

Was hält Partnerschaften zusammen? Soziologische und psychologische Ansätze zum Erfolg von Paarbeziehungen

– Ausführlicher Methodenteil –

Inhalt

1	Analysemethoden	2
1.1	<i>Skalenanalysen</i>	2
1.1.1	Exploratorische Faktorenanalysen (1. Revision).....	3
1.1.2	Konfirmatorische Faktorenanalysen (2. Revision).....	4
1.1.3	Weitere Skalenanalyseverfahren.....	9
1.2	<i>Hypothesenprüfung</i>	11
2	Operationalisierungen	17
2.1	<i>Sozialstrukturelle exogene Variablen</i>	18
2.1.1	Partnerschaftsbezogene Angaben	18
2.1.2	Soziodemographische Merkmale.....	22
2.2	<i>Konflikte</i>	27
2.3	<i>Interaktionsstile</i>	31
2.4	<i>Partnerschaftszufriedenheit</i>	35
2.5	<i>Subjektive Partnerschaftsstabilität</i>	38
2.6	<i>Das vollständige Messmodell mit allen endogenen latenten Variablen</i>	40
3	Stichprobe	43
4	Literatur	55

Das Buch ist erschienen im VS Verlag, Wiesbaden (2008)

Download dieses Dokuments: <http://www.vs-verlag.de> (Menüpunkt „OnlinePlus“)

Methode

Das folgende Kapitel beschreibt das methodische Vorgehen in der vorliegenden Studie. Konkret wird zunächst die Analysestrategie (1) dargestellt, die bei der psychometrischen Überprüfung und Revision der Skalen (1.1) sowie bei der Überprüfung der Hypothesen (1.2) verfolgt wurde. Im Anschluss daran folgt eine Beschreibung der psychometrischen Merkmale der eingesetzten Skalen (2) sowie der Stichprobe (3).

1 Analysemethoden

Die statistischen Analysen dienen zum einen dem Ziel, die Güte der eingesetzten Skalen zu überprüfen, zum anderen ermöglichen sie die Überprüfung der im Theorieteil abgeleiteten Hypothesen. Da für diese beiden Fragestellungen unterschiedliche Verfahren erforderlich sind, werden diese im Folgenden separat beschrieben.

1.1 Skalenanalysen

Da im untersuchten Datensatz eine Vielzahl von Indikatoren zur Erfassung von Wahrnehmungs- und Interaktionsmerkmalen in Partnerschaften eingesetzt wurde, war im ersten Schritt eine Reduktion des Itempools auf diejenigen Indikatoren vorzunehmen, welche die theoretischen Konstrukte (Faktoren) am besten repräsentieren. Als Selektionskriterien wurden dabei neben Aspekten der Inhaltsvalidität, d.h. einer möglichst hohen Plausibilität der Frageformulierungen für die Messung der zu erfassenden Konstrukte, auch verschiedene statistische Auswahlkriterien herangezogen. Zu diesen zählen zum einen Lage und Form der Antwortverteilungen – insbesondere Abweichungen von der Normalverteilung –, zum anderen geeignete Maßzahlen der Homogenität der Skalen im Sinne von Eindimensionalität.

Das dabei verfolgte Vorgehen ist mehrstufig: Im Rahmen einer ersten Revision werden *innerhalb der Konstrukte* mittels datenreduzierender Verfahren, d.h. hier: exploratorischer Faktorenanalysen, diejenigen Indikatoren entfernt, die dem entsprechenden „Kernkonstrukt“ nicht eindeutig zuzuordnen sind, sondern z.B. einen eigenen Faktor bilden. Im Fall des mehrdimensionalen Konstrukts „Paarinteraktion“ werden die Items eliminiert, die kein eindeutiges Ladungsmuster zeigen oder auf anderen als den theoretisch erwarteten Fakto-

ren laden. Nach diesen konstrukt^{spezifischen} Analysen wird eine konstrukt^{übergreifende} konfirmatorische Faktorenanalyse berechnet, um die Differenzierbarkeit der latenten Konstrukte zu prüfen. Ziel dieser zweiten Revision ist die Selektion derjenigen Indikatoren, die am besten und eindeutigsten das ihnen zugeordnete Konstrukt repräsentieren. Items, die die postulierte Einfachstruktur verletzen, d.h. bedeutsame Nebenladungen auf anderen als den ihnen zugeordneten Faktoren zeigen, werden dabei aus den späteren Analysen ausgeschlossen. Dieser Schritt ist unerlässlich, um bei der späteren Prüfung von „Kausalmodellen“ Fehlspezifikationen aufgrund von Varianzkonfundierungen zwischen den verschiedenen Konstrukten zu vermeiden. Aus der zweiten Revision gehen reduzierte Fassungen der Instrumente hervor, die abschließend einer herkömmlichen Skalenanalyse unterzogen werden. Zur Bewertung der psychometrischen Güte der Instrumente werden dabei vor allem Itemkennwerte wie Schwierigkeit und Trennschärfe, Schätzungen der Reliabilität sowie verschiedene Antwortverteilungscharakteristika herangezogen. Die eingesetzten Auswertungsverfahren werden im Folgenden näher beschrieben.

1.1.1 Exploratorische Faktorenanalysen (1. Revision)

Die exploratorische Faktorenanalyse (vgl. Backhaus, Erichson, Plinke, & Weiber, 2005; Überla, 1972) zählt zu den dimensionsreduzierenden Verfahren (Schnell, Hill, & Esser, 2005); sie ist dazu geeignet, auf Basis der Interkorrelationen zwischen den Indikatoren latente (nicht beobachtbare) Faktoren zu finden, welche jeweils mehrere Indikatoren optimal inhaltlich repräsentieren. Das Attribut exploratorisch bezieht sich primär darauf, dass keine Restriktionen der Faktorladungen vorgenommen werden, sondern dass prinzipiell jeder Indikator auf jedem Faktor lädt. Zudem wird die Anzahl der zu extrahierenden Faktoren nicht vorgegeben, sondern auf Basis des Kaiser-Guttman-Kriteriums empirisch ermittelt; nach diesem Kriterium wird die Faktorextraktion solange fortgesetzt, wie der Eigenwert der Faktoren größer als 1 ist.

Resultieren mehrere Faktoren aus der Faktorextraktion (z.B. im Bereich Paarinteraktion), wird die Faktorlösung anschließend im mehrdimensionalen Raum so rotiert, dass die Summe der Faktorladungen, d.h. der Korrelationen zwischen den Indikatoren und den zugehörigen Faktoren, insgesamt maximiert wird. Hierbei sind vom Anwender zwei Entscheidungen zu treffen:

1. Zunächst muss eine Annahme hinsichtlich des Fehlermodells getroffen werden, die eine Entscheidung zwischen Hauptkomponentenanalysen (principal component analysis, PCA) und Hauptachsen-Faktorenanalyse (principal axis factor analysis, PAF) impliziert: Der PCA liegt die Annahme zugrunde, dass prinzipiell die gesamte Varianz der Indikatoren durch die Faktoren aufklärbar ist, während bei der PAF zusätzlich eine nicht durch die Faktoren erklärbare Fehlervarianz modelliert wird (vgl. Bortz, 1999).
2. Daneben muss – nicht zuletzt auf Basis inhaltlicher Überlegungen – entweder Unabhängigkeit (Orthogonalität) oder Abhängigkeit der latenten Faktoren angenommen werden (vgl. zu den jeweiligen Vor- und Nachteilen Bortz, 1999).

Aus diesen beiden Entscheidungen resultieren vier mögliche faktorenanalytische Varianten. Da sich in den Analysen nur unwesentliche Abweichungen zwischen den Ergebnissen dieser Varianten zeigen, werden ausschließlich die Befunde der Hauptkomponentenanalysen mit orthogonaler (Varimax-)Rotation wiedergegeben.

Das Ziel der exploratorischen Faktorenanalysen besteht wie erwähnt darin, für das jeweils fokussierte Konstrukt eine Vorselektion der Indikatoren vorzunehmen, d.h. ihr Einsatz dient hier nicht nur der Überprüfung der Dimensionalität der Indikatoren, sondern daneben auch der Itemselektion. Zu diesem Zweck wurden die aus der ersten Skalenrevision resultierenden „besten“ – d.h. am eindeutigsten ihrem jeweiligen Faktor zuordenbaren – Items in eine zweite, konstruktübergreifende Faktorenanalyse eingegeben.

1.1.2 Konfirmatorische Faktorenanalysen (2. Revision)

Die konfirmatorische Faktorenanalyse stellt eine Verfahrensklasse innerhalb der Methode linearer Strukturgleichungsmodelle (engl. *structural equation modeling*, SEM) dar und dient der Spezifikation und dem Test von Messmodellen (Reinecke, 2005). Ihr Grundprinzip und einige Aspekte der Interpretation entsprechen weitgehend der zweiten SEM-Verfahrensklasse, der (kovarianzstrukturbasierten) Pfadanalyse, welche in der vorliegenden Studie zur Hypothesenprüfung verwendet wurde (vgl. hierzu Abschnitt 1.2). Daher sind die im Folgenden beschriebenen Punkte auf die hypothesenprüfenden Analysen übertragbar.

Im Unterschied zu älteren explorativen faktorenanalytischen Verfahren basieren konfirmatorische Faktorenanalysen nicht auf Korrelations-, sondern auf Varianz-Kovarianzmatrizen (vgl. zur Einführung Bollen, 1989; Hoyle, 1995; Reinecke, 2005). Ein wichtiges Charakteristikum des Ansatzes besteht darin, dass der Forscher vor der Durchführung der Analysen bereits ein konkretes Modell der theoretisch zu erwartenden Zusammenhänge bzw. Effekte aufgestellt hat, welches dann an die empirischen Daten angepasst wird. Technisch betrachtet handelt es sich dabei um einen iterativen Prozess, in dem die Modellparameter – also Effekte, Faktorladungen, Residuen etc. – zunächst aus den empirischen Kovarianzen (Matrix S) geschätzt werden, bevor im nächsten Schritt auf der Grundlage dieser Parameter wiederum eine so genannte modellimplizierte Kovarianzmatrix (Σ) geschätzt wird. Dieser Prozess wird solange fortgesetzt, bis die Diskrepanzfunktion $F=f(S-\Sigma)^1$ ein zuvor festgelegtes Minimierungskriterium unterschreitet (vgl. Reinecke, 2005).

Ziel der konfirmatorischen Faktorenanalyse (Einführungen finden sich z.B. bei Bollen, 1989; Reinecke, 2005) ist die Schätzung der Passung zwischen den empirischen Daten und einem theoretisch begründeten Messmodell, nach dem die hinter den Indikatoren stehenden latenten Konstrukte (z.B. Persönlichkeitsmerkmale) die beobachtbaren Antworten auf die jeweiligen Indikatoren determinieren. Wie bei der Hauptachsen-Faktorenanalyse wird dabei jedoch davon ausgegangen, dass stets eine gewisse Residualvarianz bleibt, dass die „Erklärung“ des Antwortverhaltens also nicht perfekt ist. Die erwünschte und zu testende Ladungsstruktur der Indikatoren lässt sich wie folgt charakterisieren: Die Faktorladungen auf den zugehörigen Faktoren sollen möglichst hoch ausfallen, während Nebenladungen auf anderen Faktoren per Restriktion ausgeschlossen werden. Diese Struktur wird auch als Einfachstruktur (simple structure, Thurstone, 1947) bezeichnet. Eine perfekte Einfachstruktur wird in der Praxis nicht zu finden sein; um zu beurteilen, wie gut die Daten zu einer Einfachstruktur passen, können jedoch im Rahmen der konfirmatorischen Faktorenanalyse unterschiedliche Gütemaße und Fitindizes berechnet werden, die auf der Diskrepanzfunktion basieren (vgl. Bollen, 1989; Hu & Bentler, 1995; Reinecke, 2005):

¹ Durch diese Schreibweise soll ausgedrückt werden, dass die Minimierung der Diskrepanzfunktion einer Minimierung der Differenz zwischen empirischer und modellimplizierter Kovarianzmatrix entspricht. Die genaue Diskrepanzfunktion lautet allerdings je nach eingesetzter Schätzung unterschiedlich (vgl. ausführlich hierzu Reinecke, 2005).

- Zunächst einmal kann aus der Diskrepanzfunktion F über die einfache Formel $\chi^2 = (N - 1) \cdot F$ eine inferenzstatistische Absicherung der Abweichung der Vorhersageresiduen von 0 vorgenommen werden (zur Herleitung vgl. Bollen, 1989). Da hier auf Diskrepanzen zwischen Modell und Daten getestet wird (mit der Nullhypothese als Wunschhypothese), gilt eine umgekehrte Testlogik, d.h. ein auf dem gewählten Signifikanzniveau (im vorliegenden Beitrag auf $p = .05$ festgelegt) signifikanter χ^2 -Wert indiziert eine mangelhafte Passung des gesamten Modells.

Ein entscheidender Nachteil von χ^2 besteht jedoch darin, dass dieses mit steigendem Stichprobenumfang (vgl. die oben angegebene Formel) sowie mit steigender Modellkomplexität zunimmt (Hu & Bentler, 1995). Reinecke (2005) bemerkt hierzu: „Kleinere Differenzen zwischen S und Σ können bei großen Stichproben dramatische Auswirkungen auf die χ^2 -Statistik haben, so daß fast immer eine Modellwiderlegung die Folge ist“. Insofern sind neben χ^2 noch andere Fitindizes zu berücksichtigen, um die Modellgüte beurteilen zu können.

- Ein weiteres häufig eingesetztes Maß für die Diskrepanz zwischen beobachteten und modellimplizierten Kovarianzen ist der *Root Mean Square Error of Approximation* (RMSEA). Werte unter 0.05 bedeuten hierbei geringe Abweichungen, Werte zwischen 0.05 und 0.08 eine mittlere Abweichung, höhere Werte große Abweichungen und damit einen schlechten Modellfit (Reinecke, 2005). Im Programm LISREL lässt sich hier außerdem über die Ausgabe des Konfidenzintervalls inferenzstatistisch die Hypothese prüfen, dass der berechnete RMSEA-Wert kleiner als 0.05 ist.
- Daneben gibt es weitere Fitindizes, die – analog zu R^2 in der multiplen Regression – auf den Bereich von 0 bis 1 normiert, aber primär deskriptiv zu interpretieren sind (vgl. Reinecke, 2005): Der Goodness of Fit-Index (GFI) sowie der Comparative Fit Index (CFI) sind z.B. gegenüber dem Stichprobenumfang relativ unempfindlich. Der Adjusted Goodness of Fit-Index (AGFI) bezieht im Unterschied zum GFI auch die Modellkomplexität in Form der Freiheitsgrade ein. Als Daumenregel für einen guten Modellfit werden in der Literatur Werte von mindestens .90 angegeben (vgl. Hu & Bentler, 1995).

Der Modellfit ist stets vor der Interpretation der einzelnen Koeffizienten zu prüfen: Finden sich statistisch signifikante Abweichungen zwischen der durch das Modell implizierten und der empirisch beobachteten Kovarianz-Matrix, sollte das Modell als ganzes verworfen wer-

den; von der Interpretation einzelner Pfadkoeffizienten ist in diesem Fall abzuraten (Rietz, Rudinger, & Andres, 1996). Indiziert hingegen der Modellfit bei Restriktion aller „Nebenladungen“ der Indikatoren auf Null keine signifikanten Abweichungen zwischen theoretischem Faktorenmodell und den empirischen Daten und zeigen sich gleichzeitig substantielle Faktorladungen, kann die Einfachstruktur als empirisch gegeben angesehen werden. In diesem Fall können die Pfadkoeffizienten sinnvoll interpretiert werden. Dieses Vorgehen bietet gegenüber klassischen Ansätzen der Faktorenanalyse die Möglichkeit, inferenzstatistisch zu testen, ob in einem spezifischen Modell bedeutsame Nebenladungen und damit – bei einer Operationalisierung über die verwendeten Indikatoren – Konfundierungen zwischen den darin enthaltenen Konstrukten vorliegen.

Reinecke (2005) weist darauf hin, dass die Interpretation von Modellfit und Koeffizienten aus Maximum Likelihood-(ML-)Schätzungen an diverse datenbezogene Voraussetzungen gebunden ist. Wie die Verteilungskennwerte der untersuchten Konstrukte zeigen (vgl. Abschnitt 2), ist in der vorliegenden Arbeit insbesondere die Multinormalverteilung der zugrunde liegenden Indikatoren häufig verletzt.² Zur Korrektur der hieraus resultierenden Verzerrungen wurden verschiedene alternative Schätzvarianten angefordert:

1. Bei Analysen an kleineren Substichproben ($N < 1000$) wird eine spezielle robuste ML-Schätzung der Koeffizienten in Verbindung mit einer von Satorra und Bentler (1988) entwickelten reskalierten χ^2 -Statistik verwendet. Hierbei werden nicht die Koeffizienten selbst, sondern nur deren Standardfehler sowie die χ^2 -Statistik korrigiert.
2. Bei ausreichendem Stichprobenumfang ($N \geq 1000$) besteht die optimale Lösung in der Berechnung einer verteilungsfreien Schätzung (gelegentlich in der Literatur als asymptotically distribution free [ADF] bezeichnet), die über eine mit den höheren Verteilungsmomenten gewichtete Diskrepanzfunktion gewonnen werden kann (weighted least squares, WLS). Allerdings sollten für stabile WLS-Schätzungen je nach Modellkomplexität und Verteilungseigenschaften mindestens tausend Beobachtungen vorlie-

² West et al. (1995) berichten auf der Basis von Simulationsstudien, dass χ^2 bei zunehmenden Abweichungen von der Normalverteilung zunehmend überschätzt wird, was sich in einer übermäßig konservativen Modellprüfung niederschlägt.

gen, während die Satorra-Bentler-Korrektur bereits ab etwa $N=200$ gute Schätzungen des Modellfits liefert (West et al., 1995).³

3. In einigen Fällen werden die Standardfehler der Koeffizienten über das Bootstrap-Verfahren ermittelt. Beim Bootstrapping (Efron & Tibshirani, 1993) handelt es sich um eine neuere, relativ rechenintensive Methode, in der durch wiederholtes resampling der Analysestichprobe (Ziehen von Stichproben mit Zurücklegen) verteilungsfreie Schätzer für Koeffizienten und Standardfehler gewonnen werden (Efron, 2000). Dieses Verfahren wird insbesondere in den Pfadmodellen, in denen mindestens eine dichotome abhängige Variable modelliert wird, zur Bestimmung der indirekten Effekte eingesetzt.

In der vorliegenden Arbeit enthält die konfirmatorische Faktorenanalyse allerdings auch explorative Momente, da sie der zweiten, abschließenden Skalenrevision dient: Hierzu werden Modifikationsindizes für die nicht freigesetzten Faktorladungen angefordert, die angeben, wie stark sich der Modellfit (d.h. die χ^2 -Statistik) insgesamt verbessert, wenn Nebenladungen zugelassen bzw. die entsprechenden Restriktionen auf 0 aufgegeben würden. Da substantielle Nebenladungen aufgrund der damit verbundenen Varianzkonfundierung die Interpretierbarkeit der Mediatormodelle gefährden würden, werden Itemvariablen mit zu hohen Nebenladungen ausgeschlossen. Überschreitet ein Modifikationsindex einen bestimmten Schwellenwert, der in der vorliegenden Arbeit auf 10 festgelegt wurde, muss von einer Varianzkonfundierung ausgegangen werden, und der entsprechende Indikator wird aus dem Messmodell entfernt. Diese Prozedur wird solange wiederholt, bis keine als bedeutsam angesehenen Nebenladungen mehr vorliegen. Andere Restriktionen, beispielsweise zu Messfehlerkorrelationen, werden nicht gelockert.

Die aus der konfirmatorischen Faktorenanalyse resultierenden Revisionen der Instrumente werden abschließend noch einigen „klassischen“ Verfahren der Item- und Skalenanalyse unterzogen, welche im nächsten Abschnitt näher beschrieben werden.

³ Dabei ist zu beachten, dass bei robuster ML-Schätzung mit Satorra-Bentler-Korrektur lediglich die χ^2 -Statistik sowie die Standardfehler, nicht aber die Parameterschätzer selbst korrigiert werden, während bei einer WLS-Schätzung eine optimale Nutzung der Daten inklusive deren Schiefe und Kurtosis gewährleistet ist (vgl. Reinecke, 2005).

1.1.3 Weitere Skalenganalyseverfahren

Zur Beurteilung der Güte der vorgeschlagenen Rating-Items werden auf Einzelitemebene Indizes für Itemschwierigkeit und Trennschärfe bestimmt. Auf Skalenebene wird daneben über Cronbachs Alpha die Reliabilität für jede Skala geschätzt. Im Fall der abhängigen Variable *subjektive Partnerschaftsstabilität* wird schließlich eine Skalogrammanalyse zum Test auf Guttman-kompatible Antwortmuster durchgeführt. Zur Beschreibung der Antwortverteilungen werden daneben die zentralen Verteilungsmomente – also Mittelwert, Varianz (bzw. Standardabweichung), Schiefe und Kurtosis (Bortz, 1999) – zu jedem Indikator berechnet. Alle genannten Item- und Skalenkennwerte werden im Folgenden kurz erläutert.

- Itemschwierigkeit: Der Schwierigkeitsindex (Fisseni, 1997; Mummendey, 1995) gibt an, wie groß der Anteil der Testpersonen ist, die – im Fall von Leistungstests – ein Item „richtig“ beantworten bzw. – wie im hier gegebenen Fall von Persönlichkeitstests im weiteren Sinne – einem Item zustimmen. Mit Hilfe der Itemschwierigkeit lassen sich Items, denen entweder die Mehrzahl der Testpersonen zugestimmt oder nicht zugestimmt hat, von solchen trennen, bei denen sich über alle Testpersonen Zustimmung und Ablehnung weitgehend die Waage halten. Der Schwierigkeitsindex variiert zwischen 0 (bei einer Zustimmungsskala bedeutet dies, dass alle Personen zu einer zu bewertenden Aussage maximale Ablehnung äußern, d.h. sie geben als Antwort „trifft überhaupt nicht zu“ o.Ä. an) und 1 (alle Personen stimmen dem Item maximal zu, d.h. sie geben als Antwort z.B. „trifft voll zu“ an). Üblicherweise werden Items aus einem Fragebogen ausgeschlossen, die Itemschwierigkeiten kleiner als 0.2 oder größer als 0.8 aufweisen, da diese Items aufgrund der sehr konformen Reaktionen der Testpersonen zur Unterscheidung zwischen einzelnen Individuen nur schlecht geeignet sind. Als optimal wird eine Schwierigkeit von 0.5 angesehen. Der hier berechnete Schwierigkeitsindex für mehrfach abgestufte Antworten berücksichtigt die Varianz der Items (Fisseni, 1997):⁴

⁴ Wird die Itemvarianz nicht berücksichtigt, kann die Itemschwierigkeit für zwei Items gleich ausfallen, obwohl ihre Punkteverteilungen erheblich divergieren. Es gilt aber, dass Testpersonen auf Items umso unterschiedlicher reagieren, je höher deren Varianz ist und dass Items mit einer höheren Varianz somit besser zwischen den Testpersonen differenzieren.

$$p = \frac{\sum g^2}{\sum gm^2}$$

Es bedeuten:

SI: Schwierigkeitsindex für mehrfach abgestufte Items unter Berücksichtigung der Varianz

$\sum g^2$: Quadratsumme von g, mit g = Item-Score und $\sum g$ = Test-Score (Summe aller Item-Scores)

$\sum gm^2$: Quadratsumme von gm, mit gm = maximal erreichbarer Item-Score

$\sum gm$ = maximal erreichbarer Test-Score (N * gm)

- Trennschärfe: Die Trennschärfe eines Items (Fisseni, 1997; Mummendey, 1995; Schnell et al., 2005) gibt den Grad des Zusammenhangs zwischen der Zustimmung zu diesem Item und dem Gesamt-Score der übrigen Items der entsprechenden Skala in Form einer Produkt-Moment-Korrelation (häufig als r_{it} bezeichnet) an. Mit Hilfe des Trennschärfe-Koeffizienten, der Werte zwischen 0 und 1 annehmen kann, lassen sich Aussagen darüber machen, inwieweit ein Item den Fragebogen- bzw. Skalenwert vorhersagen kann. Es gilt, dass ein Item umso besser zwischen Personen mit hoher und niedriger Ausprägung des interessierenden Merkmals differenzieren kann, je höher die Trennschärfe ausfällt. Konventionellerweise wird ein Trennschärfekoeffizient von kleiner 0.3 als niedrig, von 0.3 bis 0.5 als mittel und von größer 0.5 als hoch bezeichnet. Items mit einer Trennschärfe von kleiner als 0.3 sollten aus dem Fragebogen ausgeschlossen werden. Trennschärfekoeffizienten zeigen auf der Ebene der Einzelitems die Homogenität einer Skala bzw. eines Fragebogens an: Ist der Trennschärfekoeffizient für alle Items hoch, gilt die entsprechende Skala bzw. der Fragebogen als homogen, d.h. jedes Item trägt gleich viel zum Test- bzw. Skalen-Score bei und repräsentiert somit ungefähr gleich stark den Inhalt des Fragebogen- bzw. Skalengesamtscores.
- Cronbachs α (Alpha): Der Koeffizient Cronbachs α (Feldt & Brennan, 1989; Mummendey, 1995; Schnell et al., 2005) dient der Bestimmung der internen Konsistenz eines Fragebogens oder einer Skala auf Basis der durchschnittlichen Korrelation aller möglichen Testhälften. Mit seiner Hilfe kann über die Homogenität einer Skala entschieden werden; α ist somit als Maß der Reliabilität anzusehen. Konventionellerweise wird eine Skala als intern konsistent bzw. reliabel bezeichnet, wenn Cronbachs α einen Wert von größer 0.8 annimmt.⁵ Die Berechnungsformel lautet:

⁵ Beachtet werden muss, dass Cronbachs α mit wachsender Itemanzahl ansteigt. Aus diesem Grund muss bei der Beurteilung immer auch die Itemanzahl berücksichtigt werden.

$$\alpha = \frac{n}{n-1} \left[1 - \frac{\sum \sigma_i^2}{\sigma_x^2} \right]$$

Es bedeuten:

n : Anzahl der Items

σ_i^2 : Varianz jedes einzelnen Items

σ_x^2 : Varianz der Skala

- Guttman-Skalogrammanalyse: Die zur Messung der subjektiven Partnerschaftsstabilität eingesetzten Items sind als Guttman-Skala konzipiert, d.h. es wird ein Stufenprozess postuliert, in dem die Bejahung eines beliebigen Items theoretisch die Zustimmung zu allen vorhergehenden Items impliziert (vgl. Schnell et al., 2005). Die empirischen Antwortmuster können jedoch von diesem „idealtypischen“ Reaktionsmuster abweichen, indem mindestens ein „leichteres“ Item verneint und ein „schwereres“ bejaht wird. Die Passung der Daten zu einer Guttman-Skala lässt sich empirisch testen, indem die Anzahl derartiger unzulässiger Antwortmuster (U) zu der Gesamtzahl an Antworten in Relation gesetzt wird. Daraus ergibt sich der so genannte Reproduzierbarkeitskoeffizient $R_G = \frac{U}{N \cdot k}$, wobei

N : Anzahl der Befragten

k : Anzahl der Items der Skala.

Als Untergrenze für eine akzeptable Passung der Daten zur Guttman-Skalierung wird in der Literatur $R_G > .90$ angegeben (Schnell et al., 2005).

1.2 Hypothesenprüfung

Neben den in Abschnitt 1.1 beschriebenen Verfahren zur Überprüfung der Güte der Skalen werden zur Hypothesenprüfung verschiedene regressions- und pfadanalytische Verfahren eingesetzt. Diese werden im Folgenden detailliert beschrieben.

Das Hauptziel der vorliegenden Arbeit besteht in der Identifikation vermittelnder Prozesse, durch die bestimmte sozialstrukturelle Faktoren die subjektive Stabilität von Partnerschaften und Ehen beeinflussen. Konkret soll eruiert werden, ob und in welchem Ausmaß die Paarinteraktion in Alltags- und Konfliktsituationen, Konfliktwahrnehmungen sowie die wahrgenommene Partnerschaftszufriedenheit verschiedene mittlerweile gut erforschte sozialstrukturelle Einflüsse auf die Partnerschaftsstabilität erklären können. In der metho-

dischen Fachliteratur werden solche Drittvariablen, die auf Basis theoretischer Überlegungen und empirischer Analysen den Einfluss einer unabhängigen Variablen „erklären“, als Mediatorvariablen bezeichnet (Baron & Kenny, 1986). Teilweise finden sich in sozialwissenschaftlichen Methodenlehrbüchern abweichende Termini; so verwenden Schnell et al. (2005) für denselben Sachverhalt den Begriff der Interpretation, der sich jedoch in der sonstigen Methodenliteratur nicht durchgesetzt hat. Mediation bedeutet statistisch, dass ein im Ausgangsmodell signifikanter direkter Effekt des Prädiktors auf die Kriteriumsvariable unter Kontrolle der Mediatorvariablen insignifikant wird. In der Praxis wird eine solche vollständige Mediation häufig nicht zu beobachten sein; vielmehr deutet auch bereits eine „deutliche“ Reduktion des direkten Effekts des Prädiktors auf eine (partielle) Mediation hin.⁶ Wie stark der Koeffizient sinken muss, damit die Mediation signifikant ist, kann über geeignete inferenzstatistische Tests bestimmt werden, die jedoch in gängigen Statistikpaketen meist nicht implementiert sind; für einfache Konstellationen mit lediglich einer Mediatorvariablen sind Behelfslösungen auf Basis spezieller Syntaxprogramme vorgestellt worden (Preacher & Hayes, 2004). Diese reichen jedoch nicht mehr aus, wenn wie in der vorliegenden Arbeit multiple Mediation untersucht wird, d.h. über verschiedene, untereinander korrelierte Drittvariablen vermittelte Effekte. Aufgrund der Vielzahl potentieller vermittelnder Mechanismen ist eine mehrstufige Analysestrategie erforderlich, in der zunächst Blöcke von Mediatorvariablen identifiziert werden und anschließend eine differenziertere vergleichende Prüfung der empirisch bedeutsamen vermittelten Effekte erfolgt. Für den letzten Schritt bieten sich Pfad- bzw. Strukturgleichungsmodelle an, da hier mittels Effektzerlegung direkte und indirekte Effekte getestet werden können (Bollen, 1987).

Die allgemeine Analysestrategie

Das allgemeine Vorgehen in den Analysen besteht dabei aus drei Schritten:

1. *Screening der Effekte der Stabilitätsdeterminanten:* Als erstes werden multiple Regressionsmodelle zur Erklärung der subjektiven Partnerschaftsstabilität berechnet, in welchen jeweils mehrere thematisch zusammengehörige Einflussfaktoren enthalten sind. Zusätzlich

⁶ Es kann gezeigt werden, dass die Differenz im Regressionsgewicht (d.h. totaler Effekt minus direkter Effekt im Mediatormodell) äquivalent ist zum Produkt des Effekts vom Prädiktor auf den Mediator und des Effekts vom Mediator auf das Kriterium (vgl. Frazier, Tix, & Barron, 2004). Die Bestimmung des Standardfehlers des indirekten Effekts kann jedoch nicht über das arithmetische Mittel der Standardfehler der

werden einige grundlegende Kontrollvariablen wie Kohortenzugehörigkeit, Geschlecht, Beziehungstyp und Sozialisation in Ost- bzw. Westdeutschland in die Modelle aufgenommen. Das Ziel dieser ersten Analysen besteht darin zu klären, ob sich die erwarteten Einflüsse der verschiedenen Determinanten in der untersuchten Stichprobe und auf Basis der gewählten Operationalisierungen überhaupt finden lassen.

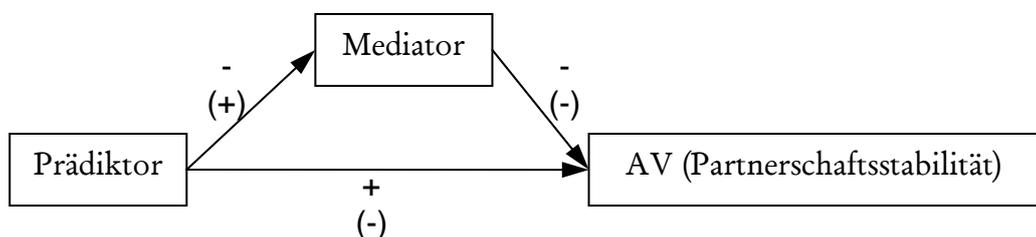
Zur Vorbereitung erfolgt ein Screening bezüglich Moderatorvariablen (zur Differenzierung von Moderator- und Mediatoreffekten vgl. Baron & Kenny, 1986): In nicht dargestellten explorativen Voranalysen wird der Einfluss der nach den Hypothesen postulierten Determinanten getrennt für verschiedene Subgruppen bestimmt, die durch die genannten Kontrollvariablen gebildet werden. Dabei werden die Daten sowohl nach Einzelvariablen (Ost- vs. Westdeutsche, Männer vs. Frauen) als auch nach Kombinationen aus diesen (z.B. ostdeutsche vs. westdeutsche Männer in Ersten) gesplittet. Falls die Effekte nur in bestimmten Subgruppen festzustellen sind, beschränkt sich die Ergebnisdarstellung aus Platzgründen auf die relevanten Gruppen.

2. *Screening auf vermittelte Effekte:* Im zweiten Schritt werden – ggf. für die im ersten Schritt identifizierten Subgruppen – hierarchische (geschachtelte) Regressionsmodelle eingesetzt, durch die eine erste Prüfung auf vermittelte Effekte erfolgt (vgl. Baron & Kenny, 1986; Frazier et al., 2004). Dazu werden zusätzlich zu den interessierenden Stabilitätsdeterminanten sukzessive Blöcke theoretisch erwartbarer Mediatorvariablen kontrolliert. Über Veränderungen der Regressionsgewichte der Prädiktorvariablen können Rückschlüsse auf vermittelnde Mechanismen gezogen werden. Hier sind zwei mögliche Szenarien zu unterscheiden: (a) der Effekt eines Prädiktors verringert sich unter Hinzunahme einer Drittvariable (Mediation) oder (b) der Effekt wird stärker (Suppression). In pfadanalytischer Betrachtung sind beide Szenarien durch Gleichheit (bei Mediation) bzw. Ungleichheit (bei Suppression) der Vorzeichen von direkten und indirekten Effekten (vgl. Abbildung 1) differenzierbar (MacKinnon, Krull, & Lockwood, 2000). Dabei werden die indirekten Effekte aus dem Produkt der beiden „Teilpfade“ berechnet. In der vorliegenden Arbeit liegt der Schwerpunkt auf Mediatoreffekten, empirisch finden sich jedoch auch vereinzelte Suppressoreffekte.

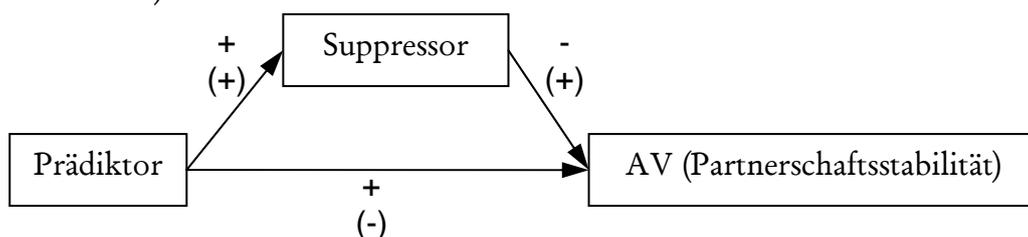
Einzeleffekte erfolgen, sondern es muss eine Gewichtung mit der Höhe der Einzeleffekte vorgenommen werden (so genannter Sobel-Test, vgl. Shrout & Bolger, 2002; Sobel, 1982).

3. *Hypothesenprüfung, inferenzstatistischer Test direkter und indirekter Effekte:* Zur detaillierten Überprüfung der direkten und indirekten (vermittelten) Effekte werden mit Hilfe der Software Mplus Pfadanalysen (vgl. zur Einführung Backhaus et al., 2005; Reinecke, 2005) berechnet. Die verwendeten Schätzverfahren und Kriterien des Modellfits entsprechen dabei den bei der Skalenanalyse eingesetzten konfirmatorischen Faktorenanalysen (zu den Einzelheiten vgl. Abschnitt 1.1.2). Der Vorteil dieser Analysen gegenüber den multiplen Regressionsanalysen besteht nicht zuletzt in der differenzierteren Prüfung vermittelter Effekte: Zum einen werden neben den Effekten auf die AV (subjektive Partnerschaftsstabilität) auch die Effekte der Prädiktoren auf die Mediatorvariablen geschätzt; zum anderen gibt Mplus für jeden Prädiktor spezifische, d.h. über eine einzelne Mediatorvariable vermittelte, indirekte Effekte (vgl. hierzu Bollen, 1987) sowie die entsprechenden Standardfehler aus, sodass jeder spezifische indirekte Effekt inferenzstatistisch überprüft werden kann.

Abbildung 1: Schematische Darstellung von Mediator- und Suppressoreffekten in Pfadmodellen



(a) Beispiel für Mediation (gleiche Vorzeichen des direkten und des indirekten Effekts)



(b) Beispiel für Suppression (ungleiche Vorzeichen des direkten und des indirekten Effekts)

Hinsichtlich des Ausmaßes statistischer Mediation lassen sich die folgenden drei Fälle unterscheiden (vgl. Baron & Kenny, 1986):

1. *Vollständige Mediation* liegt dann vor, wenn der direkte Effekt eines Prädiktors unter Kontrolle der Mediatorvariablen insignifikant wird und wenn die indirekten Effekte in der Summe ebenfalls signifikant sind.
2. Von *partieller Mediation* ist auszugehen, wenn signifikante indirekte Effekte vorliegen, der direkte Effekt des Prädiktors jedoch unter Kontrolle der Mediatorvariablen nicht insignifikant wird.
3. Wenn schließlich kein indirekter Effekt nachweisbar ist und / oder wenn der Einfluss des Prädiktors auch multivariat signifikant bleibt, liegt *keine Mediation* vor.

Eingesetzte Modellierungsarten

Wie in den entsprechenden Skalenanalysen (Abschnitt 2.5) beschrieben, wurde die abhängige Variable subjektive Partnerschaftsstabilität als Index aus dichotomen Items zusammengesetzt. Aufgrund der relativ geringen Zahl an Ausprägungen sowie der schiefen Verteilung des Index wird die Anwendbarkeit linearer Modellierungen in Zweifel gezogen. Neben linearen Regressionsanalysen wurden daher ordered logit-Modelle eingesetzt, die eine Erweiterung der logistischen Regression auf mehrstufige ordinalskalierte abhängige Variablen darstellen (vgl. Fox, 1997; McCullagh & Nelder, 1989). Durch den Vergleich der linearen und der logit-Regressionen kann nebenbei die Robustheit der Befunde gegenüber Voraussetzungsverletzungen geprüft werden. Befunde aus den ordered logit-Analysen werden aus Platzgründen nur berichtet, wenn sie deutlich von den dargestellten Ergebnissen der linearen Regressionsanalysen abweichen.

Moderatoreffekte

Im ersten Abschnitt des Ergebnisteils werden Hypothesen zur Wechselwirkung von Konflikten und Paarinteraktionsstilen mittels moderierter Regression mit Interaktionstermen geprüft. Da hier Prädiktor *und* Moderator metrisch sind, sind bei der Analyse und der Ergebnisinterpretation einige Besonderheiten zu beachten (vgl. zur Übersicht Aiken & West, 1991), die im Folgenden kurz erläutert werden. Zunächst einmal kann die Richtung eines Interaktionseffekts zwischen zwei metrischen Variablen nur bestimmt werden, wenn beide Prädiktoren vor der Multiplikation standardisiert oder zumindest um 0 herum zentriert (d.h. Subtraktion des Gesamtmittelwerts von jedem Wert) werden. Durch Standardisierung oder Zentrierung wird gleichzeitig das häufig auftretende Problem der Multikollinearität von Prädiktor und Moderator gelöst, da der Interaktionsterm zweier zentrierter Zufallsva-

riablen mit den beiden Einzelvariablen unkorreliert ist, wenn diese normalverteilt sind (vgl. McClelland & Judd, 1993). Einige Besonderheiten sind bei der Interpretation von Interaktionseffekten zu beachten (Frazier et al., 2004): Die von den üblichen Statistikpaketen ausgegebenen standardisierten Koeffizienten sind nicht interpretierbar; stattdessen sind die Einzelvariablen vor der Produktbildung zu standardisieren und das resultierende unstandardisierte Regressionsgewicht im Sinne eines standardisierten Effekts zu interpretieren. Im Fall eines signifikanten Interaktionseffekts sollte zusätzlich geprüft werden, um welche Art von Wechselwirkung es sich handelt (z.B. ordinal oder disordinal). Zur visuellen Inspektion empfiehlt sich dazu im Fall zweier metrischer Variablen ein Plot, in dem exemplarisch für verschiedene Ausprägungen der Moderatorvariablen entsprechende Regressionsgeraden des betrachteten Zusammenhangs abgebildet wird, z.B. für den Mittelwert sowie eine Standardabweichung oberhalb und unterhalb des Mittelwerts (vgl. Aiken & West, 1991).

An einigen anderen Stellen wird von Moderatoreffekten struktureller Variablen (z.B. Sozialisation in Ost-Westdeutschland, Beziehungstyp etc.) gesprochen. Hierzu ist zu beachten, dass im Ergebnisteil keine explizite Testung der entsprechenden Interaktionseffekte erfolgt, da die resultierenden Modelle aufgrund der Menge an Koeffizienten ansonsten nur noch schwierig darstellbar bzw. lesbar wären.

Implikationen des Studiendesigns

Aus dem Querschnittcharakter der Studie resultieren einige Restriktionen hinsichtlich der Modellspezifikationen. Insbesondere ergibt sich die Schwierigkeit, dass Alter bzw. Kohorte und Partnerschaftsdauer konfundiert und daher hoch miteinander korreliert sind (für Geburtsjahr und Partnerschaftsdauer beträgt der Zusammenhang $r = .817$, $p = .000$). Aufgrund der dadurch bedingten Multikollinearität ist eine gleichzeitige Kontrolle beider Variablen in den multivariaten Modellen z.T. rechnerisch nicht möglich bzw. inhaltlich nicht sinnvoll. In den meisten Analysen wird die Geburtskohorte als mehrfach gestufte Variable kontrolliert, in einigen Fällen ist die Beziehungsdauer allerdings von theoretischem Interesse und wird daher an Stelle der Kohorte in die Modelle aufgenommen. Wie bei allen Auswertungen von Querschnittdaten ist in der vorliegenden Arbeit zu berücksichtigen, dass sich die berichteten „Effekte“ nicht kausal interpretieren lassen; vielmehr besteht das Ziel der Analysen darin, erste Hinweise auf Mediatoreffekte der „psychologischen“ intervenie-

renden Variablen (Paarinteraktion, Konflikte, Partnerschaftszufriedenheit) zu gewinnen, die in zukünftigen Arbeiten zu replizieren und zu vertiefen sind.

Anmerkungen zur Ergebnisdarstellung

In der Ergebnisdarstellung wurden folgende Konventionen getroffen: Allgemein werden in allen Tabellen und Abbildungen Regressionsgewichte und Signifikanzniveaus angegeben. Aufgrund der recht unterschiedlichen Umfänge der Analysetichproben wird zwar zur Hypothesenprüfung ein einheitliches Signifikanzniveau von $p < .01$ zugrunde gelegt; allerdings wird auch im Fall einer höheren empirischen Irrtumswahrscheinlichkeit im Text auf tendenzielle Zusammenhänge bzw. Einflüsse hingewiesen. Bei Analysen in Subgruppen mit $N \leq 1000$ werden dabei Zusammenhänge mit $p < .1$, bei größeren Subgruppen mit $p < .05$ erwähnt bzw. als Trends interpretiert.

In den Tabellen sind zusätzlich zu jedem Koeffizienten t-Werte für den inferenzstatistischen Test auf $b \neq 0$ in Klammern angegeben. Um eine Interpretierbarkeit der Höhe und eine bessere Vergleichbarkeit der Regressionskoeffizienten zu gewährleisten, werden in den Tabellen für metrische Prädiktoren jeweils standardisierte Koeffizienten angegeben. Bei mehrfach gestuften qualitativen Variablen bzw. den hieraus resultierenden Sets von (binären) Dummyvariablen ist die Angabe standardisierter Koeffizienten hingegen nicht sinnvoll, da eine Dummyvariable nicht um eine Standardabweichung variieren *kann* und somit die übliche Interpretation standardisierter Regressionsgewichte nicht anwendbar ist (Fox, 1997). Dem Vorschlag von Fox (1997) folgend, werden nur für quantitative Prädiktoren standardisierte Regressionsgewichte, für Dummyvariablen hingegen unstandardisierte Koeffizienten berichtet. Zur besseren Identifizierbarkeit der Dummyvariablen sind diese in den Tabellen jeweils kursiv gedruckt.

2 Operationalisierungen

In den folgenden Abschnitten werden die Operationalisierungen der verschiedenen Konstrukte beschrieben. Bezüglich der sozialstrukturellen exogenen Variablen wird dabei vor allem auf das codierungstechnische Vorgehen eingegangen, während der Schwerpunkt bei den endogenen Variablen auf die psychometrische Qualität der Skalen sowie auf die Vermeidung von Varianzkonfundierungen zwischen den verschiedenen Modellvariablen gelegt wird. Diese unterschiedliche Schwerpunktsetzung wird damit begründet, dass die endoge-

nen im Unterschied zu den exogenen Variablen subjektive Wahrnehmungen der Partnerschaft umfassen, die mit multiplen Indikatoren erhoben wurden und somit die Anwendung testtheoretischer Gütekriterien erlauben und erfordern. Demnach kommen primär bei den vermittelnden Konstrukten und der abhängigen Variablen verschiedene faktoren- und itemanalytische Verfahren zum Einsatz.

2.1 Sozialstrukturelle exogene Variablen

Die folgenden Beschreibungen der Operationalisierungen werden nach inhaltlichen Bereichen gegliedert. Am Ende des Abschnitts werden in Tabellenform Deskriptionen der entsprechenden Verteilungen wiedergegeben.

2.1.1 Partnerschaftsbezogene Angaben

Ein Auswahlkriterium für die Stichprobe bestand darin, dass die Befragten zum Untersuchungszeitpunkt zusammen mit ihren Partnern in einem Haushalt lebten (zu weiteren Kriterien und Details der Untersuchung vgl. Rüssmann, Arránz Becker, & Kelzenberg, 2004). Die folgende Übersicht (Tabelle 1) zeigt die Verteilung der verschiedenen Partnerschaftstypen getrennt nach Herkunft aus Ost- bzw. Westdeutschland:

Tabelle 1: Beziehungstypen nach Sozialisation in Ost- bzw. Westdeutschland

Beziehungstyp	Sozialisation in				Gesamt	
	Ostdeutschland		Westdeutschland		N	%
	N	%	N	%		
voreheliche Lebensgemeinschaft	179	16.4%	58	7.5%	237	12.7%
nacheheliche Lebensgemeinschaft	58	5.3%	34	4.4%	92	4.9%
Erstehe	746	68.3%	602	77.6%	1348	72.2%
Mehrfachehe	109	10.0%	82	10.6%	191	10.2%
Gesamt	1092	100.0%	776	100.0%	1868	100.0%

Die Häufigkeitsverteilung zeigt, dass insbesondere die Erstehepaare dominieren, wohingegen speziell nachehelich kohabitierende Paare⁷ zu schwach vertreten sind, sodass hier keine separaten Analysen möglich sind.

Aus entsprechenden Items wurden zudem die Variablen *voreheliche Kohabitation* sowie *Dauer vorehelicher Kohabitation* gebildet, ebenso wie die Variablen *Gesamtkinderzahl* sowie *Anzahl gemeinsamer (leiblicher) Kinder*. Daneben wurde eine Variable zum *Timing der Erstgeburt* relativ zur Eheschließung gebildet. Zunächst wurde hierzu das Alter des ältesten Kindes bestimmt, anschließend wurde dieses von der Gesamtedauer subtrahiert. Da auch eine Eheschließung kurz nach dem Übergang zur Elternschaft noch als Indikator einer ehelichen Geburt angesehen werden kann (vgl. z.B. Wagner, 1997), wurden alle in einem Intervall von +/- einem Jahr um die Eheschließung geborenen Kinder als ehelich klassifiziert.

Die Operationalisierung des *Familienzyklus* orientierte sich an den Überlegungen von Mühlfeld (1976) und wurde primär aus dem Alter des jüngsten Kindes im Haushalt gebildet. Die entsprechende Variable unterscheidet zwischen den folgenden acht Stadien:

1. Vorbereitungs- oder Aufbauphase („Honeymoon-Phase“), in der (noch) keine Kinder vorhanden sind
2. Expansionsphase, in der Kinder vorhanden sind; das jüngste Kind ist jünger als drei Jahre
3. Phase der Primärsozialisation, in der das jüngste Kind zwischen drei und sechs Jahren alt ist (Vorschulalter)
4. Phase mit schulpflichtigen Kindern, in der das jüngste Kind zwischen sechs und zwölf Jahren alt ist (Grundschule, Wechsel zur weiterführenden Schule)
5. Phase der Adoleszenz, in der das jüngste Kind zwischen zwölf und achtzehn Jahren alt ist (weiterführende Schule bis Schulabschluss bzw. bei Haupt- und Realschülern bis Ende der Ausbildung)
6. Phase der Spätadoleszenz, in der das jüngste Kind älter als achtzehn Jahre ist und (zumindest teilweise) im elterlichen Haushalt lebt oder, falls es nicht mehr bei den Eltern wohnt, finanziell von diesen unterstützt wird (Ablösung der Kinder)
7. Kontraktionsphase („Empty-Nest“-Phase), in der das jüngste Kind älter als achtzehn Jahre ist und kein Kind mehr im elterlichen Haushalt lebt und/oder finanziell abhängig ist; mindestens einer der Partner ist noch berufstätig
8. Altersphase, in der das jüngste Kind älter als achtzehn Jahre ist und kein Kind mehr im elterlichen Haushalt lebt und/oder finanziell unterstützt wird; beide Partner sind nicht mehr berufstätig; diese Phase dauert bis zum Tod (oder bis zur Trennung) der Partner

Während sich die Variable zum Familienzyklus auch zur Erfassung nichtlinearer „Verläufe“ der Kriterien des Partnerschaftserfolgs eignet, wird zur Modellierung linearer Verände-

⁷ Hierbei werden aufgrund der kleinen Fallzahlen Geschiedene und in Trennung lebende Verheiratete nicht weiter differenziert.

rungen die metrische Variable *Alter des jüngsten Kindes* verwendet. Diese Variable wird eingesetzt, wenn kein Vergleich zu Kinderlosen angestrebt wird, sondern eher Veränderungen im Zuge des Heranwachsens von Kindern im Fokus der Betrachtung stehen. Der Vorteil dieser Operationalisierungsvariante besteht darin, dass Veränderungen eines Koeffizienten in den statistischen Mediatormodellen einfacher zu interpretieren sind als die eines Satzes von Dummyvariablen.

Die *Partnerschaftsdauer zum Zeitpunkt der Haushaltsgründung* wurde als Indikator des Timings von Institutionalisierungsprozessen der Beziehung eingesetzt. Dazu wurde die Differenz aus gesamter Partnerschaftsdauer und Dauer des Zusammenlebens gebildet. Beim *Heiratsalter* wurde die Differenz aus Lebensalter der Befragten und ihrer Ehedauer berechnet. Aus einer entsprechenden Frage wurde daneben die *Anzahl früherer fester Partnerschaften* bestimmt.

Das Ausmaß der *wahrgenommenen Einstellungsähnlichkeit* wurde mittels sieben Items erfragt, die sich inhaltlich auf die folgenden Bereiche bezogen: Umgang mit Geld, Geschmack bezüglich Kleidung, Freizeitinteressen, Ansichten zu Ordnung und Sauberkeit im Haushalt, Lebensanschauungen, politische Einstellungen, sexuelle Kompatibilität. Tabelle 2 informiert über Verteilungscharakteristika der Indikatoren:

Tabelle 2: Verteilungs- und Itemkennwerte zur Skala subjektive Einstellungsähnlichkeit, Gesamtstichprobe

wahrgenommene Übereinstimmung im Bereich:	N	\bar{x}	s	Schiefe	Kurtosis	p	r_{it}
Umgang mit Geld	2032	4.84	1.38	-.96**	-.08	.703	.42
Geschmack bezügl. Kleidung	2034	4.00	1.48	-.29**	-.76**	.504	.31
Freizeitinteressen	2035	4.23	1.56	-.48**	-.84**	.564	.33
Ordnung und Sauberkeit im Haushalt	2040	4.69	1.47	-.92**	-.20	.670	.33
Lebensanschauungen	2039	5.25	1.07	-1.69**	2.82**	.796	.43
politische Einstellungen	1992	4.68	1.37	-.87**	-.06	.660	.31
sexuelle Kompatibilität	1974	5.05	1.21	-1.38**	1.40**	.749	.40

** signifikant mit $p < 0.001$

Wie die Verteilungskennwerte der Indikatoren zeigen, werden die Items zur Einstellungsähnlichkeit deutlich linksschief beantwortet, d.h. der Schwerpunkt der Verteilung liegt im oberen Antwortbereich, was auch durch die mittleren bis leichten Itemschwierigkeiten bestätigt wird. Auch bezüglich der Wölbung lassen sich bei vier Indikatoren bedeutsame Abweichungen von einer Normalverteilung erkennen, in zwei Fällen (Kleidungs geschmack, Freizeitinteressen) in Richtung einer breitgipfligen, in den beiden übrigen Fällen in Richtung einer spitzgipfligen Verteilungsform. Die Trennschärfekoeffizienten liegen im unteren, aber akzeptablen Bereich, die interne Konsistenz der Skala berechnet sich zu $\alpha = .65$.

Zur Operationalisierung partnerschaftsspezifischen Kapitals wurde jeweils ein Index zu immateriellen und materiellen Investitionen eingesetzt. *Immaterielle Investitionen* wurden über eine Einschätzung gemessen, in welchem Ausmaß im Beziehungsverlauf eine Anpassung an die Wünsche des Partners stattgefunden hat. Drei Items wurden zum Index *materielle partnerschaftsspezifische Investitionen* aufsummiert: gemeinsamer Besitz einer Eigentumswohnung oder eines Hauses, gemeinsame Wohnungseinrichtung sowie ein gemeinsames Auto. Ein weiteres Item (Planung oder Aufbau eines gemeinsamen Betriebes) wurde aus dem Index ausgeschlossen und separat in den Analysen untersucht, da es Zusammenhänge mit umgekehrtem Vorzeichen zu anderen Konstrukten aufweist (vgl. Ergebnisteil).

Tabelle 3: Verteilungs- und Itemkennwerte zum Index *materielle Investitionen*, Gesamtstichprobe

Investition	N	p ^a	r _{it}
Besitz Eigentumswohnung oder Haus	2032	.60	.16
gemeinsame Wohnungseinrichtung	2030	.83	.25
gemeinsames Auto	2023	.76	.19

^a Die Itemschwierigkeit der dichotomen Items wird als Anteilswert der Zustimmung angegeben.

Zwei Items dienten der Erfassung des Erfolges der elterlichen Partnerschaft: *Stabilität der elterlichen Partnerschaft* gibt an, ob die Eltern sich getrennt haben oder geschieden sind (Ausprägung instabil) bzw. ob sie noch zusammen leben oder erst durch Tod eines Elternteils getrennt wurden (stabil). Daneben wurde zur Messung der *Qualität der elterlichen Partnerschaft* ein Item verwendet, in welchem von den Befragten ein Rating von 1 (sehr unglücklich) bis 6 (sehr glücklich) abzugeben war.

2.1.2 Soziodemographische Merkmale

Um lineare und nichtlineare alters- bzw. kalenderzeitbezogene Veränderungen⁸ modellieren zu können, wurde aus der Altersangabe eine mehrfach gestufte Variable Geburtskohorte mit den folgenden Kategorien gebildet: vor 1940 geboren, 1941 bis 1950 geboren, 1951 bis 1960 geboren, 1961-1970 geboren und nach 1971 geboren.

Auf der Paarebene wurde ein Index zur *Altershomogamie* gebildet. Dabei wurden Paare, bei denen die Frau älter ist als der Mann, als altershypogam klassifiziert; Paare, bei denen der Mann bis zu drei Jahre älter ist als die Frau, repräsentieren – angesichts der herrschenden Partnermarktgegebenheiten und Altersnormen, nach denen bei der Paarbildung Männer meist einige Jahre älter sind als Frauen – die Gruppe der altershomogamen Paare. Noch größere Altersdifferenzen wurden als altershypergam klassifiziert.

Daneben wurden verschiedene erwerbsbezogene Variablen eingesetzt:

- Das *Bildungsniveau* beider Partner wurde über zwei alternative Operationalisierungen gemessen: Erstens wurde der jeweilige höchste erreichte Schulabschluss in die in der Regel mindestens erforderlichen Schuljahre umcodiert (Indikator 1), und zweitens wurde ein kombinierter CASMIN-Index mit Schul- und Ausbildungsabschluss gebildet (Indikator 2) (vgl. Winkler, 1996). Ein zusätzlicher Index zur *Bildungshomogamie* differenziert auf der Paarebene nach Geschlecht, ob der Mann und die Frau jeweils niedrig (hier: bis zu 10 regulären Schuljahren) oder hoch (mehr als 10 Jahre Schulbildung) gebildet sind (4 Ausprägungen bzw. Kombinationen).
- Der *Erwerbsumfang*, über den für beide Partner Angaben vorliegen, wurde getrennt für Männer und Frauen in die Kategorien nicht erwerbstätig, geringfügig (unter 15 Stunden pro Woche) erwerbstätig, Teilzeit erwerbstätig mit mindestens 15 Stunden pro Woche und Vollzeit erwerbstätig codiert. Aufgrund kleiner Fallzahl der geringen Erwerbsumfänge bei den Männern wurde diese Variable in eine binäre Klassifikation mit den Ausprägungen unter 15 Stunden pro Woche (inkl. erwerbslos) und ab 15 Stunden pro Woche recodiert. Zusätzlich zu diesen individuellen Variablen wurde zur Analyse der paar-internen Aufgabenteilung ein *Paarindex zum Erwerbsstatus* (unter Berücksichtigung des Geschlechts) mit den folgenden Kategorien konstruiert: Mann Vollzeit, Frau nicht erwerbstätig; Mann Vollzeit, Frau Teilzeit; beide Vollzeit; Frau Vollzeit, Mann nicht erwerbstätig oder Teilzeit; beide Teilzeit oder arbeitslos bzw. nicht erwerbstätig sowie

⁸ Aufgrund des Querschnittsdesigns lassen sich diese beiden Effekte nicht differenzieren.

beide nicht erwerbstätig wegen Lebensphase (Ruhestand, Erziehungszeit, Studium etc.). Die Klassifikation wurde so gewählt, dass die Zellenbelegung auch bei Subgruppenanalysen möglichst ausreichend bleibt.

- Weiterhin wurde das *bedarfsgewichtete Haushaltseinkommen* berechnet. Da lediglich klassierte Angaben zum Haushaltseinkommen vorlagen, wurde hierzu das durch die Klassenmitte geschätzte Einkommen durch die Anzahl der Familienmitglieder dividiert, wobei Kinder mit dem Faktor .5 gewichtet wurden. Zusätzlich wurde eine Variable zur *subjektiven Zufriedenheit mit der finanziellen Situation* gebildet.
- Zur Bestimmung der paarinternen Einkommensverteilung wurde eine nach Geschlecht codierte Variable *Hauptverdiener(in)* gebildet mit den Ausprägungen Mann, Frau und beide (annähernd) gleiches Einkommen.

Bezüglich der Aufteilung von Haushaltsarbeiten unter Berücksichtigung der Genderperspektive wurden zwei Variablen gebildet: Zum einen der *Anteil (in %) der allein vom Mann übernommenen Hausarbeiten*, zum anderen der *Anteil der vom Mann übernommenen traditionell eher weiblich typisierten Arbeiten* (Kochen, Waschen, Putzen). Eine differenziertere Untersuchung der Aufgabenteilung im Haushalt war nicht möglich, da keine genauen Zeitbudgetdaten zur Verfügung standen, sondern lediglich eine Grobklassifikation dazu, welcher Partner im Regelfall die jeweilige Aufgabe übernimmt (Antwortkategorien: Befragter, Partner, dritte Person, keiner, abwechselnd).

Auf der Einstellungsebene wurde ein Indikator zur *Traditionalität von Geschlechtsrollenorientierungen* gebildet. Von den Befragten liegen Einschätzungen der Zustimmung beider Partner zur herkömmlichen Rollenverteilung des männlichen Ernährermodells in Form sechsstufiger Ratings vor. Zur Messung der traditionellen Ausrichtung des Paares wurden daneben Summenscores über beide Partner gebildet.

Als Indikator der *Homogamie der Geschlechtsrollenorientierungen* wurde die Differenz der Geschlechtsrollenorientierungen zwischen beiden Partnern bestimmt. Da nur in wenigen Fällen größere Diskrepanzen auftraten, wurde ein dichotomer Indikator mit den Ausprägungen gleiche vs. abweichende traditionale Geschlechtsrollenorientierungen gebildet.

Die *Konfessionszugehörigkeit* wurde für beide Partner aufgrund kleiner Fallzahlen in einzelnen Untergruppen die folgenden vier Kategorien codiert: Konfessionslose, Katholiken, Protes-

tanten und Angehörige sonstiger Religionsgemeinschaften. Als Indikator der *Religiosität* wurde daneben für beide Partner die Häufigkeit des Kirchgangs erfragt (Kategorien: nie, weniger als einmal pro Jahr, ein- bis zweimal pro Jahr, mehrmals pro Jahr, ein- bis dreimal im Monat, einmal pro Woche, mehr als einmal pro Woche).

Auch hier wurden auf der Paarebene Homogamieindizes konstruiert: Bezüglich der *Konfessionshomogamie* wurden Paare, bei denen beide Partner konfessionslos waren, differenziert von Paaren mit gleicher bzw. unterschiedlicher Konfession (hierunter fallen auch Paare mit dem konfessionell gebundenen und einem konfessionslosen Partner). Zur *religiösen Homogamie* wurde eine Variable mit drei Ausprägungen gebildet, die angibt, ob beide Partner wenig religiös sind (Kirchgang ein- bis zweimal im Jahr oder seltener), ob einer der Partner religiös ist (mehrmals im Jahr Kirchgang oder öfter) bzw. ob beide religiös sind. Allerdings besteht hier das Problem, dass konfessionslose Personen nicht zur Kirchgangshäufigkeit befragt wurden, was die Fallzahl in entsprechenden Analysen teilweise deutlich verringert. Um dieses Problem abzumildern, wurden in einer zweiten Operationalisierungsvariante Konfessionslose als gering religiös codiert.

Der Urbanitätsgrad des Wohnortes wurde über eine Einschätzung der Einwohnerzahl durch die Befragten erfasst (zur Verteilung vgl. Abschnitt 3). Zur groben Unterscheidung von städtischem und ländlichem Kontext wurde daraus eine Dummyvariable mit den Ausprägungen „unter 50.000 Einwohner“ (0) und „50.000 und mehr Einwohner“ (1) gebildet.

Tabelle 4 und Tabelle 5 geben abschließend eine tabellarische Übersicht der Verteilungen der in diesem Abschnitt beschriebenen Variablen, getrennt für Ost- und Westdeutsche:

Tabelle 4: Randverteilungen der kategorialen exogenen sozialstrukturellen Variablen

		Sozialisation in			
		Ostdeutschland		Westdeutschland	
		N	%	N	%
Geburtskohorte	bis 1940	215	18.50	105	12.00
	1941-1950	200	17.20	149	17.00
	1951-1960	275	23.60	244	27.80
	1961-1970	273	23.50	274	31.20
	ab 1971	201	17.30	105	12.00
	Gesamt	1164	100.00	877	100.00
Altershomogamie	altershypogam	175	15.10	152	17.40
	altershomogam	628	54.20	420	47.90
	altershypergam	356	30.70	304	34.70
	Gesamt	1159	100.00	876	100.00

		Sozialisation in			
		Ostdeutschland		Westdeutschland	
		N	%	N	%
Timing der Geburt des ersten Kindes	mehr als 1 Jahr vor der Hochzeit	93	12.60	41	6.80
	Abstand von 1 Jahr um die Hochzeit herum	289	39.20	170	28.10
	mehr als 1 Jahr nach Hochzeit	356	48.20	394	65.10
	Gesamt	738	100.00	605	100.00
Familienzyklus	Kinderlos	159	15.00	138	16.60
	Expansionsphase	78	7.40	82	9.80
	Primärsozialisation	55	5.20	64	7.70
	Schulpflichtiges Kind	102	9.60	143	17.20
	Adoleszenz	136	12.90	93	11.20
	Späte Adoleszenz	272	25.70	176	21.10
	Kontraktions-/ Altersphase	256	24.20	137	16.40
Gesamt	1058	100.00	833	100.00	
Konfessionszugehörigkeit Befragter	keine Religionsgemeinschaft	755	65.40	121	13.90
	katholisch	66	5.70	331	37.90
	evangelisch	302	26.10	389	44.60
	sonstige	32	2.80	32	3.70
	Gesamt	1155	100.00	873	100.00
Konfessionszugehörigkeit Partner	keine Religionsgemeinschaft	784	68.00	140	16.00
	katholisch	74	6.40	342	39.10
	evangelisch	265	23.00	360	41.10
	sonstige	30	2.60	33	3.80
	Gesamt	1153	100.00	875	100.00
Erwerbsstatus Mann	nicht erwerbstätig	409	35.40	199	22.90
	< 15 Stunden pro Woche / geringfügig beschäftigt	9	.80	7	.80
	Teilzeit mit 15 und mehr Stunden	24	2.10	17	2.00
	Vollzeit	712	61.70	645	74.30
	Gesamt	1154	100.00	868	100.00
Erwerbsstatus Frau	nicht erwerbstätig	503	43.50	387	44.40
	< 15 Stunden pro Woche / geringfügig beschäftigt	40	3.50	86	9.90
	Teilzeit mit 15 und mehr Stunden	137	11.90	201	23.10
	Vollzeit	476	41.20	198	22.70
	Gesamt	1156	100.00	872	100.00

		Sozialisation in			
		Ostdeutschland		Westdeutschland	
		N	%	N	%
Erwerbsstatus Paar	Mann Vollzeit, Frau nicht erwerbstätig	196	17.50	265	31.60
	Mann Vollzeit, Frau Teilzeit	123	11.00	208	24.80
	beide Vollzeit	378	33.70	156	18.60
	Frau Vollzeit, Mann nicht erwerbstätig oder Teilzeit	93	8.30	40	4.80
	beide Teilzeit oder nicht erwerbstätig	105	9.40	84	10.00
	beide nicht erwerbstätig wg. Lebensphase	227	20.20	85	10.10
	Gesamt	1122	100.00	838	100.00
Hauptverdiener(in)	Frau	217	18.90	115	13.20
	Doppelverdiener	243	21.10	101	11.60
	Mann	690	60.00	655	75.20
	Gesamt	1150	100.00	871	100.00
Konfessionshomogamie	beide konfessionslos	704	61.20	83	9.50
	heterogam	181	15.70	291	33.30
	homogam	265	23.00	499	57.20
	Gesamt	1150	100.00	873	100.00
Bildungshomogamie	Mann niedrig, Frau niedrig	773	67.70	497	57.60
	Mann hoch, Frau niedrig	119	10.40	132	15.30
	Mann niedrig, Frau hoch	102	8.90	68	7.90
	Mann hoch, Frau hoch	148	13.00	166	19.20
	Gesamt	1142	100.00	863	100.00
religiöse Homogamie	beide wenig religiös	224	68.90	463	66.30
	einer religiös (heterogam)	28	8.60	86	12.30
	beide religiös	73	22.50	149	21.30
	Gesamt	325	100.00	698	100.00
religiöse Homogamie (mit Konfessionslosen)	beide wenig religiös / konfessionslos	1043	90.00	636	72.60
	einer religiös (heterogam)	43	3.70	91	10.40
	beide religiös	73	6.30	149	17.00
	Gesamt	1159	100.00	876	100.00

Tabelle 5: Deskriptive Verteilungskennwerte für die metrischen sozialstrukturellen Variablen

	Sozialisation in									
	Ostdeutschland					Westdeutschland				
	N	Min.	Max.	\bar{x}	s	N	Min.	Max.	\bar{x}	s
Alter Mann	1163	18.00	88.00	48.02	13.65	877	19.00	80.00	47.56	11.95
Alter Frau	1161	17.00	73.00	45.64	13.69	876	18.00	70.00	45.00	11.45
Beziehungsdauer	1163	.08	52.00	21.64	13.94	876	.33	60.00	20.76	12.22
voreheliche Kohabitationsdauer (Jahre)	914	.00	45.02	1.22	3.07	769	.00	22.00	1.21	2.47
Partnerschaftsdauer bei Haushaltsgründung (Jahre)	1054	.00	31.05	1.09	2.10	794	.00	20.00	1.32	2.30
Heiratsalter (Jahre)	906	16.00	64.00	24.50	7.40	759	16.00	59.91	24.57	6.30
Gesamtkinderzahl	1163	0	8	1.64	1.13	877	0	11	1.78	1.21
Anzahl gemeinsamer Kinder	1155	0	7.00	1.34	1.11	877	0	6.00	1.54	1.15
Alter jüngstes Kind	975	.00	48.00	21.12	12.15	731	.00	50.00	17.20	11.77

Zufriedenheit mit finanzieller Situation	1160	1	6	3.78	1.39	872	1	6	4.53	1.23
Haushaltseinkommen (bedarfsgewichtet)	888	100.00	2250.00	795.95	383.98	622	125.00	2250.00	1033.18	466.17
Anteil der vom Mann übernommenen Aufgaben im Haushalt	1156	.00	100.00	39.82	26.59	869	.00	100.00	35.68	26.08
Anteil der vom Mann übernommenen weiblich typisierten Aufgaben im Haushalt	1094	.00	100.00	15.47	29.38	829	.00	100.00	12.73	27.79
Anzahl früherer Partnerschaften	810	0	12.00	1.70	1.51	850	0	10.00	1.57	1.63
Traditionalismus Mann	1136	1	6	2.07	1.52	862	1	6	2.87	1.62
Traditionalismus Frau	1148	1	6	1.97	1.47	858	1	6	2.68	1.59
Religiosität Befragter	407	1	7	3.52	1.58	756	1	7	3.63	1.58
Religiosität Partner	376	1	7	3.39	1.68	736	1	7	3.39	1.61
Schuljahre Mann	1150	7	13	10.35	1.51	869	7	13	10.60	1.71
Schuljahre Frau	1149	7	13	10.28	1.45	870	7	13	10.43	1.55
Bildungsniveau Mann (inkl. Ausbildung)	1063	1	7	5.02	1.09	786	1	7	5.18	1.21
Bildungsniveau Frau (inkl. Ausbildung)	1077	1	7	4.93	1.09	834	1	7	4.84	1.31
Index gemeinsame materielle Investitionen	1142	0	3	2.15	.86	865	0	3	2.28	.84
Anpassung an Partnerwünsche	1154	1	6	4.10	1.52	869	1	6	3.76	1.44
Qualität der elterl. Partnerschaft / Ehe	1008	1	6	4.48	1.31	823	1	6	4.20	1.39
<i>Trennung / Scheidung Eltern</i>	1117	0	1	.11 ^a		864	0	1	.09	
<i>voreheliche Kohabitation</i>	914	0	1	.41		769	0	1	.43	
<i>gleiche Geschlechtsrollenorientierungen</i>	1160	0	1	.22		853	0	1	.20	
<i>Stadt / Land</i>	1098	0	1	.30	.46	837	0	1	.30	.46

^a Der Mittelwert binär codierter Variablen (kursiv) lässt sich als Anteilswert der Zustimmung interpretieren.

Im Folgenden werden die Operationalisierungen der vermittelnden Konstrukte (Paarkonflikte, Interaktionsstile und Partnerschaftszufriedenheit) sowie der abhängigen Variablen (Partnerschaftsstabilität) dargestellt. Der Wortlaut der eingesetzten Items findet sich im Anhang des Buches.

2.2 Konflikte

Insgesamt wurden in den Interviews Angaben zur Häufigkeit von Meinungsverschiedenheiten in insgesamt 16 verschiedenen inhaltlichen Bereichen erfragt. Dabei wurden die beiden Aspekte Häufigkeit und Belastungsgrad differenziert. Nach der Belastung wurde ausschließlich in den Fällen gefragt, in denen bei der Konflikthäufigkeit nicht „nie“ angegeben wurde; andernfalls ist die Variable Belastungsgrad daher als systembedingt fehlend codiert.

Auf Basis der Überlegung, dass Konflikte, die nie auftreten, auch keine Belastungswirkung haben können, wurden alle Nennungen von „nie“ bei der Konflikthäufigkeit als minimale Ausprägung des Belastungsgrades (0) codiert. Tabelle 6 stellt die univariaten Verteilungen von Konflikthäufigkeit und -belastung einander gegenüber und zeigt die Interkorrelation beider Einschätzungen:

Tabelle 6: Häufigkeit und Belastungsgrad von Meinungsverschiedenheiten sowie Interkorrelationen zwischen den beiden Einschätzungen, Gesamtstichprobe

Konfliktbereich	N	Anteil: Nennung „gelegentlich“ oder öfter ^a	Anteil: Belastung „etwas“ oder stärker ^b	Korrelation: Häufigkeit × Belastung	
				ρ	r
finanzielle Dinge	2035	21.9	24.9	.93**	.80**
Verwandtschaft/Freunde	2038	30.9	26.3	.91**	.87**
Gespräche zu selten/zu oft	2030	16.8	16.9	.98**	.85**
zu wenig gefühlsmäßige Zuwendung	2031	20.3	11.2	.97**	.85**
mangelnde Empathie	2034	21.3	23.0	.95**	.88**
Sexualität	2000	12.0	14.8	.99**	.87**
Zeitmangel	2034	23.0	22.6	.96**	.88**
Freizeitgestaltung	2039	14.0	11.7	.98**	.88**
Vertrauen und Eifersucht	2035	9.2	10.5	.99**	.87**
Aufteilung der Haushaltsarbeiten	2038	19.1	11.3	.97**	.90**
Berufstätigkeit der Frau	2020	4.8	4.4	1.00**	.84**
Kindererziehung, -betreuung ^c	1663	16.2	55.4	.84**	.83**
Ansichten Ordnung und Sauberkeit	2037	25.9	13.8	.95**	.80**
unterschiedliche Weltanschauungen	2037	11.1	4.8	.99**	.85**
Fertilitätsentscheidungen	1976	2.9	2.8	1.00**	.90**
berufliche Dinge	2027	6.6	6.9	1.00**	.80**

** signifikant mit $p < 0.001$

^a Die Einschätzung der Konflikthäufigkeit erfolgte in den Kategorien nie, selten, gelegentlich, häufiger, oft, sehr oft

^b Die Einschätzung der Konfliktbelastung erfolgte in den Kategorien gar nicht, wenig, etwas, eher stark, stark, sehr stark

^c In dieser Zeile sind nur Paare mit Kindern aufgeführt.

Die relativen Häufigkeiten der Einschätzungen in Tabelle 6 veranschaulichen, dass in der untersuchten Stichprobe eher selten von Meinungsverschiedenheiten berichtet wird. Dies kann zum einen darauf zurückgeführt werden, dass Konflikte trotz allem (relativ) seltene Ereignisse darstellen, zum anderen steht dies in Einklang zu der hohen Zufriedenheit und

Stabilität der untersuchten Paarbeziehungen. Daneben sind auch kognitive Prozesse im Sinne eines Herabspielens von Paarkonflikten, wie sie z.B. nach der Theorie der Dissonanzreduktion (Festinger, 1957) erwartet werden können, nicht auszuschließen. Nach der Häufigkeit des Auftretens sind Konflikte in den Bereichen *Verwandtschaft / Freunde, Ansichten über Ordnung und Sauberkeit* sowie *Zeitmangel* dominant; die stärkste Belastung hingegen geht von Konflikten in den Bereichen *Kindererziehung / -betreuung, Verwandtschaft / Freunde* und *finanzielle Dinge* aus.

Daneben ist den letzten beiden Spalten zu entnehmen, dass die Einschätzung von Konflikthäufigkeit und -belastung über alle Bereiche hinweg sehr hoch korreliert, eine separate Analyse beider Aspekte erscheint daher nicht unbedingt notwendig. Um beide Facetten simultan zu berücksichtigen, wird zunächst für jeden Bereich ein Summenscore aus Häufigkeit und Belastung gebildet; in den folgenden Analysen werden nur noch diese Scores verwendet.⁹ Anschließend ist mittels einer exploratorischen Faktorenanalyse die Eindimensionalität der Skala zu prüfen, wobei gegebenenfalls einzelne Indikatoren eliminiert werden. Dieses Vorgehen kann damit begründet werden, dass die Differenzierung einzelner Konfliktbereiche nicht im Fokus der vorliegenden Arbeit steht (zu entsprechenden detaillierteren Analysen vgl. Rüssmann, 2006) und dass bereichsspezifische Auswertungen die Analysen unverhältnismäßig kompliziert hätten. Schließlich wird eine zweite Itemrevision auf Basis einer multivariaten CFA vorgenommen, und die hieraus resultierenden Indikatoren werden zu einem Gesamtkonfliktscore aggregiert.

Eine exploratorische Faktorenanalyse unter Eingabe sämtlicher Konfliktbereiche ergibt zunächst eine dreifaktorielle Lösung (Varianzaufklärung: 42.01%), bei der die Konfliktfelder *berufliche Dinge* und *Erwerbstätigkeit der Frau* bzw. *Fertilitätsentscheidungen* und *Vertrauen / Eifersucht* jeweils eigene Faktoren bilden. Werden diese vier Konfliktfelder aus der Analyse entfernt, resultiert eine einfaktorielle Lösung, durch die 37.64% der Gesamtvarianz erklärt wird. Tabelle 7 illustriert die entsprechenden Faktorladungen der einzelnen Indikatoren:

⁹ Der Score zum Konfliktfeld „Kindererziehung und -betreuung“ wird dabei bei den kinderlosen Paaren (N=336) auf 0 gesetzt.

Tabelle 7: Exploratorische Faktorenanalyse zur Skala „Paarkonflikte“, Gesamtstichprobe (N=1974)

Konfliktbereich	Faktorladung
finanzielle Dinge	.51
Verwandtschaft / Freunde	.48
Gespräche zu selten / zu oft	.61
zu wenig gefühlsmäßige Zuwendung	.73
mangelnde Empathie	.73
Sexualität	.67
Zeitmangel	.62
Freizeitgestaltung	.58
Aufteilung der Haushaltsarbeiten	.60
Ansichten Ordnung / Sauberkeit	.56

Wie die unter 2.6 beschriebene konstruktübergreifende konfirmatorische Faktorenanalyse zeigt, lassen sich Konflikte in einigen Feldern im multivariaten Kontext nicht eindeutig von den übrigen latenten Modellkonstrukten trennen. Daher werden die folgenden Konfliktbereiche aus der Skala entfernt: *Verwandtschaft / Freunde*, *Gesprächshäufigkeit*, *Sexualität*, *Zeitmangel* und *Aufteilung der Haushaltsarbeiten*.

Im letzten Schritt werden für die zweite Revision der Konfliktskala verschiedene Item- und Skalenkennwerte berechnet, um ihre psychometrische Qualität bewerten zu können. Tabelle 8 gibt einen Überblick über Verteilungskennwerte sowie Schwierigkeits- und Trennschärfeindizes der einzelnen Indikatoren:

Tabelle 8: Verteilungs- und Itemkennwerte zur Skala Paarkonflikt, Gesamtstichprobe

Konfliktbereich	N	\bar{x}	s	Schiefe	Kurtosis	p	r_{it}
finanzielle Dinge (k1)	2035	3.17	2.66	.97**	.03	.119	.36
zu wenig gefühlsmäßige Zuwendung (k2)	2030	2.79	2.70	1.29**	.61**	.105	.52
mangelnde Empathie (k3)	2032	2.98	2.59	1.06**	.20	.108	.54
Freizeitgestaltung (k4)	2039	2.22	2.15	1.79**	2.78**	.066	.41
Ansichten Ordnung / Sauberkeit (k5)	2037	2.76	2.47	1.30**	1.09**	.095	.39

** signifikant mit $p < 0.001$

Die Verteilungskennwerte zeigen, dass sämtliche Indikatoren stark rechtsschief und überwiegend spitzgipflig beantwortet wurden, d.h. es wird von einem großen Teil der Befragten eine geringe Konfliktbelastung geäußert, hohe Ausprägungen kommen nur sehr selten vor; folglich fallen die Itemschwierigkeiten sehr hoch aus (entspricht einem niedrigen p). Die

durchweg signifikanten Abweichungen von einer Normalverteilung deuten darauf hin, dass nach Möglichkeit auf entsprechende Verfahren zur Korrektur der Parameterschätzer und Standardfehler zu rekurren ist, wie sie in den Strukturgleichungsmodellen eingesetzt werden. Für die OLS-Regressionsmodelle kann immerhin auf Basis des gegebenen Stichprobenumfangs davon ausgegangen werden, dass die Standardfehler der Schätzer asymptotisch korrekt sind (vgl. Bortz, 1999). Alle Trennschärfekoeffizienten liegen im akzeptablen Bereich; die interne Konsistenz der Skala berechnet sich zu $\alpha = .69$.

2.3 Interaktionsstile

Gemäß den theoretischen Ausführungen wurde zunächst geprüft, ob sich die im Theorie-Teil beschriebenen zentralen Interaktionsstile in den erhobenen Daten empirisch finden lassen. Aufgrund des umfangreichen Itempools zum Thema Kommunikation in Partnerschaften im Datensatz wurden zum Zweck der Itemselektion zunächst exploratorische Faktorenanalysen eingesetzt; die entsprechenden Ergebnisse dieser Voranalysen werden an dieser Stelle zur Verbesserung der Übersichtlichkeit nicht wiedergegeben (vgl. zu einer umfassenden Beschreibung der verwendeten Items Rüssmann et al., 2004).¹⁰ Nach der faktorenanalytisch geleiteten Selektion nicht eindeutig zuordenbarer Indikatoren resultiert die folgende Faktorenlösung (vgl. Tabelle 9):

¹⁰ Obgleich in der vorliegenden Arbeit dieselben Daten verwendet werden wie in den Beiträgen in Hill (2004), resultiert – insbesondere im Bereich Paarinteraktion – eine differierende Dimensionierung der Indikatoren. Dies liegt primär an der unterschiedlichen Analysestrategie sowie an der Zielsetzung, die verwendeten Instrumente im konstruktübergreifenden Kontext von Varianzkonfundierungen zu bereinigen. So zeigt sich beispielsweise in entsprechenden konfirmatorischen Faktorenanalysen, dass die aus der Bindungstheorie ableitbaren Arbeitsmodelle Intimität und Vertrauen nicht eindeutig vom Faktor Partnerschaftszufriedenheit abzugrenzen sind. Aus diesem Grund wurden diese beiden Dimensionen in der vorliegenden Arbeit nicht weiter berücksichtigt.

Tabelle 9: Exploratorische Faktorenanalyse zu den Interaktionsskalen, Gesamtstichprobe (N=1891)

Indikator	Faktor ^a			
	DKV	PAK	KKV	VM
1. Konflikt: Kritisieren des Partners	.73	-.07	-.05	-.12
2. Konflikt: Aufzählen der Fehler des Partners	.68	-.04	-.07	-.03
3. Konflikt: verletzen mit spöttischen Bemerkungen	.67	-.07	-.13	.07
4. Konflikt: Unterbrechen des Partners / der Partnerin	.65	-.08	-.04	.03
5. Konflikt: Befragte/r hält Partner für einen Dummkopf	.62	-.06	-.13	.05
6. Konflikt: mindestens einer wird laut/fängt an zu schreien	.56	.02	-.15	.10
7. Alltag: positive Rückmeldung, wenn Partner/in etwas sagt	-.08	.73	.21	.06
8. Alltag: loben, wenn ihr Partner/in etwas gut gemacht hat	-.08	.72	.13	.09
9. Alltag: nachfragen, wenn nicht klar ist, was Partner/in meint	.03	.71	.05	-.15
10. Alltag: aufmerksames Zuhören gegenüber Partner/in	-.13	.70	.09	-.12
11. Konflikt: gemeinsame Lösung steht im Vordergrund	-.12	.10	.72	-.12
12. Konflikt: Respekt vor Meinung d. Partners	-.21	.13	.65	.00
13. Konflikt: zunächst Äußerung von Gefühlen und Ansichten	.07	.19	.63	-.19
14. Konflikt entschärfen: Humor, etwas Nettes sagen	-.10	.02	.61	.14
15. Konflikt: mindestens einer bleibt ruhig und diskutiert sachlich	-.15	.08	.53	.06
16. besser wahre Gefühle verbergen als Partner/in verletzen	.11	.00	.06	.80
17. Probleme lieber gar nicht erst ansprechen	-.02	-.10	-.08	.79

^a Bezeichnung der Faktoren: DKV = destruktives Konfliktverhalten; PAK = positive Alltagskommunikation; KKV = konstruktives Konfliktverhalten; VM = Konfliktvermeidung

Die Ergebnisse dieser ersten Analyse demonstrieren bereits, dass sich die im Theorieteil eingeführte Differenzierung des Kommunikationssettings eindeutig in den Daten widerspiegelt: Wie erwartet bilden Indikatoren, die Alltagsinteraktion betreffen, einen eigenen Faktor, der hier als *positive Alltagskommunikation* bezeichnet wird. Im Bereich der Konfliktinteraktion lassen sich klar die beiden Dimensionen *konstruktives* und *destruktives Konfliktverhalten* differenzieren. Hierbei ist hervorzuheben, dass es sich offensichtlich nicht um eine bipolare Dimension handelt, sondern dass das Auftreten funktionaler und dysfunktionaler Verhaltensweisen in Konflikten eher unabhängig voneinander ist. Eine Sonderrolle nimmt die *Konfliktvermeidung* ein, deren Indikatoren wiederum einen eigenen Faktor bilden; eine eindeutige Zuordnung zum Alltags- bzw. Konfliktsetting ist hier nicht möglich. Die Varianzaufklärung der beschriebenen Vierfaktorlösung beträgt 49.37%.

Im nächsten Schritt wurde konstruktübergreifend mittels konfirmatorischer Faktorenanalyse geprüft, ob eine Trennung der Interaktionsstile von den übrigen Modellkonstrukten empirisch möglich ist bzw. welche Indikatoren aufgrund von Nebenladungen eliminiert werden sollten (detaillierte Ergebnisse hierzu finden sich unter 2.6): Item 4 (Faktor: DKV)

zeigt dabei unerwünschte Nebenladungen auf den Faktoren *Paarkonflikt* und *Partnerschaftsstabilität*, Item 13 (KKV) lässt sich nicht eindeutig von den Konstrukten *positive Alltagskommunikation*, *destruktives Konfliktverhalten* sowie *Vermeidung* trennen. Item 9 (PAK) hängt ebenfalls mit dem Faktor *destruktives Konfliktverhalten* zusammen. Auf diese Indikatoren wurde daher verzichtet. Am Ende ergeben sich aus dieser zweiten Revision Neufassungen der Interaktionskalen, die im Folgenden auf ihre psychometrische Qualität hin überprüft werden.

Als erstes wird eine Kurzsкала zur positiven Alltagskommunikation beschrieben. Tabelle 10 zeigt die Verteilungen und Itemkennwerte der entsprechenden Indikatoren:

Tabelle 10: Verteilungs- und Itemkennwerte zur Skala positive Alltagskommunikation, Gesamtstichprobe

Indikator	N	\bar{x}	s	Schiefe	Kurtosis	p	r_{it}
Alltag: aufmerksames Zuhören in Unterhaltungen mit Partner/in (pak1)	2033	5.12	1.03	-1.21**	1.29**	.756	.38
Alltag: loben, wenn ihr Partner/in etwas gut gemacht hat (pak2)	2032	5.02	1.12	-1.13**	.70**	.735	.51
Alltag: positive Rückmeldung geben, wenn Partner/in etwas sagt (pak3)	2016	5.00	1.07	-1.15**	1.17**	.727	.55

** signifikant mit $p < 0.001$

Bezüglich der Lage und Form der Verteilungen fällt auf, dass die Antworten sich im oberen Bereich der sechsstufigen Skala konzentrieren: Die Mittelwerte fallen hoch aus, die Itemschwierigkeit niedrig, und die signifikanten negativen Schiefestatistiken verweisen auf nach rechts steil abfallende Verteilungen, deren Gipfel eher spitz zulaufen (positive Kurtosis). Die Trennschärfen der Items liegen im mittleren Bereich. Die interne Konsistenz der Skala liegt bei $\alpha = .66$, was angesichts der geringen Itemanzahl akzeptabel ist.

Als nächstes wird die revidierte Fassung der Skala konstruktives Konfliktverhalten besprochen, welche sich aus vier Indikatoren zusammensetzt. Tabelle 11 illustriert die zugehörigen Verteilungs- und Skaleneigenschaften:

Tabelle 11: Verteilungs- und Itemkennwerte zur Skala konstruktives Konfliktverhalten, Gesamtstichprobe

Indikator	N	\bar{x}	s	Schiefe	Kurtosis	p	r_{it}
Konflikt: gemeinsame Lösung steht im Vordergrund (kkv1)	2020	5.14	1.31	-1.66**	1.99**	.781	.42
Konflikt: Respekt vor Meinung d. Partners (kkv2)	2027	5.13	1.29	-1.59**	1.86**	.777	.43
Konflikt entschärfen: Humor, Nettes sagen, Verständnis zeige (kkv3)	2017	4.39	1.58	-.77**	-.45**	.603	.36
Konflikt: mindestens einer bleibt ruhig und diskutiert sachlich (kkv4)	2025	4.66	1.59	-.95**	-.28**	.674	.34

** signifikant mit $p < 0.001$

Die Antworten bewegen sich im Mittel deutlich oberhalb des Skalenmittelpunktes, die Itemschwierigkeit liegt niedrig (indiziert durch hohes p). Auch hier fällt die Linksschiefe der Verteilungen auf; im Unterschied zur Skala positive Alltagskommunikation sind die Verteilungsgipfel der letzten beiden Indikatoren jedoch eher flach (negative Kurtosis). Die Trennschärfekoeffizienten liegen im unteren, aber noch akzeptablen Bereich. Die interne Konsistenz der Skala berechnet sich zu $\alpha = .60$.

Weiterhin resultiert aus der konfirmatorische Faktorenanalyse die Skala destruktives Konfliktverhalten mit fünf Indikatoren. Tabelle 12 fasst die wichtigsten Item- und Skalenparameter zusammen:

Tabelle 12: Verteilungs- und Itemkennwerte zur Skala destruktives Konfliktverhalten, Gesamtstichprobe

Indikator	N	\bar{x}	s	Schiefe	Kurtosis	p	r_{it}
Konflikt: Kritisieren des Partners (dkv1)	2023	3.13	1.72	.20**	-1.22**	.353	.51
Konflikt: Aufzählen der Fehler des Partners (dkv2)	2032	2.13	1.54	1.20**	.24*	.192	.51
Konflikt: verletzen mit spöttischen Bemerkungen (dkv3)	2026	2.22	1.57	1.06**	-.13	.206	.49
Konflikt: Befragte/r hält Partner manchmal für einen Dummkopf (dkv4)	2020	2.06	1.60	1.33**	.42**	.190	.46
Konflikt: mindestens einer wird laut/fängt an zu schreien (dkv5)	2029	2.98	1.88	.40**	-1.32**	.344	.39

* signifikant mit $p < 0.01$

** signifikant mit $p < 0.001$

Die Mitte der Itemverteilungen liegt bei allen Items unterhalb des Skalenmittelpunkts (3.5), d.h. tendenziell werden geringe Ausprägungen des destruktiven Konfliktverhaltens angegeben, die Itemschwierigkeit ist folglich hoch. Die positive Schiefe der Verteilungen zeigt zudem an, dass diese links steiler und rechts flacher abfallen; die Verteilungsgipfel sind bei einigen Indikatoren breiter, bei anderen spitzer als bei einer Normalverteilung. Die Trennschärfekoeffizienten liegen im mittleren Bereich. Die interne Konsistenz der Skala ist zufrieden stellend ($\alpha = .71$).

Abschließend werden noch die Eigenschaften der Skala Konfliktvermeidung dargestellt (vgl. Tabelle 13):

Tabelle 13: Verteilungs- und Itemkennwerte zur Skala Konfliktvermeidung, Gesamtstichprobe

Indikator	N	\bar{X}	s	Schiefe	Kurtosis	p	r_{it}
Probleme gar nicht erst ansprechen (verm1)	2025	1.93	1.55	1.60**	1.30**	.170	.36
besser wahre Gefühle verbergen als Partner/in verletzen (verm2)	1989	2.53	1.73	.82**	-.65**	.260	.36

** signifikant mit $p < 0.001$

Die Mitte der Antwortverteilungen liegt deutlich unterhalb des Skalenmittelpunkts (hohe Itemschwierigkeit) und die Verteilungen sind rechtsschief. Beim ersten Item ist eine eher spitz-, beim zweiten hingegen eine breitgipflige Verteilung zu erkennen. Die Trennschärfe, die in diesem Fall der Interkorrelation der beiden Items entspricht, liegt im noch akzeptablen Bereich. Die Reliabilität der Skala ist hingegen nicht befriedigend ($\alpha = .53$); allerdings muss hier in Rechnung gestellt werden, dass die Skala nur zwei Indikatoren umfasst und daher die Berechnung von split-half-Reliabilitätsschätzern wie der internen Konsistenz ohnehin nicht unproblematisch ist. Zu exploratorischen Zwecken werden die Items trotz der nicht ganz überzeugenden Skaleneigenschaften zu einem Index aggregiert.

2.4 Partnerschaftszufriedenheit

Zur Erfassung der Partnerschaftszufriedenheit wurde auf deutsche Versionen der Relationship Assessment Scale (RAS; Hendrick, 1988) zurückgegriffen, wie sie erstmals von Hassebrauck (1991) sowie von Sander und Böcker (1993) vorgelegt worden sind. Im Kontext

der vorliegenden Arbeit ergibt sich hier allerdings das Problem, dass ein Item der ursprünglichen Skala das Ausmaß an Paarkonflikten erfasst. Da hier eine Varianzkonfundierung mit der Skala Paarkonflikte (vgl. Abschnitt 2.2) zu erwarten war, wurde dieser Indikator vor der Befragung gegen die allgemeine Einschätzung „Wie glücklich sind Sie insgesamt in Ihrer Beziehung?“ ausgetauscht, welche sich in einigen gebräuchlichen Instrumenten wie der Kansas Marital Satisfaction Scale (KMS) findet (vgl. Schumm et al., 1986). Daneben wurde das ursprünglich fünfstufige Antwortformat durch eine zehnstufige Skala ersetzt, um eine feinere Varianzdifferenzierung zu erreichen. Die Itemformulierungen lauten wie folgt:

1. Wie zufrieden sind Sie insgesamt mit Ihrer Beziehung?
(1 – sehr unzufrieden, 10 – sehr zufrieden)
2. Wie gut kommt Ihr/e Partner/in Ihren Bedürfnissen entgegen?
(1 – überhaupt nicht, 10 – voll und ganz)
3. Wie gut ist Ihre Beziehung verglichen mit anderen Beziehungen?
(1 – sehr schlecht, 10 – sehr gut)
4. Wie oft wünschen Sie sich, dass Sie diese Beziehung *nicht* eingegangen wären?
(1 – niemals, 10 – sehr oft)
5. In welchem Maß hat Ihre Beziehung Ihre ursprünglichen Erwartungen erfüllt?
(1 – überhaupt nicht, 10 – voll und ganz)
6. Wie sehr lieben Sie Ihre/n Partner/in?
(1 – überhaupt nicht, 10 – voll und ganz)
7. Wie glücklich sind Sie insgesamt in Ihrer Beziehung?
(1 – sehr unglücklich, 10 – sehr glücklich)

Da Item 4 in den Pretests (Rüssmann et al., 2004) vermutlich aufgrund der negativen Polung eine ungünstige Trennschärfe zeigte, wurden die Befragten bei einer hohen Ausprägung von den Interviewern erklärend auf die Polung des Items hingewiesen und erhielten dadurch Gelegenheit, ihre Antwort – falls notwendig – zu korrigieren. Bei den Antworten fällt auf, dass 194 der Befragten bei Item 3, welches sich inhaltlich auf den Vergleich mit anderen Partnerschaften bezieht, die Kategorie „weiß nicht“ wählten; offensichtlich fällt der Vergleich zwischen der eigenen und fremden Partnerschaften in dieser Hinsicht schwer. Um diese Fälle in den folgenden Analysen dennoch verwenden zu können, wurden diese Antworten für die folgenden Analysen durch den Mittelwert aus den übrigen Indikatoren ersetzt.

Als erstes wurden die Indikatoren mittels einer exploratorischen Faktorenanalyse auf Eindimensionalität geprüft. Tabelle 14 informiert über die resultierenden Faktorladungen der einfaktoriellen Lösung:

Tabelle 14: Exploratorische Faktorenanalyse zur Skala „Partnerschaftszufriedenheit“, Gesamtstichprobe (N=1947)

Item	Faktorladung
7. Wie glücklich ist Befragter insgesamt mit Beziehung	.87
6. Wie sehr liebt Befragte/r Partner	.84
2. Kommt Partner Bedürfnissen entgegen	.80
5. Erfüllung ursprünglicher Erwartungen	.80
1. Zufriedenheit mit Beziehung insgesamt	.76
3. Qualität d. Beziehung verglichen mit anderen Beziehungen	.75
4. Häufigkeit Wunsch Beziehung nicht eingegangen zu Isein	-.60

Auch hier zeigt sich, dass das negativ gepolte Item 4 den latenten Faktor insgesamt am schlechtesten beschreibt. Dennoch sind alle Ladungskoeffizienten größer oder gleich $|.6|$ und damit als ausreichend hoch einzustufen; die Varianzaufklärung beträgt 60.5%. Während demnach auf Basis dieser Analyse keine Itemselektion erfolgen kann, zeigt das konstruktübergreifende Messmodell (2.6) mehrere unerwünschte Nebenladungen: So laden die Items 2, 5 und 7 auf dem Faktor Paarkonflikt, die beiden letztgenannten auf dem Faktor destruktives Konfliktverhalten und Item 5 außerdem auf dem latenten Konstrukt konstruktives Konfliktverhalten. Schließlich lässt sich Indikator 4 auch dem Konstrukt Partnerschaftsstabilität zuordnen. Auf Basis dieser Befunde wurde die Skala auf die Items 1, 3 und 6 reduziert. Tabelle 15 zeigt die Verteilungseigenschaften und psychometrischen Itemkennwerte der resultierenden Skala Partnerschaftszufriedenheit:

Tabelle 15: Verteilungs- und Itemkennwerte zur Skala Partnerschaftszufriedenheit, Gesamtstichprobe

Item	N	\bar{x}	s	Schiefe	Kurtosis	p	r_{it}
2. Kommt Partner Bedürfnissen entgegen (pz1)	2031	8.17	1.73	-1.30**	2.10**	.697	.67
3. Qualität d. Beziehung verglichen mit anderen Beziehungen (pz2)	2023	8.53	1.56	-1.24**	1.76**	.751	.65
7. Wie glücklich ist Befragter insgesamt mit Beziehung (pz3)	2035	8.90	1.52	-2.10**	5.57**	.816	.67

** signifikant mit $p < 0.001$

Die Betrachtung der Mittelwerte sowie der höheren Verteilungsmomente zeigt, dass die Fragen von der Mehrzahl der Befragten überwiegend und vergleichsweise uniform in Richtung einer hohen Partnerschaftszufriedenheit beantwortet wurden, die Itemschwierigkeit

fällt entsprechend niedrig aus. Die Trennschärfen aller Items liegen im mittleren bis hohen Bereich, was die Homogenität des Instruments belegt.

Abschließend ist festzuhalten, dass die Skala in der vorliegenden Revision trotz der geringen Itemanzahl eine sehr gute interne Konsistenz aufweist ($\alpha = .81$). In Kombination mit ihrer Eindimensionalität und der empirischen Differenzierbarkeit der Indikatoren von anderen Modellkonstrukten ist die Anwendung in einem multivariaten Kontext daher ausreichend abgesichert.

2.5 Subjektive Partnerschaftsstabilität

Zur Erfassung der Partnerschaftsstabilität wurden fünf Items eingesetzt, die in Anlehnung an das von Booth et al. (1983) entwickelte Marital Instability Inventory (MII) bzw. eine deutsche Fassung hiervon (vgl. Hartmann & Simon, 1997) formuliert wurden. Darin wird ein progredienter Trennungsprozess postuliert, der sich von Trennungsgedanken über die Mobilisierung sozialer Unterstützung bis hin zu konkreten trennungsvorbereitenden Aktionen erstreckt. Verschiedene nordamerikanische empirische Studien demonstrieren die prädiktive Validität des MII bei der Trennungs- und Scheidungsvorhersage (Booth et al., 1983; Booth, Johnson, White, & Edwards, 1985, 1986; Booth & White, 1980), während bisher im deutschen Sprachraum noch kaum Befunde zur psychometrischen Qualität der Skala vorliegen.

Da neben Ehepaaren auch nichteheliche Lebensgemeinschaften untersucht werden sollen, wurde das ursprünglich für Ehepaare entwickelte Instrument leicht adaptiert. Konkret wurde zu einem der Items eine äquivalente Fassung für nichteheliche Lebensgemeinschaften entwickelt, die einen vergleichbar hohen Grad an Instabilität indizieren sollte. Dabei handelte es sich um die Frage „Haben Sie sich jemals wegen einer möglichen Scheidung von einem Anwalt beraten lassen?“, welche für nichteheliche Paare ersetzt wurde durch die Frage „Haben Sie sich jemals wegen einer möglichen Trennung um eine eigene Wohnung bemüht?“. Die eingesetzten Itemformulierungen lauten wie folgt:

1. Haben Sie jemals gedacht, dass die Partnerschaft mit Ihrer Partnerin in Schwierigkeiten ist?
2. Haben Sie jemals ernsthaft an eine Trennung gedacht?
3. Haben Sie jemals mit einem guten Freund oder einer guten Freundin über eine Trennung gesprochen?
4. Haben Sie Ihrer Partnerin jemals ernsthaft eine Trennung vorgeschlagen?

- 5a) für Ehepaare: Haben Sie sich jemals wegen einer möglichen Scheidung von einem Anwalt beraten lassen?
- 5b) für NEL: Haben Sie sich schon einmal um eine eigene Wohnung bemüht für den Fall, dass Sie aus dem gemeinsamen Haushalt ausziehen?

Die Beantwortung der Items erfolgte als dichotome „ja-nein“-Antwort (neben „weiß nicht“ und „keine Angabe“ zur Erfassung des Item-nonresponse). Aufgrund des dichotomen Antwortformats der Indikatoren war eine Berechnung von exploratorischen Faktorenanalysen für diese Skala nicht möglich; in die konfirmatorische Faktorenanalyse (2.6) wurde statt der einzelnen Indikatoren der Summenscore eingegeben, wobei die entsprechende Faktorladung zur Identifikation des Messmodells auf 1 fixiert werden musste (entspricht der Annahme von Messfehlerfreiheit, vgl. Reinecke, 2005). Um dennoch die psychometrische Qualität der Skala überprüfen zu können, wurden verschiedene Item- und Skalenkennwerte berechnet. Hierfür wurden die beiden Varianten des letzten Items in eine gemeinsame Variable codiert. Angesichts des hinter der Skala stehenden Stufenprozesses wurde mittels einer Skalogramm-Analyse (vgl. Bortz & Döring, 2002) geprüft, ob auch empirisch ein hierarchisches Antwortmuster vorliegt. Dabei resultiert ein hoher Anteil (91.1%) Guttman-konformer Antwortmuster, der Reproduzierbarkeitskoeffizient berechnet sich zu $R_G = .98$. Damit kann die Annahme des Stufenprozesses von Trennungen innerhalb der erfassten Indikatoren in der untersuchten Stichprobe als bestätigt angesehen werden.

In Tabelle 16 werden schließlich noch Schwierigkeiten und Trennschärfen der Indikatoren wiedergegeben:

Tabelle 16: Verteilungs- und Itemkennwerte zur Skala subjektive Partnerschaftsstabilität, Gesamtstichprobe

Item	N	p ^a	r _{it}
1. Gedanken, dass Partnerschaft in Schwierigkeiten ist (ps1)	2022	.39	.50
2. Gedanken an eine ernsthafte Trennung (ps2)	2032	.20	.77
3. Mit Freund/in über Trennung gesprochen (ps3)	2036	.16	.68
4. Partner ernsthaft eine Trennung vorgeschlagen (ps4)	2038	.13	.69
5. Scheidungsberatung Anwalt (Ehe) bzw. Wohnungssuche (NEL) (ps5)	2033	.03	.36

^a Die Itemschwierigkeit wurde hier als Anteilswert der ja-Antworten berechnet.

Die Antwortverteilungen der einzelnen Items verweisen noch einmal auf die relativ hohe subjektive Stabilität der untersuchten Partnerschaften. Insbesondere das letzte Item weist mit einer Zustimmung von 3.1% eine sehr schiefe Antwortverteilung auf. Auf Basis der

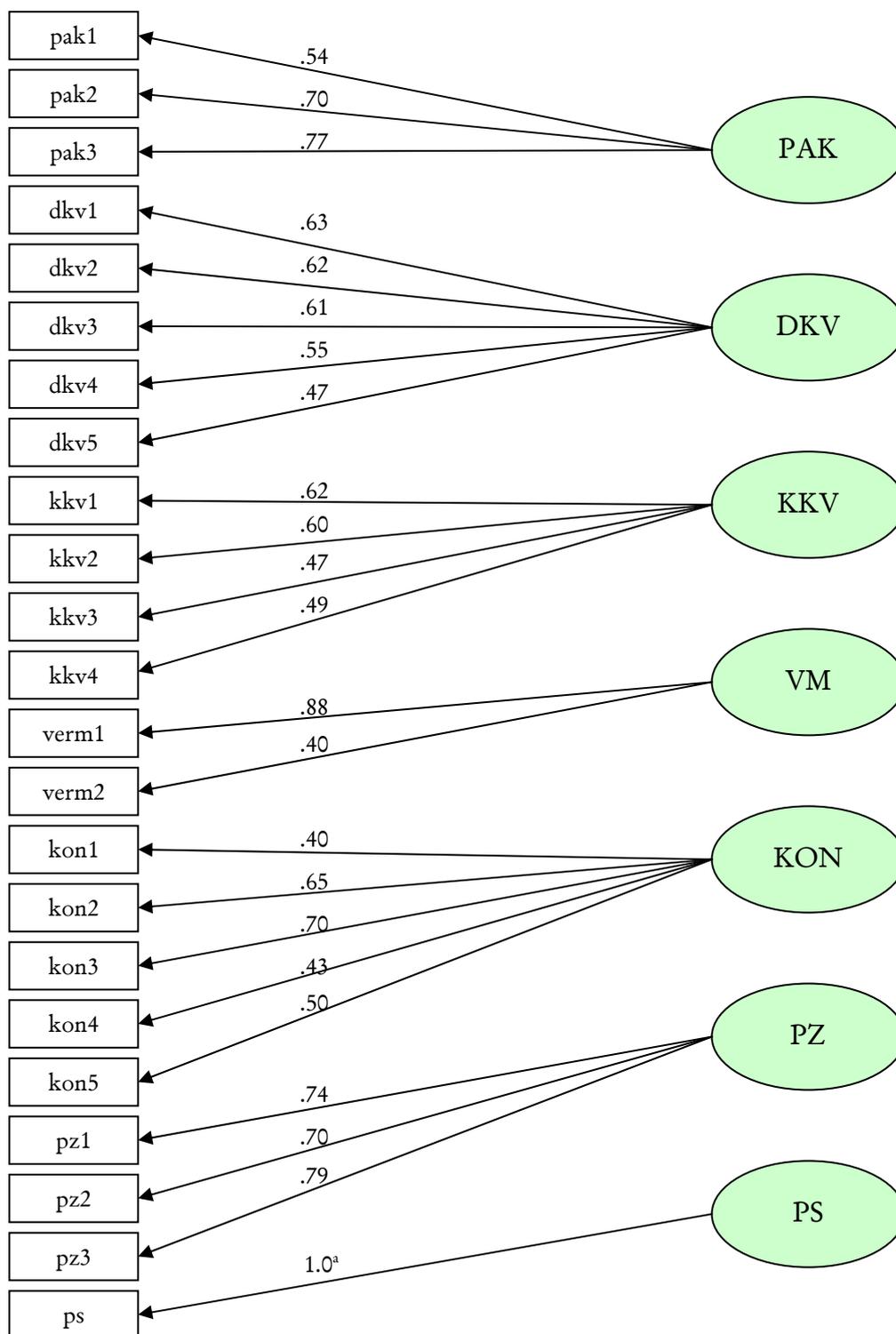
ausreichenden Trennschärfe ist jedoch durchaus ein eigenständiger Beitrag zum Gesamtscore erkennbar; deshalb wurde das Item trotz seiner hohen Schwierigkeit beibehalten. Die interne Konsistenz der Skala, die nach einer speziellen Formel für dichotome Indikatoren berechnet wurde (Kuder & Richardson, 1937), weist mit .80 einen sehr guten Wert auf.

2.6 Das vollständige Messmodell mit allen endogenen latenten Variablen

Im Folgenden wird das konstruktübergreifende Messmodell mit allen endogenen latenten Konstrukten, d.h. ohne die exogenen sozialstrukturellen Variablen, mittels einer konfirmatorischen Faktorenanalyse getestet. Zunächst wurde jedoch ein Test auf Multinormalverteilung (Mardia, 1970, 1974; Mardia & Foster, 1983) durchgeführt, obgleich bereits die univariaten Verteilungen auf deutliche Abweichungen von der Normalverteilung verweisen. Erwartungsgemäß fällt dieser multivariate Test ebenfalls signifikant aus: $\chi^2_{[df=2, N=1850]} = 10481.29, p = .000$. Die im Folgenden dargestellte konfirmatorische Faktorenanalyse wird daher mittels der WLS-Methode geschätzt, die keinerlei Verteilungsannahmen stellt.¹¹ In den letzten Abschnitten wurde bereits erwähnt, welche Indikatoren aufgrund substantieller Nebenladungen aus dem Modell entfernt wurden. Das unreduzierte „Startmodell“, in dem sämtliche aus der ersten Skalenrevision resultierenden Indikatoren eingegeben wurden, wird hier aus Gründen der Übersichtlichkeit nicht wiedergegeben. Nach dem Entfernen der nicht eindeutig zuordenbaren Indikatoren (vgl. hierzu die Abschnitte 2.2 bis 2.4) ergibt sich das auf der nächsten Seite abgebildete Gesamtmessmodell (vgl. Abbildung 2). Der Wortlaut der darin enthaltenen Indikatoren findet sich im Anhang des Buches.

¹¹ Eine zusätzlich durchgeführte konfirmatorische Faktorenanalyse mit einer robusten ML-Schätzung ergibt ähnliche Parameter; allerdings fällt hier – aufgrund der nicht normalverteilten Daten – der Modellfit geringfügig schlechter aus als bei der vorgestellten Lösung.

Abbildung 2: Konfirmatorisches Faktorenmodell (WLS-Schätzung) mit sämtlichen endogenen Konstrukten (N=1850)



^a Der Koeffizient wurde auf 1 fixiert, um das Messmodell identifizierbar zu machen.

Faktorenbezeichnungen: PAK = positive Alltagskommunikation; DKV = destruktives Konfliktverhalten; KKV = konstruktives Konfliktverhalten; VM = Konfliktvermeidung; KON = Paarkonflikte; PZ = Partnerschaftszufriedenheit; PS = Partnerschaftsstabilität

Anmerkungen: Die Itemformulieren der einzelnen Indikatoren sind den Abschnitten 2.2 bis 2.4 zu entnehmen. In der Abbildung sind standardisierte Ladungskoeffizienten angegeben; alle Koeffizienten sind signifikant mit $p < .001$. Interkorrelationen zwischen den Faktoren sind zugelassen, werden aber zur Verbesserung der Übersichtlichkeit nicht abgebildet (vgl. hierzu Tabelle 17).

Bevor die Ladungskoeffizienten interpretiert werden können, ist zunächst der Modellfit zu eruieren, der die Angemessenheit der angelegten Restriktionen (Einfachstruktur) anzeigt. Die χ^2 -Statistik deutet zunächst auf einen schlechten Modellfit hin: $\chi^2_{[df=210, N=1850]}=420.72$, $p < .001$. Allerdings ist hierbei zu berücksichtigen, dass dies auch auf die relativ große Stichprobe sowie auf die hohe Modelkomplexität zurückzuführen sein könnte. Andere Diskrepanzmaße verweisen hingegen auf einen exzellenten (RMSEA=.023, $p=1.0$) bzw. einen akzeptablen (SRMR=.070) Fit. Auch verschiedene Goodness-of-Fit-Indizes zeigen eine gute bis befriedigende Passung des Modells an (GFI=.96, AGFI=.95, CFI=.87). Vor diesem Hintergrund wird das Modell beibehalten.

Die Ladungskoeffizienten sind der Höhe nach substantiell und ausnahmslos signifikant ($p < .001$). In Anbetracht der Tatsache, dass die Modifikationsindizes keine bedeutsamen Nebenladungen nahe legen, kann davon ausgegangen werden, dass die faktorielle Struktur insgesamt akzeptabel ist und keine größere Varianzkonfundierung aufweist. Somit ist die Interpretierbarkeit von Mediatormodellen im Rahmen der Hypothesenprüfung auf Basis der revidierten Skalen als gegeben anzusehen.

Nicht in Abbildung 2 eingezeichnet sind die Interkorrelationen zwischen den latenten Faktoren. Diese sind in Tabelle 17 wiedergegeben:

Tabelle 17: Interkorrelationen der Faktoren (N=1850)

	PAK	DKV	KKV	VERM	KON	PZ
DKV	-0.31**					
KKV	0.48**	-0.52**				
VM	-0.12**	-0.003	-0.07 ⁺			
KON	-0.37**	0.59**	-0.51**	-0.001		
PZ	0.54**	-0.45**	0.59**	-0.14**	-0.59**	
PS	0.15**	-0.39**	0.30**	0.07*	-0.42**	0.41**

⁺ $p < 0.05$

* $p < 0.01$

** $p < 0.001$

Zwischen den Faktoren zeigen sich zumeist mittelstarke Zusammenhänge, die allesamt in die erwartete Richtung zeigen. Bezüglich der Konfliktvermeidung fallen hingegen die eher schwachen Zusammenhänge zu den übrigen Konstrukten auf. Interessant erscheint hierbei,

dass eine negative Korrelation zur Partnerschaftszufriedenheit, hingegen ein positiver (wenn auch sehr schwacher) Zusammenhang zur Partnerschaftsstabilität besteht.

Die in diesem Abschnitt berichtete konfirmatorische Faktorenanalyse zeigt insgesamt, dass sich die theoretisch erwarteten partnerschaftlichen Kommunikationsstile und -dimensionen anhand der eingesetzten Operationalisierungen in den Daten wieder finden lassen und dass sie daneben recht gut von der Konfliktwahrnehmung sowie von der Wahrnehmung des Partnerschaftserfolgs abgegrenzt werden können. Damit ist eine zentrale Voraussetzung für die Anwendung von Pfad- bzw. Strukturgleichungsmodellen auf die vorliegenden Daten gegeben.

Zur Überprüfung der Hypothesen wurden die Skalenscores additiv nach Maßgabe der beschriebenen Skalenanalysen gebildet (vgl. Abbildung 11). Tabelle 18 illustriert die wichtigsten Verteilungskennwerte der verschiedenen Skalen:

Tabelle 18: Deskriptive Statistiken

	N	Mini- mum	Maximum	\bar{x}	s	Schiefe	Kurtosis
positive Alltags- kommunikation (PAK)	2038	1.33	6.00	5.05	.83	-.91**	.79**
konstruktives Konfliktverhalten (KKV)	2037	1.00	6.00	4.83	.99	-.90**	.59**
destruktives Konfliktverhalten (DKV)	2034	1.00	6.00	2.50	1.13	.66**	-.14
Konfliktvermeidung (VER)	2033	1.00	6.00	2.22	1.36	1.09**	.46**
Konfliktscore (KON)	2039	1.00	11.80	2.79	1.68	1.17**	1.45**
Partnerschaftszufrieden- heit (PZ)	2040	1.67	10.00	8.53	1.37	-1.57**	3.20**
Partnerschaftsstabilität (PS)	2007	.00	5.00	4.09	1.36	-1.47**	1.03**

** signifikant mit $p < 0.001$

3 Stichprobe

Da die Zielsetzung der Datenerhebung darin bestand, sowohl sozialstrukturelle Variablen als auch partnerschaftsbezogene Wahrnehmungen simultan zu erfassen, musste die Stichprobe so gezogen werden, dass möglichst ein breites Spektrum unterschiedlicher sozialer Positionen vertreten ist. Die Entscheidung fiel aufgrund der guten Untersuchungsökonomie zu Gunsten von computergestützten Telefoninterviews (CATI); die Telefonnummern

hierfür wurden mittels eines random digit dialing-Verfahrens generiert, was den Vorteil bietet, auch nicht in Telefonverzeichnissen vorhandene Anschlüsse erreichbar zu machen. Die Befragung selbst fand an Einzelpersonen¹² statt, die sich zum Erhebungszeitpunkt in einer ehelichen oder in einer nichtehelichen Lebensgemeinschaft (also einer festen Partnerschaft mit gemeinsamem Haushalt) befanden. Bezüglich früherer Partnerschaften wurden dabei keine weiteren Screeningkriterien eingesetzt, sodass auch naheheliche Partnerschaften und Mehrfachehen in die Stichprobe einbezogen wurden. Um bei den Interviews Verständnisschwierigkeiten und kulturelle Unterschiede innerhalb der Stichprobe zu minimieren, wurde die Befragung zudem nur durchgeführt, wenn beide Partner die deutsche Staatsangehörigkeit hatten und überwiegend in Deutschland (BRD oder DDR) aufgewachsen waren. Ferner gab es eine Altersbeschränkung auf volljährige Personen bis zum Alter von 70 Jahren.

Um differenzierte Analysen in den alten und neuen Bundesländern zu ermöglichen, erfolgte die Stichprobenziehung nach Sozialisation in Ost- bzw. Westdeutschland¹³ disproportional geschichtet, indem eine 50/50%-Vorgabe gestellt wurde. Analog dazu wurde disproportional nach dem Geschlecht (je 50% Männer und Frauen) geschichtet, um mögliche geschlechtsspezifische Ausfallmechanismen zu kompensieren. Die bedingte Verteilung für die beiden genannten Merkmale in der realisierten Nettostichprobe zeigt, dass dennoch letztlich mehr Interviews mit Frauen als mit Männern realisiert werden konnten (vgl. Tabelle 19). Dies ist insbesondere auf einen unterhalb der Vorgaben liegenden Anteil von männlichen Befragten aus Westdeutschland zurückzuführen. Daneben sind die ostdeutschen Befragten gegenüber den westdeutschen in der Überzahl.

Tabelle 19: Stichprobenplan der vorliegenden Untersuchung

Geschlecht	Befragter aufgewachsen in ...				Gesamt	
	Ostdeutschland		Westdeutschland		N	%
männlich	521	25.5	282	13.8	803	39.3
weiblich	643	31.5	595	29.2	1238	60.7
Gesamt	1164	57.0	877	43.0	2041	100.0

¹² Von einer Befragung beider Partner musste abgesehen werden, da sich eine Paarbefragung aufgrund der im Rahmen eines Pretests festgestellten hohen Kosten (Rüssmann et al., 2004) mit den zur Verfügung stehenden Mitteln und im angestrebten Umfang als nicht durchführbar erwiesen hatte.

¹³ Die dazu eingesetzte Frage lautete: „In welchem Teil Deutschlands sind Sie größtenteils aufgewachsen?“ (Antworten: Ostdeutschland, Westdeutschland). Im Ausland aufgewachsene Zielpersonen wurden nicht befragt.

Im Folgenden wird die Analysestichprobe hinsichtlich einiger demographischer Merkmale näher beschrieben.¹⁴ Um dabei gleichzeitig besser einschätzen zu können, ob größere Abweichungen in den entsprechenden Verteilungen von der Grundgesamtheit vorliegen, wird ein Vergleich mit den nahezu zeitgleich erhobenen Daten der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS) 2000 vorgenommen; so beziehen sich beispielsweise auch die regelmäßig erscheinenden Datenberichte des Statistischen Bundesamtes im Bereich Familie auf ALLBUS-Daten (vgl. Statistisches Bundesamt, 2004). Im Folgenden werden Diskrepanzen in zentralen demographischen Merkmalen zwischen den ALLBUS-Daten und dem vorliegenden Datensatz getestet, d.h. die ungewichteten ALLBUS-Daten¹⁵ dienen als Referenz. Die Grundgesamtheit besteht aus in Ost- und Westdeutschland lebenden Männern und Frauen in nichtehelichen Lebensgemeinschaften oder Ehen im Alter von 18 bis 70 Jahren. Alle älteren Personen, die im ALLBUS ebenfalls befragt werden, wurden aus der Analyse entfernt, um die Daten vergleichbar zu machen. Zur Bildung der Schichten wurde aus Gründen der Vergleichbarkeit nicht die oben erwähnte Frage nach dem Ort des Aufwachsens benutzt, sondern die ebenfalls in den Daten enthaltenen Informationen zum aktuellen Wohnort (alte vs. neue Bundesländer). Dass die Verteilung dieser beiden Variablen in ca. 5% der Fälle voneinander abweicht – d.h. Migration stattgefunden hat –, zeigt Tabelle 20:

Tabelle 20: Vergleich des Ortes der Sozialisation mit dem aktuellen Wohnort

			Wohnort				Gesamt	
			Westdeutschland		Ostdeutschland		N	%
			N	%	N	%		
Befragter aufgewach- sen in ...	Ostdeutschland		50	2.7	1048	56.1	1098	58.8
	Westdeutschland		726	38.9	44	2.4	770	41.2
Gesamt			776	41.5	1092	58.5	1868	100.0

¹⁴ Weitere detaillierte Dekriptionen der Stichprobe finden sich bei Hill (2004, im Anhang) bzw. bei Rüssmann (2004).

¹⁵ Da die Stichprobenziehung beim ALLBUS 2000 auf Melderegisterdaten basierte, haben alle Personen dieselbe Auswahlwahrscheinlichkeit; daher sind die Daten auf Personenebene selbstgewichtend, d.h. es ist keine Desinggewichtung erforderlich. Bezüglich der alten und neuen Bundesländer erfolgte die Stichprobenziehung allerdings disproportional geschichtet, d.h. es wurde ein oversampling von Befragten in den neuen Bundesländern vorgenommen, um für Ostdeutschland differenzierte Auswertungen zu ermöglichen. Bei einer übergreifenden Deskription über die gesamtdeutsche Stichprobe wären die Daten daher mit einem Ost-West-Gewicht zu gewichten; bei den im Folgenden dargestellten separaten Analysen für West- und Ostdeutschland ist dies hingegen nicht erforderlich (vgl. Koch, Wasmer, Harkness, & Scholz, 2001).

Verglichen werden die beiden Stichproben hinsichtlich der Verteilungen des Alters, des Bildungsniveaus, des Einkommens, der Partnerschafts- bzw. Ehedauer, der Kinderzahl sowie der Wohnortgröße. Bezüglich des Alters der Befragten zeigt Tabelle 21 die folgenden Verteilungen:

Tabelle 21: Vergleich der Altersverteilung mit dem ALLBUS 2000

Geschlecht	Gesamt Alter Befragter (Kategorien)	Westdeutschland vorliegende Studie				Ostdeutschland vorliegende Studie			
		N	%	N	%	N	%	N	%
männlich	18 bis 29 J.	13	5.3	99	11.9	54	11.1	47	10.8
	30 bis 39 J.	63	25.7	189	22.7	92	19.0	97	22.3
	40 bis 49 J.	67	27.3	206	24.7	101	20.8	105	24.1
	50 bis 59 J.	98	40.0	175	21.0	223	46.0	83	19.1
	60 bis 70 J.	4	1.6	164	19.7	15	3.1	103	23.7
	Gesamt	245	100.0	833	100.0	485	100.0	435	100.0
weiblich	18 bis 29 J.	35	6.6	139	16.2	74	12.2	51	11.9
	30 bis 39 J.	145	27.3	216	25.2	139	22.9	104	24.4
	40 bis 49 J.	168	31.6	197	23.0	168	27.7	108	25.3
	50 bis 59 J.	175	33.0	170	19.9	219	36.1	85	19.9
	60 bis 70 J.	8	1.5	134	15.7	7	1.2	79	18.5
	Gesamt	531	100.0	856	100.0	607	100.0	427	100.0

Die hier analysierten Daten zeigen im Vergleich zum ALLBUS-Referenzdatensatz eine Unterrepräsentation der jüngsten Gruppe (nur Westdeutschland) sowie der ältesten Befragten und eine deutliche Übererfassung der 50- bis 59-Jährigen. Die Abweichung zwischen den Verteilungen ist signifikant für beide Geschlechter in Westdeutschland (Männer: $\chi^2_{[df=4, N=1078]}=75.70$, $p=.000$; Frauen: $\chi^2_{[df=4, N=1378]}=120.78$, $p=.000$) und Ostdeutschland (Männer: $\chi^2_{[df=4, N=920]}=128.04$, $p=.000$; Frauen: $\chi^2_{[df=4, N=1034]}=113.78$, $p=.000$). Demnach kann bezüglich des Alters nicht von der Repräsentativität der untersuchten Stichprobe für die Untersuchungspopulation ausgegangen werden. Immerhin zeigen Mittelwertvergleiche über das metrisch codierte Alter der Befragten nur geringe Unterschiede zwischen beiden Studien (vgl. Tabelle 22):

Tabelle 22: Mittelwertunterschiede im Alter der Befragten zwischen der vorliegenden Studie und dem ALLBUS 2000

Ost/West	Geschlecht	Studie	N	\bar{x}	s	T	df	Signifikanz
Westdeutschland	männlich	vorl. Studie	245	46.95	11.94	1.20	430.46	.230
		ALLBUS	833	45.88	13.06			
	weiblich	vorl. Studie	531	45.28	11.30	2.26	1248.72	.024
		ALLBUS	856	43.78	13.15			
Ostdeutschland	männlich	vorl. Studie	485	48.27	14.10	1.92	915.21	.055
		ALLBUS	435	46.53	13.36			
	weiblich	vorl. Studie	607	45.34	13.15	.01	1032	.990
		ALLBUS	427	45.33	12.88			

Als nächstes wird die Verteilung des Bildungsindikators „höchster Schulabschluss“ verglichen (vgl. Tabelle 23):

Tabelle 23: Vergleich der Verteilung des Schulbildungsniveaus mit dem ALLBUS 2000

Geschlecht	höchster Schulabschluss Befragter	Westdeutschland				Ostdeutschland			
		vorliegende Studie		ALLBUS		vorliegende Studie		ALLBUS	
		N	%	N	%	N	%	N	%
männlich	kein Abschluss	2	.8	16	1.9	4	.8	3	.7
	Volks- / Hauptschulabschluss	80	32.9	396	47.9	139	29.1	120	27.7
	Mittlere Reife, Realschulabschluss	82	33.7	175	21.2	220	46.0	209	48.3
	Abitur, Allgemeine bzw. Fachhochschulreife	79	32.5	239	28.9	115	24.1	101	23.3
	Gesamt		243	100.0	826	100.0	478	100.0	433
weiblich	kein Abschluss	2	.4	19	2.2	6	1.0	2	.5
	Volks- / Hauptschulabschluss	177	33.5	386	45.4	184	30.5	119	27.9
	Mittlere Reife, Realschulabschluss	238	45.1	256	30.1	305	50.6	216	50.6
	Abitur, Allgemeine bzw. Fachhochschulreife	111	21.0	189	22.2	108	17.9	90	21.1
	Gesamt		528	100.0	850	100.0	603	100.0	427

Wie Tabelle 23 zeigt, sind in der westdeutschen Teilstichprobe Volks- und Hauptschulabschluss unterrepräsentiert. Die Abweichungen vom ALLBUS 2000 sind hier insgesamt signifikant (Männer: $\chi^2_{[df=3, N=1069]}=24.02, p=.000$; Frauen: $\chi^2_{[df=3, N=1378]}=39.18, p=.000$). In der

ostdeutschen Teilstichprobe hingegen sind nur unwesentliche Abweichungen von der Referenzverteilung zu konstatieren (Männer: $\chi^2_{[df=3, N=911]}=0.51$, $p=.918$; Frauen: $\chi^2_{[df=3, N=1030]}=2.79$, $p=.425$).¹⁶ Repräsentativität der Stichprobe bezüglich des Bildungsniveaus kann demnach nur für Ostdeutschland angenommen werden.

In Tabelle 24 sind die Verteilungen des Haushaltsnettoeinkommens¹⁷ abgebildet:

Tabelle 24: Vergleich der Verteilung des Haushaltsnettoeinkommens mit dem ALLBUS 2000

Gesamt		Westdeutschland				Ostdeutschland			
Geschlecht	Haushaltsnettoeinkommen (Klassen)	vorliegende Studie		ALLBUS		vorliegende Studie		ALLBUS	
		N	%	N	%	N	%	N	%
männlich	< 1000 €	0	.0	29	4.4	20	5.3	26	7.7
	1000-1499 €	14	6.9	85	13.0	80	21.1	66	19.6
	1500-1999 €	21	10.4	130	19.9	88	23.2	108	32.1
	2000-2499 €	30	14.9	125	19.2	78	20.6	65	19.3
	2500-2999 €	46	22.8	96	14.7	35	9.2	43	12.8
	3000-3999 €	40	19.8	105	16.1	52	13.7	24	7.1
	4000 € und mehr	51	25.2	82	12.6	26	6.9	4	1.2
Gesamt		202	100.0	652	100.0	379	100.0	336	100.0
weiblich	< 1000 €	7	1.9	60	9.4	29	6.4	13	3.8
	1000-1499 €	16	4.3	90	14.1	103	22.9	73	21.6
	1500-1999 €	58	15.6	133	20.8	104	23.1	103	30.5
	2000-2499 €	100	26.9	124	19.4	91	20.2	73	21.6
	2500-2999 €	51	13.7	84	13.1	44	9.8	49	14.5
	3000-3999 €	82	22.0	96	15.0	53	11.8	19	5.6
	4000 € und mehr	58	15.6	53	8.3	26	5.8	8	2.4
Gesamt		372	100.0	640	100.0	450	100.0	338	100.0

Die Gegenüberstellung beider Datensätze verweist auf eine Übererfassung der obersten und eine tendenzielle Untererfassung der unteren Einkommensgruppen. Zwar weisen die Verteilungen insgesamt ein ähnliches Muster auf, dennoch sind die Abweichungen in Westdeutschland (Männer: $\chi^2_{[df=6, N=854]}=46.62$, $p=.000$; Frauen: $\chi^2_{[df=6, N=1012]}=68.86$, $p=.000$) und

¹⁶ Die Zuordnung zu einheitlichen Bildungsniveaus für alte und neue Bundesländer ist wegen der historischen Unterschiede in den Bildungssystemen allerdings nicht einfach. Im vorliegenden Datensatz wurden „Abschluss Polytechnische Oberschule“ sowie „Abschluss Fachoberschule“ zur Kategorie „Mittlere Reife“ (Tabelle 23) hinzugefügt, da diese im ALLBUS nicht differenziert wurden. Außerdem wurden die Kategorien „noch Schüler“ sowie „anderer Abschluss“ wegen zu geringer Fallzahlen aus der Analyse ausgeschlossen.

¹⁷ Das individuelle Einkommen konnte nicht verglichen werden, da im Rahmen der vorliegenden Untersuchung ausschließlich Daten zum Haushaltsnettoeinkommen verfügbar sind. Da in Westdeutschland nur wenige Haushalte (insgesamt $N=7$) in den drei Klassen unter 1000 Euro pro Monat vertreten waren, wurden die entsprechenden Kategorien aggregiert.

in Ostdeutschland (Männer: $\chi^2_{[df=6, N=715]}=30.14$, $p=.000$; Frauen: $\chi^2_{[df=6, N=788]}=23.60$, $p=.000$) signifikant. Hinsichtlich des Haushaltseinkommens muss also davon ausgegangen werden, dass die untersuchte Stichprobe nicht repräsentativ ist.

Als nächstes wird untersucht, wie sich die Wohnortgröße verteilt (vgl. Tabelle 25):

Tabelle 25: Vergleich der Verteilung der Wohnortgröße mit dem ALLBUS 2000

Geschlecht	Einwohner im Wohnort	vorliegende Studie		ALLBUS		vorliegende Studie		ALLBUS	
		N	%	N	%	N	%	N	%
männlich	bis 2.000	46	18.9	38	4.6	101	21.7	79	18.2
	2.000 - 4.000	27	11.1	65	7.8	51	11.0	77	17.7
	5.000 - 20.000	67	27.6	255	30.6	91	19.6	81	18.6
	20.000 - 50.000	30	12.3	163	19.6	59	12.7	62	14.3
	50.000 - 100.000	19	7.8	76	9.1	24	5.2	29	6.7
	100.000 - 500.000	32	13.2	128	15.4	72	15.5	79	18.2
	500.000 und mehr	22	9.1	108	13.0	67	14.4	28	6.4
	Gesamt	243	100.0	833	100.0	465	100.0	435	100.0
weiblich	bis 2.000	98	19.4	31	3.6	134	23.8	75	17.6
	2.000 - 4.000	81	16.0	61	7.1	64	11.4	73	17.1
	5.000 - 20.000	121	24.0	291	34.0	131	23.3	72	16.9
	20.000 - 50.000	70	13.9	141	16.5	77	13.7	56	13.1
	50.000 - 100.000	42	8.3	72	8.4	33	5.9	31	7.3
	100.000 - 500.000	56	11.1	154	18.0	68	12.1	85	19.9
	500.000 und mehr	37	7.3	106	12.4	55	9.8	35	8.2
	Gesamt	505	100.0	856	100.0	562	100.0	427	100.0

Tabelle 25 illustriert, dass in der untersuchten westdeutschen Stichprobe Bewohner kleinerer Ortschaften mit unter 4000 Einwohnern deutlich überrepräsentiert sind, während solche in stärker urbanem Wohnumfeld tendenziell eher unterrepräsentiert sind. Diese Abweichungen sind insgesamt statistisch signifikant (Männer: $\chi^2_{[df=6, N=1076]}=61.56$, $p=.000$; Frauen: $\chi^2_{[df=6, N=1361]}=137.18$, $p=.000$). Für Ostdeutschland resultieren schwächere, aber ebenfalls bedeutsame Abweichungen gegenüber dem ALLBUS (Männer: $\chi^2_{[df=6, N=900]}=24.46$, $p=.000$; Frauen: $\chi^2_{[df=6, N=989]}=26.17$, $p=.000$), auch wenn hier kein deutlicher linearer Zusammenhang zu identifizieren ist. Insgesamt muss vor diesem Hintergrund die Annahme der Repräsentativität bezüglich der Wohnortgröße verworfen werden.

Im Folgenden werden zwei partnerschaftsbezogene Merkmale¹⁸ dargestellt: Partnerschaftsdauer und Kinderzahl. Bezüglich der Partnerschaftsdauer ist anzumerken, dass im ALLBUS

¹⁸ Die getrennte Durchführung der Analysen nach Geschlecht wird dabei aus den eingangs aufgeführten Gründen beibehalten, obwohl bei einer Zufallsauswahl keine Unterschiede in den Angaben zwischen Männern und Frauen zu erwarten sind, da es sich um ein Beziehungsmerkmal handelt.

2000 lediglich das Datum der Hochzeit erfragt wird, nicht jedoch des Beginns einer nichtehelichen Partnerschaft bzw. Lebensgemeinschaft. Daher muss sich der Vergleich der Partnerschaftsdauer auf die Ehepaare in beiden Stichproben beschränken (vgl. Tabelle 26):

Tabelle 26: Vergleich der Verteilung der Ehedauer mit dem ALLBUS 2000

Geschlecht	Ehedauer (kategorisiert)	Westdeutschland				Ostdeutschland			
		vorliegende Studie		ALLBUS		vorliegende Studie		ALLBUS	
		N	%	N	%	N	%	N	%
männlich	bis 2 J.	2	.9	31	4.7	5	1.4	10	3.0
	über 2 bis 5 J.	9	4.2	44	6.6	14	3.8	20	6.0
	über 5 bis 10 J.	32	15.0	87	13.1	30	8.2	21	6.3
	über 10 bis 20 J.	64	30.0	173	26.1	74	20.2	92	27.4
	über 20 J.	106	49.8	328	49.5	244	66.5	193	57.4
	Gesamt	213	100.0	663	100.0	367	100.0	336	100.0
weiblich	bis 2 J.	7	1.5	39	5.9	9	1.8	10	2.8
	über 2 bis 5 J.	13	2.8	58	8.7	18	3.7	11	3.1
	über 5 bis 10 J.	50	10.6	77	11.6	41	8.4	32	9.1
	über 10 bis 20 J.	152	32.3	154	23.1	121	24.8	101	28.8
	über 20 J.	249	52.9	338	50.8	299	61.3	197	56.1
	Gesamt	471	100.0	666	100.0	488	100.0	351	100.0

Ogleich die Verteilungen sich ähneln, sind in der untersuchten Stichprobe in West- und Ostdeutschland tendenziell Personen in kurzen Ehen unter-, solche in Langzeitehen (über 20 Jahre) hingegen überrepräsentiert. In Westdeutschland sind die Unterschiede der Verteilungen nur bei den Frauen signifikant ($\chi^2_{[df=4, N=1137]}=37.70, p=.000$; Männer: $\chi^2_{[df=5, N=876]}=8.89, p=.064$), in Ostdeutschland hingegen weder für Männer ($\chi^2_{[df=5, N=703]}=10.87, p=.028$) noch für Frauen ($\chi^2_{[df=5, N=839]}=3.35, p=.501$). Zusätzlich durchgeführte Mittelwertvergleiche (vgl. Tabelle 27) verweisen darauf, dass die mittlere Ehedauer in der vorliegenden Studie in West- und Ostdeutschland im Verhältnis zum ALLBUS tendenziell überschätzt wird; allerdings sind die Unterschiede nur bei den westdeutschen Frauen und den ostdeutschen Männern signifikant.

Tabelle 27: Mittelwertunterschiede in der Ehedauer der Befragten zwischen der vorliegenden Studie und dem ALLBUS 2000

Ost/West	Geschlecht	Studie	N	\bar{x}	s	T	df	Signifikanz
Westdeutschland	männlich	vorl. Studie	213	22.24	11.92	1.42	379.62	.158
		ALLBUS	664	20.89	12.74			
	weiblich	vorl. Studie	471	22.87	11.53	2.31	1094.28	
		ALLBUS	666	21.15	13.46			
Ostdeutschland	männlich	vorl. Studie	367	26.75	12.94	3.45	702	.001
		ALLBUS	337	23.44	12.52			
	weiblich	vorl. Studie	486	24.99	12.47	1.39	835	
		ALLBUS	351	23.77	12.62			

Der letzte Vergleich bezieht sich auf die Verteilung der Kinderzahl (vgl. Tabelle 28):

Tabelle 28: Vergleich der Verteilung der Kinderzahl mit dem ALLBUS 2000

Geschlecht	Kinderzahl	Westdeutschland				Ostdeutschland			
		vorliegende Studie		ALLBUS		vorliegende Studie		ALLBUS	
		N	%	N	%	N	%	N	%
männlich	kinderlos	45	18.5	179	21.5	84	17.4	69	15.9
	1	33	13.6	187	22.4	134	27.8	126	29.0
	2	111	45.7	327	39.3	172	35.7	171	39.4
	3	38	15.6	99	11.9	67	13.9	51	11.8
	4 und mehr	16	6.6	41	4.9	25	5.2	17	3.9
	Gesamt		243	100.0	833	100.0	482	100.0	434
weiblich	kinderlos	79	14.9	198	23.2	98	16.2	58	13.6
	1	107	20.2	174	20.4	175	28.9	125	29.3
	2	241	45.6	298	35.0	243	40.2	167	39.1
	3	72	13.6	118	13.8	63	10.4	60	14.1
	4 und mehr	30	5.7	64	7.5	26	4.3	17	4.0
	Gesamt		529	100.0	852	100.0	605	100.0	427

Der Vergleich der Kinderzahl zeigt für Westdeutschland eine leichte Untererfassung des Anteils Kinderloser sowie eine Übererfassung des Anteils von Zwei-Kind-Familien (jeweils nur auf Basis der Angaben der Frauen). Diese Abweichungen bei den westdeutschen Frauen sind signifikant (Frauen: $\chi^2_{[df=4, N=1381]}=22.23$, $p=.000$; Männer: $\chi^2_{[df=4, N=1076]}=13.00$, $p=.011$). In Ostdeutschland finden sich nur unbedeutende Abweichungen vom ALLBUS (Männer: $\chi^2_{[df=4, N=916]}=2.91$, $p=.574$; Frauen: $\chi^2_{[df=4, N=1032]}=4.05$, $p=.399$). Mittelwertvergleiche deuten darauf hin, dass die Kinderzahl ausschließlich bei den westdeutschen Männern überschätzt wird (vgl. Tabelle 29):

Tabelle 29: Mittelwertunterschiede in der Kinderzahl der Befragten zwischen der vorliegenden Studie und dem ALLBUS 2000

Ost/West	Geschlecht	Studie	N	\bar{x}	s	T	df	Signifikanz
Westdeutschland	männlich	vorl. Studie	245	1.84	1.24	2.79	1076	.005
		ALLBUS	833	1.59	1.20			
	weiblich	vorl. Studie	531	1.80	1.20	1.93	1184.94	.054
		ALLBUS	856	1.66	1.29			
Ostdeutschland	männlich	vorl. Studie	485	1.81	3.62	1.18	575.63	.237
		ALLBUS	435	1.61	1.06			
	weiblich	vorl. Studie	607	1.61	1.10	-.96	1032	.337
		ALLBUS	245	1.84	1.24			

Alles in allem zeigen die dargestellten Vergleiche mit den ALLBUS-Referenzdaten, dass die untersuchte Stichprobe hinsichtlich einiger demographischer Merkmale möglicherweise von der zu untersuchenden Grundgesamtheit abweicht, d.h. nicht „repräsentativ“ ist. Nach Schnell (1993: 16) bezeichnet Repräsentativität das Ausmaß der „Übereinstimmung von Stichprobenhäufigkeiten mit den entsprechenden bekannten Anteilen in der Grundgesamtheit“. Dennoch ist auf Basis der folgenden Überlegungen davon auszugehen, dass die gefundenen Zusammenhänge dadurch nicht wesentlich verzerrt sind:

1. Der Begriff der Repräsentativität ist im sozialwissenschaftlichen Kontext nicht zu streng auszulegen, da in entsprechenden Umfragen aufgrund der fehlenden Auskunftspflicht systematisch fehlende Angaben aufgrund von Verweigerung ohnehin nicht auszuschließen sind. Daneben werden meist bestimmte Bevölkerungsgruppen schwerer erreichbar oder werden bewusst ausgeschlossen (z.B. Ausländer), sodass Repräsentativität streng genommen nur in Bezug auf die Population der Auskunftsfähigen, -willigen und Erreichbaren geprüft werden kann (Hartmann & Schimpl-Neimanns 1992: 316). Diese Problematik trifft im übrigen auch auf die ALLBUS-Daten selbst zu (Hartmann 1990; Hartmann & Schimpl-Neimanns 1992). Insofern sind Abweichungen vom ALLBUS nicht gleichbedeutend mit Abweichungen von der Grundgesamtheit.
2. Zudem hängt das Ergebnis des Tests auf Verteilungsabweichungen partiell von der gewählten Anzahl von Merkmalskategorien ab: Werden nur wenige Kategorien (z.B. Einkommensklassen) unterschieden, können Verteilungsunterschiede *innerhalb* dieser Klassen dadurch verborgen bleiben. In den oben wiedergegebenen Analysen wurden relativ fein differenzierte Verteilungen getestet, sodass die Abweichungen zum Teil als Resultat des

„strengen“ Tests angesehen werden können. Es erscheint unplausibel, dass dadurch Schätzungen von Zusammenhängen in bedeutendem Ausmaß verzerrt werden.

3. Schließlich ist zu beachten, dass die χ^2 -Statistik bei großen Stichproben auch im Fall geringer Verteilungsunterschiede eher signifikant wird und daher zu konservativen Testentscheidungen führt. Bei der gegebenen Stichprobengröße ($N > 2000$) sollten einzelne signifikante Abweichungen aus diesem Grund nicht überbewertet werden.

Ein weiterer Punkt betrifft die disproportionale Schichtung der vorliegenden Stichprobe nach Sozialisation in Ost- / Westdeutschland und Geschlecht, durch welche die Verteilung dieser Merkmale prinzipiell von den Verteilungen in der Grundgesamtheit abweichen. In solchen Fällen muss entweder (a) eine entsprechende Designgewichtung eingesetzt werden, welche die ungleichen Auswahlwahrscheinlichkeiten ausgleicht, oder aber (b) der Vergleich erfolgt innerhalb der vier durch die beiden dichotomen Variablen gebildeten Subgruppen separat. In der vorliegenden Arbeit wurde aus den folgenden Gründen der zweite Ansatz gewählt:

- a) Bei den Analysen zur Überprüfung der Forschungshypothesen wird auf eine Gewichtung verzichtet. Dies begründet sich zum einen dadurch, dass kein Felddatensatz mit Screeningdaten verfügbar war, mittels dessen sich die entsprechenden Gewichte hätten bestimmen lassen (vgl. zum Vorgehen hierbei Hartmann, 1997). Noch schwerer wiegen jedoch inhaltliche Bedenken gegen Gewichtung, sofern diese nicht für eine Deskription von Merkmalsverteilungen, sondern für Zusammenhangsanalysen eingesetzt wird: Eine (zweifelhafte) Grundannahme der Gewichtung besteht nämlich darin, dass eine nicht repräsentative (d.h. abweichende Häufigkeitsverteilungen aufweisende) Stichprobe durch die Gewichtung nach einigen wenigen Merkmalen mit bekannter Populationsverteilung in *allen* Merkmalen mit der Grundgesamtheit vergleichbar gemacht werden kann. Gerade in der vorliegenden Arbeit werden diverse Merkmale untersucht, für die keine Populationsdaten aus amtlichen Statistiken vorliegen, z.B. Paarinteraktion, Konflikte etc. Durch eine Gewichtung nach sozialstrukturellen Eckparametern, so die Annahme, kann Repräsentativität bezüglich der kombinierten Randverteilungen dieser Konstrukte erreicht werden. Schnell (1993) kann jedoch zeigen, dass sich zentrale demographische Merkmale nicht aus typischen Umfragedaten (z.B. Einstellungen) vorhersagen lassen, sodass dieser „Repräsentanzbeweis“ unzulässig ist. Darüber hinaus demonstriert Schnell

anhand einer Simulation, dass eine Gewichtung je nach Datenlage zu Unter- oder Überschätzung von Parametern führen kann, die verzerrte Hypothesentests begünstigt. Hinzu kommt, dass in der vorliegenden Studie die Schichtvariablen nicht das Explanandum darstellen; es liegt daher eine so genannte exogene Schichtung vor, bei der eine Gewichtung im Allgemeinen obsolet ist (vgl. Hartmann, 1997; Hartmann & Schimpl-Neimanns, 1992).

- b) In den Analysen zur Hypothesenprüfung werden alle Modelle im ersten Schritt für ost- und westdeutsche Frauen bzw. Männer *getrennt* berechnet. Diese werden aus Platzgründen jedoch lediglich dann dargestellt, wenn sich deutliche Ost-West-Unterschiede in den Analysen finden. Ist letzteres nicht der Fall, werden aggregierte Modelle wiedergegeben, in denen die Schichtungsvariablen Geschlecht bzw. Ort der Sozialisation kontrolliert werden.

Aus den in diesem Abschnitt beschriebenen deskriptiven Analysen lässt sich eine allgemeine Schlussfolgerung ziehen: Es konnte in den deskriptiven Auswertungen gezeigt werden, dass die Untersuchungstichprobe bezüglich zentraler demographischer bzw. sozialstruktureller Merkmale genügend Varianz aufweist, um systematische Auswirkungen von Variationen in den entsprechenden Variablen auf Paarinteraktion und Partnerschaftserfolg analysieren zu können. Damit ist eine wichtige Voraussetzung für die Tragfähigkeit der im Ergebnisteil dargestellten Analysen gegeben.

4 Literatur

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, London, New Delhi: Sage.
- Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2005). *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung* (11., überarbeitete ed.). Berlin, Heidelberg, New York: Springer.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182.
- Bollen, K. A. (1987). Total, direct, and indirect effects in structural equation models. In C. C. Clogg (Ed.), *Sociological methodology* (pp. 37-69).
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York, Chichester, Brisbane, Toronto, Singapur: John Wiley & Sons.
- Booth, A., Johnson, D. R., & Edwards, J. N. (1983). Measuring marital instability. *Journal of Marriage and the Family*, 45, 387-394.
- Booth, A., Johnson, D. R., White, L. K., & Edwards, J. N. (1985). Predicting divorce and permanent separation. *Journal of Family Issues*, 6(3), 331-346.
- Booth, A., Johnson, D. R., White, L. K., & Edwards, J. N. (1986). Divorce and marital instability over the life course. *Journal of Family Issues*, 7(4), 421-442.
- Booth, A., & White, L. K. (1980). Thinking about divorce. *Journal of Marriage and the Family*, 3, 605-616.
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler*. Berlin: Springer.
- Bortz, J., & Döring, N. (2002). *Forschungsmethoden und Evaluation*. Berlin: Springer.
- Efron, B. (2000). The bootstrap and modern statistics. *Journal of the American Statistical Association*, 95(452), 1293-1296.
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.
- Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement* (3. ed.). New York, London: Sage.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford: Stanford University Press.
- Fisseni, H.-J. (1997). *Lehrbuch der psychologischen Diagnostik*. Göttingen: Hogrefe.
- Fox, J. (1997). *Applied regression analysis, linear models, and related methods*. Thousand Oaks, London, New Delhi: Sage Publications.
- Frazier, P. A., Tix, A. P., & Barron, K. E. (2004). Testing moderator and mediator effects in counseling psychology research. *Journal of Counseling Psychology*, 51(1), 115-134.
- Hartmann, J. (1997). Komplexes Stichprobendesign und Ereignisanalyse. Zur Notwendigkeit einer Gewichtung bei disproportional geschichteter Stichprobenziehung. *Mannheimer Zentrum für Europäische Sozialforschung (MZES) Working Paper*(17), 1-46.
- Hartmann, J., & Simon, R. (1997). Zur Messung der Ehequalität: Theoretische Überlegungen, empirische Ergebnisse und praktische Folgerungen. In J. Kopp (Ed.), *Methodische Probleme der Familienforschung: Zu den praktischen Schwierigkeiten bei der Durchführung einer empirischen Untersuchung* (pp. 147-177). Frankfurt, New York: Campus Verlag.
- Hartmann, P. H., & Schimpl-Neimanns, B. (1992). Sind Sozialstrukturanalysen mit Umfragedaten möglich? Analysen zur Repräsentativität einer Sozialforschungsumfrage. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 44(2), 315-340.

- Hassebrauck, M. (1991). ZIP - ein Verfahren zur Erfassung der Zufriedenheit in Paarbeziehungen. *Zeitschrift für Sozialpsychologie*, 22, 256-259.
- Hendrick, S. S. (1988). A generic measure of relationship satisfaction. *Journal of Marriage and the Family*, 50, 93-98.
- Hill, P. B. (Ed.). (2004). *Interaktion und Kommunikation. Eine empirische Studie zu Alltagsinteraktionen, Konflikten und Zufriedenheit in Partnerschaften*. Würzburg: Ergon.
- Hoyle, R. H. (1995). The structural equation modeling approach: Basic concepts and fundamental Issues. In R. H. Hoyle (Ed.), *Structural equation modeling: concepts, issues, and applications* (pp. 1-15). Thousand Oaks: Sage.
- Hu, L.-T., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: concepts, issues, and applications* (pp. 76-99). Thousand Oaks: Sage.
- Koch, A., Wasmer, M., Harkness, J., & Scholz, E. (2001). *Konzeption und Durchführung der "Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften" (ALLBUS) 2000*. Mannheim: Zentrum für Umfragen, Methoden und Analysen (ZUMA).
- Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2, 151-160.
- MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science*, 1(4), 173-181.
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57(3), 519-530.
- Mardia, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. *Sankhya: The Indian Journal of Statistics*, 36(2), 115-128.
- Mardia, K. V., & Foster, K. (1983). Omnibus tests of multinormality based on skewness and kurtosis. *Communication in Statistics*, 12, 207-221.
- McClelland, G. H., & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin*, 114(2), 376-390.
- McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2. ed.). Boca Raton u.a.: Chapman & Hall.
- Mühlfeld, C. (1976). *Familiensoziologie. Eine systematische Einführung*. Hamburg: Hoffmann und Campe.
- Mummendey, H. D. (1995). *Die Fragebogenmethode* (2., überarbeitete ed.). Göttingen: Hogrefe Verlag.
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, 36(4), 717-731.
- Reinecke, J. (2005). *Strukturgleichungsmodelle in den Sozialwissenschaften*. München: Oldenbourg.
- Rietz, C., Rudinger, G., & Andres, J. (1996). Lineare Strukturgleichungsmodelle. In E. Erdfelder & R. Mausfeld & T. Meiser & G. Rudinger (Eds.), *Handbuch Quantitative Methoden* (pp. 253-268). Weinheim: Beltz.
- Rüssmann, K. (2004). Sozialstruktur und Konfliktpotential. In P. B. Hill (Ed.), *Interaktion und Kommunikation. Eine empirische Studie zu Alltagsinteraktionen, Konflikten und Zufriedenheit in Partnerschaften* (pp. 105-158). Würzburg: Ergon.
- Rüssmann, K. (2006). *Sozialstruktur und Konflikte in Partnerschaften. Eine empirische Studie zur Auswirkung von sozial- und familienstrukturellen Merkmalen auf partnerschaftliche Konflikte*. Hamburg: Verlag Dr. Kovac.
- Rüssmann, K., Arránz Becker, O., & Kelzenberg, D. (2004). Konzepte und Skalen zur Messung von Beziehungserfolg - Zur Entwicklung einschlägiger Instrumente. In P.

- B. Hill (Ed.), *Interaktion und Kommunikation. Eine empirische Studie zu Alltagsinteraktionen, Konflikten und Zufriedenheit in Partnerschaften* (pp. 105-158). Würzburg: Ergon.
- Sander, J., & Böcker, S. (1993). Die Deutsche Form der Relationship Assessment Scale (RAS): Eine kurze Skala zur Messung der Zufriedenheit in einer Partnerschaft. *Diagnostica*, 39(1), 55-62.
- Satorra, A., & Bentler, P. M. (1988). Scaling corrections for chi-square statistics in covariance structure analysis. *Proceedings of the Business and Economic Statistics Section of the American Statistical Association*.
- Schnell, R. (1993). Die Homogenität sozialer Kategorien als Voraussetzung für "Repräsentativität" und Gewichtungungsverfahren. *Zeitschrift für Soziologie*, 22(1), 16-32.
- Schnell, R., Hill, P. B., & Esser, E. (2005). *Methoden der empirischen Sozialforschung* (7., völlig überarbeitete und erweiterte ed.). München, Wien: Oldenbourg.
- Schumm, W. R., Paff-Bergen, L. A., Hatch, R. C., Obiorah, F. C., Copeland, J. M., Meens, L. D., & Bugaighis, M. A. (1986). Concurrent and discriminant validity of the Kansas Marital Satisfaction Scale. *Journal of Marriage and the Family*, 48, 381-387.
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, 7(4), 422-445.
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological methodology* (pp. 290-312). Washington, DC: American Sociological Association.
- Statistisches Bundesamt (Ed.). (2004). *Datenreport 2004. Zahlen und Fakten über die Bundesrepublik Deutschland*. Bonn: Bundeszentrale für politische Bildung.
- Thurstone, L. L. (1947). *Multiple-factor analysis*. Chicago: University of Chicago Press.
- Überla, K. (1972). *Faktorenanalyse*. Berlin, Heidelberg, New York: Springer.
- Wagner, M. (1997). *Scheidung in Ost- und Westdeutschland. Zum Verhältnis von Ehestabilität und Sozialstruktur seit den 30er Jahren*. Frankfurt/New York: Campus.
- West, S. G., Finch, J. F., & Bentler, P. M. (1995). Structural equation models with non-normal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: concepts, issues, and applications* (pp. 56-75). Thousand Oaks: Sage.
- Winkler, G. (1996). *Sozialreport 1996. Sonderheft 1/2: Zur sozialen Situation und deren subjektiver Reflexion in den neuen Bundesländern*. Bonn: Verlag JHW Dietz.



<http://www.springer.com/978-3-531-16083-2>

Was hält Partnerschaften zusammen?

Psychologische und soziologische Erklärungsansätze
zum Erfolg von Paarbeziehungen

Arránz Becker, O.

2008, VIII, 347 S., Softcover

ISBN: 978-3-531-16083-2