, dass die beschriebenen Situationen nicht so häufig auftreten.

Negative Affectivity	10	Wenn „Name des Kindes“ nicht haben kann, was er/sie will, wie oft wird er/sie dann wütend?
Orienting / Regulatory Capacity	13	Wenn Sie „Name des Kindes“ vorsingen oder mit ihr sprechen, wie oft beruhigt er/sie sich sofort?
Orienting / Regulatory Capacity	15	Wenn Sie „Name des Kindes“ sanft streicheln, wie oft beruhigt er/sie sich sofort?

3. Empirische Ergebnisse der Kreuzvalidierung der „neuen“ 9-Item-Variante im Rahmen der VIVA-Studie

Die so entstandene 9-Item-Variante wurde durch das von der Deutschen Forschungsgemeinschaft geförderte Forschungsprojekt ViVA (*Video-Based-Validity-Analysis*, Teil des DFG-Schwerpunktprogramms 1646, Projektleitung S. Weinert) mit dem kompletten *IBQ-R* verglichen. Im Rahmen der Studie wurden 38 Bamberger Familien zu Hause besucht und anhand der wortwörtlich übernommenen neun verbliebenen Items befragt. Es schloss sich eine etwa einstündige Erhebung verschiedener Maße an, die an die Erhebungen der Säuglingskohorte (Startkohorte 1) des NEPS angelehnt war. Am Ende des Besuchs wurde den Eltern die vollständige deutschsprachige Variante des *IBQ-R* von Kristen et al. (2007) ausgehändigt. Bei einem zweiten Besuch wurde der Fragebogen wieder eingesammelt und weitere Erhebungen durchgeführt. Das Alter der untersuchten Kinder betrug zwischen 7 Monaten, 4 Tagen und 9 Monaten, 5 Tagen (MW: 8 Monate, 12 Tage). Das Geschlechterverhältnis war nahezu ausgeglichen (18 Mädchen, 20 Jungen).

Aus den Items des *IBQ-R* wurden anhand der offiziellen Zuordnung die Subskalen zur Erfassung von *Negative Affectivity* und *Orienting/Regulatory Capacity* gebildet und anschließend mit den Subskalen der NEPS-Kurzform verglichen. Dabei fanden sich deutliche positive Zusammenhänge zwischen den beiden Instrumenten (vgl. Tabelle 2), während die beiden Subskalen sowohl innerhalb als auch zwischen den Instrumenten negativ miteinander korrelieren (vgl. Tabellen 2 und 3). Da dieser Zusammenhang von den Autoren des *IBQ-R* für ihre amerikanische Stichprobe nicht berichtet wird, könnte man diesen Umstand als Hinweis auf einen möglichen kulturellen Effekt deuten, insofern dass die befragten Eltern in Deutschland die beiden Subskalen in einen negativen Zusammenhang brachten. Zur Beurteilung dieser Hypothese wäre jedoch eine eingehendere Untersuchung notwendig.

Tabelle 2: Spearman-Korrelation der Skalen zwischen den Instrumenten

		NEPS	
		NA	O/R
IBQ-R	NA	.51**	-.39*
	O/R	-.33*	.70***

Tabelle 3: Spearman-Korrelation der Skalen innerhalb der Instrumente

		NEPS	IBQ-R	
		NA	NA	NA
O/R	NA	-.40**		
	O/R			-.51***

Anmerkungen: NA = *Negative Affectivity*, O/R = *Orienting / Regulatory Capacity*, Zweiseitiges Signifikanzniveau, wobei * $p < .05$, ** $p < .01$, *** $p < .001$

Ein Direktvergleich der korrespondierenden Items in den beiden Varianten ergab für die meisten Items signifikante mittelhohe Zusammenhänge (Spearman's $r = .29-.49$; MW: .40). Davon ausgenommen ist zum einen das Item 14 (vgl. Tabelle 1), für das sich kein Zusammenhang mit der Beantwortung der entsprechenden Frage im Rahmen des

kompletten *IBQ-R* fand und zum anderen die Items 10 und 12, für die es im *IBQ-R* jeweils zwei korrespondierende Items gab, die im Rahmen der Bildung der *IBQ-R-VsF* zusammengelegt wurden. In beiden Fällen fand sich nur zu einem der beiden Ursprungitems ein signifikanter Zusammenhang. Dass die Korrelationen insgesamt nicht höher ausfallen, deutet darauf hin, dass die Beantwortung der Fragen durch situative Bedingungen, Effekte des Modus (Interview vs. selbstauszufüllender Fragebogen) oder die Abfolge der Fragen beeinflusst wurde. Zusätzlich ist zu bedenken, dass aufgrund der zeitlichen Reihenfolge ein schwacher Einfluss der Beantwortung der 9-Item-Variante auf die Beantwortung der Fragen des kompletten *IBQ-R* zwar sehr unwahrscheinlich, jedoch nicht völlig auszuschließen ist, zumal dazwischen je nach Bearbeitungszeitpunkt des *IBQ-R* mindestens 1-2 Stunden und häufig 1-2 Tage vergingen.

Obwohl Item 10 gemäß offizieller Zuordnung in beide Skalen einfließen sollte (in O/R invertiert), fügte es sich in der empirischen Überprüfung lediglich in die Skala *Negative Affectivity* ein (Trennschärfe von .01 mit O/R in der NEPS-Stichprobe). Daher wurden für die weitere Berechnung der internen Konsistenz der Skala *Orienting/Regulatory Capacity* lediglich die Items 11-15 herangezogen. Die sonstigen Korrelationen der Items mit den beiden Skalen fielen gemäß ihrer vorgesehenen Zuordnung aus (Trennschärfen zwischen .18 und .51; MW: .27 und in der NEPS-Stichprobe zwischen .28 und .46; MW: .34). Zusätzlich korrelierten – passend zu den negativen Skalenkorrelationen – einige der Items auch negativ mit der jeweils anderen Skala. Cronbach's Alpha fiel für die Kürze der Skalen und den geringen Stichprobenumfang recht passabel aus (NA $\alpha = .53$; O/R $\alpha = .44$). In der NEPS-Stichprobe wurden entsprechend leicht höhere Werte erreicht (NA $\alpha = .58$; O/R $\alpha = .59$).

Insgesamt lässt sich festhalten, dass die Erfassung der beiden Subskalen auch auf Basis von neun Items sehr gut funktioniert. Erwartungsgemäß fallen die Korrelationen niedriger aus als die von Putnam et al. (2014) berichteten Zusammenhänge zwischen der kompletten Version und der *Very Short Form* ($r \approx .8$). Bedenkt man die zusätzliche Kürzung um 75 Prozent, sind die Ergebnisse jedoch durchaus akzeptabel.

4. Ausblick auf die Messung von Temperament in den kommenden Wellen der Kohorte

Die Messung frühkindlichen Temperaments mittels des hier vorgestellten und durch erste Befunde des VIVA-Projekts zusätzlich extern validierten Kurzinstruments ist in einen Gesamtrahmen der Messung und Analyse der individuellen Persönlichkeitsentwicklung sowie der sozial-emotionalen Entwicklung eingebettet. Aufgrund der großen Fülle an Befragungskonzepten, welche im Rahmen der Erhebungen der frühkindlichen Kohorte zum Einsatz kommen, ist es nicht möglich, alle 9 Items in jeder Erhebungswelle zu erheben. So werden aus der Subskala *negative affectivity* drei der fünf Items in den Wellen zwei und drei im Befragungsmodul zur kindlichen Entwicklung als Panelitems abgefragt. Für diese drei Items (6, 9 und 10 in Tabelle 2) werden in Zukunft Informationen aus insgesamt drei Wellen vorliegen. In der vierten Welle kommt dann der ebenfalls auf Arbeiten von Rothbart zurückgehende Children's Behavior Questionnaire (CBQ) zum Einsatz, der Items zur Messung der drei vorgestellten Temperamentsdimensionen enthält.

Zusätzlich werden in einer der Folgewellen die Persönlichkeitseigenschaften der Eltern ebenso gemessen, wie deren Einschätzung der Persönlichkeit ihres Kindes. Zu diesem Zweck wird eine Operationalisierung der *Big Five* (Rammstedt & John, 2007) eingesetzt.

Damit wird in den nächsten Jahren mit den Daten dieser Kohorte nicht nur eine Nachzeichnung der Entwicklung von (früh-)kindlichem Temperament möglich werden sondern darüber hinaus werden Daten erhoben, die eine Queranalyse von Temperaments- und Persönlichkeitskonstrukten möglich machen.

5. Literatur

- Allport, G. (1961). *Pattern and growth in personality*. New York: Holt, Rinehart & Winston.
- Bates, J. E. (1980). The concept of difficult temperament. *Merrill-Palmer Quarterly*, 26(4), 299–319.
- Baumgarten, F. (1933). Die 'Charaktereigenschaften'. In A. Francke (Ed.), *Beiträge zur Charakter- und Persönlichkeitsforschung*. Bern.
- Blossfeld, H.-P., Roßbach, H.-G. & von Maurice, J. (Hrsg.) (2011). *Education as a Lifelong Process – The German National Educational Panel Study (NEPS)*. *Zeitschrift für Erziehungswissenschaft: Sonderheft 14*.
- Bridges, L., Palmer, S., Morales, M., Hurtado, M., & Tsai, D. (1993). Agreement between affectively based observational and parent-report measure of temperament at infant age 6 months. *Infant Behavior and Development*, 16(4), 61–72.
- Buss, A., & Plomin, R. (1984). *Temperament: Early developing personality traits*. Hillsdale, New York: Erlbaum.
- Crockenberg, S., & Acredolo, C. (1983). Infant temperament ratings: A function of infants of mothers, or both? *Infant Behavior and Development*, 6(1), 61–72.
- Elliot, A. J., & Thrash, T. M. (2002). Approach-avoidance motivation in personality: Approach and avoidance temperaments and goals. *Journal of Personality and Social Psychology*, 82(5), 804–818.
- Eysenck, H. (1967). *The biological basis of personality*. Springfield, IL: Thomas.
- Gartstein, M. A., & Rothbart, M. K. (2003). Studying infant temperament via the Revised Infant Behavior Questionnaire. *Infant Behavior and Development*, 26(1), 64–86.
- Goldsmith, H. H., & Rothbart, M. K. (1991). Contemporary instruments for assessing early temperament by questionnaire and in the laboratory. In A. Angleitner & J. Strelau (Eds.), *Explorations in temperament: International perspectives on theory and measurement* (pp. 249–272). New York: Plenum.
- Gray, J. (1982). *The neuropsychology of anxiety*. Oxford: Oxford University Press.
- Helbig, A., Putnam, S. P., Gartstein, M. A., & Rothbart, M. K. (2009). *Development and Assessment of Short and Very Short Forms of the Infant Behavior Questionnaire - Revised*. Presentation at the biannual convention of the Society for Research in Child Development. Denver: CO.
- Kagan, J. (1994). *Galen's prophecy: Temperament in human nature*. New York: Basic Books.
- Karreman, A., Haas, S. de, Tuijl, C. van, Aken, M. A. G. van, & Deković, M. (2010). Relations among temperament, parenting and problem behavior in young children. *Infant Behavior & Development*, 33(1), 39–49.
- Kohnstamm, G. A. (1989). Temperament in Childhood: Cross-cultural and sex differences. In G. A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds.), *Temperament in Childhood* (pp. 483–508). Oxford: John Wiley & Sons.
- Kristen, S., Eisenbeis, H., Thoermer, C., & Sodian, B. (2007). *Deutsche Version des IBQ-R: Unpublished manuscript*, Ludwig-Maximilians-Universität.
- Laake, L. M., & Bridgett, D. J. (2014). Happy babies, chatty toddlers: infant positive affect facilitates early expressive, but not receptive language. *Infant Behavior & Development*, 37(1), 29–32.

- Lawson, K., & Ruff, H. (2004). Early attention and negative emotionality predict later cognitive and behavioural function. *International Journal of Behavioral Development, 28*(2), 157–165.
- Nigg, J. T., & Goldsmith, H. H. (1998). Developmental Psychopathology, Personality, and Temperament: Reflections on Recent Behavioral Genetics Research. *Human Biology, 70*(2), 387–412.
- Putnam, S. P., Helbig, A. L., Gartstein, M. A., Rothbart, M. K., & Leerkes, E. (2014). Development and assessment of short and very short forms of the infant behavior questionnaire-revised. *Journal of Personality Assessment, 96*(4), 445–458.
- Rammstedt, B., & John, O. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality, 41*(2), 203–212.
- Reznick, J., Gibbons, J., Johnson, M., & McDonough, P. (1989). Behavioral inhibition in a normative sample. In J. Reznick (Ed.), *Perspectives on behavioral inhibition. The John D. and Catherine T. MacArthur Foundation series on mental health and development* (pp. 25–49). Chicago, IL: The University of Chicago Press.
- Rothbart, M. K. (1981). Measurement of temperament in infancy. *Child Development, 52*(2), 569–578.
- Rothbart, M. K. (1986). Longitudinal observation of infant temperament. *Developmental Psychology, 22*(3), 356–365.
- Rothbart, M. K. (1989). Temperament and the Pursuit of an Integrated Developmental Psychology. In G. A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds.), *Temperament in childhood* (pp. 187–248). New York: Wiley.
- Rothbart, M. K. (2004). Temperament and the Pursuit of an Integrated Developmental Psychology. *Merrill-Palmer Quarterly, 50*(4), 492–505.
- Rothbart, M. K., & Bates, J. E. (1998). Temperament. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology: Vol. 3. Social, emotional and personality development* (5th ed., pp. 105–176). New York: Wiley.
- Vonderlin, E., Ropeter, A., & Pauen, S. (2012). Erfassung des frühkindlichen Temperaments mit dem Infant Behavior Questionnaire Revised [Assessment of temperament with the Infant Behavior Questionnaire Revised (IBQ-R) - the psychometric properties of a German version]. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie, 40*(5), 307–314.
- Webb, E. (1915). Character and Intelligence: An Attempt at an Exact Study of Character. *British Journal of Psychology, 1*(1).
- Zuckermann, M. (1991). *Psychobiology of personality*. Cambridge: Cambridge University Press.

6. Anhang

Tabelle A1 Skalenkonsistenzen der Temperamentsdimensionen

	Cronbachs α (wenn Item gelöscht)		
	Alle	8-9 Monate	5-7 Monate
Surgency/Extraversion			
Wenn Ihr Baby an- oder ausgezogen wurde, wie oft hat es sich gewunden und/oder versucht wegzurollen?	.417	.249	.551
Wie oft hat Ihr Baby sich neuen Objekten schnell angenähert?	.337	.232	.463
Wenn Ihr Baby ins Badewasser gesetzt wurde, wie oft hat es gelacht?	.293	.177	.414
Während des „Kuckuck“ Spielens, wie oft hat Ihr Baby gelacht?	.304	.281	.340
Wie oft hat Ihr Baby gelächelt oder gelacht, wenn ihr/ihm ein Spielzeug gegeben wurde?	.284	.214	.360
Negative Affectivity			
Wenn Ihr Baby müde war, wie oft hat es Anzeichen von Stress und Unbehagen gezeigt?	.520	.512	.534
Wie oft erschien Ihr Baby wütend (Weinen und Quengeln), wenn Sie es im Bettchen gelassen haben?	.505	.466	.563
Am Ende eines aufregenden Tages, wie oft ist Ihr Baby weinerlich geworden?	.623	.638	.603
Wenn Sie mit etwas Anderem beschäftigt waren und Ihr Baby nicht in der Lage war, Ihre Aufmerksamkeit zu erlangen, wie oft hat es geweint?	.491	.472	.521
Wenn Ihr Baby etwas haben wollte, wie oft war es aufgebracht, wenn es nicht das haben konnte, was es wollte?	.569	.576	.559
Orienting/Regulatory Capacity			
Wenn Sie Ihrem Baby vorgesungen oder mit ihm gesprochen haben, wie oft hat es sich sofort beruhigt?	.612	.596	.642
Wenn Ihr Baby gehalten wurde, wie oft hat es sich scheinbar amüsiert?	.609	.593	.638
Wie oft hat Ihr Baby es genossen wenn es, wie beim Wiegen und Schaukeln, sanft und rhythmisch bewegt wurde?	.577	.526	.650
Wenn Ihr Baby gewiegt oder umarmt wurde, wie oft hat es dies anscheinend genossen?	.584	.563	.621
Wenn Sie Ihr Baby sanft am Körper geklopft oder darüber gestrichen haben, wie oft hat es sich sofort beruhigt?	.545	.524	.581

Tabelle A2 Deskriptive Kennwerte und Trennschärfekoeffizienten der Temperamentsdimensionen

Surgency/Extraversion	Alle	8-9 Monate	5-7 Monate
Mittelwert: 28,25	.379	.276	.485
Varianz: 20,25	267	166	101
Standardabweichung: 4,5			
Negative Affectivity	Alle	8-9 Monate	5-7 Monate
Mittelwert: 18,5	.600	.594	.611
Varianz: 33,79	258	157	101
Standardabweichung: 5,81			
Orienting/Regulatory Capacity	Alle	8-9 Monate	5-7 Monate
Mittelwert: 29,47	.641	.617	.680
Varianz: 19,45	277	170	107
Standardabweichung: 4,42			