



WESTFÄLISCHE  
WILHELMS-UNIVERSITÄT  
MÜNSTER



 **Universität Trier**

---

**Methodendokumentation  
der kriminologischen Schülerbefragung  
in Münster 2000–2003**

**– Vier-Wellen-Panel –**

Andreas Pöge

**Schriftenreihe:  
Jugendkriminalität in der modernen Stadt – Methoden  
Nr. 9 / 2005**

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**

Fachbereich IV,  
Abteilung Soziologie,  
Universitätsring 15  
54 286 Trier

#### **Internet-Adressen**

[http://www.uni.muenster.de/Jura.krim/Abt\\_IV](http://www.uni.muenster.de/Jura.krim/Abt_IV)

<http://www.uni-trier.de/uni/fb4/soziologie/sozihome.htm>

#### **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).**

---

# Inhaltsverzeichnis

<b>1. Einleitung</b>	<b>1</b>
<b>2. Studiendesign</b>	<b>3</b>
<b>3. Erstellung der Paneldaten</b>	<b>5</b>
3.1. Das Zuordnungsverfahren mit Krypto-Codes . . . . .	5
3.2. Durchführung der Zuordnungen . . . . .	7
3.2.1. Probleme der Zuordnungsmethode . . . . .	8
3.2.2. Problemlösungsstrategie . . . . .	10
<b>4. Realisierte Zuordnungen</b>	<b>13</b>
4.1. Die verwendeten Datensätze und ihre Kurzbezeichnungen . . . . .	13
4.2. Abschätzung und Bewertung der Zuordnungsgüte . . . . .	14
<b>5. Zusammensetzung der Paneldaten</b>	<b>19</b>
5.1. Die Panelquerschnitte . . . . .	19
5.1.1. Panelquerschnitt 2000 ( $P_{Q_{t_1}}$ ) . . . . .	19
5.1.2. Panelquerschnitt 2001 ( $P_{Q_{t_2}}$ ) . . . . .	21
5.1.3. Panelquerschnitt 2002 ( $P_{Q_{t_3}}$ ) . . . . .	23
5.1.4. Panelquerschnitt 2003 ( $P_{Q_{t_4}}$ ) . . . . .	25
5.2. Die Zwei-Wellen-Panels . . . . .	27
5.2.1. Zwei-Wellen-Panel 2000/2001 ( $P_{t_{1,2}}$ ) . . . . .	27
5.2.2. Zwei-Wellen-Panel 2001/2002 ( $P_{t_{2,3}}$ ) . . . . .	29
5.2.3. Zwei-Wellen-Panel 2002/2003 ( $P_{t_{3,4}}$ ) . . . . .	32
5.2.4. Zwei-Wellen-Panel 2000/2002 ( $P_{t_{1,3}}$ ) . . . . .	35
5.2.5. Zwei-Wellen-Panel 2000/2003 ( $P_{t_{1,4}}$ ) . . . . .	37
5.2.6. Zwei-Wellen-Panel 2001/2003 ( $P_{t_{2,4}}$ ) . . . . .	39
5.3. Die Drei-Wellen-Panels . . . . .	42
5.3.1. Drei-Wellen-Panel 2000 bis 2002 ( $P_{t_{1,2,3}}$ ) . . . . .	42
5.3.2. Drei-Wellen-Panel 2001 bis 2003 ( $P_{t_{2,3,4}}$ ) . . . . .	46
5.3.3. Drei-Wellen-Panel 2000, 2001, 2003 ( $P_{t_{1,2,4}}$ ) . . . . .	48
5.3.4. Drei-Wellen-Panel 2000, 2002, 2003 ( $P_{t_{1,3,4}}$ ) . . . . .	51
5.4. Das Vier-Wellen-Panel 2000 bis 2003 ( $P_{t_{1,2,3,4}}$ ) . . . . .	54
5.5. Zusammenfassung und Erklärung . . . . .	59
5.5.1. Individualebene . . . . .	59
5.5.2. Schulebene . . . . .	60

<b>A. Anhang</b>	<b>63</b>
A.1. Häufigkeiten der Codes . . . . .	63
A.2. Mathematische Herleitung . . . . .	63
A.3. Codeblätter . . . . .	66
A.4. Häufigkeiten der Codevariablen . . . . .	70
A.5. Schul-, Klasseninformation und Befragungsteilnahme . . . . .	76
A.5.1. Münster 2001 . . . . .	76
A.5.2. Münster 2002 . . . . .	76
A.5.3. Münster 2003 . . . . .	77
<b>Literaturverzeichnis</b>	<b>79</b>

## Tabellenverzeichnis

4.1.	Bezeichnungen und Fallzahlen der Münsteraner Datensätze 2000–2003	14
4.2.	Zuordnungsgüte der Paneldaten	16
5.1.	Zusammensetzung nach Geschlecht $P_{Q_{t_1}}$	19
5.2.	Altersdurchschnitt $P_{Q_{t_1}}$	20
5.3.	Zusammensetzung $P_{Q_{t_1}}$ nach Geschlecht und Schulform (Zeilenproz.)	20
5.4.	Zusammensetzung $P_{Q_{t_1}}$ nach Geschlecht und Schulform (Spaltenproz.)	21
5.5.	Zusammensetzung nach Geschlecht $P_{Q_{t_2}}$	21
5.6.	Altersdurchschnitt $P_{Q_{t_2}}$	21
5.7.	Zusammensetzung $P_{Q_{t_2}}$ nach Geschlecht und Schulform (Zeilenproz.)	22
5.8.	Zusammensetzung $P_{Q_{t_2}}$ nach Geschlecht und Schulform (Spaltenproz.)	23
5.9.	Zusammensetzung nach Geschlecht $P_{Q_{t_3}}$	23
5.10.	Altersdurchschnitt $P_{Q_{t_3}}$	23
5.11.	Zusammensetzung $P_{Q_{t_3}}$ nach Geschlecht und Schulform (Zeilenproz.)	24
5.12.	Zusammensetzung $P_{Q_{t_3}}$ nach Geschlecht und Schulform (Spaltenproz.)	25
5.13.	Zusammensetzung nach Geschlecht $P_{Q_{t_4}}$	25
5.14.	Altersdurchschnitt $P_{Q_{t_4}}$	25
5.15.	Zusammensetzung $P_{Q_{t_4}}$ nach Geschlecht und Schulform (Zeilenproz.)	26
5.16.	Zusammensetzung $P_{Q_{t_4}}$ nach Geschlecht und Schulform (Spaltenproz.)	26
5.17.	Zusammensetzung nach Geschlecht $P_{t_{1,2}}$	27
5.18.	Altersdurchschnitt $P_{t_{1,2}}$	28
5.19.	Zusammensetzung $P_{t_{1,2}}$ nach Geschlecht und Schulform (Zeilenproz.)	28
5.20.	Zusammensetzung $P_{t_{1,2}}$ nach Geschlecht und Schulform (Spaltenproz.)	29
5.21.	Zusammensetzung nach Geschlecht $P_{t_{2,3}}$	30
5.22.	Altersdurchschnitt $P_{t_{2,3}}$	30
5.23.	Zusammensetzung $P_{t_{2,3}}$ nach Geschlecht und Schulform (Zeilenproz.)	31
5.24.	Zusammensetzung $P_{t_{2,3}}$ nach Geschlecht und Schulform (Spaltenproz.)	32
5.25.	Zusammensetzung nach Geschlecht $P_{t_{3,4}}$	32
5.26.	Altersdurchschnitt $P_{t_{3,4}}$	33
5.27.	Zusammensetzung $P_{t_{3,4}}$ nach Geschlecht und Schulform (Zeilenproz.)	34
5.28.	Zusammensetzung $P_{t_{3,4}}$ nach Geschlecht und Schulform (Spaltenproz.)	34
5.29.	Zusammensetzung nach Geschlecht $P_{t_{1,3}}$	35
5.30.	Altersdurchschnitt $P_{t_{1,3}}$	35
5.31.	Zusammensetzung $P_{t_{1,3}}$ nach Geschlecht und Schulform (Zeilenproz.)	36
5.32.	Zusammensetzung $P_{t_{1,3}}$ nach Geschlecht und Schulform (Spaltenproz.)	37
5.33.	Zusammensetzung nach Geschlecht $P_{t_{1,4}}$	37
5.34.	Altersdurchschnitt $P_{t_{1,4}}$	38
5.35.	Zusammensetzung $P_{t_{1,4}}$ nach Geschlecht und Schulform (Zeilenproz.)	38

5.36. Zusammensetzung $P_{t_{1,4}}$ nach Geschlecht und Schulform (Spaltenproz.) .	39
5.37. Zusammensetzung nach Geschlecht $P_{t_{2,4}}$ . . . . .	40
5.38. Altersdurchschnitt $P_{t_{2,4}}$ . . . . .	40
5.39. Zusammensetzung $P_{t_{2,4}}$ nach Geschlecht und Schulform (Zeilenproz.) .	41
5.40. Zusammensetzung $P_{t_{2,4}}$ nach Geschlecht und Schulform (Spaltenproz.) .	42
5.41. Zusammensetzung nach Geschlecht $P_{t_{1,2,3}}$ . . . . .	43
5.42. Altersdurchschnitt $P_{t_{1,2,3}}$ . . . . .	43
5.43. Zusammensetzung $P_{t_{1,2,3}}$ nach Geschlecht und Schulform (Zeilenproz.) .	44
5.44. Zusammensetzung $P_{t_{1,2,3}}$ nach Geschlecht und Schulform (Spaltenproz.)	45
5.45. Zusammensetzung nach Geschlecht $P_{t_{2,3,4}}$ . . . . .	46
5.46. Altersdurchschnitt $P_{t_{2,3,4}}$ . . . . .	46
5.47. Zusammensetzung $P_{t_{2,3,4}}$ nach Geschlecht und Schulform (Zeilenproz.) .	47
5.48. Zusammensetzung $P_{t_{2,3,4}}$ nach Geschlecht und Schulform (Spaltenproz.)	48
5.49. Zusammensetzung nach Geschlecht $P_{t_{1,2,4}}$ . . . . .	49
5.50. Altersdurchschnitt $P_{t_{1,2,4}}$ . . . . .	49
5.51. Zusammensetzung $P_{t_{1,2,4}}$ nach Geschlecht und Schulform (Zeilenproz.) .	50
5.52. Zusammensetzung $P_{t_{1,2,4}}$ nach Geschlecht und Schulform (Spaltenproz.)	51
5.53. Zusammensetzung nach Geschlecht $P_{t_{1,3,4}}$ . . . . .	52
5.54. Altersdurchschnitt $P_{t_{1,3,4}}$ . . . . .	52
5.55. Zusammensetzung $P_{t_{1,3,4}}$ nach Geschlecht und Schulform (Zeilenproz.) .	53
5.56. Zusammensetzung $P_{t_{1,3,4}}$ nach Geschlecht und Schulform (Spaltenproz.)	54
5.57. Zusammensetzung nach Geschlecht $P_{t_{1,2,3,4}}$ . . . . .	55
5.58. Altersdurchschnitt $P_{t_{1,2,3,4}}$ . . . . .	55
5.59. Zusammensetzung $P_{t_{1,2,3,4}}$ nach Geschlecht und Schulform (Zeilenproz.)	57
5.60. Zusammensetzung $P_{t_{1,2,3,4}}$ nach Geschlecht und Schulform (Spaltenproz.)	58
5.61. Verzerrung der Paneldaten Münster 2002/2003 . . . . .	60
5.62. Zusammensetzung der Daten Münster 2000 nach Geschlecht und Schulform (Zeilenproz.) . . . . .	61
A.1. Häufigkeiten der Codes in Münster . . . . .	63
A.2. Häufigkeiten der aufgetretenen Codes in Münster 2003 (erweiterter Code)	63
A.3. Häufigkeiten von Co001 in Münster (Haarfarbe Vater) . . . . .	70
A.4. Häufigkeiten von Co002 in Münster (Vorname Oma) . . . . .	71
A.5. Häufigkeiten von Co003 in Münster (eigene Haarfarbe) . . . . .	72
A.6. Häufigkeiten von Co004 in Münster (eigener Geburtstag) . . . . .	73
A.7. Häufigkeiten von Co005 in Münster (eigene Augenfarbe) . . . . .	74
A.8. Häufigkeiten von Co009 in Münster (eigener Vorname) . . . . .	75
A.9. Häufigkeiten von Co006 in Münster 2001 (Klassenwechsel ja/nein) . . .	76
A.10. Häufigkeiten von Co007 in Münster 2001 (Schulwechsel ja/nein) . . . .	76
A.11. Häufigkeiten von Co006 in Münster 2002 (Klassenwechsel ja/nein) . . .	76
A.12. Häufigkeiten von Co007 in Münster 2002 (Schulwechsel ja/nein) . . . .	77
A.13. Häufigkeiten von Co008 in Münster 2002 (BefragungsteilnehmerIn ja/nein)	77
A.14. Häufigkeiten von Co006 in Münster 2003 (Sitzengeblieben ja/nein) . . .	77
A.15. Häufigkeiten von Co007 in Münster 2003 (Schulwechsel ja/nein) . . . .	78



A.16. Häufigkeiten von Co008 in Münster 2003 (BefragungsteilnehmerIn ja/nein) 78



## Abbildungsverzeichnis

2.1. Erhebungsdesign der Längsschnittstudie 2000–2003 . . . . .	3
3.1. Geplantes Zuordnungsverfahren der Fragebögen in Münster 2000/2001	6
A.1. Codeblatt Münster 2000 . . . . .	66
A.2. Codeblatt Münster 2001 . . . . .	67
A.3. Codeblatt Münster 2002 . . . . .	68
A.4. Codeblatt Münster 2003 . . . . .	69



# 1. Einleitung

Die Ziele des DFG-Projektes *Kriminalität in der modernen Stadt*<sup>1</sup> liegen unter anderem in der Untersuchung der Entstehung und Entwicklung delinquenter bzw. devianter Handlungsstile, sowie deren Kontrollbedingungen bzw. Kontrollprozesse im Längsschnitt. Der methodische Zugang zu diesen Problemkomplexen erforderte ein für diesen theoretischen und inhaltlichen Kontext in Deutschland zum ersten Mal geplantes kohortenspezifisches Längsschnittdesign, das in zwei westdeutschen Städten unterschiedlicher Größe und Sozialstruktur, Münster und Duisburg, umgesetzt wurde. Die durch wiederholte Schülerbefragungen erhobenen Längsschnittinformationen (Paneldaten) sollen sowohl die Analyse von Ursache-Wirkungsbeziehungen zwischen den zentralen Untersuchungsvariablen (einschließlich möglicher Rückkopplungsprozesse) als auch die Analyse von Stabilität und Veränderung der individuellen Kriminalität ermöglichen.

Diesem Ansatz folgend wurden in den Jahren 2000, 2001, 2002 und 2003 kriminologische Jugendbefragungen in der Stadt Münster durchgeführt. Die bei den Befragungen erhobenen Daten wurden im Jahr 2003 zu einem Vier-Wellen-Panel zusammengefügt. Dieser Bericht dokumentiert zentrale Punkte der methodischen Vorgehensweise. Insbesondere werden das Studiendesign, das Zuordnungsverfahren und die realisierte Stichprobe beschrieben.

Eine weitere Analyse der Daten und Interpretationen der Ergebnisse sind nicht Gegenstand dieses als Methodendokumentation angelegten Berichtes. Ebenfalls ausgeklammert sind die Bereiche Instrumentenentwicklung, Fragebogengestaltung, Skalendokumentation und Validierung. Zeitvergleiche zwischen den Erhebungswellen werden nur auf der Ebene der Stichprobenszusammensetzung angestellt.

---

<sup>1</sup>Projektleitung: Prof. Dr. Klaus Boers, Institut für Kriminalwissenschaften, WWU Münster und Prof. Dr. Jost Reinecke, Abteilung Soziologie, Universität Trier, DFG-Förderzeichen: Bo 1234/6-1, 6-2 und Re 832/4-1, 4-2.



## 2. Studiendesign

Seit dem Jahr 2000 wurden in einjährigen Abständen an Münsteraner, seit dem Jahr 2002 auch an Duisburger Schulen Fragebogeninterviews durchgeführt. Entsprechend der methodischen Konzeption des Forschungsprojektes als Panelstudie wurden dabei dieselben Personen wiederholt befragt. Neben weiteren Querschnittsdaten, die im Jahr 2000 in Münster und im Jahr 2001 in Bocholt<sup>2</sup> zusätzlich erhoben wurden, liegt der Schwerpunkt der Längsschnittstudie auf der Population derjenigen Münsteraner Jugendlichen, die im Jahr 2000 die 7. Klasse, und denjenigen Duisburger Jugendlichen, die im Jahr 2002 die 7. und 9. Klasse besuchten.<sup>3</sup> Diese drei Populationen wurden im einjährigen Abstand verfolgt. Der genaue Erhebungsplan kann der Abbildung 2.1 entnommen werden. Da für das hier dokumentierte Vier-Wellen-Panel die Daten der Münsteraner Befragungen die Grundlage bilden, wird im Folgenden nur auf diese eingegangen.

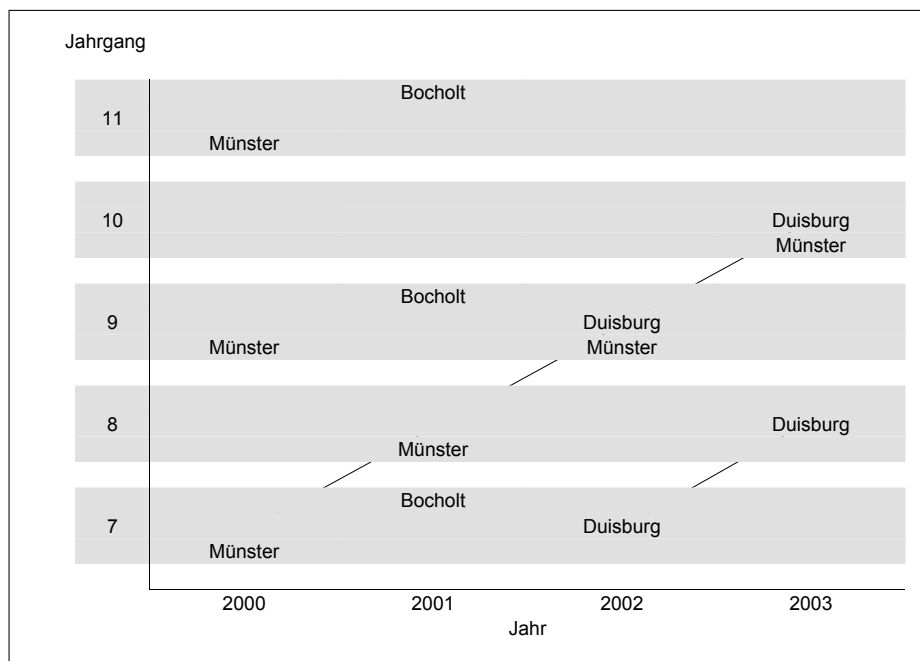


Abbildung 2.1.: Erhebungsdesign der Längsschnittstudie 2000–2003

In der *ersten* Münsteraner Schülerbefragung im Jahr 2000 wurde eine Vollerhebung der Schülerinnen und Schüler der Jahrgangsstufe 7 angestrebt.<sup>4</sup> Die *zweite* Schülerbe-

<sup>2</sup>Zur Dokumentation der Bocholter Studie aus dem Jahr 2001 siehe Wittenberg und Hilfert (2004).

<sup>3</sup>Zur Dokumentation der Duisburger Studie aus dem Jahr 2002 siehe Motzke und Brondies (2004), zur weiterführenden Studie aus dem Jahr 2003 siehe Brondies (2004).

<sup>4</sup>Ergänzt wurde die Studie durch Stichproben aus den Klassen des 9. und 11. Jahrgangs sowie den

fragung im Jahr 2001 strebte eine Wiederbefragung der Schülerinnen und Schüler aus der ersten Studie an. Alle Schülerinnen und Schüler der Jahrgangsstufe 8, die Schulen besuchten, die an der ersten Erhebung teilgenommen hatten, sollten wiederbefragt werden.<sup>5</sup> Der geschilderten Erhebungssystematik folgend sollten in der *dritten* Schülerbefragung im Jahr 2002 möglichst viele Schülerinnen und Schüler der Jahrgangsstufe 9, die Schulen besuchten, die an den ersten beiden Erhebungen teilgenommen hatten, wiederbefragt werden.<sup>6</sup> Die vorläufige Abschlussbefragung wurde in Münster mit der *vierten* Schülerbefragung im Jahr 2003 durchgeführt. Auch hier bestand das Ziel, im Zuge des Längsschnittdesigns eine möglichst hohe Anzahl an Schülerinnen und Schülern der Jahrgangsstufe 10, die die entsprechenden Schulen besuchten, wieder zu befragen.<sup>7</sup>

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

Vor den jeweiligen Haupterhebungen wurden die Eltern und Schüler über die Untersuchungsziele informiert und auf die Freiwilligkeit der Teilnahme hingewiesen. Die schriftlichen Befragungen fanden im Klassenverband statt. Der zeitliche Rahmen erstreckte sich über eine Schuldoppelstunde. In wenigen Einzelfällen kam es jedoch vor, dass die Beantwortung der Fragen mehr Zeit in Anspruch nahm, insbesondere wenn sprachliche Schwierigkeiten die Beantwortung beeinträchtigten.

---

Eingangsklassen an den Münsteraner Berufskollegs. Zur Dokumentation der Münsteraner Studie aus dem Jahr 2000 siehe Motzke und Wittenberg (2004).

<sup>5</sup>Zur Dokumentation der Münsteraner Studie aus dem Jahr 2001 siehe Wittenberg (2004a).

<sup>6</sup>Zur Dokumentation der Münsteraner Studie aus dem Jahr 2002 siehe Wittenberg (2004b).

<sup>7</sup>Zur Dokumentation der Münsteraner Studie aus dem Jahr 2003 siehe Wittenberg (2004c).



### **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 verschlüsselten Codes die Fragebögenzuordnungen zwischen den unterschiedlichen Jahren leisten sollte. Diese Vorgehensweise war mit der Datenschutzbehörde des Landes Nordrhein-Westfalen abgestimmt.

#### **3.1. Das Zuordnungsverfahren mit Krypto-Codes**

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 vom Fragebogen getrenntes Codeblatt ausfüllen, auf dem fünf persönliche Fragen gestellt wurden, aus deren Antwort jeweils ein bestimmter Buchstabe oder eine Zahl notiert werden musste. Die Fragen bezogen sich auf unveränderliche Merkmale des Probanden oder seines Umfeldes (natürliche Haarfarbe, Name des Vaters etc.). Diese Buchstaben-Zahlenkombination sollte nach einer Verschlüsselung 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. Codeblätter und zugehörige Fragebögen hatten jeweils eine eindeutige Nummer, so dass als Ergebnis die Fragebögen der verschiedenen Wellen über den Code einander zugeordnet werden sollten (siehe Abbildung 3.1).

Im Einzelnen waren die geplanten Schritte wie folgt: Während der Befragung im Klassenverband füllen die Schülerinnen und Schüler zunächst die Codeblätter und die Fragebögen sorgfältig aus. Auf einem Codeblatt und einem Fragebogen, die dieselbe Person ausfüllt, befindet sich dieselbe Nummer. Die Codeblätter und Fragebögen werden getrennt voneinander in Briefumschläge gesteckt, verschlossen und abgegeben. Codeblätter und Fragebögen gehen getrennt voneinander in die Datenerfassung, das einzige Verbindglied stellt die Nummer auf Codeblatt und Fragebogen dar. Nach der Datenerfassung werden die einzelnen Codes mit einem sicheren MD5-Verfahren verschlüsselt, d. h. in eine eindeutige, aber nicht zu entschlüsselnde 32-stellige Buchstaben- und Zahlenkombination, den so genannte Krypto-Code, umgewandelt. Die ursprünglichen Informationen wie Geburtstag, Haarfarbe etc. sind somit irreversibel verschlüsselt und können nicht reproduziert werden. Daraufhin werden die Codeblätter und die Daten mit

den unverschlüsselten Codes sowie die Original-Fragebögen vernichtet. Die verschlüsselten 32-stelligen Krypto-Codes können dann mit den Fragebogendaten anhand der Codeblatt- bzw. Fragebogennummer zusammengespielt werden, danach können diese Nummern ebenfalls vernichtet werden.

Das geschilderte Vorgehen wurde in jeder Erhebungswelle wiederholt, wodurch gewährleistet werden konnte, dass die Fragebogendaten über den Krypto-Code einander zugeordnet werden konnten, ohne dass jemals ursprüngliche Informationen aus den Codeblatt-Fragen mit Informationen aus den Fragebögen zusammengeführt werden mussten. Dieses Verfahren gewährleistet somit die Anonymität der befragten Personen.

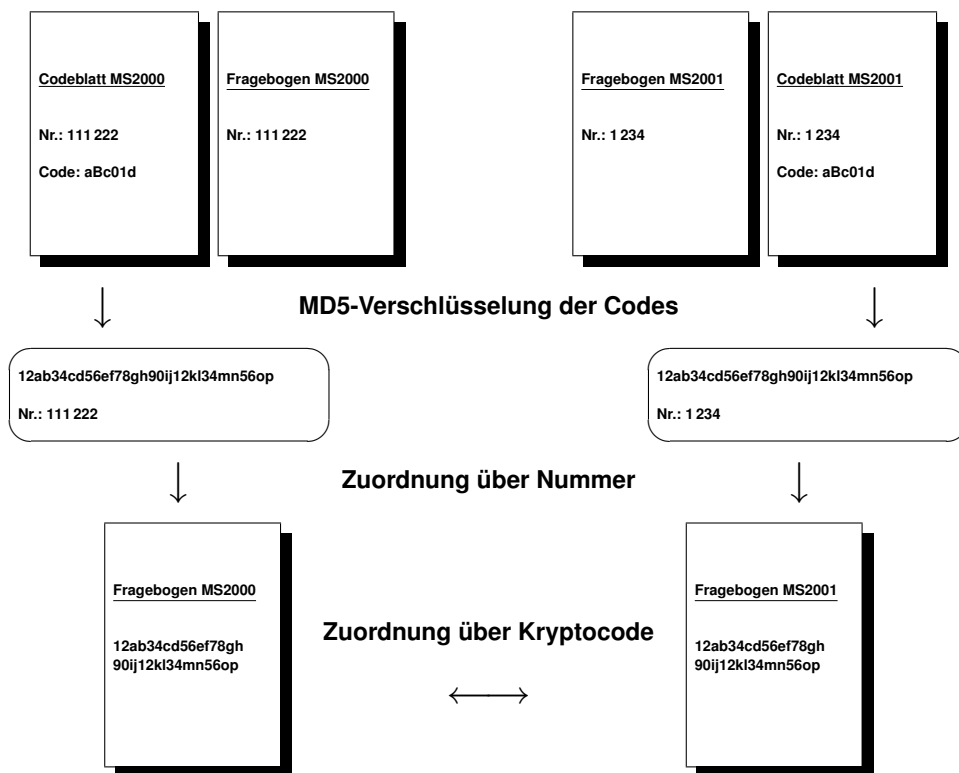


Abbildung 3.1.: Geplantes Zuordnungsverfahren der Fragebögen in Münster 2000/2001

Da die Codefragen die Privatsphäre der Befragten und deren Familien betreffen, musste trotz der geplanten Verschlüsselung dem Datenschutz dahingehend Rechnung getragen werden, dass nicht einfach die ersten Buchstaben der jeweiligen Antwort notiert werden konnten, sondern die jeweils letzten. Die Begründung hierfür lag in der Befürchtung, dass, falls die im Klassenverband ausgefüllten Codeblätter in die Hände von Personen mit genügend Information über die Schülerinnen und Schüler fielen, diese zu identifizieren sein könnten. Dann hätten die datenschutzrelevanten Informationen aus den Fragebögen im ungünstigsten Fall von unbefugten Personen den einzelnen

Schülerinnen und Schülern zugeordnet werden und diesen hieraus Nachteile entstehen können.

Während der ersten Befragung im Jahr 2000 sollte auf den Codeblättern<sup>8</sup> Folgendes handschriftlich notiert werden:

- Der letzte Buchstabe der natürlichen Haarfarbe des Vaters.
- Der erste Buchstabe des Vornamens der Oma mütterlicherseits.
- Der letzte Buchstaben der eigenen natürlichen Haarfarbe.
- Die beiden Tagesziffern des eigenen Geburtstages.
- Der letzte Buchstabe der eigenen Augenfarbe.

Im Jahr 2001 wurde das Codeblatt um die Fragen nach Klassenwechsel und Schulwechsel ergänzt, hierbei konnte eine Antwortvorgabe angekreuzt werden. Für die Befragung in 2002 wurde das Codeblatt um die Frage nach einer Befragungsteilnahme im vorigen Jahr erweitert. In 2003 wurde das Layout grundlegend geändert, so dass kein handschriftliches Ausfüllen erfolgen musste, sondern alle Antwortvorgaben zum Ankreuzen aufgeführt wurden. Außerdem wurde eine weitere Frage aufgenommen, nämlich die nach dem ersten Buchstaben des eigenen Vornamens.<sup>9</sup> Darüber hinaus wurde die Frage nach einem Klassenwechsel zum besseren Verständnis umformuliert.

### 3.2. Durchführung der Zuordnungen

Bei der technischen Durchführung der Zuordnungen der Fragebögen aus den einzelnen Wellen traten Probleme auf, die Änderungen des Verfahrens notwendig machten.

Analysiert man zunächst die Daten, die über die ausgefüllten Codeblätter zur Verfügung stehen, so fällt in Bezug auf die gebildeten Codes auf, dass nicht jede Schülerin und jeder Schüler einen eindeutigen Code aufweist.<sup>10</sup> In den vier Wellen treten nur ungefähr bei drei Vierteln der Probanden eindeutige Codes auf. Bei einem Viertel der Befragten weisen unterschiedliche Personen einen gleichen Code auf. Diese Personen sind über den Code ohne Hinzunahme weiterer Informationen nicht eindeutig zu identifizieren. Im Folgenden soll dieses Problem als Problem der *Identifizierung* bezeichnet werden.

Wählt man aus jeder Welle nur die eindeutigen Codes aus und ordnet zwischen den Wellen diese Codes einander zu, so finden sich zwischen jeweils zwei Erhebungszeitpunkten lediglich 500 bis 600 Zuordnungen. Man kann also davon ausgehen, dass

---

<sup>8</sup>Abdrucke der Codeblätter aller Befragungszeitpunkte finden sich im Anhang Abbildungen A.1 bis A.4

<sup>9</sup>Diese Frage blieb jedoch ohne Bedeutung, da in Münster auf eine weitere Befragungswelle verzichtet wurde.

<sup>10</sup>Zu den aufgetretenen Häufigkeiten des gesamten Codes siehe im Anhang Tabellen A.1 und A.2. Die Häufigkeiten der einzelnen Codevariablen finden sich ebenfalls im Anhang in den Tabellen A.3 bis A.16.

gleiche Personen in unterschiedlichen Wellen nicht denselben Code aufweisen.<sup>11</sup> Die hauptsächliche Ursache hierfür liegt darin, dass die Schülerinnen und Schüler den Code nicht exakt in jeder Welle reproduzieren konnten. Dieses Problem soll im Folgenden als Problem der *Reproduktion* bezeichnet werden.

Nachfolgend werden die Ursachen für die aufgetretenen Probleme analysiert und die schließlich angewendete Problemlösestrategie erläutert.

### 3.2.1. Probleme der Zuordnungsmethode

Die Ursachen für die aufgetretenen Zuordnungsprobleme der Identifizierung und Reproduktion sind unterschiedlicher Art und überlappen sich teilweise. Hier soll nach *exogenen* und *endogenen* Ursachen unterschieden werden. Als exogene Gründe für die Zuordnungsproblematik sind hier solche zu verstehen, die vom Forschungsteam nicht direkt und als endogene solche, die direkt durch das Team beeinflussbar waren. Exogene Gründe waren vor allem:

- unleserliche Handschriften auf den Codeblättern
- Abstimmung mit dem Datenschutz, die zu komplizierten Fragestellungen führte, welche das Ausfüllen der Codeblätter erschwerte
- mangelnde Reproduktion der Codes durch die Schüler aufgrund von Flüchtigkeitsfehlern oder absichtlichen Fehlern beim Ausfüllen der Codeblätter

Endogene Gründe:

- Layoutmängel des Codeblatts, die fehlerhaftes oder unleserliches Ausfüllen begünstigten
- zu hohe Schwierigkeit der Codefragen, die eine Reproduzierbarkeit erschwerte
- ungünstige Auswahl der Fragen bzw. Buchstaben, die eine eindeutige Identifizierung der Codes verhinderte

#### **Unleserliche Handschriften**

Ein nicht vom Forschungsteam direkt beeinflussbares Problem, welches in diesem Ausmaß nicht zu erwarten war, war das zum Teil unleserliche Ausfüllen der Codeblätter. Dies führte dazu, dass in einer Reihe von Fällen der Code nicht eindeutig zu entziffern war, nicht für die EDV zur Verfügung stand und deshalb keine Zuordnungen anzeigen konnte. Eine Lösung dieses Problems ist für die Datensätze 2000, 2001 und 2002 nicht möglich, jedoch wurde das Codeblattlayout dergestalt geändert, dass in der letzten Erhebungswelle kein handschriftliches Ausfüllen mehr nötig war, sondern nur noch Antwortvorgaben angekreuzt werden mussten.

---

<sup>11</sup>Einige Probanden aus verschiedenen Jahren, wiesen sogar einen gleichen Code auf, obwohl es sich nicht um dieselben Personen handelte. Im Vergleich zu den Größenordnungen der genannten Probleme, war dies jedoch unbedeutend.

#### **Datenschutzvorgaben**

Durch die nötige Abstimmung mit dem Datenschutz musste bei der Fragengestaltung Rücksicht genommen werden, so dass als Kompromiss verwirrende Fragestellungen in das Codeblatt eingehen mussten. So war der Wechsel von „erster Buchstabe ...“ und „letzter Buchstabe ...“ für die ausfüllenden Personen vermutlich zu unübersichtlich.

#### **Beabsichtigte Fehler und Flüchtigkeitsfehler**

Bei der Beantwortung der einzelnen Fragen auf dem Codeblatt können natürlich sowohl beabsichtigte als auch unbeabsichtigte Fehler auftreten, so dass der Code ein und desselben Schülers oder derselben Schülerin aus zwei Jahren unterschiedlich sein kann, obwohl er gleich sein müsste. Grundsätzlich kann dieses Problem bei jeder Frage auftreten. Absichtliche Falschangaben können nicht vermieden werden, leicht und eindeutig zu beantwortende Fragen sollten aber helfen, unbeabsichtigte Fehler zu vermeiden. Insofern kann man hier davon sprechen, dass zwar Flüchtigkeitsfehler und absichtliche Fehler vom Forschungsteam nicht beeinflussbar sind, durch die Art der Fragen aber die Möglichkeit besteht, das Auftreten von Fehlern, die aufgrund der Schwierigkeit der Fragen gemacht werden, zu reduzieren.

#### **Layoutmängel des Codeblattes**

Bei der Dateneingabe der handschriftlich ausgefüllten Codeblätter zeigte sich, dass ein Teil der Handschriften der Schülerinnen und Schüler unleserlich war. Ein Layout mit vorgegebenen Antwortkästchen, die nur angekreuzt werden müssen, erscheint unter diesem Gesichtspunkt als sehr viel vorteilhafter. Solch ein Layout wurde deshalb auch in der letzten Erhebungswelle angewendet. Zwar verschlechtert sich dadurch die Übersichtlichkeit des Codeblattes leicht, jedoch wiegt der Vorteil der besseren Lesbarkeit dieses kleinere Problem wieder auf.

#### **Zu hoher Schwierigkeitsgrad der Codefragen**

Da ein erheblicher Teil der Schülerinnen und Schüler nicht in der Lage war, das Codeblatt in allen Jahren gleich, d. h. also richtig auszufüllen, muss die Frage gestellt werden, ob der Schwierigkeitsgrad der Fragen zu hoch war. Möglicherweise ist ein nicht unerheblicher Anteil Schülerinnen und Schüler der 7. bis 10. Jahrgangsstufe nicht in der Lage, die oben genannten Fragen richtig zu beantworten oder sich bei der Beantwortung genügend zu konzentrieren. Durch das Forschungsteam beeinflussbar war natürlich die Auswahl der Fragen, soweit sie datenschutzrechtlich unbedenklich waren. Hier war man zu dem Entschluss gekommen, den Probanden die genannten Fragen vorlegen zu können. Zwar waren einige Fragen schwieriger zu beantworten (z. B. Haarfarbe des Vaters oder Vorname der Oma mütterlicherseits) falls persönliche Gründe die Beantwortung schwierig oder unmöglich machten (z. B. beim Fehlen einer solchen Bezugsperson). Jedoch waren die Interviewerinnen und Interviewer diesbezüglichen angewiesen worden, die problematischen Fälle vor dem Ausfüllen anzusprechen bzw. in diesen Fällen Antwortvorgaben zu nennen („O“ für Oma, „E“ für Glatze etc.).

### Ungünstige Auswahl und Anzahl an Fragen für die Identifizierung

Neben der Tatsache, dass Codes, die in beiden Jahren gleich sein müssten, unterschiedlich sind, ist das Auftreten von Mehrfachcodes das gewichtigste Problem. Wie oben schon erläutert, sind nur drei Viertel der Codes in den vier Wellen eindeutig. Ein Viertel der Codes tritt in der jeweiligen Welle mehrfach auf und kann daher nicht eindeutig identifiziert werden. Offensichtlich eignet sich der Code nicht dazu, die rund 2 000 Umfrageteilnehmerinnen und -teilnehmer eindeutig zu bestimmen. Dieser Code setzt sich, wie oben erläutert, aus Buchstaben bzw. Zahlen zusammen, die aus den Antworten der fünf Codefragen hervorgehen. Die Anzahl der Kombinationsmöglichkeiten, die sich aus diesen Zahlen und Buchstaben ergeben, ist nicht groß genug, dass mit genügender Wahrscheinlichkeit jede Schülerin und jeder Schüler einen einzigartigen Code erhält.<sup>12</sup> Festzuhalten bleibt, dass zusätzliche Informationen für eine eindeutige Identifizierung der Codeblätter bzw. Fragebögen schon in einer Welle hinzugezogen werden müssen. Eine Lösung dieses Problems kann nur in der Modifizierung der Fragen bestehen, oder in der Erhöhung der Anzahl der Codevariablen. Eine Erhöhung der Anzahl der Codevariablen hat den Nachteil, dass dadurch natürlich auch die Fehleranfälligkeit des Codes weiter steigt, eine Änderung der Fragen hat den großen Nachteil, dass der Code über die Zeit dann nicht mehr derselbe sein kann.<sup>13</sup>

#### 3.2.2. Problemlösungsstrategie

Um eine befriedigende Anzahl an Zuordnungen realisieren zu können, wurde das ursprüngliche Zuordnungsverfahren folgendermaßen abgeändert:

Um Fehler im Code zulassen zu können, musste auf die Kryptographierung der Codes als *komplette* Zeichen-Zahlenkette für die Phase der Zuordnungsfindungen verzichtet werden. Nach solch einer Verschlüsselung wären die Informationen der einzelnen Antworten auf die Codefragen unkenntlich. Die geplante Komplett-Verschlüsselung war erst nach der Bearbeitungsphase möglich. Bei der Identifizierung sowohl der einzelnen Codes als auch bei deren Zuordnung zwischen den Wellen wurden darüber hinaus die Schulinformation<sup>14</sup> und das Geschlecht der Probanden hinzugenommen.

Das schließlich 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 Fra-

<sup>12</sup>Zu diesem Problem siehe im Anhang Abschnitt A.2.

<sup>13</sup>Im letzten Jahr der Münsteraner Befragung wurde der Code um eine Frage erweitert. Dies führte zu einer drastischen Reduzierung der Mehrfachcodes (siehe im Anhang Tabelle A.2). Diese Verbesserung blieb in Münster ohne Bedeutung, konnte aber für die Duisburger Befragung gewinnbringend eingesetzt werden.

<sup>14</sup>Die Hinzunahme der Schulinformation war beim ursprünglichen Verfahren für die Identifizierung allerdings schon einkalkuliert worden.

gebogennummern 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. Auch in diesen beiden Schritten wurden als Validierung der Zuordnungen Handschriftenvergleiche durchgeführt, die offensichtlich nicht passenden Zuordnungen verworfen und vor der Durchführung des nächsten Schrittes die passenden Nummern aus dem Datensatz entfernt.

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





## 4. Realisierte Zuordnungen

Mit Hilfe des modifizierten Zuordnungsverfahrens konnten für die Paneldatensätze die in Tabelle 4.1 aufgeführten Fallzahlen realisiert werden. Für die geplanten Längsschnittanalysen sind neben den Zwei-Wellen-Panels die lückenlosen Drei-Wellen-Panels 2000 bis 2002 ( $t_{1,2,3}$ ) mit **997** Fällen und 2001 bis 2003 ( $t_{2,3,4}$ ) mit **1 075** Fällen sehr bedeutsam. Das endgültige und vollständige Resultat der Längsschnitt-Erhebungen in Münster stellt jedoch der Datensatz des Vier-Wellen-Panels der Jahre 2000 bis 2003 ( $t_{1,2,3,4}$ ) mit seinen **813** Fällen dar.<sup>15</sup>

### 4.1. Die verwendeten Datensätze und ihre Kurzbezeichnungen

Im Folgenden sollen der Einfachheit halber 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. Auch hierbei sind nur die Schulen berücksichtigt, die zu allen Zeitpunkten an den Befragungen teilnahmen („Trend netto“).  $S_{t_3}$  enthält alle (2 251 Fälle) der Schulstatistik des Schuljahres 2001/2002, wobei wiederum nur die zu allen Zeitpunkten teilnehmenden Schulen berücksichtigt sind.

Ein „Q“ mit einem Zeitpunkt als Index steht für einen ursprünglichen Querschnittsdatsatz der Einzelbefragungen. Hierbei handelt es sich um die Querschnittsdaten der Schülerinnen und Schüler, die eine Schule besuchten, die zu allen Zeitpunkten an den Befragungen teilnahm („Trend netto“) (vgl. Wittenberg 2004a, S. 12).  $Q_{t_1}$  beinhaltet alle (1 850) gültigen Fälle des Jahres 2000, wobei nur die Schulen berücksichtigt sind, die auch an allen weiteren Befragungszeitpunkten teilnahmen.

Generell werden die einzelnen Paneldaten durch ein „P“ mit einem Index, der die Zeitpunkte ausdrückt, bezeichnet. So steht  $P_{t_2,3}$  für den Paneldatensatz mit allen (1 373) gültigen Fällen, die zwischen  $t_2$  und  $t_3$  einander zugeordnet werden konnten.

Ein „ $P_Q$ “ mit einem Zeitpunkt als Index bezeichnet einen Panelquerschnitt. Diese Panelquerschnitte beinhalten *alle* vorhandenen Fälle des gesamten Datensatzes aus *einem* Zeitpunkt. So enthält  $P_{Q_{t_2}}$  alle (1 668) gültigen Fälle des Datensatzes, die für  $t_2$  vorhanden sind. Dies sind im Unterschied zum Querschnitt  $Q_{t_2}$  alle Fälle, die im Jahre 2001 befragt wurden und mindestens *einem* Fall *eines* anderen Zeitpunktes zugeordnet werden konnten.

---

<sup>15</sup>Berücksichtigt sind hier nur die gültigen Fälle, die auch den Plausibilitätskontrollen der einzelnen Querschnitte standhielten.

Tabelle 4.1.: Bezeichnungen und Fallzahlen der Münsteraner Datensätze 2000–2003

Abkürzung	Fallzahl ( $N_b$ )	Beschreibung
$S_{t_1}$	2 105	Amtliche Schulstatistik, Schuljahr 1999/2000 (Stichtag: 15. 10.) <sup>a</sup>
$S_{t_2}$	2 181	Amtliche Schulstatistik, Schuljahr 2000/2001 (Stichtag: 15. 10.) <sup>a</sup>
$S_{t_3}$	2 251	Amtliche Schulstatistik, Schuljahr 2001/2002 (Stichtag: 15. 10.) <sup>a</sup>
$S_{t_4}$	2 077	Amtliche Schulstatistik, Schuljahr 2002/2003 (Stichtag: 15. 10.) <sup>a</sup>
$Q_{t_1}$	1 850	Querschnittsdaten des ersten Befragungsjahres (2000 bzw. $t_1$ ) <sup>a</sup>
$Q_{t_2}$	1 915	Querschnittsdaten des zweiten Befragungsjahres (2001 bzw. $t_2$ ) <sup>a</sup>
$Q_{t_3}$	1 947	Querschnittsdaten des dritten Befragungsjahres (2002 bzw. $t_3$ ) <sup>a</sup>
$Q_{t_4}$	1 819	Querschnittsdaten des vierten Befragungsjahres (2003 bzw. $t_4$ ) <sup>a</sup>
$P_{Q_{t_1}}$	1 382	Panelquerschnittsdaten des Jahres 2000 <sup>b</sup>
$P_{Q_{t_2}}$	1 668	Panelquerschnittsdaten des Jahres 2001 <sup>b</sup>
$P_{Q_{t_3}}$	1 725	Panelquerschnittsdaten des Jahres 2002 <sup>b</sup>
$P_{Q_{t_4}}$	1 490	Panelquerschnittsdaten des Jahre 2003 <sup>b</sup>
$P_{t_{1,2}}$	1 271	Paneldaten der Jahre 2000 und 2001 <sup>b</sup>
$P_{t_{2,3}}$	1 373	Paneldaten der Jahre 2001 und 2002 <sup>b</sup>
$P_{t_{3,4}}$	1 406	Paneldaten der Jahre 2002 und 2003 <sup>b</sup>
$P_{t_{1,3}}$	1 052	Paneldaten der Jahre 2000 und 2002 <sup>b</sup>
$P_{t_{1,4}}$	910	Paneldaten der Jahre 2000 und 2003 <sup>b</sup>
$P_{t_{2,4}}$	1 105	Paneldaten der Jahre 2001 und 2003 <sup>b</sup>
$P_{t_{1,2,3}}$	997	Paneldaten der Jahre 2000 bis 2002 <sup>b</sup>
$P_{t_{2,3,4}}$	1 075	Paneldaten der Jahre 2001 bis 2003 <sup>b</sup>
$P_{t_{1,2,4}}$	823	Paneldaten der Jahre 2000, 2001 und 2003 <sup>b</sup>
$P_{t_{1,3,4}}$	851	Paneldaten der Jahre 2000, 2002 und 2003 <sup>b</sup>
$P_{t_{1,2,3,4}}$	813	Paneldaten der Jahre 2000 bis 2003 <sup>b</sup>

$N_b$  bedeutet beobachtetes  $N$

<sup>a</sup> nur Schulen, die an allen Befragungszeitpunkten teilnahmen

<sup>b</sup> gültige Fälle zu allen entsprechenden Befragungszeitpunkten

## 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*.<sup>16</sup> Auf Schulebene beschränkt sich hierbei

<sup>16</sup>Dazu folgendes Beispiel: Zum Zeitpunkt  $t_1$  gingen 2 105 Schülerinnen und Schüler auf die von uns befragten Schulen, 1 850 von ihnen konnten von uns (verwertbar) befragt werden, 255 nicht. Zum Zeitpunkt  $t_2$  waren es laut Schulstatistik 2 181 Probanden, von denen wir 1 915 befragen konnten, 266 nicht. Wie viele der Probanden aus  $t_1$  haben wir zu  $t_2$  wiederbefragt bzw. wie viele der Fragebögen aus  $t_1$  können Bögen aus  $t_2$  theoretisch zugeordnet werden? Nimmt man an, dass die Schulpopulation (relativ) stabil blieb, kann man zwei Extrempunkte formulieren. Zum einen könnten alle 1 850 wie-

die Grundgesamtheit zunächst auf die Jugendlichen der befragten Jahrgänge, die eine *Schule* besuchten, die zu allen Zeitpunkten der Befragung zustimmten („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* (vier Jahre) im Sinne von Zu- und Wegzügen nach bzw. aus Münster, 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.<sup>17</sup>

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.

---

derbefragt worden sein, die 255 Ausfälle ( $t_1$ ) sind in den 266 Ausfällen ( $t_2$ ) enthalten. Zum anderen könnten alle 255 Ausfälle ( $t_1$ ) zum zweiten Zeitpunkt wiederbefragt worden sein, alle 266 Ausfälle ( $t_2$ ) könnten allerdings im ersten Zeitpunkt schon befragt worden sein, so dass nur noch 1 329 Personen zu beiden Zeitpunkten befragt wurden. Die realisierten Zuordnungen von 1 271 ergeben somit eine Quote von 96 Prozent oder 68 Prozent. Beides ist unwahrscheinlich.

<sup>17</sup>Lässt man diese Betrachtungen außer Acht, so sind noch 813 Jugendliche von den ursprünglichen 1 850 im Vier-Wellen-Panel enthalten. Dies sind Jugendliche, die im beobachteten Zeitraum dreimal versetzt wurden, nie umgezogen sind, nie auf eine Schule außerhalb der Stichprobe umgeschult wurden, nie an den Befragungstagen gefehlt haben und bei jeder Befragung einen hinreichenden Code geliefert haben. Von einem sehr konservativen Standpunkt aus kann hieraus mit einer gewissen Berechtigung eine Panelmortalität von 56 % abgeleitet werden.

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 Querschnittsstichproben ( $N(Q_{t_i})$ ) und den Schulpopulationen ( $N(S_{t_i})$ ).<sup>18</sup>

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

Aus den zu erwartenden Anzahlen ( $N_e$ ) und den tatsächlich beobachteten bzw. realisierten Anzahlen der Zuordnungen ( $N_b$ ) lassen sich die in Tabelle 4.2 angeführten Quoten berechnen. Sie dürften, wie oben angedeutet, die „wahre“ Zuordnungsquote eher unterschätzen, da auf aggregierter Ebene der Schulstatistiken nicht 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

	$N_e$	$N_b$	Zuordnungsquote (%)
$P_{Q_1}$	1850	1382	74,7
$P_{Q_2}$	1915	1668	87,1
$P_{Q_3}$	1947	1725	88,6
$P_{Q_4}$	1819	1490	81,9
$P_{t_{1,2}}$	1683	1271	75,5
$P_{t_{2,3}}$	1710	1373	80,3
$P_{t_{3,4}}$	1705	1406	82,2
$P_{t_{1,3}}$	1711	1052	61,5
$P_{t_{1,4}}$	1620	910	56,2
$P_{t_{2,4}}$	1677	1105	65,9
$P_{t_{1,2,3}}$	1502	997	66,4
$P_{t_{2,3,4}}$	1497	1075	71,8
$P_{t_{1,2,4}}$	1474	823	55,8
$P_{t_{1,3,4}}$	1499	851	56,8
$P_{t_{1,2,3,4}}$	1316	813	61,8

<sup>18</sup>Anders ausgedrückt entspricht dies dem Produkt der größten Ausgangsstichprobe und den multiplizierten Rücklaufquoten der beteiligten Querschnitte.

Um die Zuordnungsquoten der realisierten Paneldatensätze<sup>19</sup> 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 Addresspanels deutlich größer wird.

Die Zuordnungsquoten für die Zwei-Wellen-Panels aus jeweils zwei aufeinanderfolgenden Zeitpunkten liegen zwischen 76 und 82 Prozent. Es ist ersichtlich, dass die Quoten im Verlauf der Befragungen besser wurden. So wurde die schlechteste Zuordnungsquote zwischen den Jahren 2000 ( $t_1$ ) und 2001 ( $t_2$ ), die beste zwischen den Jahren 2002 ( $t_3$ ) und 2003 ( $t_4$ ) verwirklicht; auch die absolute Fallzahl stieg entsprechend an. Dies mag zum einen am steigenden Alter der Befragten und der möglicherweise damit verbundenen höheren Fähigkeit, die Codefragen richtig zu beantworten bzw. sich zu konzentrieren, liegen. Zum anderen mag auch ein positive Gewöhnungseffekt bei den Probanden eine Rolle spielen. Die Quoten der beiden lückenlosen Drei-Wellen-Panels liegen bei 66 und 72 Prozent, die des Vier-Wellen-Panels erwartungsgemäß niedriger bei 62 Prozent.

Die Zuordnungsquoten der Paneldatensätze, die ein oder zwei fehlende Zeitpunkte aufweisen, sind naturgemäß niedriger. Dies liegt zum Teil daran, dass beim oben beschriebenen Zuordnungsverfahren die Validierung per Handschriftenvergleich durch veränderte Handschriften immer schwieriger wird, je größer der dazwischenliegende Zeitraum ist, zum Teil auch an den oben genannten Annahmen bei der Schätzung der zu erwartenden Zuordnungsfallzahlen. Die nicht „sichtbaren“ Veränderungen der Grundgesamtheit auf der Aggregatebene, die von der oben beschriebenen Formel 4.1 nicht exakt berücksichtigt werden können, werden natürlich im Laufe der Zeit immer gravierender.

---

<sup>19</sup>Im Falle der Panelquerschnitte bezeichnet die Zuordnungsquote den Anteil derjenigen Bögen, zu denen in irgendeinem weiteren Jahr ein passender gefunden werden konnte. Diese Quote ist in den späteren Jahren deutlich höher als zu Beginn der Befragungen. Die etwas niedrigere Quote des Jahres 2003 ( $t_4$ ) liegt unter anderem an der Tatsache, dass ein anschließender Datensatz fehlt.



## 5. Zusammensetzung der Paneldaten

In diesem Abschnitt soll die Zusammensetzung der einzelnen Paneldatensätze beschrieben werden. Zum Vergleich werden dabei die Daten der Querschnitte und der Schulstatistiken aus den entsprechenden Jahren hinzugezogen. Bei den vergleichenden Darstellungen werden die zentralen Variablen Geschlecht, Alter und Schulform verwendet.<sup>20</sup>

### 5.1. Die Panelquerschnitte

Wie oben bereits erwähnt, enthalten die Panelquerschnitte der einzelnen Jahre alle Fälle, deren zugehörige (und nach Plausibilitätskontrollen gültige) Fragebögen einem (ebenfalls gültigen) Bogen zumindest eines weiteren Zeitpunktes zugeordnet werden konnten. Sie enthalten dementsprechend nur diejenigen Fälle aus den Querschnitten nicht mehr, zu denen kein passender Bogen irgendeines weiteren Jahres gefunden werden konnte. Bei diesen Datensätzen handelt es sich also um stark konstruierte Daten, die nur unter bestimmten Analysegesichtspunkten wichtig sein können.<sup>21</sup>

#### 5.1.1. Panelquerschnitt 2000 ( $P_{Q_{t_1}}$ )

Der Panelquerschnitt des Jahres 2000 enthält 1 382 Fälle. Dies ist insofern von Bedeutung, als dass diese Anzahl die Obergrenze aller anschließenden Paneldatensätze darstellt, die das Jahr 2000 und weitere Zeitpunkte enthalten, da, wie oben beschrieben, in diesen Fällen alle überhaupt zugeordneten Fälle des Jahres 2000 enthalten sind.

Tabelle 5.1.: Zusammensetzung nach Geschlecht  $P_{Q_{t_1}}$

	männlich		weiblich		gesamt $N$ (100 %)
	$N$	%	$N$	%	
$S_{t_1}$	1 097	52,1	1 008	47,9	2 105
$Q_{t_1}$	931	50,3	919	49,7	1 850
$P_{Q_{t_1}}$	665	48,1	717	51,9	1 382

Wie in Tabelle 5.1 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine leichte Verschiebung vor. Schon im Querschnitt des Jahres 2000 sind im Vergleich zur

<sup>20</sup>Bei der Darstellung der Zusammensetzung der „echten“ Paneldaten (also ohne die Panelquerschnitte), 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 ist in unserem Fall nahezu fehlerlos, da sie extern erhoben wurde und nicht vom Antwortverhalten der Probanden abhängt.

<sup>21</sup>Wenn also genau die Population der überhaupt zuzuordnenden Fälle von Interesse ist.

Schulstatistik die Schülerinnen etwas überrepräsentiert. Im Panelquerschnitt ist dieser Unterschied deutlicher: es sind vier Prozent mehr Schülerinnen vorhanden als in der Schulstatistik ausgewiesen.

Tabelle 5.2.: Altersdurchschnitt  $P_{Q_{t_1}}$ 

	Alter	Std.-abw.	gesamt $N$
$Q_{t_1}$	13,0	0,8	1 845
$P_{Q_{t_1}}$	12,9	0,7	1 378

Das Alter der Jugendlichen, deren Fälle im Panelquerschnitt enthalten sind, weicht nur minimal vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.2). Hierbei gab es vier fehlende Angaben im Panelquerschnitt, fünf fehlende Angaben im Querschnitt. Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik bei nur einem Schuljahrgang kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

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

		Schulform <sup>a</sup> (%)				gesamt $N$
		HS	RS	GY	SO	(100 %)
männlich	$S_{t_1}$	23,8	30,6	39,6	6,0	1 097
	$Q_{t_1}$	20,3	31,1	43,0	5,6	931
	$P_{Q_{t_1}}$	19,2	30,5	46,8	3,5	665
weiblich	$S_{t_1}$	19,5	30,7	46,3	3,5	1 008
	$Q_{t_1}$	17,6	30,5	48,2	3,7	919
	$P_{Q_{t_1}}$	15,6	29,0	52,7	2,6	717
gesamt	$S_{t_1}$	21,8	30,6	42,8	4,8	2 105
	$Q_{t_1}$	19,0	30,8	45,6	4,6	1 850
	$P_{Q_{t_1}}$	17,4	29,7	49,9	3,0	1 382

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

Aus Tabelle 5.3 wird ersichtlich, dass der Panelquerschnitt 2000 im Vergleich zur Schulstatistik systematisch verzerrt ist. Im Paneldatensatz befinden sich insgesamt fast zwei Prozent weniger Sonderschülerinnen und -schüler, vier Prozent weniger Hauptschülerinnen und -schüler, dafür rund sieben Prozent mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik ausgewiesen. Der Anteil der Realschülerinnen und -schüler ist dabei annähernd gleich groß.

Über alle Schulformen hinweg sind prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.4). Der Unterschied zur Schulstatistik ist hierbei in der Sonderschule am größten und in der Realschule am geringsten. Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 5.5.



## 5.1. Die Panelquerschnitte

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

		Schulform <sup>a</sup> (%)				gesamt (%)
		HS	RS	GY	SO	
männlich	$S_{t_1}$	57,0	52,1	48,2	65,3	52,1
	$Q_{t_1}$	53,8	50,9	47,4	60,5	50,3
	$P_{Q_{t_1}}$	53,3	49,4	45,1	54,8	48,1
weiblich	$S_{t_1}$	43,0	47,9	51,8	34,7	47,9
	$Q_{t_1}$	46,2	49,1	52,6	39,5	49,7
	$P_{Q_{t_1}}$	46,7	50,6	54,9	45,2	51,9

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

### 5.1.2. Panelquerschnitt 2001 ( $P_{Q_{t_2}}$ )

Der Panelquerschnitt des Jahres 2001 enthält 1 668 Fälle. Dies sind alle überhaupt zuordenbaren Fälle des Jahres 2001.

Tabelle 5.5.: Zusammensetzung nach Geschlecht  $P_{Q_{t_2}}$

	männlich		weiblich		gesamt $N$ (100 %)
	$N$	%	$N$	%	
$S_{t_2}$	1 134	52,0	1 047	48,0	2 181
$Q_{t_2}$	978	51,1	937	48,9	1 915
$P_{Q_{t_2}}$	816	48,9	852	51,1	1 668

Wie in Tabelle 5.5 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine leichte Verschiebung vor. Schon im Querschnitt des Jahres 2001 sind im Vergleich zur Schulstatistik die Schülerinnen etwas überrepräsentiert. Im Panelquerschnitt ist dieser Unterschied deutlicher: es sind rund drei Prozent mehr Schülerinnen vorhanden als in der Schulstatistik ausgewiesen.

Tabelle 5.6.: Altersdurchschnitt  $P_{Q_{t_2}}$

	Alter	Std.-abw.	gesamt $N$
$Q_{t_2}$	14,0	0,8	1 889
$P_{Q_{t_2}}$	14,0	0,8	1 653

Das Alter der Jugendlichen, deren Fälle im Panelquerschnitt enthalten sind, weicht nicht vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.6). Hierbei gab es 15 fehlende Angaben im Panelquerschnitt, 26 fehlende Angaben im Querschnitt. Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach

Alter differenziert und die Einwohnermeldestatistik bei nur einem Schuljahrgang kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Tabelle 5.7.: Zusammensetzung  $P_{Q_{t_2}}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)				gesamt $N$ (100 %)
		HS	RS	GY	SO	
männlich	$S_{t_2}$	24,2	31,5	38,0	6,3	1 134
	$Q_{t_2}$	23,6	32,1	37,7	6,5	978
	$P_{Q_{t_2}}$	21,8	32,4	41,2	4,7	816
weiblich	$S_{t_2}$	20,4	31,9	43,6	4,1	1 047
	$Q_{t_2}$	19,7	31,8	44,8	3,6	937
	$P_{Q_{t_2}}$	19,1	30,6	47,4	2,8	852
gesamt	$S_{t_2}$	22,4	31,7	40,7	5,3	2 181
	$Q_{t_2}$	21,7	32,0	41,2	5,1	1 915
	$P_{Q_{t_2}}$	20,4	31,5	44,4	3,7	1 668

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

Aus Tabelle 5.7 wird ersichtlich, dass der Panelquerschnitt 2001 im Vergleich zur Schulstatistik systematisch verzerrt ist. Im Paneldatensatz befinden sich insgesamt rund eineinhalb Prozent weniger Sonderschülerinnen und -schüler, zwei Prozent weniger Hauptschülerinnen und -schüler, dafür rund dreieinhalb Prozent mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik ausgewiesen. Der Anteil der Realschülerinnen und -schüler ist dabei annähernd gleich groß. Die Abweichungen sind aufgrund der insgesamt größeren Stichprobe geringer als die des Panelquerschnitts 2000.

Über alle Schulformen hinweg sind prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.8). Der Unterschied zur Schulstatistik ist hierbei in der Hauptschule am größten und in der Sonderschule am geringsten. Auch diese Abweichungen sind geringer als im Panelquerschnitt des vorherigen Jahres. Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 5.5.

## 5.1. Die Panelquerschnitte

Tabelle 5.8.: Zusammensetzung  $P_{Q_{t_2}}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

		Schulform <sup>a</sup> (%)				gesamt (%)
		HS	RS	GY	SO	
männlich	$S_{t_2}$	56,1	51,7	48,6	62,6	52,0
	$Q_{t_2}$	55,5	51,3	46,8	65,3	51,1
	$P_{Q_{t_2}}$	52,2	50,3	45,4	61,3	48,9
weiblich	$S_{t_2}$	43,9	48,3	51,4	37,4	48,0
	$Q_{t_2}$	44,5	48,7	53,2	34,7	48,9
	$P_{Q_{t_2}}$	47,8	49,7	54,6	38,7	51,1

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

### 5.1.3. Panelquerschnitt 2002 ( $P_{Q_{t_3}}$ )

Der Panelquerschnitt des Jahres 2002 enthält 1 725 Fälle. Dies sind alle überhaupt zuordenbaren Fälle des Jahres 2002. Dieser Panelquerschnitt stellt damit den größten Datensatz in den erhobenen vier Zeitpunkten dar.

Tabelle 5.9.: Zusammensetzung nach Geschlecht  $P_{Q_{t_3}}$

	männlich		weiblich		gesamt $N$ (100 %)
	$N$	%	$N$	%	
$S_{t_3}$	1 170	52,0	1 081	48,0	2 251
$Q_{t_3}$	1 012	52,0	935	48,0	1 947
$P_{Q_{t_3}}$	868	50,3	857	49,7	1 725

Wie in Tabelle 5.9 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine leichte Verschiebung vor. Während sich das prozentuale Verhältnis zwischen Schülerinnen und Schülern des Querschnitts 2002 im Vergleich zur Schulstatistik nicht unterscheidet, sind im Panelquerschnitt knapp zwei Prozent mehr Schülerinnen vorhanden als in der Schulstatistik ausgewiesen.

Tabelle 5.10.: Altersdurchschnitt  $P_{Q_{t_3}}$

	Alter	Std.-abw.	gesamt $N$
$Q_{t_3}$	15,1	0,8	1 881
$P_{Q_{t_3}}$	15,0	0,8	1 689

Das Alter der Jugendlichen, deren Fälle im Panelquerschnitt enthalten sind, weicht so gut wie nicht vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.10). Hierbei gab es 36 fehlende Angaben im Panelquerschnitt, 66 fehlende Angaben im

Querschnitt. Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik bei nur einem Schuljahrgang kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Tabelle 5.11.: Zusammensetzung  $P_{Q_{t_3}}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)				gesamt $N$ (100 %)
		HS	RS	GY	SO	
männlich	$S_{t_3}$	26,1	31,9	35,0	7,1	1 170
	$Q_{t_3}$	22,9	33,0	37,9	6,1	1 012
	$P_{Q_{t_3}}$	22,1	32,8	41,0	4,0	868
weiblich	$S_{t_3}$	23,6	31,3	40,5	4,6	1 081
	$Q_{t_3}$	21,3	32,3	42,9	3,5	935
	$P_{Q_{t_3}}$	19,8	32,2	45,0	2,9	857
gesamt	$S_{t_3}$	24,9	31,6	37,6	5,9	2 251
	$Q_{t_3}$	22,1	32,7	40,3	4,9	1 947
	$P_{Q_{t_3}}$	21,0	32,5	43,0	3,5	1 725

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

Aus Tabelle 5.11 wird ersichtlich, dass der Panelquerschnitt 2002 im Vergleich zur Schulstatistik systematisch verzerrt ist. Im Paneldatensatz befinden sich insgesamt rund zweieinhalb Prozent weniger Sonderschülerinnen und -schüler, fast vier Prozent weniger Hauptschülerinnen und -schüler, dafür fast fünfeinhalb Prozent mehr Gymnasias-tinnen und Gymnasiasten als in der Schulstatistik ausgewiesen. Der Anteil der Real-schülerinnen und -schüler ist dabei annähernd gleich groß.

Über alle Schulformen hinweg sind prozentual mehr Schülerinnen und weniger Schü-ler als in der Schulstatistik vertreten (siehe Tabelle 5.12). Der Unterschied zur Schul-statistik ist hierbei in der Sonderschule am größten und im Gymnasium am geringsten. Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Ab-schnitt 5.5.

Tabelle 5.12.: Zusammensetzung  $P_{Q_{t_3}}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

		Schulform <sup>a</sup> (%)				gesamt (%)
		HS	RS	GY	SO	
männlich	$S_{t_3}$	54,5	52,5	48,3	62,4	52,0
	$Q_{t_3}$	53,8	52,5	48,9	65,3	52,0
	$P_{Q_{t_3}}$	53,0	50,8	48,0	58,3	50,3
weiblich	$S_{t_3}$	45,5	47,5	51,7	37,6	48,0
	$Q_{t_3}$	46,2	47,5	51,1	34,7	48,0
	$P_{Q_{t_3}}$	47,0	49,2	52,0	41,7	49,7

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

#### 5.1.4. Panelquerschnitt 2003 ( $P_{Q_{t_4}}$ )

Der Panelquerschnitt des Jahres 2003 enthält 1 490 Fälle. Dies sind alle überhaupt zuordenbaren Fälle des Jahres 2003. Dieser Panelquerschnitt enthält damit wegen der oben angesprochenen fehlenden Anschlusswelle wieder deutlich weniger Fälle als die Datensätze der beiden vorangegangenen Jahre, jedoch vermutlich wegen der besseren Codereproduktion deutlich mehr als im ersten Jahr. Dies ist insbesondere deshalb auffällig, da der Querschnitt dieses Jahres die wenigsten Fälle des gesamten beobachteten Zeitraums enthält.

Tabelle 5.13.: Zusammensetzung nach Geschlecht  $P_{Q_{t_4}}$

	männlich		weiblich		gesamt $N$ (100 %)
	$N$	%	$N$	%	
$S_{t_4}$	1 056	50,8	1 021	49,2	2 077
$Q_{t_4}$	915	50,4	901	49,6	1 816
$P_{Q_{t_4}}$	716	48,1	774	51,9	1 490

Wie in Tabelle 5.13 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine leichte Verschiebung vor. Während sich das prozentuale Verhältnis zwischen Schülerinnen und Schülern des Querschnitts 2003 im Vergleich zur Schulstatistik kaum unterscheidet, sind im Panelquerschnitt rund zweieinhalb Prozent mehr Schülerinnen vorhanden als in der Schulstatistik ausgewiesen. Im Querschnitt machten drei Personen keine Angaben.

Tabelle 5.14.: Altersdurchschnitt  $P_{Q_{t_4}}$

	Alter	Std.-abw.	gesamt $N$
$Q_{t_4}$	16,0	0,8	1 786
$P_{Q_{t_4}}$	16,0	0,8	1 472

Das Alter der Jugendlichen, deren Fälle im Panelquerschnitt enthalten sind, weicht nicht vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.14). Hierbei gab es 18 fehlende Angaben im Panelquerschnitt, 33 fehlende Angaben im Querschnitt. Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik bei nur einem Schuljahrgang kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Tabelle 5.15.: Zusammensetzung  $P_{Q_{t_4}}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)				gesamt $N$
		HS	RS	GY	SO	(100 %)
männlich	$S_{t_4}$	25,0	30,9	37,3	6,8	1056
	$Q_{t_4}$	23,5	31,3	40,2	5,0	915
	$P_{Q_{t_4}}$	19,6	32,8	43,9	3,8	716
weiblich	$S_{t_4}$	21,6	32,1	41,1	5,1	1021
	$Q_{t_4}$	22,1	32,3	42,3	3,3	901
	$P_{Q_{t_4}}$	20,2	32,6	45,0	2,3	774
gesamt	$S_{t_4}$	23,4	31,5	39,2	6,0	2077
	$Q_{t_4}$	22,8	31,8	41,2	4,2	1816
	$P_{Q_{t_4}}$	19,9	32,7	44,4	3,0	1490

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

Tabelle 5.16.: Zusammensetzung  $P_{Q_{t_4}}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

		Schulform <sup>a</sup> (%)				gesamt
		HS	RS	GY	SO	(%)
männlich	$S_{t_4}$	54,4	49,8	48,4	58,1	50,8
	$Q_{t_4}$	51,9	49,6	49,1	60,5	50,4
	$P_{Q_{t_4}}$	47,3	48,3	47,4	60,0	48,1
weiblich	$S_{t_4}$	45,6	50,2	51,6	41,9	49,2
	$Q_{t_4}$	48,1	50,4	50,9	39,5	49,6
	$P_{Q_{t_4}}$	52,7	51,7	52,6	40,0	51,9

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

Aus Tabelle 5.15 wird ersichtlich, dass der Panelquerschnitt 2003 im Vergleich zur Schulstatistik systematisch verzerrt ist. Im Paneldatensatz befinden sich insgesamt zwei Prozent weniger Sonderschülerinnen und -schüler, dreieinhalb Prozent weniger Hauptschülerinnen und -schüler, dafür rund fünf Prozent mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik ausgewiesen. Der Anteil der Realschülerinnen und -schüler ist dabei annähernd gleich groß.

Außer in der Sonderschule zum Zeitpunkt  $t_4$  sind über alle Schulformen hinweg prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.16). Der Unterschied zur Schulstatistik ist hierbei in der Hauptschule am größten und im Gymnasium am geringsten. Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 5.5.

## 5.2. Die Zwei-Wellen-Panels

Die Datensätze der Zwei-Wellen-Panels enthalten alle Fälle, deren zugehörige (und nach Plausibilitätskontrollen gültige) Fragebögen einem (ebenfalls gültigen) Bogen eines weiteren Zeitpunktes zugeordnet werden konnten. In diesem Bericht werden die Datensätze aller denkbaren Kombinationen der vier Zeitpunkte vorgestellt. In der Praxis werden vermutlich die Datensätze zweier benachbarter Zeitpunkte am bedeutsamsten sein.

### 5.2.1. Zwei-Wellen-Panel 2000/2001 ( $P_{t_1,2}$ )

Das Zwei-Wellen-Panel der Jahre 2000 und 2001 enthält 1 271 Fälle. Dies sind alle gültigen Fälle der Zeitpunkte  $t_1$  und  $t_2$ , deren Fragebögen einander zugeordnet werden konnten.

Tabelle 5.17.: Zusammensetzung nach Geschlecht  $P_{t_1,2}$

		männlich		weiblich		gesamt $N$
		$N$	%	$N$	%	(100 %)
$t_1$	$S_{t_1}$	1 097	52,1	1 008	47,9	2 105
	$Q_{t_1}$	931	50,3	919	49,7	1 850
	$P_{t_1,2}$	601	47,3	670	52,7	1 271
$t_2$	$S_{t_2}$	1 134	52,0	1 047	48,0	2 181
	$Q_{t_2}$	978	51,1	937	48,9	1 915
	$P_{t_1,2}$	603	47,4	668	52,6	1 271

Wie in Tabelle 5.17 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Die Unterschiede im prozentualen Verhältnis zwischen Schülerinnen und Schülern der Panelquerschnitte und Querschnitte 2000 und 2001 im Vergleich zur Schulstatistik wurden in den Abschnitten 5.1.1 und 5.1.2 dargestellt. Im Paneldatensatz sind sowohl zum Zeitpunkt  $t_1$  als auch  $t_2$  mehr Schülerinnen als in den Referenzdaten vorhanden. Die Abweichung fällt dabei mit fast fünf Prozent zum Zeitpunkt  $t_1$  noch etwas stärker aus als mit rund viereinhalb Prozent zum Zeitpunkt  $t_2$ .

Tabelle 5.18.: Altersdurchschnitt  $P_{t_1,2}$ 

		Alter	Std.-abw.	gesamt $N$
$t_1$	$Q_{t_1}$	13,0	0,8	1845
	$P_{t_1,2}$	13,0	0,7	1269
$t_2$	$Q_{t_2}$	14,0	0,8	1889
	$P_{t_1,2}$	13,9	0,7	1265

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 5.18). Hierbei gab es zu  $t_1$  2 und zu  $t_2$  6 fehlende Angaben im Paneldatensatz, zu  $t_1$  5 und zu  $t_2$  26 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 5.19.: Zusammensetzung  $P_{t_1,2}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt $N$
			HS	RS	GY	SO	(100 %)
männlich	$t_1$	$S_{t_1}$	23,8	30,6	39,6	6,0	1097
		$Q_{t_1}$	20,3	31,1	43,0	5,6	931
		$P_{t_1,2}$	19,3	30,1	47,1	3,5	601
	$t_2$	$S_{t_2}$	24,2	31,5	38,0	6,3	1134
		$Q_{t_2}$	23,6	32,1	37,7	6,5	978
		$P_{t_1,2}$	19,4	31,0	46,1	3,5	603
weiblich	$t_1$	$S_{t_1}$	19,5	30,7	46,3	3,5	1008
		$Q_{t_1}$	17,6	30,5	48,2	3,7	919
		$P_{t_1,2}$	15,4	29,3	52,8	2,5	670
	$t_2$	$S_{t_2}$	20,4	31,9	43,6	4,1	1047
		$Q_{t_2}$	19,7	31,8	44,8	3,6	937
		$P_{t_1,2}$	15,7	29,0	52,7	2,5	668
gesamt	$t_1$	$S_{t_1}$	21,8	30,6	42,8	4,8	2105
		$Q_{t_1}$	19,0	30,8	45,6	4,6	1850
		$P_{t_1,2}$	17,2	29,7	50,1	3,0	1271
	$t_2$	$S_{t_2}$	22,4	31,7	40,7	5,3	2181
		$Q_{t_2}$	21,7	32,0	41,2	5,1	1915
		$P_{t_1,2}$	17,5	30,0	49,6	3,0	1271

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

Aus Tabelle 5.19 wird ersichtlich, dass der Paneldatensatz 2000/2001 im Vergleich zur Schulstatistik systematisch verzerrt ist. Die Unterschiede im prozentualen Verhältnis der Querschnitte 2000 und 2001 im Vergleich zur Schulstatistik wurden in den



Abschnitten 5.1.1 und 5.1.2 dargestellt. Im Paneldatensatz befinden sich zu  $t_1$  und  $t_2$  insgesamt rund zwei Prozent weniger Sonderschülerinnen und -schüler, rund viereinhalb ( $t_1$ ) und fast fünf Prozent ( $t_2$ ) weniger Hauptschülerinnen und -schüler, dafür rund sieben ( $t_1$ ) und fast neun Prozent ( $t_2$ ) mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik ausgewiesen. Der Anteil der Realschülerinnen und -schüler ist dabei annähernd gleich groß.

Über alle Schulformen hinweg sind zu beiden Zeitpunkten prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.20). Der Unterschied zur Schulstatistik ist hierbei zu beiden Zeitpunkten in der Sonderschule am größten, zum Zeitpunkt  $t_1$  im Gymnasium und zum Zeitpunkt  $t_2$  in der Realschule am geringsten. Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 5.5.

Tabelle 5.20.: Zusammensetzung  $P_{1,2}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt
			HS	RS	GY	SO	(%)
männlich	$t_1$	$S_{t_1}$	57,0	52,1	48,2	65,3	52,1
		$Q_{t_1}$	53,8	50,9	47,4	60,5	50,3
		$P_{1,2}$	53,0	48,0	44,4	55,3	47,3
	$t_2$	$S_{t_2}$	56,1	51,7	48,6	62,6	52,0
		$Q_{t_2}$	55,5	51,3	46,8	65,3	51,1
		$P_{1,2}$	52,7	49,1	44,1	55,3	47,4
weiblich	$t_1$	$S_{t_1}$	43,0	47,9	51,8	34,7	47,9
		$Q_{t_1}$	46,2	49,1	52,6	39,5	49,7
		$P_{1,2}$	47,0	52,0	55,6	44,7	52,7
	$t_2$	$S_{t_2}$	43,9	48,3	51,4	37,4	48,0
		$Q_{t_2}$	44,5	48,7	53,2	34,7	48,9
		$P_{1,2}$	47,3	50,9	55,9	44,7	52,6

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

### 5.2.2. Zwei-Wellen-Panel 2001/2002 ( $P_{2,3}$ )

Das Zwei-Wellen-Panel der Jahre 2001 und 2002 enthält 1 373 Fälle. Dies sind alle gültigen Fälle der Zeitpunkte  $t_2$  und  $t_3$ , deren Fragebögen einander zugeordnet werden konnten.

Tabelle 5.21.: Zusammensetzung nach Geschlecht  $P_{t_2,3}$ 

		männlich		weiblich		gesamt $N$
		$N$	%	$N$	%	(100 %)
$t_2$	$S_{t_2}$	1134	52,0	1047	48,0	2181
	$Q_{t_2}$	978	51,1	937	48,9	1915
	$P_{t_2,3}$	658	47,9	715	52,1	1373
$t_3$	$S_{t_3}$	1170	52,0	1081	48,0	2251
	$Q_{t_3}$	1012	52,0	935	48,0	1947
	$P_{t_2,3}$	660	48,1	713	51,9	1373

Wie in Tabelle 5.21 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Die Unterschiede im prozentualen Verhältnis zwischen Schülerinnen und Schülern der Panelquerschnitte und Querschnitte 2001 und 2002 im Vergleich zur Schulstatistik wurden in den Abschnitten 5.1.2 und 5.1.3 dargestellt. Im Paneldatensatz sind sowohl zum Zeitpunkt  $t_2$  als auch  $t_3$  mehr Schülerinnen als in den Referenzdaten vorhanden. Die Abweichung fällt dabei mit rund vier Prozent zu beiden Zeitpunkten ungefähr gleich aus.

Tabelle 5.22.: Altersdurchschnitt  $P_{t_2,3}$ 

		Alter	Std.-abw.	gesamt $N$
$t_2$	$Q_{t_2}$	14,0	0,8	1889
	$P_{t_2,3}$	14,0	0,8	1362
$t_3$	$Q_{t_3}$	15,1	0,8	1881
	$P_{t_2,3}$	15,0	0,7	1347

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 5.22). Hierbei gab es zu  $t_2$  11 und zu  $t_3$  26 fehlende Angaben im Paneldatensatz, zu  $t_2$  26 und zu  $t_3$  66 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 5.23.: Zusammensetzung  $P_{t_2,3}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt $N$
			HS	RS	GY	SO	(100 %)
männlich	$t_2$	$S_{t_2}$	24,2	31,5	38,0	6,3	1 134
		$Q_{t_2}$	23,6	32,1	37,7	6,5	978
		$P_{t_2,3}$	21,4	32,8	42,1	3,6	658
	$t_3$	$S_{t_3}$	26,1	31,9	35,0	7,1	1 170
		$Q_{t_3}$	22,9	33,0	37,9	6,1	1 012
		$P_{t_2,3}$	22,1	32,3	42,0	3,6	660
weiblich	$t_2$	$S_{t_2}$	20,4	31,9	43,6	4,1	1 047
		$Q_{t_2}$	19,7	31,8	44,8	3,6	937
		$P_{t_2,3}$	19,3	30,2	48,3	2,2	715
	$t_3$	$S_{t_3}$	23,6	31,3	40,5	4,6	1 081
		$Q_{t_3}$	21,3	32,3	42,9	3,5	935
		$P_{t_2,3}$	18,9	30,6	48,1	2,4	713
gesamt	$t_2$	$S_{t_2}$	22,4	31,7	40,7	5,3	2 181
		$Q_{t_2}$	21,7	32,0	41,2	5,1	1 915
		$P_{t_2,3}$	20,3	31,5	45,3	2,9	1 373
	$t_3$	$S_{t_3}$	24,9	31,6	37,6	5,9	2 251
		$Q_{t_3}$	22,1	32,7	40,3	4,9	1 947
		$P_{t_2,3}$	20,5	31,4	45,2	3,0	1 373

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

Aus Tabelle 5.23 wird ersichtlich, dass der Paneldatensatz 2001/2002 im Vergleich zur Schulstatistik systematisch verzerrt ist. Die Unterschiede im prozentualen Verhältnis der Querschnitte 2001 und 2002 im Vergleich zur Schulstatistik wurden in den Abschnitten 5.1.2 und 5.1.3 dargestellt. Im Paneldatensatz befinden sich insgesamt rund zweieinhalb ( $t_2$ ) und fast drei Prozent ( $t_3$ ) weniger Sonderschülerinnen und -schüler, rund zwei ( $t_2$ ) und fast viereinhalb Prozent ( $t_3$ ) weniger Hauptschülerinnen und -schüler, dafür rund viereinhalb ( $t_2$ ) und rund siebeneinhalb Prozent ( $t_3$ ) mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik ausgewiesen. Der Anteil der Realschülerinnen und -schüler ist dabei annähernd gleich groß.

Über alle Schulformen hinweg sind zu beiden Zeitpunkten prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.24). Der Unterschied zur Schulstatistik ist hierbei zum Zeitpunkt  $t_2$  in der Hauptschule und zum Zeitpunkt  $t_3$  in der Sonderschule am größten, zum Zeitpunkt  $t_2$  in der Realschule und zum Zeitpunkt  $t_3$  in der Hauptschule am geringsten. Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 5.5.

Tabelle 5.24.: Zusammensetzung  $P_{t_{2,3}}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt
			HS	RS	GY	SO	(%)
männlich	$t_2$	$S_{t_2}$	56,1	51,7	48,6	62,6	52,0
		$Q_{t_2}$	55,5	51,3	46,8	65,3	51,1
		$P_{t_{2,3}}$	50,5	50,0	44,5	60,0	47,9
	$t_3$	$S_{t_3}$	54,5	52,5	48,3	62,4	52,0
		$Q_{t_3}$	53,8	52,5	48,9	65,3	52,0
		$P_{t_{2,3}}$	52,0	49,4	44,7	58,5	48,1
weiblich	$t_2$	$S_{t_2}$	43,9	48,3	51,4	37,4	48,0
		$Q_{t_2}$	44,5	48,7	53,2	34,7	48,9
		$P_{t_{2,3}}$	49,5	50,0	55,5	40,0	52,1
	$t_3$	$S_{t_3}$	45,5	47,5	51,7	37,6	48,0
		$Q_{t_3}$	46,2	47,5	51,1	34,7	48,0
		$P_{t_{2,3}}$	48,0	50,6	55,3	41,5	51,9

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

### 5.2.3. Zwei-Wellen-Panel 2002/2003 ( $P_{t_{3,4}}$ )

Das Zwei-Wellen-Panel der Jahre 2002 und 2003 ist mit 1 406 Fällen der größte Zwei-Wellen-Paneldatensatz. Enthalten sind alle gültigen Fälle der Zeitpunkte  $t_3$  und  $t_4$ , deren Fragebögen einander zugeordnet werden konnten.

Wie in Tabelle 5.25 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine leichte Verschiebung vor. Die Unterschiede im prozentualen Verhältnis zwischen Schülerinnen und Schülern der Panelquerschnitte und Querschnitte 2002 und 2003 im Vergleich zur Schulstatistik wurden in den Abschnitten 5.1.3 und 5.1.4 dargestellt. Im Paneldatensatz sind sowohl zum Zeitpunkt  $t_3$  als auch  $t_4$  mehr Schülerinnen als in den Referenzdaten vorhanden. Die Abweichung fällt dabei zum Zeitpunkt  $t_3$  mit rund dreieinhalb Prozent etwas stärker aus als zum Zeitpunkt  $t_4$  mit rund zweieinhalb Prozent. Im Querschnitt machten zu  $t_4$  drei Personen keine Angaben.

Tabelle 5.25.: Zusammensetzung nach Geschlecht  $P_{t_{3,4}}$ 

		männlich		weiblich		gesamt $N$
		$N$	%	$N$	%	(100 %)
$t_3$	$S_{t_3}$	1170	52,0	1081	48,0	2251
	$Q_{t_3}$	1012	52,0	935	48,0	1947
	$P_{t_{3,4}}$	681	48,4	725	51,6	1406
$t_4$	$S_{t_4}$	1056	50,8	1021	49,2	2077
	$Q_{t_4}$	915	50,4	901	49,6	1816
	$P_{t_{3,4}}$	676	48,1	730	51,9	1406

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 5.26). Hierbei gab es zu  $t_3$  26 und zu  $t_4$  13 fehlende Angaben im Paneldatensatz, zu  $t_3$  66 und zu  $t_4$  33 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 5.26.: Altersdurchschnitt  $P_{t_{3,4}}$

		Alter	Std.-abw.	gesamt $N$
$t_3$	$Q_{t_3}$	15,1	0,8	1 881
	$P_{t_{3,4}}$	15,0	0,8	1 380
$t_4$	$Q_{t_4}$	16,0	0,8	1 786
	$P_{t_{3,4}}$	16,0	0,7	1 393

Aus Tabelle 5.27 wird ersichtlich, dass der Paneldatensatz 2002/2003 im Vergleich zur Schulstatistik systematisch verzerrt ist. Die Unterschiede im prozentualen Verhältnis der Querschnitte 2002 und 2003 im Vergleich zur Schulstatistik wurden in den Abschnitten 5.1.3 und 5.1.4 dargestellt. Im Paneldatensatz befinden sich zu beiden Zeitpunkten insgesamt rund drei Prozent weniger Sonderschülerinnen und -schüler, fast sechs ( $t_3$ ) und vier Prozent ( $t_4$ ) weniger Hauptschülerinnen und -schüler, dafür rund siebeneinhalb ( $t_3$ ) und rund fünfeinhalb Prozent ( $t_4$ ) mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik ausgewiesen. Der Anteil der Realschülerinnen und -schüler ist dabei annähernd gleich groß.

Tabelle 5.27.: Zusammensetzung  $P_{t_3,4}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt $N$
			HS	RS	GY	SO	(100 %)
männlich	$t_3$	$S_{t_3}$	26,1	31,9	35,0	7,1	1170
		$Q_{t_3}$	22,9	33,0	37,9	6,1	1012
		$P_{t_3,4}$	18,5	33,0	44,9	3,5	681
	$t_4$	$S_{t_4}$	25,0	30,9	37,3	6,8	1056
		$Q_{t_4}$	23,5	31,3	40,2	5,0	915
		$P_{t_3,4}$	19,2	32,7	44,5	3,6	676
weiblich	$t_3$	$S_{t_3}$	23,6	31,3	40,5	4,6	1081
		$Q_{t_3}$	21,3	32,3	42,9	3,5	935
		$P_{t_3,4}$	19,7	32,6	45,5	2,2	725
	$t_4$	$S_{t_4}$	21,6	32,1	41,1	5,1	1021
		$Q_{t_4}$	22,1	32,3	42,3	3,3	901
		$P_{t_3,4}$	19,6	33,0	45,2	2,2	730
gesamt	$t_3$	$S_{t_3}$	24,9	31,6	37,6	5,9	2251
		$Q_{t_3}$	22,1	32,7	40,3	4,9	1947
		$P_{t_3,4}$	19,1	32,8	45,2	2,8	1406
	$t_4$	$S_{t_4}$	23,4	31,5	39,2	6,0	2077
		$Q_{t_4}$	22,8	31,8	41,2	4,2	1816
		$P_{t_3,4}$	19,4	32,9	44,9	2,8	1406

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

Tabelle 5.28.: Zusammensetzung  $P_{t_3,4}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt
			HS	RS	GY	SO	(%)
männlich	$t_3$	$S_{t_3}$	54,5	52,5	48,3	62,4	52,0
		$Q_{t_3}$	53,8	52,5	48,9	65,3	52,0
		$P_{t_3,4}$	46,8	48,8	48,1	60,0	48,4
	$t_4$	$S_{t_4}$	54,4	49,8	48,4	58,1	50,8
		$Q_{t_4}$	51,9	49,6	49,1	60,5	50,4
		$P_{t_3,4}$	47,6	47,8	47,7	60,0	48,1
weiblich	$t_3$	$S_{t_3}$	45,5	47,5	51,7	37,6	48,0
		$Q_{t_3}$	46,2	47,5	51,1	34,7	48,0
		$P_{t_3,4}$	53,2	51,2	51,9	40,0	51,6
	$t_4$	$S_{t_4}$	45,6	50,2	51,6	41,9	49,2
		$Q_{t_4}$	48,1	50,4	50,9	39,5	49,6
		$P_{t_3,4}$	52,4	52,2	52,3	40,0	51,9

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

Außer in der Sonderschule zum Zeitpunkt  $t_4$  sind über alle Schulformen hinweg zu beiden Zeitpunkten prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.28). Der Unterschied zur Schulstatistik ist hierbei zu beiden Zeitpunkten in der Hauptschule am größten und im Gymnasium am geringsten. Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 5.5.

#### 5.2.4. Zwei-Wellen-Panel 2000/2002 ( $P_{t_1,t_3}$ )

Das Zwei-Wellen-Panel der Jahre 2000 und 2002 enthält 1 052 Fälle. Dies sind alle gültigen Fälle der Zeitpunkte  $t_1$  und  $t_3$ , deren Fragebögen einander zugeordnet werden konnten.

Tabelle 5.29.: Zusammensetzung nach Geschlecht  $P_{t_1,t_3}$

		männlich		weiblich		gesamt $N$
		$N$	%	$N$	%	(100 %)
$t_1$	$S_{t_1}$	1097	52,1	1008	47,9	2105
	$Q_{t_1}$	931	50,3	919	49,7	1850
	$P_{t_1,t_3}$	488	46,4	564	53,6	1052
$t_3$	$S_{t_3}$	1170	52,0	1081	48,0	2251
	$Q_{t_3}$	1012	52,0	935	48,0	1947
	$P_{t_1,t_3}$	489	46,5	563	53,5	1052

Wie in Tabelle 5.29 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine deutliche Verschiebung vor. Die Unterschiede im prozentualen Verhältnis zwischen Schülerinnen und Schülern der Panelquerschnitte und Querschnitte 2000 und 2002 im Vergleich zur Schulstatistik wurden in den Abschnitten 5.1.1 und 5.1.3 dargestellt. Im Paneldatensatz sind sowohl zum Zeitpunkt  $t_1$  als auch  $t_3$  mehr Schülerinnen als in den Referenzdaten vorhanden. Die Abweichung fällt dabei mit rund fünfeinhalb Prozent zu beiden Zeitpunkten ungefähr gleich aus.

Tabelle 5.30.: Altersdurchschnitt  $P_{t_1,t_3}$

		Alter	Std.-abw.	gesamt $N$
$t_1$	$Q_{t_1}$	13,0	0,8	1845
	$P_{t_1,t_3}$	12,9	0,7	1050
$t_3$	$Q_{t_3}$	15,1	0,8	1881
	$P_{t_1,t_3}$	14,8	0,7	1036

Das Alter der Jugendlichen, deren Fälle im Paneldatensatz enthalten sind, weicht zu beiden Zeitpunkten nicht gravierend vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.30). Hierbei gab es zu  $t_1$  2 und zu  $t_3$  16 fehlende Angaben im Paneldatensatz, zu  $t_1$  5 und zu  $t_3$  66 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 5.31.: Zusammensetzung  $P_{t_1,3}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt $N$
			HS	RS	GY	SO	(100 %)
männlich	$t_1$	$S_{t_1}$	23,8	30,6	39,6	6,0	1097
		$Q_{t_1}$	20,3	31,1	43,0	5,6	931
		$P_{t_1,3}$	18,0	31,6	48,8	1,6	488
	$t_3$	$S_{t_3}$	26,1	31,9	35,0	7,1	1170
		$Q_{t_3}$	22,9	33,0	37,9	6,1	1012
		$P_{t_1,3}$	18,8	31,5	48,1	1,6	489
weiblich	$t_1$	$S_{t_1}$	19,5	30,7	46,3	3,5	1008
		$Q_{t_1}$	17,6	30,5	48,2	3,7	919
		$P_{t_1,3}$	14,9	27,8	55,3	2,0	564
	$t_3$	$S_{t_3}$	23,6	31,3	40,5	4,6	1081
		$Q_{t_3}$	21,3	32,3	42,9	3,5	935
		$P_{t_1,3}$	15,6	27,7	54,5	2,1	563
gesamt	$t_1$	$S_{t_1}$	21,8	30,6	42,8	4,8	2105
		$Q_{t_1}$	19,0	30,8	45,6	4,6	1850
		$P_{t_1,3}$	16,3	29,6	52,3	1,8	1052
	$t_3$	$S_{t_3}$	24,9	31,6	37,6	5,9	2251
		$Q_{t_3}$	22,1	32,7	40,3	4,9	1947
		$P_{t_1,3}$	17,1	29,5	51,5	1,9	1052

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

Aus Tabelle 5.31 wird ersichtlich, dass der Paneldatensatz 2000/2002 im Vergleich zur Schulstatistik systematisch verzerrt ist. Die Unterschiede im prozentualen Verhältnis der Querschnitte 2000 und 2002 im Vergleich zur Schulstatistik wurden in den Abschnitten 5.1.1 und 5.1.3 dargestellt. Im Paneldatensatz befinden sich drei ( $t_1$ ) und vier Prozent ( $t_3$ ) weniger Sonderschülerinnen und -schüler, fünfeinhalb ( $t_1$ ) und fast acht Prozent ( $t_3$ ) weniger Hauptschülerinnen und -schüler, dafür neuneinhalb ( $t_1$ ) und fast vierzehn Prozent ( $t_3$ ) mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik ausgewiesen. Der Anteil der Realschülerinnen und -schüler ist dabei annähernd gleich groß.

Über alle Schulformen hinweg sind zu beiden Zeitpunkten prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.32). Der Unterschied zur Schulstatistik ist hierbei zu beiden Zeitpunkten in der Sonderschule am größten und in der Realschule am geringsten. Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 5.5.



Tabelle 5.32.: Zusammensetzung  $P_{1,3}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt
			HS	RS	GY	SO	(%)
männlich	$t_1$	$S_{t_1}$	57,0	52,1	48,2	65,3	52,1
		$Q_{t_1}$	53,8	50,9	47,4	60,5	50,3
		$P_{1,3}$	51,2	49,5	43,3	42,1	46,4
	$t_3$	$S_{t_3}$	54,5	52,5	48,3	62,4	52,0
		$Q_{t_3}$	53,8	52,5	48,9	65,3	52,0
		$P_{1,3}$	51,1	49,7	43,4	40,0	46,5
weiblich	$t_1$	$S_{t_1}$	43,0	47,9	51,8	34,7	47,9
		$Q_{t_1}$	46,2	49,1	52,6	39,5	49,7
		$P_{1,3}$	48,8	50,5	56,7	57,9	53,6
	$t_3$	$S_{t_3}$	45,5	47,5	51,7	37,6	48,0
		$Q_{t_3}$	46,2	47,5	51,1	34,7	48,0
		$P_{1,3}$	48,9	50,3	56,6	60,0	53,5

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

### 5.2.5. Zwei-Wellen-Panel 2000/2003 ( $P_{1,4}$ )

Das Zwei-Wellen-Panel der Jahre 2000 und 2003 stellt mit 910 Fällen den kleinsten der Zwei-Wellen-Paneldatensätze dar. Enthalten sind alle gültigen Fälle der Zeitpunkte  $t_1$  und  $t_4$ , deren Fragebögen einander zugeordnet werden konnten.

Tabelle 5.33.: Zusammensetzung nach Geschlecht  $P_{1,4}$

		männlich		weiblich		gesamt $N$
		$N$	%	$N$	%	(100 %)
$t_1$	$S_{t_1}$	1 097	52,1	1 008	47,9	2 105
	$Q_{t_1}$	931	50,3	919	49,7	1 850
	$P_{1,4}$	403	44,3	507	55,7	910
$t_4$	$S_{t_4}$	1 056	50,8	1 021	49,2	2 077
	$Q_{t_4}$	915	50,4	901	49,6	1 816
	$P_{1,4}$	405	44,5	505	55,5	910

Wie in Tabelle 5.33 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine deutliche Verschiebung vor. Die Unterschiede im prozentualen Verhältnis zwischen Schülerinnen und Schülern der Panelquerschnitte und Querschnitte 2000 und 2003 im Vergleich zur Schulstatistik wurden in den Abschnitten 5.1.1 und 5.1.4 dargestellt. Im Paneldatensatz sind sowohl zum Zeitpunkt  $t_1$  als auch  $t_4$  mehr Schülerinnen als in den Referenzdaten vorhanden. Die Abweichung fällt dabei zum Zeitpunkt  $t_1$  mit fast acht Prozent noch etwas stärker aus als zum Zeitpunkt  $t_4$  mit rund sechseinhalb Prozent. Im Querschnitt machten zu  $t_4$  drei Personen keine Angaben.

Tabelle 5.34.: Altersdurchschnitt  $P_{t_1,4}$ 

		Alter	Std.-abw.	gesamt $N$
$t_1$	$Q_{t_1}$	13,0	0,8	1845
	$P_{t_1,4}$	12,8	0,7	907
$t_4$	$Q_{t_4}$	16,0	0,8	1786
	$P_{t_1,4}$	15,8	0,7	906

Das Alter der Jugendlichen, deren Fälle im Paneldatensatz enthalten sind, weicht zu beiden Zeitpunkten nicht gravierend vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.34). Hierbei gab es zu  $t_1$  3 und zu  $t_4$  4 fehlende Angaben im Paneldatensatz, zu  $t_1$  5 und zu  $t_4$  33 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 5.35.: Zusammensetzung  $P_{t_1,4}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt $N$
			HS	RS	GY	SO	(100 %)
männlich	$t_1$	$S_{t_1}$	23,8	30,6	39,6	6,0	1097
		$Q_{t_1}$	20,3	31,1	43,0	5,6	931
		$P_{t_1,4}$	13,4	32,3	52,6	1,7	403
	$t_4$	$S_{t_4}$	25,0	30,9	37,3	6,8	1056
		$Q_{t_4}$	23,5	31,3	40,2	5,0	915
		$P_{t_1,4}$	14,3	32,6	51,4	1,7	405
weiblich	$t_1$	$S_{t_1}$	19,5	30,7	46,3	3,5	1008
		$Q_{t_1}$	17,6	30,5	48,2	3,7	919
		$P_{t_1,4}$	15,6	28,0	54,6	1,8	507
	$t_4$	$S_{t_4}$	21,6	32,1	41,1	5,1	1021
		$Q_{t_4}$	22,1	32,3	42,3	3,3	901
		$P_{t_1,4}$	16,4	27,7	54,1	1,8	505
gesamt	$t_1$	$S_{t_1}$	21,8	30,6	42,8	4,8	2105
		$Q_{t_1}$	19,0	30,8	45,6	4,6	1850
		$P_{t_1,4}$	14,6	29,9	53,7	1,8	910
	$t_4$	$S_{t_4}$	23,4	31,5	39,2	6,0	2077
		$Q_{t_4}$	22,8	31,8	41,2	4,2	1816
		$P_{t_1,4}$	15,5	29,9	52,9	1,8	910

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

Aus Tabelle 5.35 wird ersichtlich, dass der Paneldatensatz 2000/2003 im Vergleich zur Schulstatistik systematisch verzerrt ist. Die Unterschiede im prozentualen Verhält-

nis der Querschnitte 2000 und 2003 im Vergleich zur Schulstatistik wurden in den Abschnitten 5.1.1 und 5.1.4 dargestellt. Im Paneldatensatz befinden sich drei ( $t_1$ ) und rund vier Prozent ( $t_4$ ) weniger Sonderschülerinnen und -schüler, rund sieben ( $t_1$ ) und fast acht Prozent ( $t_4$ ) weniger Hauptschülerinnen und -schüler, dafür fast elf ( $t_1$ ) und rund dreizehneinhalb Prozent ( $t_4$ ) mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik ausgewiesen. Der Anteil der Realschülerinnen und -schüler ist dabei annähernd gleich groß.

Über alle Schulformen hinweg sind zu beiden Zeitpunkten prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.36). Der Unterschied zur Schulstatistik ist hierbei zu beiden Zeitpunkten in der Sonderschule am größten und in der Realschule am geringsten. Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 5.5.

Tabelle 5.36.: Zusammensetzung  $P_{1,4}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt
			HS	RS	GY	SO	(%)
männlich	$t_1$	$S_{t_1}$	57,0	52,1	48,2	65,3	52,1
		$Q_{t_1}$	53,8	50,9	47,4	60,5	50,3
		$P_{1,4}$	40,6	47,8	43,4	43,8	44,3
	$t_4$	$S_{t_4}$	54,4	49,8	48,4	58,1	50,8
		$Q_{t_4}$	51,9	49,6	49,1	60,5	50,4
		$P_{1,4}$	41,1	48,5	43,2	43,8	44,5
weiblich	$t_1$	$S_{t_1}$	43,0	47,9	51,8	34,7	47,9
		$Q_{t_1}$	46,2	49,1	52,6	39,5	49,7
		$P_{1,4}$	59,4	52,2	56,6	56,3	55,7
	$t_4$	$S_{t_4}$	45,6	50,2	51,6	41,9	49,2
		$Q_{t_4}$	48,1	50,4	50,9	39,5	49,6
		$P_{1,4}$	58,9	51,5	56,8	56,3	55,5

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

### 5.2.6. Zwei-Wellen-Panel 2001/2003 ( $P_{2,4}$ )

Das Zwei-Wellen-Panel der Jahre 2001 und 2003 enthält 1 105 Fälle. Dies sind alle gültigen Fälle der Zeitpunkte  $t_2$  und  $t_4$ , deren Fragebögen einander zugeordnet werden konnten.

Wie in Tabelle 5.37 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine deutliche Verschiebung vor. Die Unterschiede im prozentualen Verhältnis zwischen Schülerinnen und Schülern der Querschnitte 2001 und 2003 im Vergleich zur Schulstatistik wurden in den Abschnitten 5.1.2 und 5.1.4 dargestellt. Im Paneldatensatz sind sowohl zum Zeitpunkt  $t_2$  als auch  $t_4$  mehr Schülerinnen als in den Referenzdaten vorhanden. Die Abweichung fällt dabei zum Zeitpunkt  $t_2$  mit fast sieben Prozent noch

etwas stärker aus als zum Zeitpunkt  $t_4$  mit rund fünfeinhalb Prozent. Im Querschnitt machten zu  $t_4$  drei Personen keine Angaben.

Tabelle 5.37.: Zusammensetzung nach Geschlecht  $P_{t_2,4}$ 

		männlich		weiblich		gesamt $N$
		$N$	%	$N$	%	(100 %)
$t_2$	$S_{t_2}$	1134	52,0	1047	48,0	2181
	$Q_{t_2}$	978	51,1	937	48,9	1915
	$P_{t_2,4}$	499	45,2	606	54,8	1105
$t_4$	$S_{t_4}$	1056	50,8	1021	49,2	2077
	$Q_{t_4}$	915	50,4	901	49,6	1816
	$P_{t_2,4}$	500	45,2	605	54,8	1105

Das Alter der Jugendlichen, deren Fälle im Paneldatensatz enthalten sind, weicht zu beiden Zeitpunkten nicht gravierend vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.38). Hierbei gab es zu  $t_2$  10 und zu  $t_4$  9 fehlende Angaben im Paneldatensatz, zu  $t_2$  26 und zu  $t_4$  33 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 5.38.: Altersdurchschnitt  $P_{t_2,4}$ 

		Alter	Std.-abw.	gesamt $N$
$t_2$	$Q_{t_2}$	14,0	0,8	1889
	$P_{t_2,4}$	13,9	0,7	1095
$t_4$	$Q_{t_4}$	16,0	0,8	1786
	$P_{t_2,4}$	15,9	0,7	1096

Aus Tabelle 5.39 wird ersichtlich, dass der Paneldatensatz 2001/2003 im Vergleich zur Schulstatistik systematisch verzerrt ist. Die Unterschiede im prozentualen Verhältnis der Panelquerschnitte und Querschnitte 2001 und 2003 im Vergleich zur Schulstatistik wurden in den Abschnitten 5.1.2 und 5.1.4 dargestellt. Im Paneldatensatz befinden sich rund drei ( $t_2$ ) und rund dreieinhalb Prozent ( $t_4$ ) weniger Sonderschülerinnen und -schüler, vier ( $t_2$ ) und rund viereinhalb Prozent ( $t_4$ ) weniger Hauptschülerinnen und -schüler, dafür fast siebeneinhalb ( $t_2$ ) und fast achteinhalb Prozent ( $t_4$ ) mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik ausgewiesen. Der Anteil der Realschülerinnen und -schüler ist dabei annähernd gleich groß.

Über alle Schulformen hinweg sind zu beiden Zeitpunkten prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.40). Der Unterschied zur Schulstatistik ist hierbei zu beiden Zeitpunkten in der Hauptschule am größten und zum Zeitpunkt  $t_2$  in der Realschule, zum Zeitpunkt  $t_4$  in der Sonderschule

am geringsten. Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 5.5.

Tabelle 5.39.: Zusammensetzung  $P_{t_2,4}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt $N$
			HS	RS	GY	SO	(100 %)
männlich	$t_2$	$S_{t_2}$	24,2	31,5	38,0	6,3	1134
		$Q_{t_2}$	23,6	32,1	37,7	6,5	978
		$P_{t_2,4}$	16,8	33,3	47,1	2,8	499
	$t_4$	$S_{t_4}$	25,0	30,9	37,3	6,8	1056
		$Q_{t_4}$	23,5	31,3	40,2	5,0	915
		$P_{t_2,4}$	18,8	32,0	46,4	2,8	500
weiblich	$t_2$	$S_{t_2}$	20,4	31,9	43,6	4,1	1047
		$Q_{t_2}$	19,7	31,8	44,8	3,6	937
		$P_{t_2,4}$	19,6	29,7	49,0	1,7	606
	$t_4$	$S_{t_4}$	21,6	32,1	41,1	5,1	1021
		$Q_{t_4}$	22,1	32,3	42,3	3,3	901
		$P_{t_2,4}$	18,8	30,7	48,6	1,8	605
gesamt	$t_2$	$S_{t_2}$	22,4	31,7	40,7	5,3	2181
		$Q_{t_2}$	21,7	32,0	41,2	5,1	1915
		$P_{t_2,4}$	18,4	31,3	48,1	2,2	1105
	$t_4$	$S_{t_4}$	23,4	31,5	39,2	6,0	2077
		$Q_{t_4}$	22,8	31,8	41,2	4,2	1816
		$P_{t_2,4}$	18,8	31,3	47,6	2,3	1105

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

Tabelle 5.40.: Zusammensetzung  $P_{t_2,4}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt
			HS	RS	GY	SO	(%)
männlich	$t_2$	$S_{t_2}$	56,1	51,7	48,6	62,6	52,0
		$Q_{t_2}$	55,5	51,3	46,8	65,3	51,1
		$P_{t_2,4}$	41,4	48,0	44,2	58,3	45,2
	$t_4$	$S_{t_4}$	54,4	49,8	48,4	58,1	50,8
		$Q_{t_4}$	51,9	49,6	49,1	60,5	50,4
		$P_{t_2,4}$	45,2	46,2	44,1	56,0	45,2
weiblich	$t_2$	$S_{t_2}$	43,9	48,3	51,4	37,4	48,0
		$Q_{t_2}$	44,5	48,7	53,2	34,7	48,9
		$P_{t_2,4}$	58,6	52,0	55,8	41,7	54,8
	$t_4$	$S_{t_4}$	45,6	50,2	51,6	41,9	49,2
		$Q_{t_4}$	48,1	50,4	50,9	39,5	49,6
		$P_{t_2,4}$	54,8	53,8	55,9	44,0	54,8

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

### 5.3. Die Drei-Wellen-Panels

Die Datensätze der Drei-Wellen-Panels enthalten alle Fälle, deren zugehörige (und nach Plausibilitätskontrollen gültige) Fragebögen zwei (ebenfalls gültigen) Bögen zwei weiterer Zeitpunkte zugeordnet werden konnten. In diesem Bericht werden die Datensätze aller vier denkbaren Kombinationen von drei der insgesamt vier Zeitpunkte vorgestellt. In der Praxis werden vermutlich die beiden Datensätze dreier benachbarter Zeitpunkte am bedeutsamsten sein.

#### 5.3.1. Drei-Wellen-Panel 2000 bis 2002 ( $P_{t_1,2,3}$ )

Das Drei-Wellen-Panel der Jahre 2000, 2001 und 2002 enthält 997 Fälle. Dies sind alle gültigen Fälle der Zeitpunkte  $t_1$ ,  $t_2$  und  $t_3$ , deren Fragebögen einander zugeordnet werden konnten.

Wie in Tabelle 5.41 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine deutliche Verschiebung vor. Die Unterschiede im prozentualen Verhältnis zwischen Schülerinnen und Schülern der Querschnitte 2000, 2001 und 2002 im Vergleich zur Schulstatistik wurden in den Abschnitten 5.1.1, 5.1.2 und 5.1.3 dargestellt. Im Paneldatensatz sind sowohl zum Zeitpunkt  $t_1$ ,  $t_2$  als auch  $t_3$  mehr Schülerinnen als in den Referenzdaten vorhanden. Die Abweichung fällt dabei mit knapp sechseinhalb Prozent zu allen Zeitpunkten annähernd gleich aus.

Tabelle 5.41.: Zusammensetzung nach Geschlecht  $P_{t_1,2,3}$

		männlich		weiblich		gesamt $N$
		$N$	%	$N$	%	(100 %)
$t_1$	$S_{t_1}$	1097	52,1	1008	47,9	2105
	$Q_{t_1}$	931	50,3	919	49,7	1850
	$P_{t_1,2,3}$	455	45,6	542	54,4	997
$t_2$	$S_{t_2}$	1134	52,0	1047	48,0	2181
	$Q_{t_2}$	978	51,1	937	48,9	1915
	$P_{t_1,2,3}$	456	45,7	541	54,3	997
$t_3$	$S_{t_3}$	1170	52,0	1081	48,0	2251
	$Q_{t_3}$	1012	52,0	935	48,0	1947
	$P_{t_1,2,3}$	456	45,7	541	54,3	997

Das Alter der Jugendlichen, deren Fälle im Paneldatensatz enthalten sind, weicht zu allen drei Zeitpunkten nicht gravierend vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.42). Hierbei gab es zu  $t_1$  1, zu  $t_2$  3 und zu  $t_3$  15 fehlende Angaben im Paneldatensatz, zu  $t_1$  5, zu  $t_2$  26 und zu  $t_3$  66 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 5.42.: Altersdurchschnitt  $P_{t_1,2,3}$

		Alter	Std.-abw.	gesamt $N$
$t_1$	$Q_{t_1}$	13,0	0,8	1845
	$P_{t_1,2,3}$	12,9	0,7	996
$t_2$	$Q_{t_2}$	14,0	0,8	1889
	$P_{t_1,2,3}$	13,8	0,7	994
$t_3$	$Q_{t_3}$	15,1	0,8	1881
	$P_{t_1,2,3}$	14,8	0,7	982

Tabelle 5.43.: Zusammensetzung  $P_{t_1,2,3}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)				gesamt $N$	
		HS	RS	GY	SO	(100 %)	
männlich	$t_1$	$S_{t_1}$	23,8	30,6	39,6	6,0	1097
		$Q_{t_1}$	20,3	31,1	43,0	5,6	931
		$P_{t_1,2,3}$	17,8	31,4	49,0	1,8	455
	$t_2$	$S_{t_2}$	24,2	31,5	38,0	6,3	1134
		$Q_{t_2}$	23,6	32,1	37,7	6,5	978
		$P_{t_1,2,3}$	18,0	31,8	48,5	1,8	456
	$t_3$	$S_{t_3}$	26,1	31,9	35,0	7,1	1170
		$Q_{t_3}$	22,9	33,0	37,9	6,1	1012
		$P_{t_1,2,3}$	18,2	31,8	48,2	1,8	456
weiblich	$t_1$	$S_{t_1}$	19,5	30,7	46,3	3,5	1008
		$Q_{t_1}$	17,6	30,5	48,2	3,7	919
		$P_{t_1,2,3}$	15,1	28,2	54,8	1,8	542
	$t_2$	$S_{t_2}$	20,4	31,9	43,6	4,1	1047
		$Q_{t_2}$	19,7	31,8	44,8	3,6	937
		$P_{t_1,2,3}$	15,7	27,7	54,7	1,8	541
	$t_3$	$S_{t_3}$	23,6	31,3	40,5	4,6	1081
		$Q_{t_3}$	21,3	32,3	42,9	3,5	935
		$P_{t_1,2,3}$	15,9	27,9	54,3	1,8	541
gesamt	$t_1$	$S_{t_1}$	21,8	30,6	42,8	4,8	2105
		$Q_{t_1}$	19,0	30,8	45,6	4,6	1850
		$P_{t_1,2,3}$	16,3	29,7	52,2	1,8	997
	$t_2$	$S_{t_2}$	22,4	31,7	40,7	5,3	2181
		$Q_{t_2}$	21,7	32,0	41,2	5,1	1915
		$P_{t_1,2,3}$	16,8	29,6	51,9	1,8	997
	$t_3$	$S_{t_3}$	24,9	31,6	37,6	5,9	2251
		$Q_{t_3}$	22,1	32,7	40,3	4,9	1947
		$P_{t_1,2,3}$	17,0	29,7	51,6	1,8	997

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule



Aus Tabelle 5.43 wird ersichtlich, dass der Paneldatensatz 2000/2001/2002 im Vergleich zur Schulstatistik systematisch verzerrt ist. Die Unterschiede im prozentualen Verhältnis der Panelquerschnitte und Querschnitte 2000, 2001 und 2002 im Vergleich zur Schulstatistik wurden in den Abschnitten 5.1.1, 5.1.2 und 5.1.3 dargestellt. Im Paneldatensatz befinden sich drei ( $t_1$ ), dreieinhalb ( $t_2$ ) und rund vier Prozent ( $t_3$ ) weniger Sonderschülerinnen und -schüler, rund fünfeinhalb ( $t_1$  und  $t_2$ ) und fast acht Prozent ( $t_3$ ) weniger Hauptschülerinnen und -schüler, dafür fast neuneinhalb ( $t_1$ ), rund elf ( $t_2$ ) und vierzehn Prozent ( $t_3$ ) mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik ausgewiesen. Der Anteil der Realschülerinnen und -schüler ist dabei zu allen Zeitpunkten annähernd gleich groß.

Über alle Schulformen hinweg sind zu allen drei Zeitpunkten prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.44). Der Unterschied zur Schulstatistik ist hierbei zu allen drei Zeitpunkten in der Sonderschule am größten und in der Realschule am geringsten. Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 5.5.

Tabelle 5.44.: Zusammensetzung  $P_{t_1,2,3}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt
			HS	RS	GY	SO	(%)
männlich	$t_1$	$S_{t_1}$	57,0	52,1	48,2	65,3	52,1
		$Q_{t_1}$	53,8	50,9	47,4	60,5	50,3
		$P_{t_1,2,3}$	49,7	48,3	42,9	44,4	45,6
	$t_2$	$S_{t_2}$	56,1	51,7	48,6	62,6	52,0
		$Q_{t_2}$	55,5	51,3	46,8	65,3	51,1
		$P_{t_1,2,3}$	49,1	49,2	42,7	44,4	45,7
	$t_3$	$S_{t_3}$	54,5	52,5	48,3	62,4	52,0
		$Q_{t_3}$	53,8	52,5	48,9	65,3	52,0
		$P_{t_1,2,3}$	49,1	49,0	42,8	44,4	45,7
weiblich	$t_1$	$S_{t_1}$	43,0	47,9	51,8	34,7	47,9
		$Q_{t_1}$	46,2	49,1	52,6	39,5	49,7
		$P_{t_1,2,3}$	50,3	51,7	57,1	55,6	54,4
	$t_2$	$S_{t_2}$	43,9	48,3	51,4	37,4	48,0
		$Q_{t_2}$	44,5	48,7	53,2	34,7	48,9
		$P_{t_1,2,3}$	50,9	50,8	57,3	55,6	54,3
	$t_3$	$S_{t_3}$	45,5	47,5	51,7	37,6	48,0
		$Q_{t_3}$	46,2	47,5	51,1	34,7	48,0
		$P_{t_1,2,3}$	50,9	51,0	57,2	55,6	54,3

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

### 5.3.2. Drei-Wellen-Panel 2001 bis 2003 ( $P_{t_2,3,4}$ )

Das Drei-Wellen-Panel der Jahre 2001, 2002 und 2003 stellt mit 1 075 Fällen den größten Datensatz aller vier Drei-Wellen-Panels dar. Er enthält alle gültigen Fälle der Zeitpunkte  $t_2$ ,  $t_3$  und  $t_4$ , deren Fragebögen einander zugeordnet werden konnten.

Tabelle 5.45.: Zusammensetzung nach Geschlecht  $P_{t_2,3,4}$

		männlich		weiblich		gesamt $N$
		$N$	%	$N$	%	(100 %)
$t_2$	$S_{t_2}$	1 134	52,0	1 047	48,0	2 181
	$Q_{t_2}$	978	51,1	937	48,9	1 915
	$P_{t_2,3,4}$	485	45,1	590	54,9	1 075
$t_3$	$S_{t_3}$	1 170	52,0	1 081	48,0	2 251
	$Q_{t_3}$	1 012	52,0	935	48,0	1 947
	$P_{t_2,3,4}$	488	45,4	587	54,6	1 075
$t_4$	$S_{t_4}$	1 056	50,8	1 021	49,2	2 077
	$Q_{t_4}$	915	50,4	901	49,6	1 816
	$P_{t_2,3,4}$	486	45,2	589	54,8	1 075

Wie in Tabelle 5.45 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine deutliche Verschiebung vor. Die Unterschiede im prozentualen Verhältnis zwischen Schülerinnen und Schülern der Querschnitte 2001, 2002 und 2003 im Vergleich zur Schulstatistik wurden in den Abschnitten 5.1.2, 5.1.3 und 5.1.4 dargestellt. Im Paneldatensatz sind sowohl zum Zeitpunkt  $t_2$ ,  $t_3$  als auch  $t_4$  mehr Schülerinnen als in den Referenzdaten vorhanden. Die Abweichung fällt dabei mit knapp sieben Prozent zum Zeitpunkt  $t_2$  am stärksten aus, gefolgt von Zeitpunkt  $t_3$  mit rund sechseinhalb Prozent und Zeitpunkt  $t_4$  mit fünfeinhalb Prozent. Im Querschnitt machten zu  $t_4$  drei Personen keine Angaben.

Tabelle 5.46.: Altersdurchschnitt  $P_{t_2,3,4}$

		Alter	Std.-abw.	gesamt $N$
$t_2$	$Q_{t_2}$	14,0	0,8	1 889
	$P_{t_2,3,4}$	13,9	0,7	1 067
$t_3$	$Q_{t_3}$	15,1	0,8	1 881
	$P_{t_2,3,4}$	14,9	0,7	1 059
$t_4$	$Q_{t_4}$	16,0	0,8	1 786
	$P_{t_2,3,4}$	15,9	0,7	1 068

Das Alter der Jugendlichen, deren Fälle im Paneldatensatz enthalten sind, weicht zu allen drei Zeitpunkten nicht gravierend vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.46). Hierbei gab es zu  $t_2$  8, zu  $t_3$  16 und zu  $t_4$  7 fehlende Angaben im Paneldatensatz, zu  $t_2$  26, zu  $t_3$  66 und zu  $t_4$  33 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 5.47.: Zusammensetzung  $P_{t_2,3,4}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt N	
			HS	RS	GY	SO	(100 %)	
männlich	$t_2$	$S_{t_2}$	24,2	31,5	38,0	6,3	1 134	
		$Q_{t_2}$	23,6	32,1	37,7	6,5	978	
		$P_{t_2,3,4}$	16,7	33,0	47,6	2,7	485	
	$t_3$	$S_{t_3}$	26,1	31,9	35,0	7,1	1 170	
		$Q_{t_3}$	22,9	33,0	37,9	6,1	1 012	
		$P_{t_2,3,4}$	17,6	32,2	47,5	2,7	488	
	$t_4$	$S_{t_4}$	25,0	30,9	37,3	6,8	1 056	
		$Q_{t_4}$	23,5	31,3	40,2	5,0	915	
		$P_{t_2,3,4}$	18,3	31,7	47,3	2,7	486	
	weiblich	$t_2$	$S_{t_2}$	20,4	31,9	43,6	4,1	1 047
			$Q_{t_2}$	19,7	31,8	44,8	3,6	937
			$P_{t_2,3,4}$	19,2	29,8	49,5	1,5	590
$t_3$		$S_{t_3}$	23,6	31,3	40,5	4,6	1 081	
		$Q_{t_3}$	21,3	32,3	42,9	3,5	935	
		$P_{t_2,3,4}$	18,6	30,5	49,4	1,5	587	
$t_4$		$S_{t_4}$	21,6	32,1	41,1	5,1	1 021	
		$Q_{t_4}$	22,1	32,3	42,3	3,3	901	
		$P_{t_2,3,4}$	18,7	30,9	48,9	1,5	589	
gesamt		$t_2$	$S_{t_2}$	22,4	31,7	40,7	5,3	2 181
			$Q_{t_2}$	21,7	32,0	41,2	5,1	1 915
			$P_{t_2,3,4}$	18,0	31,3	48,7	2,0	1 075
	$t_3$	$S_{t_3}$	24,9	31,6	37,6	5,9	2 251	
		$Q_{t_3}$	22,1	32,7	40,3	4,9	1 947	
		$P_{t_2,3,4}$	18,1	31,3	48,6	2,0	1 075	
	$t_4$	$S_{t_4}$	23,4	31,5	39,2	6,0	2 077	
		$Q_{t_4}$	22,8	31,8	41,2	4,2	1 816	
		$P_{t_2,3,4}$	18,5	31,3	48,2	2,0	1 075	

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

Aus Tabelle 5.47 wird ersichtlich, dass der Paneldatensatz 2001/2002/2003 im Vergleich zur Schulstatistik systematisch verzerrt ist. Die Unterschiede im prozentualen Verhältnis der Panelquerschnitte und Querschnitte 2001, 2002 und 2003 im Vergleich zur Schulstatistik wurden in den Abschnitten 5.1.2, 5.1.3 und 5.1.4 dargestellt. Im Paneldatensatz befinden sich rund dreieinhalb ( $t_2$ ) und rund vier Prozent ( $t_3$  und  $t_4$ ) weniger Sonderschülerinnen und -schüler, rund viereinhalb ( $t_2$ ), rund sechseinhalb ( $t_3$ ) und

knapp fünf Prozent ( $t_4$ ) weniger Hauptschülerinnen und -schüler, dafür acht ( $t_2$ ), elf ( $t_3$ ) und neun Prozent ( $t_4$ ) mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik ausgewiesen. Der Anteil der Realschülerinnen und -schüler ist dabei zu allen Zeitpunkten annähernd gleich groß.

Außer in der Sonderschule zum Zeitpunkt  $t_4$  sind über alle Schulformen hinweg prozentual zu allen drei Zeitpunkten mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.48). Der Unterschied zur Schulstatistik ist hierbei zu allen drei Zeitpunkten in der Hauptschule am größten und in der Sonderschule am geringsten. Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 5.5.

Tabelle 5.48.: Zusammensetzung  $P_{t_2,3,4}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt
			HS	RS	GY	SO	(%)
männlich	$t_2$	$S_{t_2}$	56,1	51,7	48,6	62,6	52,0
		$Q_{t_2}$	55,5	51,3	46,8	65,3	51,1
		$P_{t_2,3,4}$	41,8	47,6	44,2	59,1	45,1
	$t_3$	$S_{t_3}$	54,5	52,5	48,3	62,4	52,0
		$Q_{t_3}$	53,8	52,5	48,9	65,3	52,0
		$P_{t_2,3,4}$	44,1	46,7	44,4	59,1	45,4
	$t_4$	$S_{t_4}$	54,4	49,8	48,4	58,1	50,8
		$Q_{t_4}$	51,9	49,6	49,1	60,5	50,4
		$P_{t_2,3,4}$	44,7	45,8	44,4	59,1	45,2
weiblich	$t_2$	$S_{t_2}$	43,9	48,3	51,4	37,4	48,0
		$Q_{t_2}$	44,5	48,7	53,2	34,7	48,9
		$P_{t_2,3,4}$	58,2	52,4	55,8	40,9	54,9
	$t_3$	$S_{t_3}$	45,5	47,5	51,7	37,6	48,0
		$Q_{t_3}$	46,2	47,5	51,1	34,7	48,0
		$P_{t_2,3,4}$	55,9	53,3	55,6	40,9	54,6
	$t_4$	$S_{t_4}$	45,6	50,2	51,6	41,9	49,2
		$Q_{t_4}$	48,1	50,4	50,9	39,5	49,6
		$P_{t_2,3,4}$	55,3	54,2	55,6	40,9	54,8

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

### 5.3.3. Drei-Wellen-Panel 2000, 2001, 2003 ( $P_{t_1,2,4}$ )

Das Drei-Wellen-Panel der Jahre 2000, 2001 und 2003 stellt mit 823 Fällen den kleinsten Drei-Wellen-Paneldatensatz dar. Enthalten sind alle gültigen Fälle der Zeitpunkte  $t_1$ ,  $t_2$  und  $t_4$ , deren Fragebögen einander zugeordnet werden konnten.

Wie in Tabelle 5.49 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine deutliche Verschiebung vor. Die Unterschiede im prozentualen Verhältnis zwischen Schülerinnen und Schülern der Querschnitte 2000, 2001 und 2003 im Vergleich zur

Schulstatistik wurden in den Abschnitten 5.1.1, 5.1.2 und 5.1.4 dargestellt. Im Paneldatensatz sind sowohl zum Zeitpunkt  $t_1$ ,  $t_2$  als auch  $t_4$  mehr Schülerinnen als in den Referenzdaten vorhanden. Die Abweichung fällt dabei mit rund achteinhalb Prozent zu den Zeitpunkten  $t_1$  und  $t_2$  etwas stärker aus als mit rund sieben Prozent zum Zeitpunkt  $t_4$ . Im Querschnitt machten zu  $t_4$  drei Personen keine Angaben.

Tabelle 5.49.: Zusammensetzung nach Geschlecht  $P_{t_1,2,4}$

		männlich		weiblich		gesamt $N$
		$N$	%	$N$	%	(100 %)
$t_1$	$S_{t_1}$	1097	52,1	1008	47,9	2105
	$Q_{t_1}$	931	50,3	919	49,7	1850
	$P_{t_1,2,4}$	356	43,3	467	56,7	823
$t_2$	$S_{t_2}$	1134	52,0	1047	48,0	2181
	$Q_{t_2}$	978	51,1	937	48,9	1915
	$P_{t_1,2,4}$	358	43,5	465	56,5	823
$t_4$	$S_{t_4}$	1056	50,8	1021	49,2	2077
	$Q_{t_4}$	915	50,4	901	49,6	1816
	$P_{t_1,2,4}$	360	43,7	463	56,3	823

Das Alter der Jugendlichen, deren Fälle im Paneldatensatz enthalten sind, weicht zu allen drei Zeitpunkten nicht gravierend vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.50). Hierbei gab es zu  $t_1$  1, zu  $t_2$  4 und zu  $t_4$  2 fehlende Angaben im Paneldatensatz, zu  $t_1$  5, zu  $t_2$  26 und zu  $t_4$  33 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 5.50.: Altersdurchschnitt  $P_{t_1,2,4}$

		Alter	Std.-abw.	gesamt $N$
$t_1$	$Q_{t_1}$	13,0	0,8	1845
	$P_{t_1,2,4}$	12,8	0,7	822
$t_2$	$Q_{t_2}$	14,0	0,8	1889
	$P_{t_1,2,4}$	13,8	0,6	819
$t_4$	$Q_{t_4}$	16,0	0,8	1786
	$P_{t_1,2,4}$	15,8	0,6	821

Aus Tabelle 5.51 wird ersichtlich, dass der Paneldatensatz 2000/2001/2003 im Vergleich zur Schulstatistik systematisch verzerrt ist. Die Unterschiede im prozentualen Verhältnis der Panelquerschnitte und Querschnitte 2000, 2001 und 2003 im Vergleich zur Schulstatistik wurden in den Abschnitten 5.1.1, 5.1.2 und 5.1.4 dargestellt. Im Paneldatensatz befinden sich rund drei ( $t_1$ ), rund dreieinhalb ( $t_2$ ) und rund viereinhalb

Prozent ( $t_4$ ) weniger Sonderschülerinnen und -schüler, siebeneinhalb ( $t_1$ ), knapp acht ( $t_2$ ) und achteinhalb Prozent ( $t_4$ ) weniger Hauptschülerinnen und -schüler, dafür elfeinhalb ( $t_2$ ), knapp dreizehneinhalb ( $t_3$ ) und fast vierzehneinhalb Prozent ( $t_4$ ) mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik ausgewiesen. Der Anteil der Realschülerinnen und -schüler ist dabei zu allen Zeitpunkten annähernd gleich groß.

Tabelle 5.51.: Zusammensetzung  $P_{1,2,4}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt $N$ (100 %)	
			HS	RS	GY	SO		
männlich	$t_1$	$S_{t_1}$	23,8	30,6	39,6	6,0	1097	
		$Q_{t_1}$	20,3	31,1	43,0	5,6	931	
		$P_{1,2,4}$	13,2	32,3	53,1	1,4	356	
	$t_2$	$S_{t_2}$	24,2	31,5	38,0	6,3	1134	
		$Q_{t_2}$	23,6	32,1	37,7	6,5	978	
		$P_{1,2,4}$	13,1	32,7	52,8	1,4	358	
	$t_4$	$S_{t_4}$	25,0	30,9	37,3	6,8	1056	
		$Q_{t_4}$	23,5	31,3	40,2	5,0	915	
		$P_{1,2,4}$	13,9	32,5	52,2	1,4	360	
	weiblich	$t_1$	$S_{t_1}$	19,5	30,7	46,3	3,5	1008
			$Q_{t_1}$	17,6	30,5	48,2	3,7	919
			$P_{1,2,4}$	15,2	27,8	55,2	1,7	467
$t_2$		$S_{t_2}$	20,4	31,9	43,6	4,1	1047	
		$Q_{t_2}$	19,7	31,8	44,8	3,6	937	
		$P_{1,2,4}$	15,5	27,7	55,1	1,7	465	
$t_4$		$S_{t_4}$	21,6	32,1	41,1	5,1	1021	
		$Q_{t_4}$	22,1	32,3	42,3	3,3	901	
		$P_{1,2,4}$	15,8	28,1	54,4	1,7	463	
gesamt		$t_1$	$S_{t_1}$	21,8	30,6	42,8	4,8	2105
			$Q_{t_1}$	19,0	30,8	45,6	4,6	1850
			$P_{1,2,4}$	14,3	29,8	54,3	1,6	823
	$t_2$	$S_{t_2}$	22,4	31,7	40,7	5,3	2181	
		$Q_{t_2}$	21,7	32,0	41,2	5,1	1915	
		$P_{1,2,4}$	14,5	29,9	54,1	1,6	823	
	$t_4$	$S_{t_4}$	23,4	31,5	39,2	6,0	2077	
		$Q_{t_4}$	22,8	31,8	41,2	4,2	1816	
		$P_{1,2,4}$	14,9	30,0	53,5	1,6	823	

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

Über alle Schulformen hinweg sind zu allen drei Zeitpunkten prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.52). Der Unterschied zur Schulstatistik ist hierbei zu allen drei Zeitpunkten in der Sonder-

schule am größten und in der Realschule am geringsten. Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 5.5.

Tabelle 5.52.: Zusammensetzung  $P_{t_1,2,4}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt	
			HS	RS	GY	SO	(%)	
männlich	$t_1$	$S_{t_1}$	57,0	52,1	48,2	65,3	52,1	
		$Q_{t_1}$	53,8	50,9	47,4	60,5	50,3	
		$P_{t_1,2,4}$	39,8	46,9	42,3	38,5	43,3	
	$t_2$	$S_{t_2}$	56,1	51,7	48,6	62,6	52,0	
		$Q_{t_2}$	55,5	51,3	46,8	65,3	51,1	
		$P_{t_1,2,4}$	39,5	47,6	42,5	38,5	43,5	
	$t_4$	$S_{t_4}$	54,4	49,8	48,4	58,1	50,8	
		$Q_{t_4}$	51,9	49,6	49,1	60,5	50,4	
		$P_{t_1,2,4}$	40,7	47,4	42,7	38,5	43,7	
	weiblich	$t_1$	$S_{t_1}$	43,0	47,9	51,8	34,7	47,9
			$Q_{t_1}$	46,2	49,1	52,6	39,5	49,7
			$P_{t_1,2,4}$	60,2	53,1	57,7	61,5	56,7
$t_2$		$S_{t_2}$	43,9	48,3	51,4	37,4	48,0	
		$Q_{t_2}$	44,5	48,7	53,2	34,7	48,9	
		$P_{t_1,2,4}$	60,5	52,4	57,5	61,5	56,5	
$t_4$		$S_{t_4}$	45,6	50,2	51,6	41,9	49,2	
		$Q_{t_4}$	48,1	50,4	50,9	39,5	49,6	
		$P_{t_1,2,4}$	59,3	52,6	57,3	61,5	56,3	

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

#### 5.3.4. Drei-Wellen-Panel 2000, 2002, 2003 ( $P_{t_1,3,4}$ )

Das Drei-Wellen-Panel der Jahre 2000, 2002 und 2003 enthält 851 Fälle. Dies sind alle gültigen Fälle der Zeitpunkte  $t_1$ ,  $t_3$  und  $t_4$ , deren Fragebögen einander zugeordnet werden konnten.

Tabelle 5.53.: Zusammensetzung nach Geschlecht  $P_{t_1,3,4}$ 

		männlich		weiblich		gesamt $N$
		$N$	%	$N$	%	(100 %)
$t_1$	$S_{t_1}$	1097	52,1	1008	47,9	2105
	$Q_{t_1}$	931	50,3	919	49,7	1850
	$P_{t_1,3,4}$	375	44,1	476	55,9	851
$t_3$	$S_{t_3}$	1170	52,0	1081	48,0	2251
	$Q_{t_3}$	1012	52,0	935	48,0	1947
	$P_{t_1,3,4}$	377	44,3	474	55,7	851
$t_4$	$S_{t_4}$	1056	50,8	1021	49,2	2077
	$Q_{t_4}$	915	50,4	901	49,6	1816
	$P_{t_1,3,4}$	380	44,7	471	55,3	851

Wie in Tabelle 5.53 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine deutliche Verschiebung vor. Die Unterschiede im prozentualen Verhältnis zwischen Schülerinnen und Schülern der Querschnitte 2000, 2002 und 2003 im Vergleich zur Schulstatistik wurden in den Abschnitten 5.1.1, 5.1.3 und 5.1.4 dargestellt. Im Paneldatensatz sind sowohl zum Zeitpunkt  $t_1$ ,  $t_3$  als auch  $t_4$  mehr Schülerinnen als in den Referenzdaten vorhanden. Die Abweichung fällt dabei mit acht Prozent zum Zeitpunkt  $t_1$  etwas stärker aus als mit rund siebeneinhalb Prozent zum Zeitpunkt  $t_3$  und rund sechs Prozent zum Zeitpunkt  $t_4$ . Im Querschnitt machten zu  $t_4$  drei Personen keine Angaben.

Das Alter der Jugendlichen, deren Fälle im Paneldatensatz enthalten sind, weicht zu allen drei Zeitpunkten nicht gravierend vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.54). Hierbei gab es zu  $t_1$  2, zu  $t_3$  9 und zu  $t_4$  2 fehlende Angaben im Paneldatensatz, zu  $t_1$  5, zu  $t_3$  66 und zu  $t_4$  33 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 5.54.: Altersdurchschnitt  $P_{t_1,3,4}$ 

		Alter	Std.-abw.	gesamt $N$
$t_1$	$Q_{t_1}$	13,0	0,8	1845
	$P_{t_1,3,4}$	12,8	0,7	849
$t_3$	$Q_{t_3}$	15,1	0,8	1881
	$P_{t_1,3,4}$	14,8	0,6	842
$t_4$	$Q_{t_4}$	16,0	0,8	1786
	$P_{t_1,3,4}$	15,8	0,6	849

Aus Tabelle 5.55 wird ersichtlich, dass der Paneldatensatz 2000/2002/2003 im Vergleich zur Schulstatistik systematisch verzerrt ist. Die Unterschiede im prozentualen Verhältnis der Panelquerschnitte und Querschnitte 2000, 2002 und 2003 im Vergleich



zur Schulstatistik wurden in den Abschnitten 5.1.1, 5.1.3 und 5.1.4 dargestellt. Im Paneldatensatz befinden sich rund drei ( $t_1$ ) und rund viereinhalb Prozent ( $t_3$  und  $t_4$ ) weniger Sonderschülerinnen und -schüler, rund siebeneinhalb ( $t_1$ ), rund zehn ( $t_3$ ) und knapp achteinhalb Prozent ( $t_4$ ) weniger Hauptschülerinnen und -schüler, dafür rund elfeinhalb ( $t_1$ ), knapp sechzehneinhalb ( $t_3$ ) und fast vierzehneinhalb Prozent ( $t_4$ ) mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik ausgewiesen. Der Anteil der Realschülerinnen und -schüler ist dabei zu allen Zeitpunkten annähernd gleich groß.

Tabelle 5.55.: Zusammensetzung  $P_{t_1,3,4}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt $N$	
			HS	RS	GY	SO	(100 %)	
männlich	$t_1$	$S_{t_1}$	23,8	30,6	39,6	6,0	1097	
		$Q_{t_1}$	20,3	31,1	43,0	5,6	931	
		$P_{t_1,3,4}$	13,3	32,3	53,1	1,3	375	
	$t_3$	$S_{t_3}$	26,1	31,9	35,0	7,1	1170	
		$Q_{t_3}$	22,9	33,0	37,9	6,1	1012	
		$P_{t_1,3,4}$	14,3	32,1	52,3	1,3	377	
	$t_4$	$S_{t_4}$	25,0	30,9	37,3	6,8	1056	
		$Q_{t_4}$	23,5	31,3	40,2	5,0	915	
		$P_{t_1,3,4}$	14,5	32,6	51,6	1,3	380	
	weiblich	$t_1$	$S_{t_1}$	19,5	30,7	46,3	3,5	1008
			$Q_{t_1}$	17,6	30,5	48,2	3,7	919
			$P_{t_1,3,4}$	14,9	27,5	55,7	1,9	476
$t_3$		$S_{t_3}$	23,6	31,3	40,5	4,6	1081	
		$Q_{t_3}$	21,3	32,3	42,9	3,5	935	
		$P_{t_1,3,4}$	15,2	27,6	55,3	1,9	474	
$t_4$		$S_{t_4}$	21,6	32,1	41,1	5,1	1021	
		$Q_{t_4}$	22,1	32,3	42,3	3,3	901	
		$P_{t_1,3,4}$	15,5	27,6	55,0	1,9	471	
gesamt		$t_1$	$S_{t_1}$	21,8	30,6	42,8	4,8	2105
			$Q_{t_1}$	19,0	30,8	45,6	4,6	1850
			$P_{t_1,3,4}$	14,2	29,6	54,5	1,6	851
	$t_3$	$S_{t_3}$	24,9	31,6	37,6	5,9	2251	
		$Q_{t_3}$	22,1	32,7	40,3	4,9	1947	
		$P_{t_1,3,4}$	14,8	29,6	53,9	1,6	851	
	$t_4$	$S_{t_4}$	23,4	31,5	39,2	6,0	2077	
		$Q_{t_4}$	22,8	31,8	41,2	4,2	1816	
		$P_{t_1,3,4}$	15,0	29,8	53,5	1,6	851	

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

Über alle Schulformen hinweg sind zu allen drei Zeitpunkten prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.56).

Der Unterschied zur Schulstatistik ist hierbei zu allen drei Zeitpunkten in der Sonderschule am größten und in der Realschule am geringsten. Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 5.5.

Tabelle 5.56.: Zusammensetzung  $P_{1,3,4}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt
			HS	RS	GY	SO	(%)
männlich	$t_1$	$S_{t_1}$	57,0	52,1	48,2	65,3	52,1
		$Q_{t_1}$	53,8	50,9	47,4	60,5	50,3
		$P_{t_{1,3,4}}$	41,3	48,0	42,9	35,7	44,1
	$t_3$	$S_{t_3}$	54,5	52,5	48,3	62,4	52,0
		$Q_{t_3}$	53,8	52,5	48,9	65,3	52,0
		$P_{t_{1,3,4}}$	42,9	48,0	42,9	35,7	44,3
	$t_4$	$S_{t_4}$	54,4	49,8	48,4	58,1	50,8
		$Q_{t_4}$	51,9	49,6	49,1	60,5	50,4
		$P_{t_{1,3,4}}$	43,0	48,8	43,1	35,7	44,7
weiblich	$t_1$	$S_{t_1}$	43,0	47,9	51,8	34,7	47,9
		$Q_{t_1}$	46,2	49,1	52,6	39,5	49,7
		$P_{t_{1,3,4}}$	58,7	52,0	57,1	64,3	55,9
	$t_3$	$S_{t_3}$	45,5	47,5	51,7	37,6	48,0
		$Q_{t_3}$	46,2	47,5	51,1	34,7	48,0
		$P_{t_{1,3,4}}$	57,1	52,0	57,1	64,3	55,7
	$t_4$	$S_{t_4}$	45,6	50,2	51,6	41,9	49,2
		$Q_{t_4}$	48,1	50,4	50,9	39,5	49,6
		$P_{t_{1,3,4}}$	57,0	51,2	56,9	64,3	55,3

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

#### 5.4. Das Vier-Wellen-Panel 2000 bis 2003 ( $P_{t_{1,2,3,4}}$ )

Der Datensatz des Vier-Wellen-Panels enthält alle Fälle, deren zugehörige (und nach Plausibilitätskontrollen gültige) Fragebögen drei (ebenfalls gültigen) Bögen drei weiterer Zeitpunkte zugeordnet werden konnten. Das Vier-Wellen-Panel der Jahre 2000 bis 2003 enthält 813 Fälle. Dies sind alle gültigen Fälle der Zeitpunkte  $t_1$ ,  $t_2$ ,  $t_3$  und  $t_4$ , deren Fragebögen einander zugeordnet werden konnten.

Wie in Tabelle 5.57 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine deutliche Verschiebung vor. Die Unterschiede im prozentualen Verhältnis zwischen Schülerinnen und Schülern der Querschnitte 2000, 2001, 2002 und 2003 im Vergleich zur Schulstatistik wurden in den Abschnitten 5.1.1, 5.1.2, 5.1.3 und 5.1.4 dargestellt. Im Paneldatensatz sind sowohl zum Zeitpunkt  $t_1$ ,  $t_2$ ,  $t_3$  als auch  $t_4$  mehr Schülerinnen als in den Referenzdaten vorhanden. Die Abweichung fällt dabei mit rund achteinhalb

Prozent zu den Zeitpunkten  $t_1$ ,  $t_2$  und  $t_3$  etwas stärker aus als mit knapp sieben Prozent zum Zeitpunkt  $t_4$ . Im Querschnitt machten zu  $t_4$  drei Personen keine Angaben.

Tabelle 5.57.: Zusammensetzung nach Geschlecht  $P_{t_1,2,3,4}$

		männlich		weiblich		gesamt $N$
		$N$	%	$N$	%	(100 %)
$t_1$	$S_{t_1}$	1097	52,1	1008	47,9	2105
	$Q_{t_1}$	931	50,3	919	49,7	1850
	$P_{t_1,2,3,4}$	353	43,4	460	56,6	813
$t_2$	$S_{t_2}$	1134	52,0	1047	48,0	2181
	$Q_{t_2}$	978	51,1	937	48,9	1915
	$P_{t_1,2,3,4}$	354	43,5	459	56,5	813
$t_3$	$S_{t_3}$	1170	52,0	1081	48,0	2251
	$Q_{t_3}$	1012	52,0	935	48,0	1947
	$P_{t_1,2,3,4}$	355	43,7	458	56,3	813
$t_4$	$S_{t_4}$	1056	50,8	1021	49,2	2077
	$Q_{t_4}$	915	50,4	901	49,6	1816
	$P_{t_1,2,3,4}$	357	43,9	456	56,1	813

Das Alter der Jugendlichen, deren Fälle im Paneldatensatz enthalten sind, weicht zu allen vier Zeitpunkten nicht gravierend vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.58). Hierbei gab es zu  $t_1$  1, zu  $t_2$  3, zu  $t_3$  8 und zu  $t_4$  2 fehlende Angaben im Paneldatensatz, zu  $t_1$  5, zu  $t_2$  26, zu  $t_3$  66 und zu  $t_4$  33 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 5.58.: Altersdurchschnitt  $P_{t_1,2,3,4}$

		Alter	Std.-abw.	gesamt $N$
$t_1$	$Q_{t_1}$	13,0	0,8	1845
	$P_{t_1,2,3,4}$	12,8	0,7	812
$t_2$	$Q_{t_2}$	14,0	0,8	1889
	$P_{t_1,2,3,4}$	13,8	0,6	810
$t_3$	$Q_{t_3}$	15,1	0,8	1881
	$P_{t_1,2,3,4}$	14,8	0,6	805
$t_4$	$Q_{t_4}$	16,0	0,8	1786
	$P_{t_1,2,3,4}$	15,8	0,6	811

Aus Tabelle 5.59 wird ersichtlich, dass der Paneldatensatz 2000 bis 2003 im Vergleich zur Schulstatistik systematisch verzerrt ist. Die Unterschiede im prozentualen Verhältnis der Panelquerschnitte und Querschnitte 2000, 2001, 2002 und 2003 im Ver-

gleich zur Schulstatistik wurden in den Abschnitten 5.1.1, 5.1.2, 5.1.3 und 5.1.4 dargestellt.

5.4. Das Vier-Wellen-Panel 2000 bis 2003 ( $P_{t_1,2,3,4}$ )

Tabelle 5.59.: Zusammensetzung  $P_{t_1,2,3,4}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt $N$
			HS	RS	GY	SO	(100 %)
männlich	$t_1$	$S_{t_1}$	23,8	30,6	39,6	6,0	1097
		$Q_{t_1}$	20,3	31,1	43,0	5,6	931
		$P_{t_1,2,3,4}$	13,0	32,3	53,3	1,4	353
	$t_2$	$S_{t_2}$	24,2	31,5	38,0	6,3	1134
		$Q_{t_2}$	23,6	32,1	37,7	6,5	978
		$P_{t_1,2,3,4}$	13,0	32,8	52,8	1,4	354
	$t_3$	$S_{t_3}$	26,1	31,9	35,0	7,1	1170
		$Q_{t_3}$	22,9	33,0	37,9	6,1	1012
		$P_{t_1,2,3,4}$	13,5	32,4	52,7	1,4	355
	$t_4$	$S_{t_4}$	25,0	30,9	37,3	6,8	1056
		$Q_{t_4}$	23,5	31,3	40,2	5,0	915
		$P_{t_1,2,3,4}$	13,7	32,5	52,4	1,4	357
weiblich	$t_1$	$S_{t_1}$	19,5	30,7	46,3	3,5	1008
		$Q_{t_1}$	17,6	30,5	48,2	3,7	919
		$P_{t_1,2,3,4}$	15,2	27,6	55,4	1,7	460
	$t_2$	$S_{t_2}$	20,4	31,9	43,6	4,1	1047
		$Q_{t_2}$	19,7	31,8	44,8	3,6	937
		$P_{t_1,2,3,4}$	15,5	27,5	55,3	1,7	459
	$t_3$	$S_{t_3}$	23,6	31,3	40,5	4,6	1081
		$Q_{t_3}$	21,3	32,3	42,9	3,5	935
		$P_{t_1,2,3,4}$	15,5	27,7	55,0	1,7	458
	$t_4$	$S_{t_4}$	21,6	32,1	41,1	5,1	1021
		$Q_{t_4}$	22,1	32,3	42,3	3,3	901
		$P_{t_1,2,3,4}$	15,8	27,9	54,6	1,8	456
gesamt	$t_1$	$S_{t_1}$	21,8	30,6	42,8	4,8	2105
		$Q_{t_1}$	19,0	30,8	45,6	4,6	1850
		$P_{t_1,2,3,4}$	14,3	29,6	54,5	1,6	813
	$t_2$	$S_{t_2}$	22,4	31,7	40,7	5,3	2181
		$Q_{t_2}$	21,7	32,0	41,2	5,1	1915
		$P_{t_1,2,3,4}$	14,4	29,8	54,2	1,6	813
	$t_3$	$S_{t_3}$	24,9	31,6	37,6	5,9	2251
		$Q_{t_3}$	22,1	32,7	40,3	4,9	1947
		$P_{t_1,2,3,4}$	14,6	29,8	54,0	1,6	813
	$t_4$	$S_{t_4}$	23,4	31,5	39,2	6,0	2077
		$Q_{t_4}$	22,8	31,8	41,2	4,2	1816
		$P_{t_1,2,3,4}$	14,9	29,9	53,6	1,6	813

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

Im Paneldatensatz befinden sich rund drei ( $t_1$ ), rund dreieinhalb ( $t_2$ ) und rund vierein-

halb Prozent ( $t_3$  und  $t_4$ ) weniger Sonderschülerinnen und -schüler, rund siebeneinhalb ( $t_1$ ), acht ( $t_2$ ), knapp zehneinhalb ( $t_3$ ) und achteinhalb Prozent ( $t_4$ ) weniger Hauptschülerinnen und -schüler, dafür rund elfeinhalb ( $t_1$ ), dreizehneinhalb ( $t_2$ ), knapp sechzehneinhalb ( $t_3$ ) und fast vierzehneinhalb Prozent ( $t_4$ ) mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik ausgewiesen. Der Anteil der Realschülerinnen und -schüler ist dabei zu allen Zeitpunkten annähernd gleich groß.

Tabelle 5.60.: Zusammensetzung  $P_{t_1,2,3,4}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)				gesamt
			HS	RS	GY	SO	(%)
männlich	$t_1$	$S_{t_1}$	57,0	52,1	48,2	65,3	52,1
		$Q_{t_1}$	53,8	50,9	47,4	60,5	50,3
		$P_{t_1,2,3,4}$	39,7	47,3	42,4	38,5	43,4
	$t_2$	$S_{t_2}$	56,1	51,7	48,6	62,6	52,0
		$Q_{t_2}$	55,5	51,3	46,8	65,3	51,1
		$P_{t_1,2,3,4}$	39,3	47,9	42,4	38,5	43,5
	$t_3$	$S_{t_3}$	54,5	52,5	48,3	62,4	52,0
		$Q_{t_3}$	53,8	52,5	48,9	65,3	52,0
		$P_{t_1,2,3,4}$	40,3	47,5	42,6	38,5	43,7
	$t_4$	$S_{t_4}$	54,4	49,8	48,4	58,1	50,8
		$Q_{t_4}$	51,9	49,6	49,1	60,5	50,4
		$P_{t_1,2,3,4}$	40,5	47,7	42,9	38,5	43,9
weiblich	$t_1$	$S_{t_1}$	43,0	47,9	51,8	34,7	47,9
		$Q_{t_1}$	46,2	49,1	52,6	39,5	49,7
		$P_{t_1,2,3,4}$	60,3	52,7	57,6	61,5	56,6
	$t_2$	$S_{t_2}$	43,9	48,3	51,4	37,4	48,0
		$Q_{t_2}$	44,5	48,7	53,2	34,7	48,9
		$P_{t_1,2,3,4}$	60,7	52,1	57,6	61,5	56,5
	$t_3$	$S_{t_3}$	45,5	47,5	51,7	37,6	48,0
		$Q_{t_3}$	46,2	47,5	51,1	34,7	48,0
		$P_{t_1,2,3,4}$	59,7	52,5	57,4	61,5	56,3
	$t_4$	$S_{t_4}$	45,6	50,2	51,6	41,9	49,2
		$Q_{t_4}$	48,1	50,4	50,9	39,5	49,6
		$P_{t_1,2,3,4}$	59,5	52,3	57,1	61,5	56,1

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

Über alle Schulformen hinweg sind zu allen vier Zeitpunkten prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.60). Der Unterschied zur Schulstatistik ist hierbei zu allen vier Zeitpunkten in der Sonderschule am größten und in der Realschule am geringsten. Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 5.5.

## 5.5. Zusammenfassung und Erklärung

Bei der Analyse der oben beschriebenen Zusammensetzungen 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 Panel- und 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.

### 5.5.1. 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 Sonderschülerinnen und -schüler sowie der Anteil der Hauptschülerinnen und -schüler ist geringer als in der Schulstatistik, teilweise sogar sehr viel niedriger. Dafür ist der Anteil der Gymnasias-tinnen und Gymnasiasten entsprechend höher, da der Anteil der Realschülerinnen und -schüler annähernd gleich groß ist wie in der Schulstatistik.

Diese Abweichungen sind 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 -schülern. 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.

Am Beispiel des Paneldatensatzes  $P_{t_3,t_4}$  (Münster 2002/2003) soll erläutert werden, welche Auswirkungen das angewendete Zuordnungsverfahren in Bezug auf die Repräsentativität der Daten hat.

In Tabelle 5.61 finden sich die Paneldaten aufgeschlüsselt nach den zugelassenen Fehlern in der Codezuordnung. Zum Vergleich sind dort die Münsteraner Querschnittsdaten aus 2003 und die Daten der Schulstatistik desselben Jahres aufgeführt.

Tabelle 5.61.: Verzerrung der Paneldaten Münster 2002/2003

	Codezuordnung $P_{t_3,t_4}$			gesamt	gesamt	gesamt
	0 Fehler	1 Fehler	2 Fehler	$P_{t_3,t_4}$	$Q_{t_4}$	$S_{t_4}$
N	637	523	246	1406	1819	2077
Geschlecht						
Anteil Mädchen (%)	59	49	41	52	50	49
Anteil Jungen (%)	41	51	59	48	50	51
Schulform						
Anteil Hauptschüler (%)	16	21	27	19	23	23
Anteil Realschüler (%)	33	33	33	33	32	31
Anteil Gymnasiasten (%)	50	44	34	45	41	39
Anteil Sonderschüler (%)	2	3	5	3	4	6

$P$  bedeutet Panel

$Q$  bedeutet Querschnitt

$S$  bedeutet Schulstatistik (nur befragte Schulen)

Es wird deutlich, dass zwar der größte Teil ( $N = 637$ ) der Schülerinnen und Schüler bei der Beantwortung des Codeblattes in 2002 und 2003 keinen Fehler machte, der Anteil derjenigen, die einen Fehler ( $N = 523$ ) machten, jedoch kaum kleiner ist. Der Anteil der Probanden, die zwei Fehler ( $N = 246$ ) machten ist zwar deutlich geringer, jedoch ebenfalls noch erheblich. Darüber hinaus wird ersichtlich, dass die Anzahl der Fehler bei der Beantwortung des Codeblattes mit Geschlecht und Schulbildung der Befragten korrelieren. Während die Verzerrungen auf Ebene des aggregierten Paneldatensatzes – der Zuordnungen bis zu zwei Fehler beinhaltet – nur leicht sind, 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.

### 5.5.2. Schulebene

Für die Teilnahme an den Befragungen war die Zustimmung der Schulkonferenzen der einzelnen Schulen notwendig. Schon an der ersten Befragung im Jahr 2000 nahmen nicht alle Schulen teil, während der weiteren Befragungen fiel eine weitere Schulen aus. Anhand der Daten des Jahres 2000 soll kurz die Systematik der Verzerrungen auf Schulebene dargestellt werden.



Wie in Tabelle 5.62 ersichtlich, liegt schon durch den Ausfall der Schulen im Jahr 2000 eine erhebliche Verschiebung der Daten in Bezug auf die Schulform vor. Sowohl die einzige Waldorf- als auch die einzige Gesamtschule Münsters nahm nicht an allen Zeitpunkten teil. Wegen des Ausfalls auf Schulebene sind im Vergleich zwischen reduzierter Schulstatistik und gesamter Schulstatistik mehr Haupt-, Real- und Sonderschüler und weniger Gymnasiasten vorhanden. Durch die Ausfälle auf der Individualebene verringern sich die Verzerrungen jedoch wieder. So ist die Verteilung nach Schulform sowohl im Querschnitt als auch im Panelquerschnitt der gesamten Schulstatistik ähnlicher als der reduzierten.

Tabelle 5.62.: Zusammensetzung der Daten Münster 2000 ( $t_1$ ) nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)						gesamt <i>N</i> (100 %)
		HS	RS	GY	SO	GS	WS	
männlich <sup>b</sup>	<i>S<sub>g</sub></i>	17,9	26,6	43,6	4,5	6,3	1,0	1458
	<i>S</i>	23,8	30,6	39,6	6,0	–	–	1097
	<i>Q</i>	20,3	31,1	43,0	5,6	–	–	931
	<i>P<sub>Q</sub></i>	19,2	30,5	46,8	3,5	–	–	665
weiblich <sup>b</sup>	<i>S<sub>g</sub></i>	14,4	25,5	50,1	2,6	6,7	1,0	1365
	<i>S</i>	19,5	30,7	46,3	3,5	–	–	1008
	<i>Q</i>	17,6	30,5	48,2	3,7	–	–	919
	<i>P<sub>Q</sub></i>	15,6	29,0	52,7	2,6	–	–	717
gesamt <sup>b</sup>	<i>S<sub>g</sub></i>	16,2	26,1	46,8	3,6	6,5	1,0	2823
	<i>S</i>	21,8	30,6	42,8	4,8	–	–	2105
	<i>Q</i>	19,0	30,8	45,6	4,6	–	–	1850
	<i>P<sub>Q</sub></i>	17,4	29,7	49,9	3,0	–	–	1382

<sup>a</sup>HS=Hauptschule, RS=Realschule, GY=Gymnasium, SO=Sonderschule

GS=Gesamtschule, WS=Waldorfschule

<sup>b</sup>*S<sub>g</sub>* 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.



## A. Anhang

### A.1. Häufigkeiten der Codes

Tabelle A.1.: Häufigkeiten der Codes in Münster

Häufigkeit	2000		2001		2002		2003	
	Anzahl Codes	gesamt	Anzahl Codes	gesamt	Anzahl Codes	gesamt	Anzahl Codes	gesamt
1	1497	1497	1538	1538	1533	1533	1447	1447
2	208	416	161	322	176	352	154	308
3	28	84	21	63	24	72	22	66
4	3	12	1	4	6	24	4	16
5	1	5	–	–	1	5	1	5
6	–	–	1	6	–	–	–	–
Summe <sup>a</sup>	1737	2014	1722	1933	1740	1986	1628	1842

<sup>a</sup> Die Gesamtzahlen für die Codes können von den einzelnen Stichprobengrößen abweichen. Zum einen wurden Codeblätter von Schulen, die nicht an allen Befragungszeitpunkten teilnahmen, zum anderen Codeblätter, deren Fragebögen durch die Plausibilitätskontrollen fielen, verwendet.

Tabelle A.2.: Häufigkeiten der aufgetretenen Codes in Münster 2003 (erweiterter Code)

	Anzahl Codes	Häufigkeit	gesamt
	1800	1	1800
	21	2	42
Summe	1821		1842

### A.2. Mathematische Herleitung

Fasst man die eindeutige Identifizierung der Codes als mathematisches Problem auf, so lässt sich folgende Fragestellung formulieren:

Bei einem Zufallsexperiment mit „Zurücklegen“ wird aus einer Menge der Größe  $N$  von unterschiedlichen Codes eine bestimmte Anzahl  $m$  von Codes gezogen. Die Frage ist nun, wie groß die Codemenge sein muss, damit die Wahrscheinlichkeit  $P(N, m)$ , dass kein Code mehrfach gezogen wird, mindestens einen bestimmten Wert, z. B. 0,95, übersteigt. Die Formel, die diese Wahrscheinlichkeit beschreibt, lautet:

$$P(N, m) = \frac{N!}{(N - m)! \cdot N^m}. \quad (\text{A.1})$$

Für große  $N$  ist die direkte Verwendung der Fakultätsfunktion unzweckmäßig, da die dann auftretenden Zahlen derartig groß sind, dass sie durch gebräuchliche Rechenmaschinen nicht mehr handhabbar sind. Die Fakultätsfunktion lässt sich jedoch nach der Formel von STIRLING (Forster 1983, S. 159 ff.) approximativ für natürliche Zahlen  $n$  durch

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

beschreiben, wobei  $e$  die EULERSche Konstante ist. Unter Benutzung dieser Formel erhält man eine Näherung  $P'$  von  $P$ :

$$\begin{aligned} P(N, m) \approx P'(N, m) &= \frac{1}{e^m} \cdot \left(\frac{N}{N-m}\right)^{N-m+\frac{1}{2}} \\ &= \exp\left(\left(N-m+\frac{1}{2}\right) \cdot (\ln(N) - \ln(N-m)) - m\right). \end{aligned}$$

Ein direktes Auflösen der Formel

$$P'(N, m) = 0,95,$$

die dem zu lösenden Problem entspricht, nach  $N$ , was zu einer Funktion  $N(m)$  führt, ist nicht möglich. Durch numerische Berechnung der Lösungen für verschiedene  $m$  lässt sich aber eine Näherungslösung finden. Sie lautet

$$N(m) \approx 9,7478x^2 - 9,4138x + 1,1389. \quad (\text{A.2})$$

Der relative Fehler dieser Approximation gegenüber der exakten Lösung beträgt im Bereich von  $m = 1000$  bis  $m = 10000$  höchstens 0,003% und ist somit hinreichend genau.

Der Fehler, der durch die Benutzung von  $P'$  statt von  $P$  entsteht, liegt in eben diesem Bereich für  $m$  und dem entsprechenden Bereich für  $N(m)$  bei höchstens 1,6%, was bedeutet, dass nicht notwendig die gewünschte Wahrscheinlichkeit von 95% erreicht wird, sondern dass die Wahrscheinlichkeit im Bereich  $0,95 \pm 1,6\%$ , d. h. zwischen 93,50% und 96,52% liegt.

### Zwischenergebnis


In der Münsteraner Umfrage von 2000 waren ca. 2 000 Schülerinnen und Schüler befragt worden, die auch ein Codeblatt ausfüllen mussten. Nach der Formel A.2 hätte die Codemenge  $N$  ca. 39 Millionen Codes enthalten müssen, damit bei einer Wahrscheinlichkeit von 95 Prozent von 2 000 Codes keiner mehrfach vorgekommen wäre. Die Codemenge  $N$  ist in unserem praktischen Fall die Menge aller möglichen Codekombinationen und lässt sich dementsprechend durch Multiplikation der Merkmalsausprägungen der einzelnen Codevariablen ausrechnen.

Die Betrachtung der aufgetretenen Häufigkeiten der Merkmalsausprägungen in den einzelnen Codevariablen zeigt, dass die so entstandene Codemenge viel zu gering ist,

um eine eindeutige Identifizierung der Codes zu ermöglichen. Allein die Merkmalsausprägungen der Variablen „Geburtstag“ (Co004) und „letzter Buchstabe Name der Oma“ (Co002) sind annähernd gleichverteilt (siehe im Anhang Tabellen A.3 bis A.7). Bei der Variable „Haarfarbe Vater“ (Co001) entfallen 2000 95 Prozent und 2001 94 Prozent der Antworten auf vier Nennungen (braun(n), schwar(z) gra(u), blon(d)), bei „eigener Haarfarbe“ (Co003) entfallen 2000 96 Prozent, 2001 97 Prozent auf nur drei Angaben (blon(d), brau(n), schwar(z)), bei „eigener Augenfarbe“ (Co005) sind es 2000 und 2001 sogar 98 Prozent auf nur zwei Merkmalsausprägungen (brau(n), bla(u)).

Die aufgetretenen Merkmalsausprägungen, deren Häufigkeiten kumuliert über 95 Prozent liegen, ergeben eine Kombinationsanzahl von ca. 18 000. Es ist also offensichtlich, dass allein die Informationen aus den Codes zur eindeutigen Identifizierung nicht ausreichen, was nicht an der Anzahl der Fragen, sondern an den aufgetretenen Häufigkeiten der Merkmalsausprägungen liegt.

### A.3. Codeblätter



WESTFÄLISCHE WILHELMS-UNIVERSITÄT MÜNSTER  
RECHTSWISSENSCHAFTLICHE FAKULTÄT  
Prof. Dr. Klaus Boers

Bearbeitungsvermerk

**Kriminologische Befragung Münsteraner Schüler**  
**Beiblatt zur Erstellung des persönlichen Codes**

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.

Das Codewort wird von uns getrennt vom Fragebogen eingegeben und anschließend mit einem Computerprogramm verschlüsselt. Danach wird dieses Blatt vernichtet.

Wichtig ist, daß Du dasselbe Codewort nächstes Jahr noch weißt. Aus diesem Grund haben wir nachfolgend fünf Fragen formuliert, die Dir bei der nächsten Befragung helfen sollen, dasselbe Wort wieder zu erinnern.

Hier nun die fünf Fragen zur Erstellung Deines persönlichen Codes:

	Bitte trage hier ins erste Feld den letzten Buchstaben der natürlichen Haarfarbe Deines Vaters (oder einer Person, die für Dich einem Vater am nächsten kommt) ein. (z.B. Glatz <b>e</b> , braun <b>n</b> , schwarz <b>z</b> , blond <b>d</b> , grau <b>u</b> usw.)
	Bitte trage ins zweite Feld den ersten Buchstaben des Vornamens Deiner Oma mütterlicherseits (oder einer Person, die für Dich wie eine Oma ist) ein. (z.B. <b>P</b> aula, <b>C</b> hristine, <b>S</b> abine, <b>J</b> utta, <b>J</b> osefine usw.)
	Bitte trage ins dritte Feld den letzten Buchstaben Deiner natürlichen Haarfarbe ein. (z.B. Glatz <b>e</b> , braun <b>n</b> , schwarz <b>z</b> , blond <b>d</b> , grau <b>u</b> usw.)
	Bitte trage ins vierte und fünfte Feld die beiden Tagesziffer Deines Geburtstages ein (z.B. <b>0</b> <b>1</b> , <b>0</b> <b>2</b> , <b>0</b> <b>3</b> , ... , <b>2</b> <b>0</b> , <b>3</b> <b>1</b> ).
	Bitte trage ins sechste Feld den letzten Buchstaben Deiner Augenfarbe ein. (z.B. braun <b>n</b> , blau <b>u</b> , grün <b>n</b> , grau <b>u</b> usw.)

Abbildung A.1.: Codeblatt Münster 2000



WESTFÄLISCHE WILHELMS-UNIVERSITÄT MÜNSTER  
**Prof. Dr. Klaus Boers**  
 Direktor des Instituts für Kriminalwissenschaften, Abteilung Kriminologie  
**PD 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 Münsteraner Schüler**  
**Beiblatt zur Erstellung des persönlichen Codes**

Januar 2001

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 den selben Vorgang wiederholen, können deine beiden 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 nachfolgend fünf Fragen formuliert, die dir helfen sollen, dich an dasselbe Wort zu erinnern.

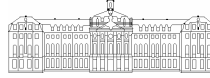
**Hier nun die fünf Fragen zur Erstellung deines persönlichen Codes:**

	Bitte trage hier ins erste Feld den letzten Buchstaben der natürlichen Haarfarbe deines Vaters (oder einer Person, die für dich einem Vater am nächsten kommt) ein. (z.B. Glatz <b>z</b> , braun <b>n</b> , schwarz <b>z</b> , blond <b>d</b> , grau <b>n</b> usw.)
	Bitte trage ins zweite Feld den ersten Buchstaben des Vornamens deiner Oma mütterlicherseits (oder einer Person, die für dich wie eine Oma ist) ein. (z.B. <b>P</b> aula, <b>C</b> hristine, <b>S</b> abine, <b>J</b> utta, <b>J</b> osefine usw.)
	Bitte trage ins dritte Feld den letzten Buchstaben deiner natürlichen Haarfarbe ein. (z.B. Glatz <b>z</b> , braun <b>n</b> , schwarz <b>z</b> , blond <b>d</b> , grau <b>n</b> usw.)
	Bitte trage ins vierte und fünfte Feld den Tag deines Geburtstages ein (z.B. <b>01</b> , <b>02</b> , <b>03</b> , ... , <b>30</b> , <b>31</b> ).
	Bitte trage ins sechste Feld den letzten Buchstaben deiner Augenfarbe ein. (z.B. braun <b>n</b> , blau <b>n</b> , grün <b>n</b> , grau <b>n</b> usw.)

Hast du im letzten Jahr die Klasse gewechselt?  ja  nein

Hast du im letzten Jahr die Schule gewechselt?  ja  nein

Abbildung A.2.: Codeblatt Münster 2001



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 Münsteraner Schüler**  
**Beiblatt zur Erstellung des persönlichen Codes**

Januar/Februar 2002

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 den selben 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 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 letzten Buchstaben der natürlichen Haarfarbe deines Vaters (oder einer Person, die für dich einem Vater am nächsten kommt) ein. (z.B. Glatz <input type="checkbox"/> , braun <input type="checkbox"/> , schwarz <input type="checkbox"/> , blond <input type="checkbox"/> , grau <input type="checkbox"/> usw.)
	Bitte trage ins zweite Feld den ersten Buchstaben des Vornamens deiner Oma mütterlicherseits (oder einer Person, die für dich wie eine Oma ist) ein. (z.B. Paula <input type="checkbox"/> , Christine <input type="checkbox"/> , Sabine <input type="checkbox"/> , Jutta <input type="checkbox"/> , Josefine usw.)
	Bitte trage ins dritte Feld den letzten Buchstaben deiner natürlichen Haarfarbe ein. (z.B. Glatz <input type="checkbox"/> , braun <input type="checkbox"/> , schwarz <input type="checkbox"/> , blond <input type="checkbox"/> , grau <input type="checkbox"/> usw.)
	Bitte trage ins vierte und fünfte Feld den Tag deines Geburtstages ein (z.B. Geburtstag am 7. Januar = <input type="checkbox"/> <input type="checkbox"/> , am 12. Mai = <input type="checkbox"/> <input type="checkbox"/> , am 31. Oktober = <input type="checkbox"/> <input type="checkbox"/> )
	Bitte trage ins sechste Feld den letzten Buchstaben deiner Augenfarbe ein. (z.B. braun <input type="checkbox"/> , blau <input type="checkbox"/> , grün <input type="checkbox"/> , grau <input type="checkbox"/> usw.)

Hast du im letzten Jahr an der Befragung teilgenommen?  ja  nein

Hast du im letzten Jahr die Klasse gewechselt?  ja  nein

Hast du im letzten Jahr die Schule gewechselt?  ja  nein

Abbildung A.3.: Codeblatt Münster 2002



**Beiblatt zur Erstellung des persönlichen Codes** Januar/Februar 2003

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 den selben 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:**

<b>1</b>	Bitte kreuze den <b>ersten</b> Buchstaben des Vornamens deiner <b>Oma mütterlicherseits</b> (oder einer Person, die für dich wie eine Oma ist) an. (z. B. Paula, Christine, Sabine, Uta, usw.) <table style="width: 100%; text-align: center; border-collapse: collapse;"> <tr><td><input type="checkbox"/> a</td><td><input type="checkbox"/> b</td><td><input type="checkbox"/> c</td><td><input type="checkbox"/> d</td><td><input type="checkbox"/> e</td><td><input type="checkbox"/> f</td><td><input type="checkbox"/> g</td><td><input type="checkbox"/> h</td><td><input type="checkbox"/> i</td><td><input type="checkbox"/> j</td><td><input type="checkbox"/> k</td><td><input type="checkbox"/> l</td><td><input type="checkbox"/> m</td><td><input type="checkbox"/> n</td><td><input type="checkbox"/> o</td></tr> <tr><td><input type="checkbox"/> p</td><td><input type="checkbox"/> q</td><td><input type="checkbox"/> r</td><td><input type="checkbox"/> s</td><td><input type="checkbox"/> t</td><td><input type="checkbox"/> u</td><td><input type="checkbox"/> v</td><td><input type="checkbox"/> w</td><td><input type="checkbox"/> x</td><td><input type="checkbox"/> y</td><td><input type="checkbox"/> z</td><td><input type="checkbox"/> ä</td><td><input type="checkbox"/> ö</td><td><input type="checkbox"/> ü</td><td><input type="checkbox"/> ß</td></tr> </table>	<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	
<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																		
<b>2</b>	Bitte kreuze den <b>ersten</b> Buchstaben deines <b>Vornamens</b> an (z. B. Michael, Thomas, Ute usw.) <table style="width: 100%; text-align: center; border-collapse: collapse;"> <tr><td><input type="checkbox"/> a</td><td><input type="checkbox"/> b</td><td><input type="checkbox"/> c</td><td><input type="checkbox"/> d</td><td><input type="checkbox"/> e</td><td><input type="checkbox"/> f</td><td><input type="checkbox"/> g</td><td><input type="checkbox"/> h</td><td><input type="checkbox"/> i</td><td><input type="checkbox"/> j</td><td><input type="checkbox"/> k</td><td><input type="checkbox"/> l</td><td><input type="checkbox"/> m</td><td><input type="checkbox"/> n</td><td><input type="checkbox"/> o</td></tr> <tr><td><input type="checkbox"/> p</td><td><input type="checkbox"/> q</td><td><input type="checkbox"/> r</td><td><input type="checkbox"/> s</td><td><input type="checkbox"/> t</td><td><input type="checkbox"/> u</td><td><input type="checkbox"/> v</td><td><input type="checkbox"/> w</td><td><input type="checkbox"/> x</td><td><input type="checkbox"/> y</td><td><input type="checkbox"/> z</td><td><input type="checkbox"/> ä</td><td><input type="checkbox"/> ö</td><td><input type="checkbox"/> ü</td><td><input type="checkbox"/> ß</td></tr> </table>	<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	
<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																		
<b>3</b>	Bitte kreuze den <b>Tag</b> deines <b>Geburtsdatums</b> an. (z.B. Geburtstag am 7. Januar = <input type="checkbox"/> 7, am 12. Mai = <input type="checkbox"/> 12, am 31. Oktober = <input type="checkbox"/> 31) <table style="width: 100%; text-align: center; border-collapse: collapse;"> <tr><td><input type="checkbox"/> 1</td><td><input type="checkbox"/> 2</td><td><input type="checkbox"/> 3</td><td><input type="checkbox"/> 4</td><td><input type="checkbox"/> 5</td><td><input type="checkbox"/> 6</td><td><input type="checkbox"/> 7</td><td><input type="checkbox"/> 8</td><td><input type="checkbox"/> 9</td><td><input type="checkbox"/> 10</td><td><input type="checkbox"/> 11</td><td><input type="checkbox"/> 12</td><td><input type="checkbox"/> 13</td><td><input type="checkbox"/> 14</td><td><input type="checkbox"/> 15</td></tr> <tr><td><input type="checkbox"/> 16</td><td><input type="checkbox"/> 17</td><td><input type="checkbox"/> 18</td><td><input type="checkbox"/> 19</td><td><input type="checkbox"/> 20</td><td><input type="checkbox"/> 21</td><td><input type="checkbox"/> 22</td><td><input type="checkbox"/> 23</td><td><input type="checkbox"/> 24</td><td><input type="checkbox"/> 25</td><td><input type="checkbox"/> 26</td><td><input type="checkbox"/> 27</td><td><input type="checkbox"/> 28</td><td><input type="checkbox"/> 29</td><td><input type="checkbox"/> 30</td><td><input type="checkbox"/> 31</td></tr> </table>	<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
<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																	
<b>4</b>	Bitte kreuze den <b>letzten</b> Buchstaben der natürlichen Haarfarbe deines <b>Vaters</b> (oder einer Person, die für dich einem Vater am nächsten kommt) an. (z. B. Glatz, braun, schwarz, usw.) <table style="width: 100%; text-align: center; border-collapse: collapse;"> <tr><td><input type="checkbox"/> a</td><td><input type="checkbox"/> b</td><td><input type="checkbox"/> c</td><td><input type="checkbox"/> d</td><td><input type="checkbox"/> e</td><td><input type="checkbox"/> f</td><td><input type="checkbox"/> g</td><td><input type="checkbox"/> h</td><td><input type="checkbox"/> i</td><td><input type="checkbox"/> j</td><td><input type="checkbox"/> k</td><td><input type="checkbox"/> l</td><td><input type="checkbox"/> m</td><td><input type="checkbox"/> n</td><td><input type="checkbox"/> o</td></tr> <tr><td><input type="checkbox"/> p</td><td><input type="checkbox"/> q</td><td><input type="checkbox"/> r</td><td><input type="checkbox"/> s</td><td><input type="checkbox"/> t</td><td><input type="checkbox"/> u</td><td><input type="checkbox"/> v</td><td><input type="checkbox"/> w</td><td><input type="checkbox"/> x</td><td><input type="checkbox"/> y</td><td><input type="checkbox"/> z</td><td><input type="checkbox"/> ä</td><td><input type="checkbox"/> ö</td><td><input type="checkbox"/> ü</td><td><input type="checkbox"/> ß</td></tr> </table>	<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	
<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																		
<b>5</b>	Bitte kreuze den <b>letzten</b> Buchstaben deiner natürlichen <b>Haarfarbe</b> an. (z. B. Glatz, braun, schwarz, blond, grau usw.) <table style="width: 100%; text-align: center; border-collapse: collapse;"> <tr><td><input type="checkbox"/> a</td><td><input type="checkbox"/> b</td><td><input type="checkbox"/> c</td><td><input type="checkbox"/> d</td><td><input type="checkbox"/> e</td><td><input type="checkbox"/> f</td><td><input type="checkbox"/> g</td><td><input type="checkbox"/> h</td><td><input type="checkbox"/> i</td><td><input type="checkbox"/> j</td><td><input type="checkbox"/> k</td><td><input type="checkbox"/> l</td><td><input type="checkbox"/> m</td><td><input type="checkbox"/> n</td><td><input type="checkbox"/> o</td></tr> <tr><td><input type="checkbox"/> p</td><td><input type="checkbox"/> q</td><td><input type="checkbox"/> r</td><td><input type="checkbox"/> s</td><td><input type="checkbox"/> t</td><td><input type="checkbox"/> u</td><td><input type="checkbox"/> v</td><td><input type="checkbox"/> w</td><td><input type="checkbox"/> x</td><td><input type="checkbox"/> y</td><td><input type="checkbox"/> z</td><td><input type="checkbox"/> ä</td><td><input type="checkbox"/> ö</td><td><input type="checkbox"/> ü</td><td><input type="checkbox"/> ß</td></tr> </table>	<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	
<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																		
<b>6</b>	Bitte kreuze den <b>letzten</b> Buchstaben deiner <b>Augenfarbe</b> an. (z. B. braun, grün, grau, usw.) <table style="width: 100%; text-align: center; border-collapse: collapse;"> <tr><td><input type="checkbox"/> a</td><td><input type="checkbox"/> b</td><td><input type="checkbox"/> c</td><td><input type="checkbox"/> d</td><td><input type="checkbox"/> e</td><td><input type="checkbox"/> f</td><td><input type="checkbox"/> g</td><td><input type="checkbox"/> h</td><td><input type="checkbox"/> i</td><td><input type="checkbox"/> j</td><td><input type="checkbox"/> k</td><td><input type="checkbox"/> l</td><td><input type="checkbox"/> m</td><td><input type="checkbox"/> n</td><td><input type="checkbox"/> o</td></tr> <tr><td><input type="checkbox"/> p</td><td><input type="checkbox"/> q</td><td><input type="checkbox"/> r</td><td><input type="checkbox"/> s</td><td><input type="checkbox"/> t</td><td><input type="checkbox"/> u</td><td><input type="checkbox"/> v</td><td><input type="checkbox"/> w</td><td><input type="checkbox"/> x</td><td><input type="checkbox"/> y</td><td><input type="checkbox"/> z</td><td><input type="checkbox"/> ä</td><td><input type="checkbox"/> ö</td><td><input type="checkbox"/> ü</td><td><input type="checkbox"/> ß</td></tr> </table>	<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	
<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 WWU Münster, Kriminalwissenschaften/Universität Trier, Soziologie – Münsteraner Schülerbefragung 2003

Abbildung A.4.: Codeblatt Münster 2003

## A.4. Häufigkeiten der Codevariablen

Tabelle A.3.: Häufigkeiten von Co001 in Münster (Haarfarbe Vater)

	2000		2001		2002		2003	
	N	%	N	%	N	%	N	%
a	3	0,1	—	—	1	0,1	2	0,1
ä	—	—	—	—	—	—	1	0,1
b	7	0,3	7	0,4	15	0,8	31	1,7
d	220	10,9	214	11,1	232	11,7	226	12,1
e	56	2,8	75	3,9	71	3,6	59	3,2
f	—	—	—	—	—	—	2	0,1
g	2	0,1	2	0,1	3	0,2	12	0,6
h	—	—	—	—	—	—	1	0,1
i	—	—	—	—	2	0,1	2	0,1
j	—	—	—	—	1	0,1	2	0,1
k	—	—	—	—	1	0,1	2	0,1
l	1	0,0	1	0,1	2	0,1	—	—
m	—	—	—	—	1	0,1	6	0,3
n	741	36,8	726	37,6	729	36,7	677	36,3
ö	—	—	—	—	—	—	1	0,1
r	—	—	1	0,1	2	0,1	4	0,2
s	4	0,2	11	0,6	14	0,7	30	1,6
ß	—	—	2	0,1	—	—	6	0,3
t	18	0,9	14	0,7	16	0,8	20	1,1
u	221	11,0	219	11,3	205	10,3	120	6,4
w	—	—	1	0,1	2	0,1	2	0,1
x	—	—	1	0,1	—	—	2	0,1
z	738	36,6	654	33,8	673	33,9	607	32,6
fehlend	3	0,1	5	0,3	16	0,8	44	2,4
gesamt	2014	100,0	1933	100,0	1986	100,0	1864	100,0

#### A.4. Häufigkeiten der Codevariablen

Tabelle A.4.: Häufigkeiten von Co002 in Münster (Vorname Oma)

	MS2000		2001		2002		2003	
	N	%	N	%	N	%	N	%
a	236	11,7	186	9,6	215	10,8	199	10,7
ä	—	—	7	0,4	5	0,3	4	0,2
b	39	1,9	44	2,3	34	1,7	32	1,7
c	75	3,7	68	3,5	82	4,1	63	3,4
d	32	1,6	29	1,5	27	1,4	24	1,3
e	226	11,2	201	10,4	226	11,4	193	10,4
f	37	1,8	27	1,4	29	1,5	26	1,4
g	107	5,3	105	5,4	104	5,2	106	5,7
h	200	9,9	181	9,4	170	8,6	174	9,3
i	90	4,5	72	3,7	139	7,0	113	6,1
j	110	5,5	104	5,4	63	3,2	55	3,0
k	69	3,4	63	3,3	63	3,2	74	4,0
l	84	4,2	83	4,3	81	4,1	91	4,9
m	291	14,4	258	13,3	280	14,1	262	14,1
n	18	0,9	25	1,3	28	1,4	31	1,7
o	97	4,8	148	7,7	65	3,1	30	1,6
ö	—	—	—	—	1	0,1	2	0,1
p	30	1,5	36	1,9	38	1,9	25	1,3
q	—	—	1	0,1	—	—	—	—
r	77	3,8	77	4,0	70	3,5	71	3,8
s	50	2,5	52	2,7	56	2,8	48	2,6
ß	—	—	1	0,1	—	—	—	—
t	39	1,9	44	2,3	39	2,0	36	1,9
u	39	1,9	50	2,6	50	2,5	46	2,5
v	9	0,4	8	0,4	7	0,4	10	0,5
w	33	1,6	38	2,0	42	2,1	43	2,3
x	—	—	—	—	2	0,1	3	0,2
y	1	0,0	2	0,1	2	0,1	5	0,3
z	8	0,4	9	0,5	3	0,2	10	0,5
fehlend	17	0,8	14	0,7	65	3,3	86	4,6
gesamt	2014	100,0	1933	100,0	1986	100,0	1864	100,0

Tabelle A.5.: Häufigkeiten von Co003 in Münster (eigene Haarfarbe)

	2000		2001		2002		2003	
	N	%	N	%	N	%	N	%
a	–	–	–	–	1	0,1	–	–
b	31	1,5	21	1,1	19	1,0	26	1,4
c	–	–	–	–	–	–	1	0,1
d	953	47,3	858	44,4	842	42,4	764	41,0
e	3	0,1	3	0,2	15	0,8	11	0,6
g	–	–	–	–	1	0,1	5	0,3
h	–	–	1	0,1	1	0,1	3	0,2
j	–	–	1	0,1	1	0,1	1	0,1
k	–	–	1	0,1	–	–	4	0,2
l	1	0,0	–	–	–	–	–	–
m	–	–	–	–	1	0,1	2	0,1
n	836	41,5	823	42,6	859	43,3	784	42,1
o	2	0,1	2	0,1	–	–	–	–
ö	–	–	–	–	–	–	4	0,2
r	4	0,2	4	0,2	2	0,1	4	0,2
s	4	0,2	3	0,2	2	0,1	10	0,5
ß	–	–	–	–	–	–	2	0,1
t	29	1,4	27	1,4	32	1,6	31	1,7
u	2	0,1	5	0,3	8	0,4	14	0,8
v	–	–	–	–	–	–	1	0,1
w	–	–	1	0,1	–	–	1	0,1
x	–	–	–	–	–	–	3	0,2
y	–	–	–	–	–	–	1	0,1
z	147	7,3	183	9,5	192	9,7	148	7,9
fehlend	2	0,1	–	–	10	0,5	35	1,9
gesamt	2014	100,0	1933	100,0	1986	100,0	1864	100,0

#### A.4. Häufigkeiten der Codevariablen

Tabelle A.6.: Häufigkeiten von Co004 in Münster (eigener Geburtstag)

	2000		2001		2002		2003	
	N	%	N	%	N	%	N	%
0	2	0,1	–	–	–	–	–	–
1	55	2,7	54	2,8	56	2,8	54	2,9
2	57	2,8	58	3,0	57	2,9	55	3,0
3	70	3,5	68	3,5	75	3,8	65	3,5
4	74	3,7	79	4,1	74	3,7	69	3,7
5	79	3,9	66	3,4	70	3,5	57	3,1
6	74	3,7	74	3,8	76	3,8	74	4,0
7	71	3,5	62	3,2	72	3,6	59	3,2
8	71	3,5	66	3,4	68	3,4	69	3,7
9	74	3,7	74	3,8	77	3,9	58	3,1
10	64	3,2	65	3,4	74	3,7	72	3,9
11	67	3,3	68	3,5	70	3,5	64	3,4
12	61	3,0	65	3,4	70	3,5	63	3,4
13	76	3,8	59	3,1	68	3,4	68	3,6
14	54	2,7	54	2,8	40	2,0	49	2,6
15	65	3,2	66	3,4	69	3,5	58	3,1
16	74	3,7	64	3,3	64	3,2	61	3,3
17	64	3,2	48	2,5	54	2,7	54	2,9
18	77	3,8	72	3,7	75	3,8	75	4,0
19	71	3,5	73	3,8	72	3,6	62	3,3
20	61	3,0	58	3,0	58	2,9	50	2,7
21	58	2,9	51	2,6	57	2,9	48	2,6
22	54	2,7	48	2,5	60	3,0	47	2,5
23	68	3,4	62	3,2	58	2,9	56	3,0
24	61	3,0	71	3,7	78	3,9	67	3,6
25	65	3,2	54	2,8	58	2,9	55	3,0
26	68	3,4	57	2,9	61	3,1	55	3,0
27	69	3,4	70	3,6	62	3,1	62	3,3
28	49	2,4	52	2,7	47	2,4	46	2,5
29	61	3,0	56	2,9	51	2,6	52	2,8
30	57	2,8	54	2,8	54	2,7	51	2,7
31	43	2,1	42	2,2	52	2,6	46	2,5
32	–	–	1	0,1	–	–	–	–
34	–	–	2	0,1	–	–	–	–
38	–	–	2	0,1	–	–	–	–
44	–	–	1	0,1	–	–	–	–
46	–	–	1	0,1	–	–	–	–
56	–	–	1	0,1	–	–	–	–
58	–	–	1	0,1	1	0,1	–	–
62	–	–	1	0,1	–	–	–	–
68	–	–	1	0,1	–	–	–	–
69	–	–	–	–	1	0,1	–	–
71	–	–	1	0,1	–	–	–	–
72	–	–	1	0,1	–	–	–	–
74	–	–	1	0,1	–	–	–	–
78	–	–	1	0,1	–	–	–	–
85	–	–	2	0,1	–	–	–	–
86	–	–	1	0,1	–	–	–	–
99	–	–	3	0,2	–	–	–	–
111	–	–	–	–	1	0,1	–	–
1910	–	–	–	–	1	0,1	–	–
1986	–	–	1	0,1	–	–	–	–
fehlend	–	–	–	–	–	–	43	2,3
gesamt	2014	100,0	1933	100,0	1986	100,0	1864	100,0

Tabelle A.7.: Häufigkeiten von Co005 in Münster (eigene Augenfarbe)

	2000		2001		2002		2003	
	N	%	N	%	N	%	N	%
l	–	–	1	0,1	–	–	–	–
a	2	0,1	1	0,1	3	0,2	2	0,1
ä	–	–	–	–	–	–	3	0,2
b	11	0,5	8	0,4	15	0,8	29	1,6
d	–	–	3	0,2	2	0,1	4	0,2
e	1	0,0	–	–	–	–	2	0,1
f	–	–	–	–	–	–	2	0,1
g	6	0,3	7	0,4	6	0,3	8	0,4
h	–	–	1	0,1	–	–	–	–
i	–	–	–	–	–	–	4	0,2
j	–	–	–	–	1	0,1	2	0,1
k	–	–	1	0,1	1	0,1	4	0,2
m	–	–	–	–	1	0,1	2	0,1
n	1150	57,1	1114	57,6	1146	57,7	1022	54,8
o	–	–	–	–	–	–	1	0,1
ö	–	–	–	–	–	–	1	0,1
p	–	–	–	–	–	–	1	0,1
r	–	–	–	–	3	0,2	–	–
s	2	0,1	4	0,2	2	0,1	6	0,3
ß	–	–	–	–	–	–	4	0,2
t	2	0,1	2	0,1	6	0,3	2	0,1
u	831	41,3	778	40,2	781	39,3	701	37,6
ü	–	–	1	0,1	–	–	2	0,1
v	–	–	–	–	1	0,1	1	0,1
w	–	–	2	0,1	–	–	1	0,1
x	–	–	–	–	–	–	2	0,1
y	1	0,0	–	–	–	–	4	0,2
z	6	0,3	9	0,5	5	0,3	10	0,5
fehlend	2	0,1	1	0,1	13	0,7	44	2,4
gesamt	2014	100,0	1933	100,0	1986	100,0	1864	100,0

#### A.4. Häufigkeiten der Codevariablen

Table A.8.: Häufigkeiten von Co009 in Münster (eigener Vorname)

	2000		2001		2002		2003	
	N	%	N	%	N	%	N	%
a	—	—	—	—	—	—	180	9,7
ä	—	—	—	—	—	—	3	0,2
b	—	—	—	—	—	—	43	2,3
c	—	—	—	—	—	—	111	6,0
d	—	—	—	—	—	—	102	5,5
e	—	—	—	—	—	—	45	2,4
f	—	—	—	—	—	—	50	2,7
g	—	—	—	—	—	—	26	1,4
h	—	—	—	—	—	—	43	2,3
i	—	—	—	—	—	—	37	2,0
j	—	—	—	—	—	—	188	10,1
k	—	—	—	—	—	—	92	4,9
l	—	—	—	—	—	—	89	4,8
m	—	—	—	—	—	—	226	12,1
n	—	—	—	—	—	—	80	4,3
o	—	—	—	—	—	—	20	1,1
ö	—	—	—	—	—	—	4	0,2
p	—	—	—	—	—	—	63	3,4
q	—	—	—	—	—	—	1	0,1
r	—	—	—	—	—	—	62	3,3
s	—	—	—	—	—	—	203	10,9
ß	—	—	—	—	—	—	2	0,1
t	—	—	—	—	—	—	103	5,5
u	—	—	—	—	—	—	10	0,5
v	—	—	—	—	—	—	27	1,4
w	—	—	—	—	—	—	9	0,5
x	—	—	—	—	—	—	1	0,1
y	—	—	—	—	—	—	13	0,7
z	—	—	—	—	—	—	4	0,2
fehlend	—	—	—	—	—	—	27	1,4
gesamt	—	—	—	—	—	—	1864	100,0

## A.5. Schul-, Klasseninformation und Befragungsteilnahme

### A.5.1. Münster 2001

*Tabelle A.9.: Häufigkeiten von Co006 in Münster 2001 (Klassenwechsel ja/nein)*

		N	%	gültige %
gültig	ja	324	16,8	18,2
	nein	1456	75,3	81,8
	gesamt	1780	92,1	100,0
fehlend		153	7,9	
gesamt		1933	100,0	

*Tabelle A.10.: Häufigkeiten von Co007 in Münster 2001 (Schulwechsel ja/nein)*

		N	%	gültige %
gültig	ja	166	8,6	9,3
	nein	1610	83,3	90,7
	gesamt	1776	91,9	100,0
fehlend		157	8,1	
gesamt		1933	100,0	

### A.5.2. Münster 2002

*Tabelle A.11.: Häufigkeiten von Co006 in Münster 2002 (Klassenwechsel ja/nein)*

		N	%	gültige %
gültig	ja	242	12,2	12,7
	nein	1669	84,0	87,3
	gesamt	1911	96,2	100,0
fehlend		75	3,8	
gesamt		1986	100,0	



## A.5. Schul-, Klasseninformation und Befragungsteilnahme

Tabelle A.12.: Häufigkeiten von Co007 in Münster 2002 (Schulwechsel ja/nein)

		N	%	gültige %
gültig	ja	142	7,2	7,4
	nein	1773	89,3	92,6
	gesamt	1915	96,4	100,0
fehlend		71	3,6	
gesamt		1986	100,0	

Tabelle A.13.: Häufigkeiten von Co008 in Münster 2002 (BefragungsteilnehmerIn ja/nein)

		N	%	gültige %
gültig	ja	1706	85,9	88,4
	nein	223	11,2	11,6
	gesamt	1929	97,1	100,0
fehlend		57	2,9	
gesamt		1986	100,0	

### A.5.3. Münster 2003

Tabelle A.14.: Häufigkeiten von Co006 in Münster 2003 (Sitzengeblieben ja/nein)

		N	%	gültige %
gültig	ja	55	3,0	3,6
	nein	1464	78,5	96,4
	gesamt	1519	81,5	100,0
fehlend		345	18,5	
gesamt		1864	100,0	

*Tabelle A.15.: Häufigkeiten von Co007 in Münster 2003 (Schulwechsel ja/nein)*

		<i>N</i>	%	gültige %
gültig	ja	71	3,8	4,7
	nein	1450	77,8	95,3
	gesamt	1521	81,6	100,0
fehlend		343	18,4	
gesamt		1986	100,0	

*Tabelle A.16.: Häufigkeiten von Co008 in Münster 2003 (BefragungsteilnehmerIn ja/nein)*

		<i>N</i>	%	gültige %
gültig	ja	1456	78,1	92,7
	nein	115	6,2	7,1
	gesamt	1571	84,3	100,0
fehlend		293	15,7	
gesamt		1864	100,0	

## Literaturverzeichnis

- Brondies, M. (2004). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2003* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 6). Münster, Trier.
- Böttger, A., Ehret, B., Othold, F., Prein, G., Schumann, K. F. & Seus, L. (2003). Methoden der Untersuchung. In K. F. Schumann (Hg.), *Delinquenz im Lebensverlauf. Bremer Längsschnittstudie zum Übergang von der Schule in den Beruf bei ehemaligen Hauptschülern* (Bd. 2, S. 35–66). Weinheim, München: Juventa.
- Forster, O. (1983). *Analysis I, Differential- und Integralrechnung einer Veränderlichen* (vierte Aufl.). Braunschweig, Wiesbaden: Vieweg.
- Motzke, K. & Brondies, M. (2004). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 5). Münster, Trier.
- Motzke, K. & Wittenberg, J. (2004). *Methodendokumentation der kriminologischen Schülerbefragung in Münster 2000* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 1). Münster, Trier.
- Wittenberg, J. (2004a). *Methodendokumentation der kriminologischen Schülerbefragung in Münster 2001* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 2). Münster, Trier.
- Wittenberg, J. (2004b). *Methodendokumentation der kriminologischen Schülerbefragung in Münster 2002* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 4). Münster, Trier.
- Wittenberg, J. (2004c). *Methodendokumentation der kriminologischen Schülerbefragung in Münster 2003* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 7). Münster, Trier.
- Wittenberg, J. & Hilfert, N. (2004). *Methodendokumentation der kriminologischen Schülerbefragung in Bocholt 2001* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 3). Münster, Trier.