



WESTFÄLISCHE
WILHELMS-UNIVERSITÄT
MÜNSTER



Universität Bielefeld

**Methodendokumentation
der kriminologischen Schülerbefragung
in Duisburg 2002 und 2003**

– Zwei-Wellen-Panel –

Daniela Pollich

**Schriftenreihe:
Jugendkriminalität in der modernen Stadt – Methoden
Nr. 14 / 2007**

ISSN 1610-2819

Informationen zur Schriftenreihe

Jugendkriminalität in der modernen Stadt – Methoden

Herausgeber:

Prof. Dr. Klaus Boers

Institut für Kriminalwissenschaften

Abteilung Kriminologie

Bispinghof 24/25

48 143 Münster

und

Prof. Dr. Jost Reinecke

Fakultät für Soziologie

Universität Bielefeld

Postfach 10 01 31

33 501 Bielefeld

Internet-Adressen

<http://www.jura.uni-muenster.de/kriminologie>

<http://www.uni-bielefeld.de/soz/krimstadt>

ISSN 1610-2819

Übersicht der bisherigen Titel der Reihe:

1. Motzke, Katharina / Wittenberg, Jochen (1/2004): Methodendokumentation der kriminologischen Schülerbefragung in Münster 2000.
2. Wittenberg, Jochen (2/2004): Methodendokumentation der kriminologischen Schülerbefragung in Münster 2001.
3. Wittenberg, Jochen / Hilfert, Nicole (3/2004): Methodendokumentation der kriminologischen Schülerbefragung in Bocholt 2001.
4. Wittenberg, Jochen (4/2004): Methodendokumentation der kriminologischen Schülerbefragung in Münster 2002.
5. Motzke, Katharina / Brondies, Marc (5/2004): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002.
6. Brondies, Marc (6/2004): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2003.
7. Wittenberg, Jochen (7/2004): Methodendokumentation der kriminologischen Schülerbefragung in Münster 2003.
8. Brondies, Marc (8/2004): Methodendokumentation der Lehrerbefragung an Münsteraner und Duisburger Schulen 2003. Erhebung durchgeführter Präventionsmaßnahmen.
9. Pöge, Andreas (9/2005): Methodendokumentation der kriminologischen Schülerbefragung in Münster 2000–2003 (Vier-Wellen-Panel).
10. Hilfert, Nicole (10/2005): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2004.
11. Kunadt, Susann (11/2006): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2005.
12. Bentrup, Christina (12/2007): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2006.
13. Pöge, Andreas (13/2007): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002 bis 2005 (Vier-Wellen-Panel).
14. **Pollich, Daniela (14/2007): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002 und 2003 (Zwei-Wellen-Panel).**

Inhaltsverzeichnis

1	Einleitung	1
2	Studiendesign	3
3	Erstellung der Paneldaten	7
4	Realisierte Zuordnungen	13
4.1	Die verwendeten Datensätze und ihre Kurzbezeichnungen	13
4.2	Abschätzung und Bewertung der Zuordnungsgüte	14
5	Beschreibung des Zwei-Wellen-Panels	17
5.1	Darstellung der Zusammensetzung	17
5.2	Zusammenfassung und Erklärung	21
5.2.1	Individualebene	21
5.2.2	Schulebene	23
	Literaturverzeichnis	25
A	Anhang	27
A.1	Codeblätter	27
A.2	Häufigkeiten der Codevariablen	30

Tabellenverzeichnis

3.1	Durchgeführte Kontrollschritte	9
3.2	Anzahl der handschriftlichen Kontrollen und deren Ergebnisse	11
4.1	Bezeichnungen und Fallzahlen der Datensätze aus 2002 und 2003	14
4.2	Zuordnungsgüte der Paneldaten in Duisburg und Münster	16
5.1	Zusammensetzung $P_{t_{1,2}}$ nach Geschlecht	17
5.2	Altersdurchschnitt $P_{t_{1,2}}$	18
5.3	Zusammensetzung $P_{t_{1,2}}$ nach Geschlecht und Schulform (Zeilenproz.)	18
5.4	Zusammensetzung $P_{t_{1,2}}$ nach Geschlecht und Schulform (Spaltenproz.)	20
5.5	Verzerrung der Paneldaten ($P_{t_{1,2}}$ zu t_1) nach Geschlecht und Anzahl der Fehler in der Codezuordnung (Zeilenproz.)	22
5.6	Verzerrung der Paneldaten ($P_{t_{1,2}}$ zu t_1) nach Schulform und Anzahl der Fehler in der Codezuordnung (Zeilenproz.)	22
5.8	Zusammensetzung $P_{t_{1,2}}$ nach Schulform (Zeilenproz.)	24
5.7	Zusammensetzung $P_{t_{1,2}}$ nach Geschlecht (Zeilenproz.)	24
A.1	Häufigkeiten von Co001 (erster Buchstabe Vorname Vater)	30
A.2	Häufigkeiten von Co002 (erster Buchstabe Vorname Mutter)	31
A.3	Häufigkeiten von Co003 (erster Buchstabe eigener Vorname)	32
A.4	Häufigkeiten von Co004 (Geburtstag)	33
A.5	Häufigkeiten von Co005 (Haarfarbe)	34
A.6	Häufigkeiten von Co006 (Augenfarbe)	35
A.7	Häufigkeiten von Co007 aus 2003 (Befragungsteilnahme im Vorjahr)	36
A.8	Häufigkeiten von Co008 aus 2003 (Schulwechsel)	36
A.9	Häufigkeiten von Co009 aus 2003 (Sitzengeblieben)	36
A.10	Häufigkeiten der Gesamtcodes (5 Stellen) aus 2002	37
A.11	Häufigkeiten der Gesamtcodes (5 Stellen) aus 2003	37
A.12	Häufigkeiten der Gesamtcodes (6 Stellen) aus 2003	37
A.13	Teilnehmende Schulen „Trend netto“ Schulstatistik und Querschnitte	38
A.14	Zusammensetzung $P_{t_{1,2}}$ nach Geschlecht und Schulform (absolut)	39

Abbildungsverzeichnis

2.1	Erhebungsdesign der Längsschnittstudie 2000–2007	4
A.1	Codeblatt Duisburg 2002	28
A.2	Codeblatt Duisburg 2003	29

1 Einleitung

Die Ziele des kriminologisch und kriminalsoziologisch ausgerichteten DFG-Projektes *Kriminalität in der modernen Stadt*¹ liegen darin, Entstehung sowie Verläufe devianter und delinquenter Handlungsstile näher zu beleuchten. Das umfassende Analysemodell, in welches die Untersuchungen eingebettet sind, berücksichtigt dabei sowohl die Makro- als auch die Meso- und Mikroebene sozialer Phänomene. Die erhobenen Daten stammen aus zwei hinsichtlich ihrer Größe und Sozialstruktur recht verschiedenen deutschen Städten, nämlich Münster und Duisburg.

Besondere Möglichkeiten der Auswertung kausaler Zusammenhänge ergeben sich durch das kombinierte *Panel-Kohorten-Design* der Studie. Dieses Design ermöglicht die längsschnittliche Analyse von Ursache-Wirkungsbeziehungen zwischen den zentralen Untersuchungsvariablen sowie eine Analyse von Stabilität und Veränderung der delinquenten Aktivität auf individueller Ebene und zusätzlich den Vergleich verschiedener Altersjahrgänge und Städte.

So wurden in Münster zwischen 2000 und 2003 jährliche Befragungen durchgeführt, in Duisburg werden seit 2002 bis voraussichtlich 2009 jedes Jahr Schülerinnen und Schüler befragt. Neben einer während der gesamten Zeit durchgehend von der siebten Jahrgangsstufe bis über die Beendigung ihrer Schullaufbahn hinaus befragten Duisburger Kohorte wurden in den Jahren 2002 und 2003 zudem Daten einer älteren Vergleichskohorte erhoben. Die betreffenden Jugendlichen besuchten in den Erhebungsjahren die neunte bzw. zehnte Klasse (zur Veranschaulichung des Erhebungsdesigns siehe Abbildung 2.1). Aus forschungsökonomischen Gründen wurde diese ältere Kohorte nicht weiter befragt; aus den entsprechenden Daten entstand ein Zwei-Wellen-Panel, welches in erster Linie dazu dient, eventuelle Kohorteneffekte der jüngeren Duisburger Kohorte kontrollieren zu können.

Der vorliegende Bericht dokumentiert zentrale Punkte der Vorgehensweise bei der Erstellung des Duisburger Zwei-Wellen-Panels der älteren Kohorte von 2002 und 2003.

Der eingesetzte Fragebogen zu Opferwerdung, selbstberichteter Delinquenz, Erziehungsstilen, Konfliktverhalten, Kriminalitätseinstellungen sowie zu Lebens-, Freizeit-, und Konsumstilen war für eine schriftliche, anonyme Befragung der Schülerinnen und Schüler konzipiert. So weit wie möglich wurden Fragen verwendet, die bereits in anderen

¹ Projektleitung: Prof. Dr. Klaus Boers, Institut für Kriminalwissenschaften, WWU Münster und Prof. Dr. Jost Reinecke, Fakultät für Soziologie, Universität Bielefeld, DFG-Förderzeichen: Bo 1234/6-1, 6-2, 6-3, 6-5, 6-6 und Re 832/4-1, 4-2, 4-3, 6-2, 6-3.

Jugendstudien eingesetzt wurden. Die Fragebögen der einzelnen Jahre waren aufgrund des Panelansatzes der Studie weitgehend identisch.

Detaillierte Darstellungen des Erhebungsinstruments, der verwendeten Skalen oder eine Validierung eben dieser sind in diesem Bericht nicht vorgesehen. Ebenfalls außen vor gelassen werden eine genaue Beschreibung der Erhebungen der beiden einbezogenen Querschnittsdatensätze sowie weitergehende inhaltliche Analysen und Ergebnisse. Insbesondere werden das Studiendesign, das Zuordnungsverfahren und der entstandene Zwei-Wellen-Paneldatensatz beschrieben. Der Datensatz der jüngeren Duisburger Kohorte wird lediglich an einigen Stellen zu Vergleichszwecken herangezogen; selbiges gilt für die Daten des Münsteraner Vier-Wellen-Panels.

2 Studiendesign

Beginnend mit dem Jahr 2000 wurden in Münster bis zum Jahr 2003 in Jahresabständen schulische Befragungen durchgeführt.² Die Münsteraner Schülerinnen und Schüler wurden auf diesem Wege von der siebten bis zur zehnten Jahrgangsstufe befragt. Seit dem Jahr 2002 werden auch in Duisburg jährlich Jugendliche befragt. Dabei enthielt die Erhebung der ersten Welle sowohl Schülerinnen und Schüler aus der siebten als auch aus der neunten Jahrgangsstufe. Während die Schülerinnen und Schüler der damals siebten Jahrgangsstufe inzwischen bis zur zwölften Jahrgangsstufe (2007) begleitet wurden, wurden diejenigen Jugendlichen, die 2002 die neunte Klasse besucht hatten nur noch ein weiteres Mal in 2003 befragt.³ Aus diesen Erhebungen entstanden bisher ein Vier-Wellen-Paneldatensatz für die Münsteraner Schüler,⁴ ein Fünf-Wellen-Paneldatensatz für die jüngere Duisburger Kohorte⁵ sowie ein Zwei-Wellen-Paneldatensatz für die ältere Duisburger Kohorte.

Neben den wiederholt befragten Kohorten wurden auch mehrere Querschnittsdatensätze erhoben. So wurden in Münster im Jahr 2000 auch die neunte und elfte Jahrgangsstufe befragt, was einen weiteren Vergleich zwischen verschiedenen befragten Altersstufen ermöglicht. Zudem erfolgte im Jahr 2001 eine Querschnittserhebung der siebten, neunten und elften Jahrgangsstufen in Bocholt, um auch einen Stadt-Land-Vergleich zu ermöglichen.⁶ Außerdem wurden im Jahr 2003 Zusatzerhebungen in Münster und in Duisburg durchgeführt, in welchen Lehrer zur Durchführung von Präventionsmaßnahmen befragt wurden.⁷

Der genaue Erhebungsplan kann der Abbildung 2.1 entnommen werden. Daraus wird ersichtlich, dass die möglichen Vergleiche sich auf verschiedene Altersgruppen innerhalb eines Querschnitts, auf gleiche Altersgruppen zu verschiedenen Erhebungszeitpunkten sowie auf Vergleiche innerhalb von Lebensverläufen beziehen können.

2 Zur Dokumentation der einzelnen Querschnittserhebungen in Münster siehe genauer Motzke & Wittenberg, 2004; Wittenberg, 2004a; Wittenberg, 2004b; Wittenberg, 2004c.

3 Zur Dokumentation der einzelnen Querschnittserhebungen in Duisburg siehe genauer Motzke & Brondies, 2004; Brondies, 2004a; Hilfert, 2005; Kunadt, 2006; Bentrup, 2007.

4 Zur Dokumentation des Münsteraner Vier-Wellen-Panels siehe Pöge, 2005a.

5 Für die Duisburger Erhebung ist bislang eine Dokumentation des Vier-Wellen-Panels verfügbar (siehe hierzu Pöge, 2007).

6 Zur genauen Dokumentation der Bocholter Erhebung siehe Wittenberg & Hilfert, 2004.

7 Eine detailliertere Dokumentation dieser Erhebung findet sich in Brondies, 2004b.

Nur das gewählte Paneldesign, also die jährliche Wiederbefragung *derselben* Personen kann die Analysierbarkeit sowohl auf der Aggregatebene als auch auf der individuellen Ebene gewährleisten (vgl. Engel & Reinecke, 1994, S. 5 f.).

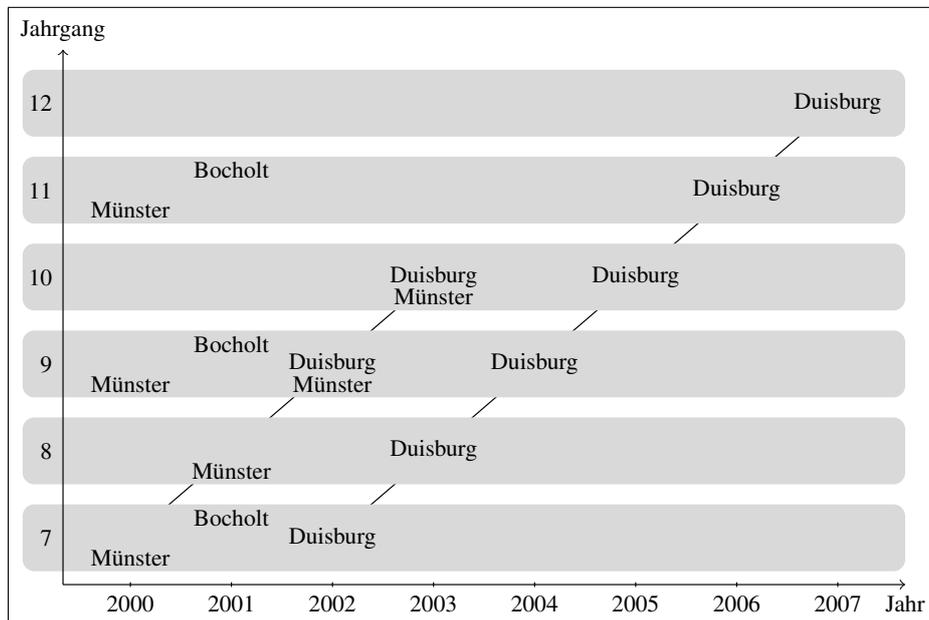


Abbildung 2.1: Erhebungsdesign der Längsschnittstudie 2000–2007

In der *ersten* Duisburger Schülerbefragung im Jahr 2002 wurde eine Vollerhebung der Schülerinnen und Schüler der siebten und neunten Jahrgangsstufen angestrebt, um eine möglichst umfassende Ausgangsgrundlage für den Paneldatensatz zu schaffen und Effekte der Panelmortalität nach Möglichkeit einzudämmen.

Die *zweite* Duisburger Schülerbefragung im Jahr 2003 strebte wie schon die erste Erhebungswelle nochmals eine Vollerhebung der Schülerinnen und Schüler der inzwischen achten und zehnten Jahrgangsstufen an. Nachdem die ältere der beiden Duisburger Kohorten nach 2003 nicht weiter befragt wurde, war es das Ziel der *dritten* und *vierten* Schülerbefragungen der Jahre 2004 und 2005, möglichst viele derjenigen Schülerinnen und Schüler der neunten bzw. zehnten Jahrgangsstufe, die bereits an der ersten und/oder zweiten Erhebung teilgenommen hatten, wieder im Klassenverband zu befragen. Mit dem Erreichen der elften Jahrgangsstufe im Jahre 2006 wurde es zunehmend schwieriger, die Befragten in der Schule anzutreffen, da viele von ihnen bereits ihre schulische Laufbahn beendet hatten. Für die Studie bedeutete dies mit der fünften Erhebungswelle eine teilweise Umstellung auf eine postalische Befragungsweise, um möglichst viele der bereits befragten Jugendlichen wieder zu erreichen und damit den weiteren Aufbau des Paneldatensatzes der jüngeren Duisburger Kohorte nicht zu gefährden.

Die bislang aktuellste Befragung stammt aus dem Jahr 2007, wobei sich der Schwerpunkt der Erhebungsmethode – mit zunehmender Beendigung der Schulausbildung in

der betreffenden Kohorte – weiter in Richtung einer postalischen Befragung verschob. Um eine maximale Anzahl der bereits im Paneldatensatz befindlichen Jugendlichen wieder befragen zu können, wurde zudem erstmals eine *face-to-face*-Befragung derjenigen Teilnehmer durchgeführt, die weder in der Schule noch auf postalischem Wege erreichbar waren.

Der vorliegende Bericht konzentriert sich ausschließlich auf das zu Vergleichszwecken erstellte Duisburger Parallelpanel der älteren Kohorte aus den Erhebungsjahren 2002 und 2003.

3 Erstellung der Paneldaten

Mit der Durchführung einer Panelstudie sind wegen der Wiederbefragung derselben Personen spezifische Probleme verbunden. Wie können Personen in den Befragungswellen der verschiedenen Jahre unter Berücksichtigung der Anonymität identifiziert werden, und wie können die von ein und derselben Person ausgefüllten Fragebögen für die spätere Analyse zusammengebracht werden? In diesem Forschungsprojekt entschied man sich für ein Verfahren, welches mit Hilfe eines persönlichen Codes die Fragebogenzuordnungen zwischen den unterschiedlichen Jahren leisten sollte.

Um eine Zuordnung der Fragebögen aus den verschiedenen Erhebungswellen zu ermöglichen, wurden Codes verwendet, die über Codeblätter abgefragt wurden. Jeder befragte Schüler sollte beim Interview ein Codeblatt ausfüllen, auf dem fünf bzw. ab der Befragung in 2003 sechs persönliche Fragen gestellt wurden, aus deren Antwort jeweils ein bestimmter Buchstabe oder eine Zahl notiert werden sollte. Die Fragen bezogen sich auf unveränderliche Merkmale des Probanden oder seines Umfeldes (natürliche Haarfarbe, Name des Vaters etc.). Diese Buchstaben-Zahlenkombination sollte den gesamten Code bilden. In jeder Erhebungswelle wurden die Codeblätter den Schülern zum Ausfüllen wieder vorgelegt, so dass bei einem stabilen Antwortverhalten der Schüler die Codeblätter, die von ein und derselben Person in den verschiedenen Wellen ausgefüllt wurden, denselben Code aufweisen mussten.

Während der Befragungen sollte auf den Codeblättern⁸ Folgendes notiert werden:

- *Co001*: Der erste Buchstabe des Vornamens des Vaters
- *Co002*: Der erste Buchstabe des Vornamens der Mutter
- *Co003*: Der erste Buchstabe des eigenen Vornamens (nur 2003)
- *Co004*: Die beiden Tagesziffern des eigenen Geburtstages
- *Co005*: Der letzte Buchstabe der eigenen Haarfarbe
- *Co006*: Der letzte Buchstabe der eigenen Augenfarbe

Aufgrund der Erfahrungen mit dem Code und dem Zuordnungsverfahren aus den Münsteraner Befragungen (vgl. Pöge, 2005a, S. 7 ff.; Pöge, 2005b) wurden auf den Duisburger Codeblättern zum einen einfachere Fragen mit einer größeren Antwortvarianz gestellt und zum anderen im Jahr 2003 der Code – ebenso wie in Münster – um eine Frage erweitert und das Layout des gesamten Blattes überarbeitet. So wurde zum Beispiel nicht mehr nach dem *letzten* Buchstaben der Haarfarbe des Vaters, sondern dem *ersten*

⁸ Abdrucke der Codeblätter aller Befragungszeitpunkte finden sich im Anhang Abbildungen A.1 und A.2.

Buchstaben des Vornamens des Vaters gefragt. Das Layout wurde dergestalt geändert, dass kein handschriftliches Ausfüllen erfolgen musste, sondern alle Antwortvorgaben zum Ankreuzen aufgeführt wurden.

Die verbesserte Zuordenbarkeit, die aus der Hinzunahme einer sechsten Codefrage resultierte, konnte für den hier vorzustellenden Datensatz noch nicht zur Geltung kommen. Erst ab der Zuordnung der Datensätze von 2003 und 2004 konnte aus dieser Möglichkeit Nutzen gezogen werden, hierzu muss nämlich die entsprechende Frage in beiden zu verbindenden Datensätzen enthalten sein. Da die Daten der älteren Duisburger Kohorte in 2002 und 2003 erhoben wurden, spielt die zusätzliche Codefrage bei der hier durchgeführten Zuordnung noch keine Rolle.

Ab dem Befragungsjahr 2003 wurden zusätzlich auch die folgenden Fragen gestellt:

- *Co007*: Befragungsteilnahme im Vorjahr (ja/nein)
- *Co008*: Schulwechsel im vergangenen Jahr (ja/nein)
- *Co009*: Nichtversetzung im vergangenen Jahr (ja/nein)

Da diese Fragen nicht voraussetzen, dass sie in beiden Panelwellen enthalten sind, die zusammengeführt werden sollen, sondern nur in der jeweils neueren von beiden relevant sind, konnte schon bei der hier dargestellten Zuordnung der Datensätze von 2002 und 2003 davon profitiert werden. Zusätzlich zu den fünf Codefragen und drei Zusatzfragen standen für die Fragebogenzuordnungen das Geschlecht der Befragten zur Verfügung, sowie die besuchte Schule, die mit einer Kennnummer erhoben wurde.

Bei der technischen Durchführung der Zuordnungen der Fragebögen aus den einzelnen Wellen traten prinzipiell die schon aus der Münsteraner Studie bekannten Probleme auf, sie konnten durch die Modifikationen und Verbesserungen jedoch etwas abgemildert werden.

Als generelle Problembereiche des Zuordnungsverfahrens waren schon in Münster die Probleme der *Identifizierung* und *Reproduktion* der Codes aufgetreten. Trotz des erweiterten Codes bleibt der Problembereich der Identifizierung hier von Bedeutung, da aufgrund der Hinzunahme einer Codefrage zwischen 2002 und 2003 noch nicht von dieser profitiert werden konnte. Doch durch die im Gegensatz zur Münsteraner Erhebung geänderten Codefragen konnte das Problem bereits etwas abgeschwächt werden. In Münster waren in allen vier Wellen nur bei rund drei Vierteln der Schülerinnen und Schülern eindeutige Codes aufgetreten – bei rund einem Viertel der Befragten wiesen unterschiedliche Personen einen gleichen Code auf und waren somit ohne Hinzunahme weiterer Informationen nicht eindeutig zu identifizieren (vgl. Pöge, 2005a, S. 63). In Duisburg hingegen wies die jüngere Kohorte bei dem fünfstelligen Code aus dem Jahr 2002 immerhin rund 90 Prozent, beim sechsstelligen Code, der in allen weiteren Befragungsjahren Anwendung fand, fast durchgehend eindeutige persönliche Codes auf (vgl. Pöge, 2007, S. 6).

Tabelle 3.1: *Durchgeführte Kontrollschritte*

	Schritt	Codevariablen	Zusatzvariablen
ohne Fehler	S1	ohne Fehler	ohne Fehler
	S2	ohne Fehler	keine Bedingung
ein Fehler	S3	ein Fehler	ohne Fehler
	S4	ein Fehler	keine Bedingung
zwei Fehler	S5	zwei Fehler	ohne Fehler

In der hier dargestellten älteren Kohorte konnten hinsichtlich des fünfstelligen Codes aus dem Jahr 2002 wiederum über 90 Prozent eindeutiger Codes verzeichnet werden. Betrachtet man den sechsstelligen Code aus 2003, so weisen nahezu alle Schülerinnen und Schüler eindeutige Codes auf. Da aus den bereits dargelegten Gründen auch aus der Erhebung von 2003 nur fünf Codefragen in die Zuordnung einbezogen werden können, sinkt der Anteil der eindeutigen Codes auf etwas über 90 Prozent. Dieser Anteil bleibt zwar hinter der Zahl nur einmal vorkommender Codes der späteren Duisburger Befragungswellen der jüngeren Kohorte zurück (vgl. Pöge, 2007, S. 101), liegt aber dennoch über dem Anteil eindeutiger persönlicher Codes im Münsteraner Panel (vgl. Pöge, 2005a, S. 63; siehe hierzu im Anhang Tabellen A.10 bis A.12).

Das Problem der Reproduktion, dass also Befragte in den verschiedenen Wellen einen nicht übereinstimmenden Code angeben, existiert auch in Duisburg noch in einem nicht zu vernachlässigendem Maße, konnte aber gleichwohl reduziert werden. In Münster fanden sich zwischen jeweils zwei Erhebungszeitpunkten lediglich rund 500 bis 600 Zuordnungen mit exakt übereinstimmendem Code, was bei allen realisierten Zuordnungen einen Prozentsatz von 40 bis 50 Prozent ausmachte (vgl. Pöge, 2005a, S. 7 f.; Pöge, 2005b, S. 66). In der jüngeren Duisburger Kohorte hingegen liegt die Quote der Zuordnungen mit einem exakt übereinstimmendem Code recht stabil bei rund 60 Prozent aller gefundenen Zuordnungen (vgl. Pöge, 2007, S. 6 f.). Im hier beschriebenen Datensatz beträgt der Anteil der Zuordnungen ohne Fehler im Code (1 182 Bogenpaare) sogar knapp 70 Prozent aller insgesamt einander zugeordneten Bögen (1 734 Bogenpaare; siehe Tabelle 3.2). Dieser hohe Prozentanteil eindeutiger Codes an den gesamten Zuordnungen resultiert nicht zuletzt daraus, dass in den späteren Kontrollschritten mit mehr zugelassenen Fehlern der Anteil erfolgreicher Zuordnungen teilweise drastisch unter dem Anteil der anderen Duisburger Wellen der jüngeren Kohorte liegt. Dennoch lässt sich festhalten, dass die Reproduktion des Codes mit Hilfe des verbesserten Layouts und der vereinfachten Fragen deutlich besser funktioniert hat als in der Münsteraner Erhebung.

Das zur Zuordnung der Bögen aus den unterschiedlichen Befragungsjahren angewendete *fehlertolerante Verfahren mit manuellem Handschriftenvergleich* bestand aus drei Schritten: In einem *ersten Schritt* wurden maschinell alle exakt übereinstimmenden Codes aus zwei Erhebungswellen herausgefunden. Die zusammengehörigen Fragebögen und Codeblätter wurden daraufhin einer manuellen Handschriftenkontrolle unterzogen, wobei die offensichtlich nicht passenden Zuordnungen aussortiert wurden. Die passenden Fragebogennummern wurden aus dem Datensatz genommen, so dass sie für die nachfolgenden Schritte nicht mehr zur Verfügung standen. Im *zweiten Schritt* wurde nach Codeübereinstimmungen unter Zulassung eines Fehlers und im *dritten Schritt* unter Zulassung von zwei Fehlern gesucht und die zugehörigen Bogennummern herausgeschrieben. Bei jedem Kontrollschritt mussten zunächst neben den Codes (mit der jeweils zugelassenen Fehleranzahl) auch das Geschlecht und die Schule übereinstimmen; zudem mussten die Schülerinnen und Schüler 2003 angegeben haben, dass sie im letzten Jahr an der Befragung teilgenommen hatten sowie nicht die Schule gewechselt hatten oder sitzen geblieben waren (S1, S3, S5). Anschließend wurden in jedem Kontrollschritt – unter Beibehaltung der vorgegebenen Fehlerzahl im persönlichen Code – diese zu erfüllenden Zusatzbedingungen weggelassen (S2, S4). Eine Übersicht über die durchgeführten Zuordnungsschritte und die Bedingungen, die der Code erfüllen musste, findet sich in Tabelle 3.1.

Das geschilderte Verfahren ist codegeleitet und hierarchisch. Zum einen ist das erste Kriterium für eine Zuordnung nach wie vor der Code, denn es werden nur offensichtliche Fehlzuordnungen ausgesondert. Zum anderen wird durch die schrittweise Durchführung gewährleistet, dass weniger Fehler im Code bei der Zuordnung bevorzugt werden.

Anstatt wie bei den Zuordnungen des Münsteraner Panels und der jüngeren Duisburger Kohorte auf geschulte studentische Hilfskräfte zurückzugreifen, die in gleich bleibenden Zweiergruppen arbeiteten, wurde hier die Handschriftenkontrolle versuchsweise durch eine Studierendengruppe im Rahmen einer Praxisübung ihrer empirischen Methodenausbildung durchgeführt. Dies hatte zur Folge, dass die Kontrollen im Kursverband stattfanden und zum Teil sehr viele Personen gleichzeitig mit den Abgleichen von Fragebögen befasst waren. Dies führte teilweise auffällig oft dazu, dass einzelne Bögen nicht gefunden werden konnten (siehe die zum Teil stark erhöhte Anzahl fehlender Bögen in Tabelle 3.2). Auch die Beurteilung zweier Bögen als nicht von demselben Befragten stammend ist besonders hinsichtlich der späteren Kontrollschritte im Vergleich zur jüngeren Duisburger Kohorte erhöht (vgl. Pöge, 2007, S. 7). Jedoch muss hervorgehoben werden, dass beispielsweise die Kontrollen dieser jüngeren Kohorte in den Jahren 2002 und 2003 noch stärker in hierarchische Teilschritte untergliedert war, was ebenfalls zu einer höheren Zuordnungsquote führen könnte. Die geringere Trefferquote bei den Zuordnungen wird hier jedoch – obgleich auch eine tatsächlich geringere Übereinstimmung der Bögen nicht völlig ausgeschlossen werden kann – hauptsächlich darauf zurückgeführt, dass die Gruppe der Studierenden, die mit den Kontrollen betraut waren,

Tabelle 3.2: Anzahl der handschriftlichen Kontrollen und deren Ergebnisse

		passt	passt nicht	fehlt	Kontrollen
ohne Fehler	S1	962	34	17	1013
	S2	220	90	9	319
	gesamt	1182	124	26	1332
ein Fehler	S3	392	31	18	441
	S4	85	1055	145	1285
	gesamt	477	1086	163	1726
zwei Fehler	S5	73	190	19	282
Summe		1732	1400	208	3340

(a) absolut durchgeführte Kontrollen pro Zuordnungsschritt

		passt	passt nicht	fehlt	Kontrollen
ohne Fehler	S1	95,0	3,4	1,6	1013
	S2	69,0	28,2	2,8	319
	gesamt	88,7	9,3	2,0	1332
ein Fehler	S3	88,9	7,0	4,1	441
	S4	6,6	82,1	11,4	1285
	gesamt	27,6	62,9	9,4	1726
zwei Fehler	S5	25,9	67,4	6,7	282
Summe		51,9	41,9	6,2	3340

(b) prozentuale Ergebnisse (Zeilenprozentuierung)

zum einen recht unerfahren war, zum anderen stark fluktuierte und daher die Urteile sehr viel heterogener ausfielen als es bei konstant arbeitenden Hilfskräften der Fall gewesen wäre. Letztlich ist auch der recht frühe Abbruch der Kontrollen (zum Vergleich mit anderen Zuordnungsprozeduren siehe Pöge, 2007, S. 9) auf die geringe Erfolgsquote der letzten Zuordnungsschritte im Zusammenspiel mit am Ende mangelnden personellen Ressourcen zurückzuführen. Daher muss der offensichtliche Qualitätsgewinn, den der Einsatz einer konstanten Anzahl geschulter und sorgfältig instruierter Kräfte zur Durchführung der Handschriftenabgleiche mit sich bringt, an dieser Stelle nochmals betont werden.

4 Realisierte Zuordnungen

Mit Hilfe des dargestellten Zuordnungsverfahrens konnte ein Zwei-Wellen-Paneldatensatz mit 1 732 Fällen erstellt werden. Wie diese Zuordnung im Einzelnen zu bewerten ist, ist den nachfolgenden Kapiteln zu entnehmen.

4.1 Die verwendeten Datensätze und ihre Kurzbezeichnungen

Für die weiteren Ausführungen relevante Daten wie die zugrunde liegenden Querschnittsdatsätze oder die betreffenden Schulstatistiken sind in Tabelle 4.1 aufgeführt. Der Einfachheit halber sollen im Folgenden Kurzbezeichnungen für die verwendeten Daten benutzt werden (siehe Tabelle 4.1).

Die Bezeichnung „S“ mit indiziertem Zeitpunkt steht für die Daten der amtlichen Schulstatistik zu einem bestimmten Zeitpunkt. Hierbei sind nur die Schulen berücksichtigt, die zu *beiden* Zeitpunkten an den Befragungen teilnahmen. Nicht berücksichtigt wurden Schulen, die nur an einer der beiden Erhebungswellen teilnahmen, oder die nur eine Befragung der Schülerinnen und Schüler der jüngeren Kohorte erlaubten.

Ein „Q“ mit einem Zeitpunkt als Index steht für einen ursprünglichen Querschnittsdatsatz der Einzelbefragungen. Hierbei handelt es sich wiederum um die Querschnittsdaten derjenigen Schülerinnen und Schüler, die eine Schule besuchten, die zu *beiden* Zeitpunkten an den Befragungen teilnahm. Damit geschehen die Bewertungen der Zusammensetzung des Paneldatensatzes auf der Basis nur derjenigen Schulen, die durch zweimalige Teilnahme an der Befragung ein durchgehendes Vorkommen ihrer Schülerinnen und Schüler im Zwei-Wellen-Panel erst ermöglichen. Diese Auswahl der Daten der Schulstatistiken und der Querschnitte als Bewertungsgrundlage für den Paneldatensatz wird als „Trend netto“ bezeichnet (vgl. Wittenberg, 2004a, S. 12). Im Gegensatz zur Darstellung der Häufigkeiten, mit der die einzelnen Codefragen beantwortet wurden (siehe im Anhang Tabellen A.1 bis A.9), wurden hier die diejenigen Fälle aus den Datensätzen eliminiert, die einer Plausibilitätskontrolle nicht standgehalten haben.

Die Paneldaten werden durch ein „P“ mit dem Index, der die Erhebungszeitpunkte ausdrückt, bezeichnet. So wird der hier beschriebene Zwei-Wellen-Paneldatensatz mit $P_{1,2}$ bezeichnet.

Tabelle 4.1: Bezeichnungen und Fallzahlen der Datensätze aus 2002 und 2003

Abkürzung	Fallzahl (N_b)	Beschreibung
S_{t_1}	3 041	Amtliche Schulstatistik <i>Netto</i> , Schuljahr 2001/2002 (Stichtag: 15. 10.) ^a
S_{t_2}	2 725	Amtliche Schulstatistik <i>Netto</i> , Schuljahr 2002/2003 (Stichtag: 15. 10.) ^a
Q_{t_1}	2 559	Querschnittsdaten des ersten Befragungsjahres <i>Netto</i> (2002 bzw. t_1) ^a
Q_{t_2}	2 427	Querschnittsdaten des zweiten Befragungsjahres <i>Netto</i> (2003 bzw. t_2) ^a
$P_{t_{1,2}}$	1 732	Paneldaten der Jahre 2002 und 2003

N_b bedeutet beobachtetes N

^a nur Schulen, die an allen Befragungszeitpunkten teilnahmen

4.2 Abschätzung und Bewertung der Zuordnungsgüte

Die Bewertung des angewendeten Zuordnungsverfahrens nach der realisierten Zuordnungsquote gestaltet sich komplizierter als auf den ersten Blick erkennbar.

Um die Zuordnungsquote bestimmen zu können, muss die Anzahl derjenigen Schülerinnen und Schüler bestimmt bzw. geschätzt werden, die zwischen den Befragungswellen hätten zugeordnet werden *können*.⁹ Auf Schulebene beschränkt sich hierbei die Grundgesamtheit zunächst auf die Jugendlichen der befragten Jahrgänge, die eine *Schule* besuchten, die zu allen Zeitpunkten der Befragung zustimmte („Trend netto“). Diese Anzahl kann relativ genau den jeweiligen Schulstatistiken entnommen werden (siehe Tabelle 4.1). Auch die Schulstatistiken selbst sind meist nicht exakt, da die ausgewiesenen Belegungszahlen zu den *Statistikstichtagen* durch zwischenzeitliche Ab- und Zugänge der Schulen von den tatsächlichen Belegungen an den *Befragungsstichtagen* abweichen können; diese Abweichungen scheinen jedoch vernachlässigbar klein. Darüber hinaus problematisch sind Veränderungen der Schulpopulation über den *gesamten Befragungszeitraum* (zwei Jahre) im Sinne von Zu- und Wegzügen nach bzw. aus Duisburg, Schulwechsler in bzw. von Schulen, die nicht an den Befragungen teilnahmen, Sitzenbleiber etc. Durch die Wahl der Erhebungsmethode (Fragebogeninterviews im Klassenverband) ist die erreichbare Probandenzahl auf diejenigen beschränkt, die an den jeweiligen Befragungstagen anwesend waren. Da die Fragebogenzuordnungen über die oben geschilderte Methode unter Zuhilfenahme von Codeblättern vorgenommen wurden, reduziert sich die Zahl weiter auf diejenigen Befragungsteilnehmerinnen und -teilnehmer, die ihr Codeblatt verwertbar ausgefüllt hatten.¹⁰

9 Erläuternde Beispiele hierzu sind in Pöge, 2005a, S. 14 f. und Pöge, 2007, S. 14 zu finden.

10 Lässt man diese Betrachtungen außer Acht, so sind noch 1 732 Jugendliche von den ursprünglichen 2 559 im Zwei-Wellen-Panel enthalten. Dies sind Jugendliche, die im beobachteten Zeitraum versetzt wurden, nicht umgezogen sind, nicht auf eine Schule außerhalb der Stichprobe umgeschult wurden, nicht an den Befragungstagen gefehlt haben und bei beiden Befragungen einen hinreichenden Code

Da als einzige Referenzdaten die aggregierten Schulstatistiken der Befragungsjahre zur Verfügung stehen und diese keinen befriedigenden Aufschluss über die oben genannten Probleme geben können, kann die Zahl der Probanden, die *tatsächlich mehrfach befragt wurden*, also hätten zugeordnet werden können, nur geschätzt werden.

Um die zu erwartende Zahl (N_e) der möglichen Zuordnungen zwischen den jeweiligen Zeitpunkten schätzen zu können, seien folgende Annahmen zugelassen:

1. In der größten offiziellen Schulpopulation der Befragungszeitpunkte sind alle Schülerinnen und Schüler der anderen Zeitpunkte enthalten
2. Die Differenzen zwischen den Stichproben und den jeweiligen Schulstatistiken beruhen auf zufälligen, stichprobenneutralen Ausfällen

Annahme 1 ist notwendig, um eine Referenzgröße über die Anzahl aller Probanden zu bestimmen, die überhaupt zu irgendeinem Zeitpunkt hätten teilnehmen können. Problematisch ist die Annahme deshalb, da Veränderungen der Schulpopulation stattgefunden haben, die auf Aggregatebene nicht mehr erkennbar sind. Es wird vernachlässigt, dass zu jedem Zeitpunkt ein gewisser Teil an Schülerinnen und Schülern nicht in der größten Population enthalten waren.

Annahme 2 ist problematisch, da in der Realität die Befragungsteilnahme vermutlich nicht zufällig, sondern von bestimmten Faktoren abhängig ist. So haben erste Analysen gezeigt, dass der Rücklauf in den Querschnitten unter anderem von Geschlecht und Schulform abhängt. Es kann also vermutet werden, dass auch der Anteil der Personen, die zu *mehreren* Zeitpunkten nicht teilnahmen, nicht zufällig verteilt ist, sondern von denselben Faktoren abhängt.

Trifft man also die oben genannten Annahmen, kann folgender Satz formuliert werden: Die zu erwartende Anzahl der Befragungsteilnehmer (N_e) zu bestimmten Zeitpunkten (t_i) ergibt sich aus der größten Schulpopulation ($\max_{i=1, \dots, k}(N(S_{t_i}))$) multipliziert mit dem Produkt der Quotienten aus den jeweiligen Anzahlen der realisierten Querschnittstichproben ($N(Q_{t_i})$) und den Schulpopulationen ($N(S_{t_i})$).¹¹

$$N_e = \max_{i=1, \dots, k} (N(S_{t_i})) \cdot \prod_{i=1}^k \frac{N(Q_{t_i})}{N(S_{t_i})} \quad (4.1)$$

Aus der zu erwartenden Anzahl (N_e) und der tatsächlich beobachteten bzw. realisierten Anzahl der Zuordnungen (N_b) lässt sich die in Tabelle 4.2 angeführte Quote für die ältere Duisburger Kohorte berechnen. Sie dürfte, wie oben angedeutet, die „wahre“ Zuordnungsquote eher unterschätzen, da auf aggregierter Ebene der Schulstatistiken nicht

geliefert haben. Von einem sehr konservativen Standpunkt aus kann hieraus mit einer gewissen Berechtigung eine Panelmortalität von 32 Prozent abgeleitet werden.

¹¹ Anders ausgedrückt entspricht dies dem Produkt der größten Ausgangsstichprobe und den multipli-

sichtbare Veränderungen (Sitzenbleiber, Umzüge etc., die sich pro Jahr (aber nicht im Panel) ausgleichen), nicht berücksichtigt werden. Dieser Anteil liegt schätzungsweise bei mindestens zehn Prozent. Eine Überschätzung der Quoten durch nicht berücksichtigte, systematische Ausfälle erscheint im Vergleich dazu weniger bedeutsam.

Tabelle 4.2: Zuordnungsgüte der Paneldaten in Duisburg und Münster

	Erhebung	N_e	N_b	Zuordnungsquote (%)
$P_{t_{1,2}}$	Duisburg 2002/2003 ältere Kohorte	2279	1732	76,0
$P_{t_{1,2}}$	Duisburg 2002/2003 jüngere Kohorte	3075	2472	80,4
$P_{t_{1,2}}$	Münster 2000/2001	1683	1271	75,5

Die Zuordnungsquote für das hier dargestellte Zwei-Wellen-Panel liegt bei 76 Prozent. Aus Vergleichsgründen enthält die Tabelle 4.2 ebenfalls die Zuordnungsquoten der ersten beiden Zeitpunkte der jüngeren Duisburger Kohorte sowie der Münsteraner Studie (vgl. Pöge, 2007, S. 16). Es wird ersichtlich, dass durch die Vereinfachung der Codefragen und die Verbesserung des Layouts des Codeblattes lediglich ein sehr geringfügig höherer Anteil an Bögen einander erfolgreich zugeordnet werden konnte, als dies bei der Münsteraner Erhebung möglich war. Zudem bleibt die hier realisierte Zuordnungsquote hinter der zeitgleich in der jüngeren Duisburger Kohorte erreichten zurück. Da in Duisburg hinsichtlich des Codes gleiche Voraussetzungen herrschten und die vorgenommenen Veränderungen von Codefragen und -layout in der jüngeren Kohorte den gewünschten Effekt erzielten, kann die geringere Erfolgsquote der Zuordnung der älteren Kohorte nur auf die verschiedenen Modalitäten bei der Handschriftenkontrolle (siehe hierzu Abschnitt 3) zurückgeführt werden.

Um die Zuordnungsquoten des realisierten Paneldatensatzes bewerten zu können, können Quoten, die in anderen Panelstudien erreicht wurden, herangezogen werden. Ein übliches Verfahren bei Paneluntersuchungen stellt die Zuordnung der Fragebögen über erhobene Adressen der Schülerinnen und Schüler dar. Hierbei tritt erfahrungsgemäß schon bei der *ersten Adresserhebung* ein Ausfall von ca. 50 Prozent auf (exemplarisch hierzu siehe Böttger et al., 2003, S. 35 ff.), der im Laufe solcher Adresspanels deutlich größer wird.

zierten Rücklaufquoten der beteiligten Querschnitte. Die detaillierte mathematische Herleitung der Formel ist Pöge (2005a, S. 63 ff.) zu entnehmen.

5 Beschreibung des Zwei-Wellen-Panels

5.1 Darstellung der Zusammensetzung

In diesem Abschnitt soll die Zusammensetzung des Zwei-Wellen-Panels der älteren Duisburger Kohorte beschrieben werden. Zum Vergleich werden dabei die Daten der Querschnitte und der Schulstatistiken aus den entsprechenden Jahren (beide „Trend netto“) hinzugezogen. Für die vergleichenden Darstellungen werden die zentralen Variablen Geschlecht, Alter und Schulform verwendet.¹² Der Datensatz des Zwei-Wellen-Panels enthält alle Fälle, deren zugehörige (und nach Plausibilitätskontrollen gültige) Fragebögen einem (ebenfalls gültigen) Bogen eines weiteren Zeitpunktes zugeordnet werden konnten.

Tabelle 5.1: Zusammensetzung $P_{t_{1,2}}$ nach Geschlecht

		männlich		weiblich		gesamt N
		N	%	N	%	(100 %)
t_1	S_{t_1}	1478	48,6	1563	51,4	3041
	Q_{t_1}	1258	49,2	1300	50,8	2558 ^a
	$P_{t_{1,2}}$	799	46,1	933	53,9	1732
t_2	S_{t_2}	1310	48,1	1415	51,9	2725
	Q_{t_2}	1178	48,5	1249	51,5	2427
	$P_{t_{1,2}}$	795	45,9	937	54,1	1732

^a Die reduzierte Fallzahl resultiert aus einer fehlenden Angabe beim Geschlecht.

Wie in Tabelle 5.1 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine leichte Verschiebung vor. Im Paneldatensatz sind sowohl zum Zeitpunkt t_1 als auch t_2 prozentual mehr Schülerinnen als in den Referenzdaten vorhanden. Die Abweichung fällt dabei mit knapp drei Prozentpunkten zu beiden Zeitpunkten eher moderat aus.

¹² Bei der Darstellung der Zusammensetzung der Paneldaten können hierbei beim Geschlecht und Alter Inkonsistenzen über die Zeit auftreten. Dies liegt darin begründet, dass bei *jeder* Frage fehlerhafte bzw. inkonsistente Antworten auftreten. So natürlich auch bei so „eindeutig“ zu beantwortenden Fragen wie denen nach Alter und Geschlecht. Einzig die Informationen über Schule und somit auch die Schulform sind in unserem Fall nahezu fehlerlos, da sie extern erhoben wurden und nicht vom Antwortverhalten der Probanden abhängen.

Tabelle 5.2: Altersdurchschnitt $P_{t_1,2}$

		\bar{x} Alter	Std.-abw.	gesamt N
t_1	Q_{t_1}	15,1	0,8	2510
	$P_{t_1,2}$	15,0	0,8	1714
t_2	Q_{t_2}	16,0	0,8	2378
	$P_{t_1,2}$	16,0	0,7	1713

Das Alter der Jugendlichen, deren Fälle im Paneldatensatz enthalten sind, weicht zu beiden Zeitpunkten nur minimal vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.2). Hierbei gab es zu t_1 18 und zu t_2 19 fehlende Angaben im Paneldatensatz, zu t_1 und zu t_2 jeweils 49 fehlende Angaben im Querschnitt. Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Tabelle 5.3: Zusammensetzung $P_{t_1,2}$ nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform ^a (%)					gesamt N
			HS	RS	GS	GY	SO	(100 %)
männlich	t_1	S_{t_1}	33,3	14,6	38,4	13,7	—	1478
		Q_{t_1}	32,2	15,4	35,8	15,3	1,3	1258
		$P_{t_1,2}$	27,8	17,5	35,5	18,8	0,4	799
	t_2	S_{t_2}	30,5	14,9	39,5	15,1	—	1310
		Q_{t_2}	29,2	15,0	38,6	16,6	0,5	1178
		$P_{t_1,2}$	27,7	17,4	35,8	18,7	0,4	795
weiblich	t_1	S_{t_1}	25,7	17,3	34,9	22,1	—	1563
		Q_{t_1}	24,7	18,2	31,5	25,8	—	1300
		$P_{t_1,2}$	20,5	18,2	32,0	29,3	—	933
	t_2	S_{t_2}	23,3	16,5	37,1	23,0	—	1415
		Q_{t_2}	22,8	16,7	35,9	24,5	0,1	1249
		$P_{t_1,2}$	20,7	17,8	33,7	27,7	—	937
gesamt	t_1	S_{t_1}	29,4	16,0	36,6	18,0	—	3041
		Q_{t_1}	28,3	16,8	33,6	20,7	0,6	2559 ^b
		$P_{t_1,2}$	23,8	17,9	33,7	24,4	0,2	1732
	t_2	S_{t_2}	26,8	15,7	38,3	19,2	—	2725
		Q_{t_2}	25,9	15,9	37,2	20,7	0,3	2427
		$P_{t_1,2}$	23,9	17,6	34,7	23,6	0,2	1732

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, SO=Sonderschule

^b Die Abweichung in der Summe der Fallzahlen des Querschnitts von 2002 resultiert aus einer fehlenden Angabe beim Geschlecht.

Aus Tabelle 5.3 wird ersichtlich, dass der Paneldatensatz 2002/2003 im Vergleich zur Schulstatistik systematisch verzerrt ist. Im Paneldatensatz befinden sich zu t_1 in der insgesamt Betrachtung rund fünfeinhalb Prozentpunkte weniger Hauptschülerinnen und Hauptschüler als in der Schulstatistik, zu t_2 knapp drei Prozentpunkte. Ebenfalls leicht unterrepräsentiert sind die Gesamtschülerinnen und Gesamtschüler. Hier sind zu t_1 knapp drei Prozentpunkte und zu t_2 etwa dreieinhalb Prozentpunkte Unterschied zur Schulstatistik festzustellen. Leicht überrepräsentiert sind hingegen die Realschülerinnen und Realschüler, die Abweichung von der Schulstatistik beträgt hier zu beiden Erhebungszeitpunkten knapp zwei Prozentpunkte. Stärker überrepräsentiert sind die Schülerinnen und Schüler, die das Gymnasium besuchen: zu t_1 sind etwa sechseinhalb Prozentpunkte mehr enthalten als durch die Schulstatistik vorgegeben, zu t_2 sind dies noch rund viereinhalb Prozentpunkte. Das Phänomen, dass laut Schulstatistik keine Sonderschüler im Paneldatensatz enthalten sein dürften, in den Querschnittsdatensätzen als auch im Zwei-Wellen-Panel jedoch einige Fälle verzeichnet sind, erklärt sich durch die gewählte Darstellungsweise der Datengrundlage („Trend netto“). Da laut offizieller Schulstatistik im Jahr 2003 keine Sonderschüler die zehnte Jahrgangsstufe besuchten, fallen die Sonderschulen aus dem Pool derjenigen Schulen weg, die zu beiden Zeitpunkten an der Befragung teilnahmen. Da faktisch aber 2003 auch Zehntklässler an Sonderschulen erreicht werden konnten, welche in der Schulstatistik offensichtlich nicht verzeichnet waren (vgl. Brondies, 2004a, S. 16) kam es dazu, dass in beiden Querschnitten Sonderschüler gleicher Schulen befragt werden konnten. Da auch im Paneldatensatz einige Sonderschüler einander zugeordnet werden konnten, wurden sie sowohl in den Querschnitten („Trend netto“) als auch beim Paneldatensatz in der Darstellung berücksichtigt, obwohl sie laut Schulstatistik nicht existieren dürften. Eine Übersicht über alle Schulen, die hier zur Erstellung der Vergleichgrundlage für den Paneldatensatz aus der der Schulstatistik und den Querschnitten einbezogen worden sind, ist in Tabelle A.13 im Anhang zu finden.

Betrachtet man nun die Verzerrung nach Geschlecht und Schulform gleichzeitig (siehe Tabelle 5.3 und zu den absoluten Häufigkeiten Tabelle A.14 im Anhang), so fällt auf, dass zwar generell mehr Jungen als Mädchen die Hauptschule besuchen, dass aber gleichzeitig die im Vergleich zur Schulstatistik zu schwache Besetzung der Hauptschulen stärker auf das Fehlen männlicher Hauptschüler zurückgeht. Die Betrachtung des Geschlechterverhältnisses an Hauptschulen lässt erkennen, dass der Anteil der Mädchen im Vergleich zur Schulstatistik sogar höher ist (siehe Tabelle 5.4). Analoges gilt für die Gesamtschulen. Hier ist zwar das Geschlechterverhältnis laut Schulstatistik nahezu ausgeglichen, im Paneldatensatz ist jedoch der Jungenanteil niedriger als der Mädchenanteil, was auf eine zu geringe Anzahl insbesondere männlicher Gesamtschüler im Zwei-Wellen-Panel hindeutet.

Die insgesamt leicht überrepräsentierten Realschülerinnen und Realschüler weisen zwar – nur auf die Realschulen bezogen – einen etwas höheren Mädchenanteil auf, jedoch

verändert sich diese Zusammensetzung auch im Paneldatensatz kaum. Der erhöhte Anteil von Jugendlichen, die die Realschule besuchen, scheint von beiden Geschlechtern fast gleichermaßen verursacht; es lässt sich lediglich zu t_1 eine leichte Tendenz zu einem etwas erhöhten Jungenanteil ausmachen. Schon bei der Betrachtung der Verzerrung nach Schulform ohne Berücksichtigung des Geschlechts weist das Gymnasium ein recht starkes Übergewicht auf. Schlüsselte man dieses nun nach dem Geschlecht der Jugendlichen auf, dass insbesondere die Gruppe der Mädchen hierfür verantwortlich ist. Schon in der Schulstatistik fällt auf, dass der Mädchenanteil an Gymnasien wesentlich höher ist als der Jungenanteil; betrachtet man den Paneldatensatz, so wird dieser Effekt dort sogar noch etwas verstärkt. Weiterhin sind die Sonderschüler stark nach Geschlecht verzerrt, fast alle von ihnen sind männlich. Da jedoch die Fallzahlen äußerst gering sind (siehe hierzu Tabelle A.14 im Anhang), ist eine aussagekräftige Interpretation hier nicht möglich.

Tabelle 5.4: Zusammensetzung $P_{t_1,2}$ nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform ^a (%)					gesamt N
			HS	RS	GS	GY	SO	(100 %)
männlich	t_1	S_{t_1}	55,0	44,4	51,0	36,9	—	48,6
		Q_{t_1}	56,0	45,1	52,3	36,5	100,0	49,2
		$P_{t_1,2}$	53,8	45,2	48,7	35,5	100,0	46,1
	t_2	S_{t_2}	54,7	45,5	49,7	37,8	—	48,1
		Q_{t_2}	54,7	46,0	50,3	39,0	85,7	48,5
		$P_{t_1,2}$	53,1	45,2	47,4	36,4	100,0	45,9
weiblich	t_1	S_{t_1}	45,0	55,6	49,0	63,1	—	51,4
		Q_{t_1}	44,0	54,9	47,7	63,5	0,0	50,8
		$P_{t_1,2}$	46,2	54,8	51,3	64,5	0,0	53,9
	t_2	S_{t_2}	45,3	54,5	50,3	62,2	—	51,9
		Q_{t_2}	45,3	54,0	49,7	61,0	14,3	51,5
		$P_{t_1,2}$	46,9	54,8	52,6	63,6	0,0	54,1

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, SO=Sonderschule

5.2 Zusammenfassung und Erklärung

Bei der Analyse der oben beschriebenen Zusammensetzung der Paneldaten im Vergleich mit den Daten der Querschnitte und der Schulstatistik ist die Unterscheidung nach Individual- und Schulebene sinnvoll. Analog zu den Methodenberichten der Querschnittserhebungen wurde auf eine Darstellung der Zusammensetzung der gesamten Münsteraner Schulbevölkerung zu Gunsten einer detaillierten Darstellung der Zusammensetzung der Schulpopulationen an allen Zeitpunkten teilnehmender Schulen verzichtet („Trend netto“). Die oben beschriebenen Verzerrungen bezogen sich immer auf Vergleiche zwischen Zwei-Wellen-Panel, Querschnittsdatensätzen und der Schulstatistik nach „Trend netto“ und sind deshalb durch Ausfälle auf Individualebene begründet. An dieser Stelle soll zusätzlich kurz auf die Ausfälle auf der Schulebene eingegangen werden.

5.2.1 Individualebene

Zusammenfassend lässt sich sagen, dass die Paneldaten in Bezug auf die Zusammensetzung nach Geschlecht und Schulform (teilweise leicht, teilweise erheblich) verzerrt sind. Es finden sich im Vergleich zur Schulstatistik über alle Paneldatensätze hinweg mehr Mädchen als Jungen in den Daten. Der Anteil der Hauptschülerinnen und Hauptschüler sowie der Anteil der Gesamtschülerinnen und Gesamtschüler ist geringer als in der Schulstatistik, teilweise deutlich niedriger. Dafür ist der Anteil der Realschülerinnen und Realschüler etwas höher als in der Schulstatistik, der Anteil der Gymnasiastinnen und Gymnasiasten sogar recht stark erhöht.

Diese Abweichungen sind in den meisten Fällen in geringerer Stärke schon in den Querschnittsdaten zu erkennen. Generell setzen sich also die Verzerrungen der Querschnittsdaten im Vergleich zur Schulstatistik in den Paneldaten verstärkt fort. Dies lässt den Schluss zu, dass dieselben Gründe, die zu den beschriebenen Verschiebungen führten, auch (Mit-) Ursachen für die Panelverzerrungen sind. Hierzu zählen die gewollten und ungewollten Ausfälle in den Querschnitten aufgrund von Krankheit, Schwänzen, Verweigerungen etc. Man muss also annehmen, dass bei Jungen, Sonder- und Hauptschülern diese Ausfallgründe öfter auftreten als bei Mädchen, Gymnasiastinnen und Gymnasiasten sowie Realschülerinnen und Realschülern. Es scheinen innerhalb dieser Personengruppen auch über die Jahre unterschiedliche Probanden wegen der genannten Ursachen auszufallen. Fielen immer dieselben Personen aus, würde der Trend im Panel nicht verstärkt. Eine weitere Ursache für die Verzerrungen der Paneldatensätze liegt in der Art der Zuordnung der Fragebögen. Das oben erläuterte Verfahren über wiederholte Codes weist den Nachteil auf, dass das gleiche Ausfüllen der Codefragen unterschiedlichen Personengruppen unterschiedlich stark Probleme bereitete.

Tabelle 5.5: Verzerrung der Paneldaten ($P_{t_{1,2}}$ zu t_1) nach Geschlecht und Anzahl der Fehler in der Codezuordnung (Zeilenprozentuierung)

		männlich		weiblich		gesamt N (100 %)
		N	%	N	%	
$P_{t_{1,2}}$	0	507	42,9	675	57,1	1 182
	1	252	52,8	225	47,2	477
	2	40	54,8	33	45,2	73
gesamt		799	46,1	933	53,9	1 732

Tabelle 5.6: Verzerrung der Paneldaten ($P_{t_{1,2}}$ zu t_1) nach Schulform und Anzahl der Fehler in der Codezuordnung (Zeilenprozentuierung)

		Schulform ^a (%)					gesamt N (100 %)
		HS	RS	GS	GY	SO	
$P_{t_{3,4}}$	0	22,2	19,1	31,8	26,7	0,2	1 182
	1	26,2	15,5	37,9	20,1	0,2	477
	2	35,6	13,7	35,6	15,1	—	73
gesamt		23,8	17,9	33,7	24,4	0,2	1 732

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, SO=Sonderschule

In den Tabellen 5.5 und 5.6 finden sich die Paneldaten aufgeschlüsselt nach den zugelassenen Fehlern in der Codezuordnung.

Es wird deutlich, dass zwar der größte Teil ($N = 1 182$) der Schülerinnen und Schüler bei der Beantwortung des Codeblattes in 2002 und 2003 keinen Fehler machte, der Anteil derjenigen, die einen Fehler ($N = 477$) machten bereits kleiner ist, jedoch immer noch ins Gewicht fällt. Der Anteil der Probanden, die zwei Fehler ($N = 73$) machten ist zwar deutlich geringer, jedoch wurde hier auch nicht die ganze Bandbreite möglicher Kontrollschritte mit zwei Fehlern im Code ausgeschöpft.

Darüber hinaus wird ersichtlich, dass die Anzahl der Fehler bei der Beantwortung des Codeblattes mit Geschlecht und Schulbildung der Befragten zusammenhängen. Während die Verzerrungen auf Ebene des aggregierten Paneldatensatzes – der Zuordnungen bis zu zwei Fehler beinhaltet – meist nur relativ leicht sind (siehe Abschnitt 5.1, lassen sich bei der Aufschlüsselung nach Fehlern deutliche Verschiebungen erkennen. Augenscheinlich war es den weiblichen Befragungsteilnehmern eher möglich, in zwei Jahren fehlerfrei zu antworten; der Anteil der Mädchen ist bei den Zuordnungen mit keinem Fehler deutlich größer als derjenige der Jungen. Auch die Schulbildung hat einen (erwartungsgemäßen) Effekt. Je höher die Schulbildung, desto weniger Fehler wurden bei der Beantwortung gemacht. Vermutlich spielen kognitive Fähigkeiten – bei aller Einfachheit der Fragen – ebenso eine Rolle wie Konzentrationsfähigkeit und -wille.

Mit der Überlagerung der beiden Phänomene, also der schul- bzw. geschlechtsspezifischen Ausfallquote, die schon in den Querschnitten beobachtet werden konnte, wie der augenscheinlich schul- und geschlechtsspezifischen Fähigkeit, die Codeblätter richtig auszufüllen, ist die beobachtete Verzerrung der Paneldaten zu erklären.

5.2.2 Schulebene

Für die Teilnahme an den Befragungen war die Zustimmung der Schulkonferenzen der einzelnen Schulen notwendig. Schon an der ersten Befragung im Jahr 2002 nahmen nicht alle Schulen teil, während der zweiten Befragungen fiel eine weitere Schule aus. Anhand der Daten der Jahre 2002 und 2003 soll kurz die Systematik der Verzerrungen auf Geschlechter- und Schulebene dargestellt werden.

Tabelle 5.7 zeigt, dass hinsichtlich der Geschlechterverteilung im Vergleich zwischen der kompletten Schulstatistik und der Schulstatistik „Trend netto“ zu beiden Zeitpunkten allenfalls leichte Verzerrungen vorliegen. Anders ist dies jedoch im Vergleich mit dem Zwei-Wellen-Panel. Hier lässt sich sowohl zu t_1 als auch zu t_2 konstatieren, dass der Mädchenanteil etwas zu Lasten des Jungenanteils angestiegen ist. In beiden Panelwellen liegt diese Verschiebung zwischen drei und vier Prozentpunkten.

Wie in Tabelle 5.8 ersichtlich, liegt durch den Ausfall einiger Schulen eine erhebliche Verschiebung der Daten in Bezug auf die Schulform vor. Wegen des Ausfalls auf Schulebene sind zunächst in der *reduzierten* Schulstatistik („Trend netto“) im Vergleich zur *gesamten* Schulstatistik Hauptschülerinnen und Hauptschüler *überrepräsentiert*, Gymnasiastinnen und Gymnasiasten dagegen *unterrepräsentiert*. Jugendliche, die Realschulen oder Gesamtschulen besuchen, sind annähernd gleich verteilt.

Die Ausfälle auf Individualebene, die zwischen reduzierter Schulstatistik und Querschnitts- bzw. Paneldaten deutlich werden, folgen einem genau entgegengesetzten Trend: hier sind die Hauptschulen stark unterrepräsentiert, die Gymnasien überrepräsentiert. Durch die Ausfälle auf der Individualebene verringern sich also die Effekte der Verzerrungen auf der Schulebene hinsichtlich Hauptschule und Gymnasium wieder; sie verkehren sich sogar in die gegenläufige Richtung. So sind die Verteilungen der Hauptschüler und Gymnasiasten im Querschnitt der gesamten Schulstatistik ähnlicher als der reduzierten. Im Falle der etwas überrepräsentierten Realschulen und unterrepräsentierten Gesamtschulen lässt sich dieses Phänomen nicht beobachten. Die Abweichungen sind hier konsistent und jeweils beim Zwei-Wellen-Panel am stärksten ausgeprägt.

Tabelle 5.8: Zusammensetzung $P_{t_{1,2}}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)					gesamt N
		HS	RS	GS	GY	SO	(100 %)
t_1	$S_{g_{t_1}}$	24,1	15,8	37,1	22,9	0,2	5209
	S_{t_1}	29,4	16,0	36,6	18,0	—	3041
	Q_{t_1}	28,3	16,8	33,6	20,7	0,6	2559
	$P_{t_{1,2}}$	23,8	17,9	33,7	24,5	0,2	1732
t_2	$S_{g_{t_2}}$	22,3	15,2	38,5	24,1	—	4672
	S_{t_2}	26,8	15,7	38,3	19,2	—	2725
	Q_{t_2}	25,9	15,9	37,2	20,7	0,3	2427
	$P_{t_{1,2}}$	23,9	17,6	34,7	23,6	0,2	1732

^a HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, SO=Sonderschule
 S_g bezeichnet die Daten der gesamten Schulstatistik

Tabelle 5.7: Zusammensetzung $P_{t_{1,2}}$ nach Geschlecht (Zeilenprozentuierung)

		männlich		weiblich		gesamt N
		N	%	N	%	(100 %)
t_1	$S_{g_{t_1}}$	2588	49,7	2621	50,3	5209
	S_{t_1}	1478	48,6	1563	51,4	3041
	Q_{t_1}	1258	49,2	1300	50,8	2558 ^a
	$P_{t_{1,2}}$	799	46,1	933	53,9	1732
t_2	$S_{g_{t_1}}$	2323	49,7	2349	50,3	4672
	S_{t_2}	1310	48,1	1415	51,9	2725
	Q_{t_2}	1178	48,5	1249	51,5	2427
	$P_{t_{1,2}}$	795	45,9	937	54,1	1732

^a Die reduzierte Fallzahl resultiert aus einer fehlenden Angabe beim Geschlecht.
 S_g bezeichnet die Daten der gesamten Schulstatistik

Bei Längsschnittuntersuchungen liegt das Augenmerk nicht auf Repräsentativität der Stichproben, da dies für Untersuchungen von Veränderungen über die Zeit nicht relevant ist. Dennoch ist es unbestreitbar vorteilhafter, dass sich die genannten Verzerrungseffekte auf Individual- und Schulebene mit Blick auf die Repräsentativität der Daten gegenseitig eher aufheben als verstärken.

Literaturverzeichnis

- Bentrup, C. (2007). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2006* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ No. 12). Münster, Bielefeld.
- Brondies, M. (2004a). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2003* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ No. 6). Münster, Trier.
- Brondies, M. (2004b). *Methodendokumentation der Lehrerbefragung an Münsteraner und Duisburger Schulen 2003. Erhebung durchgeführter Präventionsmaßnahmen* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ No. 8). Münster, Trier.
- Böttger, A., Ehret, B., Othold, F., Prein, G., Schumann, K. F., & Seus, L. (2003). Methoden der Untersuchung. In K. F. Schumann (Ed.), *Delinquenz im Lebensverlauf. Bremer Längsschnittstudie zum Übergang von der Schule in den Beruf bei ehemaligen Hauptschülern* (Vol. 2, pp. 35–66). Weinheim, München: Juventa.
- Engel, U., & Reinecke, J. (1994). *Panelanalyse. Grundlagen-Techniken-Beispiele*. New York: Walter de Gruyter.
- Hilfert, N. (2005). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2004* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ No. 10). Münster, Trier.
- Kunadt, S. (2006). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2005* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ No. 11). Münster, Bielefeld.
- Motzke, K., & Brondies, M. (2004). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ No. 5). Münster, Trier.
- Motzke, K., & Wittenberg, J. (2004). *Methodendokumentation der kriminologischen Schülerbefragung in Münster 2000* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ No. 1). Münster, Trier.
- Pöge, A. (2005a). *Methodendokumentation der kriminologischen Schülerbefragung in Münster 2000–2003 (Vier-Wellen-Panel)* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ No. 9). Münster, Trier.
- Pöge, A. (2005b). Persönliche Codes bei Längsschnittstudien: Ein Erfahrungsbericht. *ZA-Information*, 56, 50-69.

- Pöge, A. (2007). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002 bis 2005 (Vier-Wellen-Panel)*. (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ No. 13). Münster, Bielefeld.
- Wittenberg, J. (2004a). *Methodendokumentation der kriminologischen Schülerbefragung in Münster 2001* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ No. 2). Münster, Trier.
- Wittenberg, J. (2004b). *Methodendokumentation der kriminologischen Schülerbefragung in Münster 2002* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ No. 4). Münster, Trier.
- Wittenberg, J. (2004c). *Methodendokumentation der kriminologischen Schülerbefragung in Münster 2003* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ No. 7). Münster, Trier.
- Wittenberg, J., & Hilfert, N. (2004). *Methodendokumentation der kriminologischen Schülerbefragung in Bocholt 2001* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ No. 3). Münster, Trier.

A Anhang

A.1 Codeblätter



WESTFÄLISCHE WILHELMS-UNIVERSITÄT MÜNSTER
Prof. Dr. Klaus Boers
 Direktor des Instituts für Kriminalwissenschaften, Abteilung Kriminologie
Prof. Dr. Jost Reinecke
 Direktor der Arbeitsstelle Sozialindikatoren, Institut für Soziologie
 Bispinghof 24/25 48143 Münster Tel. 0251-83 22749
 kriminologie@uni-muenster.de

Kriminologische Befragung Duisburger Schülerinnen und Schüler
Beiblatt zur Erstellung des persönlichen Codes

Januar/Februar 2002

Liebe Schülerin, lieber Schüler,

wir wollen im nächsten Jahr die Befragung wiederholen, weil es sehr wichtig ist zu erfahren, was sich in der Zwischenzeit verändert hat. Damit nächstes Jahr deine beiden Fragebögen einander zugeordnet werden können, ohne deinen Namen zu kennen, benötigen wir ein Codewort. Dadurch kann niemand in Erfahrung bringen, dass du die beiden Fragebögen ausgefüllt hast. Wir möchten Dich deshalb bitten, auf diesem Zusatzblatt zum Fragebogen dieses Codewort zu bilden.

Wichtig ist also, dass du dasselbe Codewort nächstes Jahr noch weißt. Aus diesem Grund haben wir nachfolgend fünf Fragen formuliert, die dir helfen sollen, dich an dasselbe Wort zu erinnern.

*Wenn du eine der Fragen überhaupt nicht beantworten kannst,
 lass das entsprechende Feld bitte leer.*

Hier nun die fünf Fragen zur Erstellung deines persönlichen Codes:

	Bitte trage hier ins erste Feld den ersten Buchstaben des Vornamens deines Vaters (oder einer Person, die für dich einem Vater am nächsten kommt) ein. (z.B. Anton, Bernd, Hans-Peter, Fritz usw.)
	Bitte trage ins zweite Feld den ersten Buchstaben des Vornamens deiner Mutter (oder einer Person, die für dich einer Mutter am nächsten kommt) ein. (z.B. Anna, Beate, Jutta, Maria usw.)
	Bitte trage ins dritte und vierte Feld den Tag deines Geburtsdatums ein. (z.B. Geburtstag am 7. Januar = 0 7, am 12. Mai = 1 2, am 31. Oktober = 3 1)
	Bitte trage ins fünfte Feld den letzten Buchstaben deiner natürlichen Haarfarbe ein. (z.B. Glatze, braun, schwarz, blond, grau usw.)
	Bitte trage ins sechste Feld den letzten Buchstaben deiner Augenfarbe ein. (z.B. braun, blau, grün, grau usw.)

Abbildung A.1: Codeblatt Duisburg 2002

Beiblatt zur Erstellung des persönlichen Codes

Liebe Schülerin, lieber Schüler,

da wir deinen Fragebogen dem des letzten Jahres zuordnen wollen, bitten wir dich, deinen persönlichen Code zu wiederholen.

Da wir letztes Jahr das Codewort getrennt vom Fragebogen eingegeben, mit einem Computerprogramm verschlüsselt und danach vernichtet haben, ist es wichtig, dass du dich an dein Codewort erinnerst. Denn nur wenn wir mit diesem Codewort denselben Vorgang wiederholen, können deine Fragebögen einander zugeordnet werden und dies, ohne dass jemand herausfinden kann, wer diesen Fragebogen ausgefüllt hat.

Wichtig ist also, dass du dasselbe Codewort noch weißt. Aus diesem Grund haben wir letztes Jahr die nachfolgenden Fragen formuliert, die dir helfen sollen, dich an dasselbe Wort zu erinnern.

Wenn du eine der Fragen überhaupt nicht beantworten kannst, kreuze bitte kein Feld an!

Hier nun die sechs Fragen zur Erstellung deines persönlichen Codes:

1	Bitte kreuze den ersten Buchstaben des Vornamens deines Vaters (oder einer Person, die für dich einem Vater am nächsten kommt) an. (z. B. <u>A</u> nton, <u>B</u> ernd, <u>H</u> ans-Peter usw.) <input type="checkbox"/> a <input type="checkbox"/> b <input type="checkbox"/> c <input type="checkbox"/> d <input type="checkbox"/> e <input type="checkbox"/> f <input type="checkbox"/> g <input type="checkbox"/> h <input type="checkbox"/> i <input type="checkbox"/> j <input type="checkbox"/> k <input type="checkbox"/> l <input type="checkbox"/> m <input type="checkbox"/> n <input type="checkbox"/> o <input type="checkbox"/> p <input type="checkbox"/> q <input type="checkbox"/> r <input type="checkbox"/> s <input type="checkbox"/> t <input type="checkbox"/> u <input type="checkbox"/> v <input type="checkbox"/> w <input type="checkbox"/> x <input type="checkbox"/> y <input type="checkbox"/> z <input type="checkbox"/> ä <input type="checkbox"/> ö <input type="checkbox"/> ü <input type="checkbox"/> ß
2	Bitte kreuze den ersten Buchstaben des Vornamens deiner Mutter (oder einer Person, die für dich einer Mutter am nächsten kommt) an. (z. B. <u>A</u> нна, <u>B</u> eatrice, <u>J</u> utta, <u>M</u> aria, usw.) <input type="checkbox"/> a <input type="checkbox"/> b <input type="checkbox"/> c <input type="checkbox"/> d <input type="checkbox"/> e <input type="checkbox"/> f <input type="checkbox"/> g <input type="checkbox"/> h <input type="checkbox"/> i <input type="checkbox"/> j <input type="checkbox"/> k <input type="checkbox"/> l <input type="checkbox"/> m <input type="checkbox"/> n <input type="checkbox"/> o <input type="checkbox"/> p <input type="checkbox"/> q <input type="checkbox"/> r <input type="checkbox"/> s <input type="checkbox"/> t <input type="checkbox"/> u <input type="checkbox"/> v <input type="checkbox"/> w <input type="checkbox"/> x <input type="checkbox"/> y <input type="checkbox"/> z <input type="checkbox"/> ä <input type="checkbox"/> ö <input type="checkbox"/> ü <input type="checkbox"/> ß
3	Bitte kreuze den ersten Buchstaben deines Vornamens an (z. B. <u>M</u> ichael, <u>T</u> homas, <u>U</u> te usw.) <input type="checkbox"/> a <input type="checkbox"/> b <input type="checkbox"/> c <input type="checkbox"/> d <input type="checkbox"/> e <input type="checkbox"/> f <input type="checkbox"/> g <input type="checkbox"/> h <input type="checkbox"/> i <input type="checkbox"/> j <input type="checkbox"/> k <input type="checkbox"/> l <input type="checkbox"/> m <input type="checkbox"/> n <input type="checkbox"/> o <input type="checkbox"/> p <input type="checkbox"/> q <input type="checkbox"/> r <input type="checkbox"/> s <input type="checkbox"/> t <input type="checkbox"/> u <input type="checkbox"/> v <input type="checkbox"/> w <input type="checkbox"/> x <input type="checkbox"/> y <input type="checkbox"/> z <input type="checkbox"/> ä <input type="checkbox"/> ö <input type="checkbox"/> ü <input type="checkbox"/> ß
4	Bitte kreuze den Tag deines Geburtsdatums an (z.B. Geburtstag am 7. Januar = <input type="checkbox"/> , am 12. Mai = <input checked="" type="checkbox"/> , am 31. Oktober = <input checked="" type="checkbox"/>) <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 <input type="checkbox"/> 7 <input type="checkbox"/> 8 <input type="checkbox"/> 9 <input type="checkbox"/> 10 <input type="checkbox"/> 11 <input type="checkbox"/> 12 <input type="checkbox"/> 13 <input type="checkbox"/> 14 <input type="checkbox"/> 15 <input type="checkbox"/> 16 <input type="checkbox"/> 17 <input type="checkbox"/> 18 <input type="checkbox"/> 19 <input type="checkbox"/> 20 <input type="checkbox"/> 21 <input type="checkbox"/> 22 <input type="checkbox"/> 23 <input type="checkbox"/> 24 <input type="checkbox"/> 25 <input type="checkbox"/> 26 <input type="checkbox"/> 27 <input type="checkbox"/> 28 <input type="checkbox"/> 29 <input type="checkbox"/> 30 <input type="checkbox"/> 31
5	Bitte kreuze den letzten Buchstaben deiner natürlichen Haarfarbe an. (z. B. braun <input type="checkbox"/> , Glatz <input type="checkbox"/> , schwarz <input type="checkbox"/> , usw.) <input type="checkbox"/> a <input type="checkbox"/> b <input type="checkbox"/> c <input type="checkbox"/> d <input type="checkbox"/> e <input type="checkbox"/> f <input type="checkbox"/> g <input type="checkbox"/> h <input type="checkbox"/> i <input type="checkbox"/> j <input type="checkbox"/> k <input type="checkbox"/> l <input type="checkbox"/> m <input type="checkbox"/> n <input type="checkbox"/> o <input type="checkbox"/> p <input type="checkbox"/> q <input type="checkbox"/> r <input type="checkbox"/> s <input type="checkbox"/> t <input type="checkbox"/> u <input type="checkbox"/> v <input type="checkbox"/> w <input type="checkbox"/> x <input type="checkbox"/> y <input type="checkbox"/> z <input type="checkbox"/> ä <input type="checkbox"/> ö <input type="checkbox"/> ü <input type="checkbox"/> ß
6	Bitte kreuze den letzten Buchstaben deiner Augenfarbe an. (z. B. braun <input type="checkbox"/> , grün <input type="checkbox"/> , grau <input type="checkbox"/> , usw.) <input type="checkbox"/> a <input type="checkbox"/> b <input type="checkbox"/> c <input type="checkbox"/> d <input type="checkbox"/> e <input type="checkbox"/> f <input type="checkbox"/> g <input type="checkbox"/> h <input type="checkbox"/> i <input type="checkbox"/> j <input type="checkbox"/> k <input type="checkbox"/> l <input type="checkbox"/> m <input type="checkbox"/> n <input type="checkbox"/> o <input type="checkbox"/> p <input type="checkbox"/> q <input type="checkbox"/> r <input type="checkbox"/> s <input type="checkbox"/> t <input type="checkbox"/> u <input type="checkbox"/> v <input type="checkbox"/> w <input type="checkbox"/> x <input type="checkbox"/> y <input type="checkbox"/> z <input type="checkbox"/> ä <input type="checkbox"/> ö <input type="checkbox"/> ü <input type="checkbox"/> ß

Hast du im letzten Jahr an der Befragung teilgenommen? ja nein
 Hast du im letzten Jahr die Schule gewechselt? ja nein
 Bist du im letzten Jahr sitzen geblieben? ja nein

Seite 2 Westfälische Wilhelms-Universität Münster, Kriminalwissenschaften / Universität Trier, Soziologie – Duisburger Schülerbefragung 2003

Abbildung A.2: Codeblatt Duisburg 2003

A.2 Häufigkeiten der Codevariablen

Tabelle A.1: Häufigkeiten von Co001 (erster Buchstabe Vorname Vater)^a

	2002		2003	
	N	%	N	%
a	216	8,2	189	7,8
ä	—	—	—	—
b	106	4,0	98	4,0
c	47	1,8	46	1,9
d	130	4,9	133	5,5
e	57	2,2	61	2,5
f	95	3,6	80	3,3
g	95	3,6	86	3,5
h	339	12,8	299	12,3
i	34	1,3	51	2,1
j	199	7,5	162	6,7
k	143	5,4	137	5,6
l	29	1,1	33	1,4
m	249	9,4	226	9,3
n	59	2,2	56	2,3
o	23	0,9	18	0,7
ö	4	0,2	1	0,0
p	98	3,7	89	3,7
q	—	—	2	0,1
r	219	8,3	196	8,1
s	86	3,3	77	3,2
ß	—	—	1	0,0
t	62	2,3	57	2,3
u	84	3,2	75	3,1
ü	2	0,1	2	0,1
v	48	1,8	41	1,7
w	167	6,3	148	6,1
x	—	—	1	0,0
y	24	0,9	18	0,7
z	15	0,6	14	0,6
fehlend	13	0,5	34	1,4
gesamt	2643	100,1	2431	100,0

^aBei sämtlichen Darstellungen der Codevariablen kann es bei der Angabe der prozentualen Anteile zu Rundungsfehlern kommen, was in einigen Fälle zu einer Spaltensumme knapp über oder unter 100 % führt.

Tabelle A.2: Häufigkeiten von Co002 (erster Buchstabe Vorname Mutter)

	2002		2003	
	N	%	N	%
a	229	8,7	197	8,1
ä	1	0,0	—	—
b	226	8,6	204	8,4
c	147	5,6	132	5,4
d	93	3,5	84	3,5
e	168	6,4	158	6,5
f	66	2,5	66	2,7
g	169	6,4	164	6,7
h	190	7,2	172	7,1
i	66	2,5	103	4,2
j	88	3,3	43	1,8
k	95	3,6	86	3,5
l	44	1,7	45	1,9
m	302	11,4	269	11,1
n	83	3,1	84	3,5
o	12	0,5	10	0,4
ö	2	0,1	2	0,1
p	80	3,0	68	2,8
q	—	—	—	—
r	115	4,4	106	4,4
s	232	8,8	207	8,5
ß	—	—	—	—
t	29	1,1	29	1,2
u	104	3,9	96	4,0
ü	5	0,2	6	0,2
v	31	1,2	24	1,0
w	13	0,5	14	0,6
x	3	0,1	2	0,1
y	19	0,7	15	0,6
z	22	0,8	22	0,9
fehlend	9	0,3	23	0,1
gesamt	2643	100,1	2431	100,1

Tabelle A.3: Häufigkeiten von Co003 (erster Buchstabe eigener Vorname)

	2002		2003	
	N	%	N	%
a	—	—	212	8,7
ä	—	—	1	0,0
b	—	—	69	2,8
c	—	—	153	6,3
d	—	—	184	7,6
e	—	—	67	2,8
f	—	—	66	2,7
g	—	—	37	1,5
h	—	—	59	2,4
i	—	—	44	1,8
j	—	—	211	8,7
k	—	—	114	4,7
l	—	—	48	2,0
m	—	—	294	12,1
n	—	—	95	3,9
o	—	—	20	0,8
ö	—	—	13	0,5
p	—	—	75	3,1
q	—	—	—	—
r	—	—	70	2,9
s	—	—	359	14,8
ß	—	—	—	—
t	—	—	109	4,5
u	—	—	2	0,1
ü	—	—	4	0,2
v	—	—	41	1,7
w	—	—	9	0,4
x	—	—	3	0,1
y	—	—	29	1,2
z	—	—	13	0,5
fehlend	—	—	30	1,2
gesamt	—	—	2431	100,0

Tabelle A.4: Häufigkeiten von Co004 (Geburtstag)

	2002		2003	
	N	%	N	%
0	1	0,0	—	—
1	84	3,2	82	3,4
2	99	3,7	84	3,5
3	67	2,5	68	2,8
4	89	3,4	77	3,2
5	101	3,8	85	3,5
6	104	3,9	81	3,3
7	95	3,6	90	3,7
8	97	3,7	81	3,3
9	78	3,0	71	2,9
10	82	3,1	80	3,3
11	101	3,8	83	3,4
12	85	3,2	77	3,2
13	90	3,4	86	3,5
14	85	3,2	84	3,5
15	86	3,3	79	3,2
16	80	3,0	62	2,6
17	73	2,8	67	2,8
18	90	3,4	77	3,2
19	86	3,3	78	3,2
20	94	3,6	77	3,2
21	88	3,3	85	3,5
22	87	3,3	80	3,3
23	84	3,2	73	3,0
24	77	2,9	73	3,0
25	97	3,7	88	3,6
26	91	3,4	81	3,3
27	85	3,2	77	3,2
28	68	2,6	63	2,6
29	73	2,8	67	2,8
30	77	2,9	72	3,0
31	41	1,6	42	1,7
94	1	0,0	—	—
fehlend	7	0,3	61	2,5
gesamt	2643	100,1	2431	100,2

Tabelle A.5: Häufigkeiten von Co005 (Haarfarbe)

	2002		2003	
	<i>N</i>	%	<i>N</i>	%
a	—	—	1	0,0
ä	—	—	—	—
b	45	1,7	110	4,5
c	—	—	1	0,0
d	936	35,4	847	34,8
e	7	0,3	19	0,8
f	—	—	1	0,0
g	1	0,0	3	0,1
h	—	—	—	—
i	—	—	—	—
j	—	—	—	—
k	—	—	1	0,0
l	—	—	—	—
m	2	0,1	3	0,1
n	1233	46,7	1094	45,0
o	—	—	—	—
ö	—	—	—	—
p	—	—	—	—
q	—	—	—	—
r	—	—	—	—
s	10	0,4	30	1,2
ß	—	—	3	0,1
t	40	1,5	36	1,5
u	3	0,1	—	—
ü	—	—	—	—
v	—	—	—	—
w	—	—	1	0,0
x	—	—	—	—
y	—	—	—	—
z	359	13,6	259	10,7
fehlend	7	0,3	22	0,9
Gesamt	2643	100,1	2431	99,7

Tabelle A.6: Häufigkeiten von Co006 (Augenfarbe)

	2002		2003	
	<i>N</i>	%	<i>N</i>	%
a	—	—	2	0,1
ä	—	—	2	0,1
b	29	1,1	87	3,6
c	—	—	1	0,0
d	—	—	4	0,2
e	—	—	3	0,1
f	—	—	—	—
g	8	0,3	28	1,2
h	—	—	—	—
i	—	—	1	0,0
j	—	—	—	—
k	—	—	—	—
l	1	0,0	4	0,2
m	2	0,1	5	0,2
n	1726	65,3	1505	61,9
o	1	0,0	1	0,0
ö	—	—	1	0,0
p	—	—	—	—
q	—	—	—	—
r	—	—	—	—
s	1	0,0	5	0,2
ß	—	—	2	0,1
t	1	0,0	2	0,1
u	856	32,4	732	30,1
ü	—	—	—	—
v	—	—	—	—
w	1	0,0	—	—
x	1	0,0	—	—
y	—	—	—	—
z	9	0,3	19	0,8
fehlend	7	0,3	27	1,1
Gesamt	2643	99,8	2431	100,0

Tabelle A.7: Häufigkeiten von Co007 aus 2003 (Befragungsteilnahme im Vorjahr ja/nein)

		<i>N</i>	% (gültig)
gültig	ja	1891	77,8
	nein	356	14,6
	k. A.	184	7,6
gesamt		2431	100,0

Tabelle A.8: Häufigkeiten von Co008 aus 2003 (Schulwechsel ja/nein)

		<i>N</i>	% (gültig)
gültig	ja	71	2,9
	nein	2077	85,4
	k. A.	283	11,6
gesamt		2431	99,9

Tabelle A.9: Häufigkeiten von Co009 aus 2003 (Sitzengeblieben ja/nein)

		<i>N</i>	% (gültig)
gültig	ja	74	3,0
	nein	2073	85,3
	k. A.	284	11,7
gesamt		2431	100,0

A.2 Häufigkeiten der Codevariablen

Tabelle A.10: Häufigkeiten der Gesamtcodes (5 Stellen) aus 2002

		<i>N</i>	% (gesamt)	% (gültig)
gültig	1	2413	91,3	91,5
	2	208	7,9	7,9
	3	15	0,6	0,6
	gesamt	2636	99,7	100,0
fehlend		7	0,3	
gesamt		2643	100,0	

Tabelle A.11: Häufigkeiten der Gesamtcodes (5 Stellen) aus 2003

		<i>N</i>	% (gesamt)	% (gültig)
gültig	1	2242	92,2	92,6
	2	176	7,2	7,3
	3	3	0,1	0,1
	gesamt	2421	99,6	100,0
fehlend		10	0,4	
gesamt		2431	100,0	

Tabelle A.12: Häufigkeiten der Gesamtcodes (6 Stellen) aus 2003

		<i>N</i>	% (gesamt)	% (gültig)
gültig	1	2401	98,8	99,2
	2	20	0,8	0,8
	gesamt	2421	99,6	100,0
fehlend		10	0,4	
gesamt		2431	100,0	

Tabelle A.13: *Teilnehmende Schulen „Trend netto“ Schulstatistik und Querschnitte*

	$S_{t_{1,2}}$	$Q_{t_{1,2}}$
HS	Frankenschule Gemeinschaftshauptschule Ludgerusstraße Gebrüder-Grimm-Schule Anne-Frank-Schule Bronkhorstschule Heinrich-Böll-Schule Gemeinschaftshauptschule Werhstraße Gemeinschaftshauptschule Wiesbadener Straße Gemeinschaftshauptschule In den Haesen Gemeinschaftshauptschule Gneisenaustraße Gottfried-Könzgen-Schule Gemeinschaftshauptschule Hitzestraße Gemeinschaftshauptschule Friedrich-Ebert-Straße Schule Im Angerbogen	Frankenschule Gemeinschftshauptschule Ludgerusstraße Gebrüder-Grimm-Schule Anne-Frank-Schule Bronkhorstschule Heinrich-Böll-Schule Gemeinschaftshauptschule Werhstraße Gemeinschaftshauptschule Wiesbadener Straße Gemeinschaftshauptschule In den Haesen Gemeinschaftshauptschule Gneisenaustraße Gottfried-Könzgen-Schule Gemeinschaftshauptschule Hitzestraße Gemeinschaftshauptschule Friedrich-Ebert-Straße Schule Im Angerbogen
RS	August-Thyssen-Realschule Gustav-Stresemann-Realschule Karl-Lehr-Realschule Realschule Süd	August-Thyssen-Realschule Gustav-Stresemann-Realschule Karl-Lehr-Realschule Realschule Süd
GS	Gottfried-Wilhelm-Leibniz-Gesamtschule Gesamtschule Meiderich Theodor-König-Gesamtschule Gesamtschule Ruhrort Gesamtschule DU-Mitte Lise-Meitner-Gesamtschule Gesamtschule Süd	Gottfried-Wilhelm-Leibniz-Gesamtschule Gesamtschule Meiderich Theodor-König-Gesamtschule Gesamtschule Ruhrort Gesamtschule DU-Mitte Lise-Meitner-Gesamtschule Gesamtschule Süd
GY	Albert-Einstein-Gymnasium Kopernikus-Gymnasium Max-Planck-Gymnasium Mercator-Gymnasium Steinbart-Gymnasium St. Hildegardis-Gymnasium	Albert-Einstein-Gymnasium Kopernikus-Gymnasium Max-Planck-Gymnasium Mercator-Gymnasium Steinbart-Gymnasium St. Hildegardis-Gymnasium
SO	— —	Christian-Zeller-Schule Alfred-Adler-Schule

Tabelle A.14: Zusammensetzung $P_{1,2}$ nach Geschlecht und Schulform (absolut)

			Schulform ^a (%)					gesamt N
			HS	RS	GS	GY	SO	(100 %)
männlich	t_1	S_{t_1}	492	216	568	202	—	1478
		Q_{t_1}	405	194	450	193	16	1258
		$P_{t_1,2}$	222	140	284	150	3	799
	t_2	S_{t_2}	399	195	518	198	—	1310
		Q_{t_2}	344	177	455	196	6	1178
		$P_{t_1,2}$	220	138	285	149	3	795
weiblich	t_1	S_{t_1}	402	271	545	345	—	1563
		Q_{t_1}	318	236	410	336	—	1300
		$P_{t_1,2}$	191	170	299	273	—	933
	t_2	S_{t_2}	330	234	525	326	—	1415
		Q_{t_2}	285	208	449	306	1	1249
		$P_{t_1,2}$	194	167	316	260	—	937
gesamt	t_1	S_{t_1}	894	487	1113	547	—	3041
		Q_{t_1}	723	430	861	529	16	2559 ^b
		$P_{t_1,2}$	413	310	583	423	3	1732
	t_2	S_{t_2}	729	429	1043	524	—	2725
		Q_{t_2}	629	385	904	502	7	2427
		$P_{t_1,2}$	414	305	601	409	3	1732

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, SO=Sonderschule

^b Die Abweichung in der Summe der Fallzahlen des Querschnitts von 2002 resultiert aus einer fehlenden Angabe beim Geschlecht.