

## **Kommentarband zum Abschlußbericht:**

# **Die „G8 und mehr“ Umfrage der Landeselternschaft der Gymnasien NRW e.V.**

von

**Prof. Dr. Rainer Dollase**

Uni Bielefeld Abt. Psychologie, AE 12 „Psychologie in Bildung und Erziehung“

Steinhagen, August 2016

## **Gliederung:**

### **Vorwort**

- 1. Zusammenfassung der Ergebnisse der Umfrage „G8 und mehr“**
- 2. Vorgeschichte der Fragestellungen zu „G8 und mehr“**
- 3. Wissenschaftliche Verknüpfung der Umfrage „G8 und mehr“ mit der Schul- und Lernzeitforschung I**
- 4. Bedienungsanleitung für die Rezeption von Meinungsumfragen - Häufig gestellte Fragen**
- 5. Diskussion: Beitrag der Ergebnisse der Umfrage „G8 und mehr“ zu bildungspolitischen, wissenschaftlichen und theoretischen Fragen**
  - 5.1. Die Verlässlichkeit der Ergebnisse
  - 5.2. Darf Politik getroffene Entscheidungen revidieren?
  - 5.3. Zeitdiskriminierung - ein neues Phänomen
  - 5.4. Sind Gesamtschulen und gebundene Ganztagschulen noch zeitgemäß?
  - 5.5. Die lebensalterlich frühe Entstehung der G8/G9 Meinung bei Eltern
  - 5.6. Die Rolle der Kompetenzorientierung in der G8/G9 Problematik
  - 5.7. Ein Gesamtfazit: Die Reformfalle der differentiellen Schulzeit bei Gymnasien und Gesamtschulen
- 6. Anmerkung zur Auswertung und Darstellung der Ergebnisse im Abschlußbericht**

## Vorwort

Bei umfangreichen empirischen Untersuchungen, an denen tausende Menschen teilnehmen, erhalten die Verantwortlichen stets zahlreiche Anfragen, Anmerkungen und Kommentare, die eigentlich alle ausführlich beantwortet werden müssten, will man ein Mindestmaß an Höflichkeit einhalten.

An dieser Studie haben dankenswerter Weise über 50 000 Menschen teilgenommen - entsprechend hoch war die Zahl der Anmerkungen, aber auch die Reaktionen von Medien und Politik, die aus Sicht des Autors beantwortet, aber auch korrigiert werden müssen.

Diesem Umstand trägt dieser „Kommentarband“ (gewissermaßen eine „Bedienungsanleitung für die Rezeption von Meinungsumfragen“) Rechnung. Alle bis jetzt bekannt gewordenen Fragen und Kritiken werden beantwortet bzw. kommentiert.

Der Abschlußbericht - über die website der Landeselternschaft der Gymnasien NRW e.V. erhältlich - enthält nur die unkommentierten Fakten der Umfrage. Deshalb sind in diesem Kommentarband auch die Abschnitte enthalten, die man von jeder wissenschaftlichen Umfrage erwartet: Herkunft der Fragestellung, Forschungsanschlüsse, methodische Anmerkungen und vor allem eine Diskussion und Interpretation der gefundenen Fakten.

Rainer Dollase, Steinhagen, im August 2016

# I. Überblick und Zusammenfassung der Untersuchung „G8 und mehr“ (management survey)

(Dieses Kapitel ist identisch mit dem entsprechenden Kapitel des Abschlußberichtes, August 2016)

## I.1. Fragestellung

Die Landeselternschaft der Gymnasien in NRW e.V. wollte die Meinung ihrer Elternschaft, sowie weiterer am Schulleben Beteiligter zu G8 und G9 und einigen weiteren Fragen zur zeitlichen Schulorganisation bzw. zur Situation des Kindes erfahren. Deswegen der Titel „G8 und mehr“.

## I.2. Was wurde gemacht?

Es wurden 3 Befragungen durchgeführt:

1. Eine postalische Umfrage nach einem Zufallsprinzip. Von allen 462 Mitgliedsgymnasien der Landeselternschaft wurden die Schulpflegschaftsvorsitzenden angeschrieben. Sie mussten nach einer Zufallskennziffer je 5 Erziehungsberechtigte auswählen. Diese bekamen einen Fragebogen und sollten ihn ausgefüllt an die Universität Bielefeld zurückschicken.
2. Eine Onlinebefragung für Eltern, die nicht an der postalischen Umfrage teilnehmen konnten, um dem Wunsch nach größerer Beteiligung zu entsprechen und Vergleiche zwischen der postalischen und der Onlinebefragung festzustellen.
3. Die zweite Onlinebefragung war für verschiedene Statusgruppen, wie zum Beispiel Lehrer, Eltern ehemaliger Gymnasialkinder, außerschulisches Lehrpersonal, Schüler, Direktoren von Gymnasien et cetera konzipiert worden. Sie unterschied sich von der ersten Onlineumfrage dadurch, dass die Fragen über eigene Kinder weggelassen wurden. In der zweiten Onlinebefragung konnte man sich nur zur bildungspolitischen G8 und G9 Problematik und ihren Folgen äußern.

## I.3. Rücklauf

Die Gesamtbeteiligung an der Umfrage war für bildungswissenschaftliche Umfragen riesig. An den beiden Onlinebefragungen nahmen im ersten Fall 38 148 und im zweiten. Fall 16 496 Personen teil. Es konnten so insgesamt 54 644 Personen mobilisiert und aktiviert werden. In der Zufallsstichprobe erhielten wir 1310 ausgefüllte Fragebögen zurück, was einer sehr hohen Rücksendequote von 56,7 % entspricht. Die Onlinefragebögen mussten vor der statistischen Verarbeitung noch durchgesehen werden, da nicht alle Fragebögen vollständig ausgefüllt wurden. Dennoch blieben in der ersten Onlineumfrage 26.389 Personen und in der zweiten 11 846 Personen übrig.

#### **1.4. Alleinstellungsmerkmale gegenüber bisherigen Umfragen zu G8/G9**

1. Die Untersuchung richtete sich gezielt an Gymnasialeltern in NRW.
2. Eine der zwei Onlinebefragungen richtete sich auch an andere Statusgruppen - Lehrer, Schüler, Direktoren, Grundschulleitern etc.
3. Erhebung von Zusatzinformationen („... und mehr“) zum Kind und zur zeitlichen Gestaltung von Schule.
4. Die Untersuchung wurde multimethodal und methodenkritisch gestaltet.
5. Die Auswertung geht über die Grundauszählung hinaus.

#### **1.5. Methodenvergleich online vs. Zufallsstichprobe**

Mit der multimethodalen Anlage und den methodenkritischen Vorkehrungen (z.B. Verwendung von Distraktoren) und dem Mehr-Stichprobenprinzip konnte die Zuverlässigkeit und Gültigkeit der Ergebnisse eindrucksvoll bestätigt werden.

Überraschenderweise sind die Unterschiede zwischen den Ergebnissen der Onlinebefragung und der postalischen Zufallsstichprobe geringer als erwartet. In den zentralen Fragen zu G8/G9 gibt es prozentuale Unterschiede - die aber niemals in andere quantitative Dimensionen verweisen. In vielen Fragen zur zeitlichen Schulorganisation und zur Situation des Kindes gibt es nur minimale Unterschiede zwischen Onlinebefragung und postalischer Befragung nach dem Zufallsprinzip.

#### **1.6. Hauptergebnis I: - G9 wird von 79 % bzw. 88 % der Eltern gewünscht.**

Die grundsätzliche Frage, ob G8 oder G9, wird in allen 3 Umfragen mit deutlichen Mehrheiten für G9 beantwortet. Die Prozentsätze der Befürworter von G9 liegen in der ersten Onlineumfrage bei 88 %, in der zweiten Onlineumfrage bei ehemaligen Gymnasialeltern auch bei 88 % und in der Zufallsstichprobe der postalischen Befragung bei 79 %. Damit werden frühere repräsentative Untersuchungen von FORSA und EMNID auch für Gymnasialeltern und für das Jahr 2016 bestätigt. Die kompromisslosen Befürworter von G9 sind doppelt so häufig wie die kompromisslosen G8 Befürworter. Von den G9 Befürwortern sind etwa 60 % kompromisslos gegen G8 und nur ungefähr 30 % von den G8 Befürwortern kompromisslos gegen G9. Der hohe Prozentsatz kompromissloser G9 Befürwortung deutet ebenso wie die hohe Teilnahme an den Studien auf eine stark ausgeprägte Kampagnenfähigkeit der G9 Befürworter hin.

Es gibt keine Statusgruppe, in der die Befürwortung für G9 unter den Wert von 70,2 % fällt. Den niedrigsten Wert haben die „Schulleiter von Gymnasien“ die G9 allerdings auch mit einer großen Mehrheit befürworten, nämlich mit 70,2 %. Den höchsten Wert an Zustimmung erreichen die „außerschulischen Lehrkräfte“ für G9 mit 96,7 %.

**I.7. Hauptergebnis 2: - Das „G9 Gymnasium“, auch eine „Rückkehr zu G9 wie früher“ wird allen Reformversuchen von G8 gegenüber deutlich vorgezogen.**

Wenn man die Befragten nach der Übergangsempfehlung für einen Grundschulabgänger mit Gymnasialeignung fragt, also eine Reihe von Distraktoren einführt, oder wenn man Veränderungsvorschläge von G8 bewerten lässt und dabei auch die Rückkehr zu G9 als Option vorgibt, gibt es immer noch deutliche Mehrheiten, die bei rund 60 % liegen, die zurück zu einem „Gymnasium G9 wie früher“ wollen. Besonders wenige Personen wählen hierbei die Gesamtschule (2 % - 4 %) oder das „verbesserte G8 Gymnasium“ mit ungefähr 10 % - 20 %. Aktuelle politische Lösungen finden also nur bei Minderheiten Zustimmung. Das gilt auch in allen anderen Statusgruppen: überall ist das „Gymnasium G9 wie früher“ beliebter als ein „verbessertes G8-Gymnasium“ oder gar die „Gesamtschule“.

**I.8. Hauptergebnis 3: - 60 % der Befragten befürworten die Wahlfreiheit von Schulen oder einzelnen Schülern als Lösung der G8/G9 Problematik. 55 % wünschen eine bundeseinheitliche Lösung.**

Die beiden Wahlmöglichkeiten - Wahl für jede Schule und Wahl für jeden einzelnen Schüler für G8 oder G9 - fanden jeweils um die 40 % Zustimmung. Beide Optionen haben davon 19,5 % angekreuzt. Das erlaubt die rechnerische Zusammenfassung, dass rund 60 % für Wahlfreiheit in irgendeiner Form votieren. Angst vor dem „vielen Hin und Her“ haben 13,4 % der Befragten.

**I.9. Hauptergebnis 4: - G8 Anhänger wie auch G9 Anhänger sind überzeugte Gymnasialeltern und würden keine Gesamtschule wählen.**

Gefragt, warum sie, die Eltern, das Gymnasium gegenüber der Gesamtschule gewählt haben, wird der „Wert auf die höhere Qualität der Ausbildung“ in erster Linie genannt. Und zwar mit Prozentsätzen um und über 60 %. Sodann folgen die Gründe „wegen der sozialen Situation“, „wegen des besseren Rufes“ und „wegen der örtlichen Entfernung“. Die Tatsache der Schulzeitverkürzung auf 8 Jahre bis zum Abitur spielt so gut wie keine Rolle.

**I.10. Hauptergebnis 5: - Die G9 Anhängerschaft bildet sich schon bei Grundschuleltern bzw. in den frühen gymnasialen Schuljahren.**

Die Meinung zu G9 bildet sich bei den Eltern offenbar schon sehr früh im Lebensweg des Kindes. G8 und G9 Anhängerquoten der Eltern ändern sich mit dem Alter bzw. der Klassenstufe ihrer Kinder. G9 Anhänger nehmen vom 5. auf das 12. Schuljahr leicht ab, bleiben aber immer noch

deutlich über den G8 Anhängerquoten. Die G8 Befürworter sind sowohl bei den Eltern der jüngeren als auch der älteren Kindern in der Minderzahl. Die Ursache für den Alters- bzw. Klassenstufentrend könnte sein: die Eltern nehmen wahr, dass ihr Kind mit G8 gut zurecht kommt oder aber sie beteiligen sich nicht an der Umfrage, wenn das Kind schon älter ist, weil sie denken, dass dieses keinen Sinn mehr macht. In der 5. Klasse liegen die Prozentsätze bei 91,8 % pro G9 und 8,2 % für G8 - und in der Klasse 12 bei 78,1 % für G9 und 21,9 % für G8.

**1.11. Hauptergebnis 6: - G8 Anhänger gaben an, dass ihre Kinder durch G8 weniger belastet sind, erfolgreicher in der Schule sind, weniger Nachhilfe brauchen, sich weniger anstrengen müssen.**

Die Belastung der Kinder von G8 Anhängern ist geringer, weshalb deren Eltern offenbar zu G8 neigen. Im einzelnen: Belastung und Zeitaufwand ist geringer, Nachhilfe wird seltener in Anspruch genommen, der Schulerfolg ist besser (allerdings ist das Schulversagen in beiden Gruppen gleich). Die „Anstrengung“ ist ähnlich, aber die Klagen über die „Menge an Stoff“ und die „Knappheit der Zeit für Stoffvertiefung“ ist bei den G9 Befürwortern deutlich höher. Der höhere Belastungsgrad ist somit typisch für den Nachwuchs der G9 Anhänger. Es besteht ein signifikanter und relevanter Zusammenhang zwischen Belastung und G9 Entscheidung von  $d = 0.66$  (international gebräuchliches Effektmaß, bedeutet in etwa mehr als eine halbe Standardabweichung zwischen G8 und G9 Anhängern).

**1.12. Hauptergebnis 7: - G9 wird öfter für Mädchen gewünscht als für Jungen.**

Obwohl Mädchen auch in dieser Untersuchung öfter Schulerfolg haben als Jungen, sind sie signifikant mehr belastet, weshalb Eltern von Töchtern häufiger für G9 votieren.

**1.13. Hauptergebnis 8: - Alle Formen des gebundenen Ganztags werden von den Eltern schlecht bewertet - gewünscht wird Zeitflexibilität.**

G8 und G9 Anhänger sind sich einig in der Ablehnung aller Formen von gebundenen Ganztagsangeboten. Parallel dazu nehmen sie beide in ähnlicher Weise Einschränkungen auch der außerunterrichtlichen Tätigkeiten wahr, wenngleich die G8 Anhänger sich seltener eingeschränkt fühlen als die G9 Anhänger.

## 2. Vorgeschichte der Fragestellung zu „G8 und mehr“

Im Jahre 2004 wurde wegen des “verantwortungsvollen Umgangs mit der Lebenszeit des Kindes“ (Originalwortlaut der Website des Schulministeriums, 2016) in Nordrhein-Westfalen das Abitur nach acht Jahren politisch eingeführt. Im Jahre 2013 konnte die Zeit des Doppeljahrgangs – ein Teil der Schüler und Schülerinnen machte nach acht Jahren, ein anderer Teil nach neun Jahren Abitur – beendet werden.

Der „verantwortungsvolle Umgang mit der Lebenszeit des Kindes“ ist eher eine fantasievolle Umschreibung, um zu kaschieren, dass es bei „G8“ auch um eine Maßnahme zur Überwindung des demographischen Knicks geht. Die Maßnahmen im einzelnen:

1. frühe Einschulung
2. Verkürzung der gymnasialen Zeit
3. Abschaffung der Wehrpflicht
4. Förderung der Zuwanderung
5. Verlängerung der Lebensarbeitszeit.

Das sind die wesentlichen Maßnahmen, die in den vergangenen 10 - 15 Jahren ergriffen wurden, um einem Arbeitskräftemangel, insbesondere einem qualifizierten Facharbeitermangel, vorzubeugen.

Die Zeitgeschichte hat gezeigt, dass alle diese Maßnahmen nicht ohne Probleme zu erreichen sind. Die frühe Einschulung hat gegenüber der traditionellen Kindergarten - Erziehung im Durchschnitt keinen Vorteil, im Gegenteil, die Kinder sind beim Wechsel in die weiterführende Schule benachteiligt, die Abschaffung der Wehrpflicht hat zu einem absehbaren Mangel an Menschen geführt, die sich für den Soldatenberuf interessieren, die Zuwanderung erzeugt aktuell europaweit eine Reihe von Integrations- und Akzeptanzproblemen, die Verlängerung der Lebensarbeitszeit ist von der letzten großen Koalition in Deutschland erst einmal ausgesetzt worden, auch das wegen vieler Widerstände, die in anderen Ländern der EU zu beobachten sind.



Die Verkürzung der Zeit bis zum Abitur ist von Anfang an ebenfalls umstritten gewesen. Es haben sich mehrere Initiativen gebildet, zum Beispiel „G 8“ und „G9 jetzt NRW“ im Jahr 2010, – und nicht nur in NRW, sondern im gesamten Bundesgebiet hat sich Widerstand gegen eine Verkürzung der Zeit bis zum Abitur etabliert. In Nordrhein-Westfalen hat „G9 jetzt“ rund 100.000 Unterschriften gesammelt. Eine Sammlung von Unterschriften für einen Volksentscheid ist aktuell in der Vorbereitung. Der Widerstand im Lande Nordrhein-Westfalen wird auch dadurch verschärft, dass andere Länder, zum Beispiel Niedersachsen oder Hessen, zu einem G9 bzw. zu einer Wahl zwischen G8 und dG9 zurückgekehrt sind. Die Hoffnung macht sich breit, dass auch in NRW dem eindeutigen Elternwunsch nachgegeben wird, vor allem, wenn sich der Konflikt bis in die Landtagswahl 2017 hineinziehen wird.

Es hat nicht an Umfragen gefehlt, in denen die große Mehrheit demonstriert wurde, die für eine Rückkehr zu G9 ist. Im Jahre 2010, 2012, 2013 und 2014 wurden von Allensbach, EMNID, FORSA und JAKO jeweils beeindruckende Mehrheiten gefunden, die für eine Rückkehr zu G9 votieren. Die Zahlen schwanken zwischen 64 %, 72 %, 80 % und 76 % pro G9.

Für alle neuen und künftigen Untersuchungen bleibt also die Erwartung bestehen, dass sich große Mehrheiten für G9 entscheiden werden. Diese Erwartung wird von den politischen Akteuren nicht unbedingt geteilt. Am 18.12.2014 wurden zehn Empfehlungen als Ergebnis des „Runden Tisches“ verabschiedet, die die Klagen über zu große Belastungen der Kinder und Jugendlichen durch G8 mildern sollten. Daran knüpfen die politischen Akteure die Erwartung, dass damit die Debatte befriedet sei. Dem ist so nicht gewesen, wie die öffentliche Diskussion über G8 und G9 seither gezeigt hat. Die Stimmen für eine Rückkehr zu G9 waren immer und sind bis heute sehr stark vernehmlich.

In dieser Situation war und ist nur verständlich, dass sich die Landeselternschaft der Gymnasien in NRW e.V. der Meinung ihrer Mitglieder versichern musste. Trat sie anfänglich für G8 ein, so konnte der Unmut bei der Mehrheit der Eltern im Lande nicht länger übersehen werden, so dass die Idee, bei den Mitgliedern der Landeselternschaft eine Meinungsumfrage durchzuführen, prominent wurde. Im November 2015 wurde auf der

Herbst-Mitgliederversammlung die Durchführung einer Studie zu G8 und G9 beschlossen, um ein Meinungsbild für den Verband zu gewinnen. Eine Mitgliederbefragung ist in kritischen Fragen für jeden Verband sinnvoll und unterstützenswert - so auch für die Landeselternschaft.

Eine solche neue Studie musste sich in einer Reihe von Punkten von bisherigen Untersuchungen unterscheiden. Das sei am Beispiel der FORSA Studie von 2014 und der Emnid Studie für JAKO 2014 kurz erläutert. Meinungsforschungsinstitute arbeiten überwiegend mit telefonischen Befragungen, d.h. sie wählen per Zufall eine Telefonnummer aus dem Bundesgebiet aus, fragen die Teilnehmer, ob sie an einer Befragung teilnehmen wollen und wenn ja, kann gefragt werden, ob sie Eltern sind, ob sie Eltern von Kindern an Gymnasien sind, ob sie in Nordrhein-Westfalen wohnen etc. Die FORSA Studie war für NRW repräsentativ, hat 1003 Menschen befragt, darunter aber auch viele, die aktuell kein Kind im Gymnasium haben, sondern Menschen aller Altersstufen, Berufe usw. Die EMNID Studie für JAKO hat 3001 Menschen befragt, darunter 2098 Mütter. Sie ist bundesweit repräsentativ und hat Eltern - egal welcher Schulform - befragt. In der FORSA Umfrage gab es 76 %, die für G9 votierten, in der Emnid Studie 79 %.

Ein wichtiges Unterscheidungsmerkmal ist damit deutlich: es handelt sich bei der vorliegenden Umfrage primär um eine postalische Mitgliederbefragung nach dem Zufallsprinzip - also ist sie für diese Grundgesamtheit repräsentativ - der Landeselternschaft der Gymnasien in NRW e.V. an insgesamt 462 Gymnasien. Über die erste Onlinebefragung wurden allerdings nicht nur Mitglieder, sondern auch Gymnasialeltern überall erreicht und über die zweite Onlinebefragung auch andere Personengruppen.

Dass sich der Elternwille so weit von der politischen Entscheidung entfernt hat, hängt auch damit zusammen, dass nahezu alle Parteien (bis auf die Piraten, seit August 2016 auch die FDP nicht) die Regelung G8 unterstützen. Es gibt also keine starke parlamentarische Kraft, die den Elternwunsch vertritt. In solchen Situationen bilden sich häufig außerparlamentarische Initiativen und eine außerparlamentarische Opposition, deren Meinung die etablierten Parteien oftmals verwundert. Parteien bedenken allerdings nicht,

dass sie bei niedrigen Wahlbeteiligungen nur verschwindende Prozentsätze von LandesbürgerInnen vertreten – und keinesfalls Mehrheiten. Für jeden ist klar, dass wir eine Vertreter-Demokratie haben und dass dieses Prinzip sich auch bewährt hat. Dennoch muss man davon ausgehen, dass ein allzu großer Dissens zwischen Regierungshandeln und öffentlicher Meinung der Politik insgesamt schadet. Es bilden sich Splitterparteien. Stabile parlamentarische Mehrheiten sind nur schwer zu erreichen.

Bedenkliche antidemokratische Impulse sind auch eine Gefahr für die Repräsentativ-Demokratie, insofern man das Wahlvolk für mehr oder minder unmündig hält. Zum Beispiel steht in einem von der Bertelsmann-Stiftung herausgegebenen Band, ein Artikel „Kunst des Reformierens“ in dem folgendes geschrieben steht (S.24):

*„Um ihrer politischen Verantwortung gerecht zu werden, muss eine Regierung sich im Zweifelsfall auch gegen den empirischen und kontingenten Volkswillen durchsetzen.“<sup>1</sup>*

Das Prinzip ist gefährlich, da die Wahlbürger und -bürgerinnen den Eindruck gewinnen könnten, dass ihre Meinung nunmehr nur von sekundärer Bedeutung wäre. In einer lebendigen Demokratie muss auch darauf geachtet werden, dass Mehrheiten (und nicht nur Minderheiten) eine stabile Repräsentation in der Legislative haben und dass sich die Exekutive auch tatsächlich am Mehrheitswillen orientiert. Eine Abkopplung davon ist weder historisch noch aktuell irgendwo auf der Welt von Vorteil gewesen. Mißverständlich können Sätze wie diese sein „ Zum Glück leben wir in einer repräsentativen Demokratie und nicht in einer Demoskopie.“ - oder „Hier entscheiden immer noch wir und nicht das Volk.“ wenn sie denn irgendwo gesagt worden sein sollten. Auch die leichtfertige Abqualifizierung von Mehrheitsmeinungen als „Populismus“ trägt nicht zur Versachlichung der Debatten zu schwierigen Zeitproblemen bei.

Für die aktuelle Umfrage „G8 und mehr“ und die Bedeutung der Ergebnisse muss der zeitgeschichtliche Hintergrund insgesamt bedacht werden - das wird hier natürlich nicht geleistet, sondern nur darauf verwiesen.

---

<sup>1</sup> Die Kunst des Reformierens. Konzeptionelle Überlegungen zu einer erfolgreichen Regierungsstrategie, in: Bertelsmann Stiftung (Hg.), Zukunft Regieren, Gütersloh (Beiträge für eine gestaltungsfähige Politik 3/2009) (Friedbert W. Rüb zusammen mit Karen Alnor und Florian Spohr).

### 3. Wissenschaftliche Verknüpfung der Umfrage „G8 und mehr“ mit der Schul- und Lernzeitforschung

Die Aufgabenstellung der vorliegenden Studie ist also eine „angewandte Fragestellung“, die aufgrund ihres Zustandekommens sehr stark von den Informationswünschen der auftraggebenden Institution geprägt ist. Die Landeselternschaft hat ein berechtigtes Interesse mithilfe von Mitgliederbefragungen das Meinungsbild ihrer Mitglieder sorgfältig und gründlich, vor allem aber auch methodenkritisch zu erfassen.

Eine „wissenschaftliche Fragestellung“ leitet hingegen ihre Fragen aus dem bisherigen wissenschaftlichen Forschungsstand ab. Das ist eine andere Art der Genese von Fragestellungen für Untersuchungen. In ihr werden nicht Fragen der Praxis, der Verbände beantwortet, sondern Fragen, die die Wissenschaft selber aufgeworfen hat.

Im vorliegenden Fall werden Fragen, die die Praxis bzw. der Verband entwickelt hat, untersucht. Auch in diesem Fall gibt es die Möglichkeit, die Ergebnisse auf dem Hintergrund wissenschaftlicher Ergebnisse zu diskutieren. Im Folgenden sollen einige dieser Argumentationsstränge kurz angesprochen werden, ohne dass hier ein vollständiges Referat der entsprechenden Forschungsgebiete erfolgt.

1. Die **Motivations- und Entscheidungsforschung** ist für die Einstellung zu bildungspolitischen Optionen relevant. Ein wesentliches Ergebnis dieser Forschung ist, dass die Erwartung von Effekten und der Wert des tatsächlichen Effektes eine der wichtigsten Entscheidungsparameter für Menschen ist (sog. Erwartung mal Wert Modelle)<sup>2</sup>. Entscheidend ist also, um es praktisch auszudrücken, welche Erwartungen Eltern mit der Verkürzung der gymnasialen Schulzeit verbinden und welchen Wert sie den möglichen Folgen bemessen. Man kann zum Beispiel nicht ohne weiteres erwarten, dass kollektive und staatliche Zwecke einen individuellen Wert für Eltern haben – ihnen wird es in erster Linie um ihren eigenen Nachwuchs gehen.

---

<sup>2</sup> Brunstein, J. B. & Maier, G. (1996). Persönliche Ziele: Ein Überblick zum Stand der Forschung. *Psychologische Rundschau*, 47, 146-160.

**2. Unsicherheitsforschung:** Erwartungen sind für Menschen nicht unbedingt sicher. Jeder weiß, dass mit Entscheidungen, die man im Augenblick für richtig hält, auch Nachteile in der Zukunft verbunden sein können. Um diese Unsicherheit zu überwinden, suchen Menschen den Rückhalt in der Meinung anderer. Sie versuchen, die Erwartungen dadurch sicherer zu machen, dass sie sich vergewissern, dass ihre eigene Meinung auch mit der Meinung der anderen übereinstimmt. Insofern ist die G9 Bewegung auch ein gewisser Selbstläufer, der Sicherheit schafft, weil es für jedes Individuum deutlich wird, dass auch die anderen die Problematik G8 versus G9 ähnlich sehen. Und umgekehrt für G8 Anhänger in Parteien und Verbänden: Sie wissen objektiv nicht, ob Maßnahmen jenen Effekt erreichen, den sie sich erwarten. Sie hoffen es und finden Bestätigung in den Durchhalteparolen der Politik. Politik ist die Realisation von vermuteten Effekten. Auch mit Blick auf andere Länder schwindet diese Unsicherheit objektiv nicht: wenn dort bislang G8 zu keinen Problemen geführt hat, gibt es objektiv keine Sicherheit, dass es hier bei uns auch so ist. Kulturen wachsen in Jahrhunderten und Jahrzehnten und stellen je unterschiedliche Rahmenbedingungen für das Gelingen und Misslingen zur Verfügung.

Wir denken, wir meinen, wir handeln alle in einer Situation der Unsicherheit. In modernen psychologischen Veröffentlichungen wird diese Grundtatsache menschlichen Erlebens und Handelns von der Unsicherheit thematisiert und auch empirisch mit Gewalt, Terrorismus, Extremismus in Verbindung gebracht (Hogg und Blaylock ,2012, s.u.) Ob G8 oder G9 für die Zukunft eines Landes besser ist oder nicht, weiß niemand. Das wird sich erst nach Jahrzehnten zeigen oder aber auch nicht. Wir wissen nur, dass fast alle lieber G9 hätten. Und darauf muß reagiert werden. Bestehende Lösungen haben immer den Vorzug, dass man weiß, was man hat.

Zitat „Das Leben ist ein unsicheres Unternehmen. Wir können nie sicher darüber sein, was wirklich in der Vergangenheit passiert war, was in der Zukunft passieren wird, wie andere sich verhalten werden und wie wir uns selber verhalten sollten. Im Angesicht dieser unnachgiebigen Unsicherheit verbleiben die Menschen unerschrocken; sie setzen sich beharrlich Ziele, machen Pläne und verfolgen Handlungen.“<sup>3</sup>

---

<sup>3</sup> Hogg und Blaylock (Ed.) 2012 (xv) „Extremism and the Psychology of Uncertainty“ Wiley Blackwell

**3. Entwicklungspsychologisches Wissen** sammeln alle Eltern aufgrund der Erfahrung mit sich selbst, aber auch aus der Beobachtung eigener und anderer Kinder. Besonders verkannt und stellenweise auch beängstigend ist für viele Eltern, dass Entwicklungen nicht geradlinig verlaufen. Jeder kann von Kindern und Jugendlichen berichten, die in der Pubertät besonders schwierig waren, die Schulschwierigkeiten hatten, die einen „Pubertätsdurchhänger“ hatten und die sich erst später, nach einigen Umwegen, doch noch zu leistungsfähigen und stabilen Erwachsenen entwickelt haben. Die Verkürzung der gymnasialen Zeit führt zu dem Eindruck, als wolle man nur diejenigen auswählen, die eine geradlinige Entwicklung nehmen. Gleichzeitig wird das Sitzenbleiben abgeschafft, gleichzeitig kann man auf dem G8 nur unter besonderen Bedingungen mittlere Reife machen, so dass die allgegenwärtige Erwartung von eventuellen Jugendproblemen durch die G8 Entscheidung weiter verunsichert wird.

**4. Lernzeitforschung:** In gewisser Weise verwirrend, wenn nicht unglaubwürdig sind Untersuchungen, die dann auch noch zeigen, das man nach 8 Schuljahren genauso gut sein kann wie nach 9 Schuljahren. Jeder kann also dann schlussfolgern, dass ein Jahr Lernzeit offenbar zu nichts führt. Das kann nicht sein, weil jeder Mensch in seiner eigenen Erfahrung und an seiner eigenen Entwicklung lernen kann, wie viel man innerhalb eines Jahres noch zusätzlich lernen kann. Die G8 Entscheidung verunsichert tiefgreifend dieses allgemein menschliche, sichere Wissen. Die Untersucher, die Ergebnisse dieser Art präsentieren, müssen sich fragen lassen, ob die G8 Absolventen all das wissen, was man im 9. Schuljahr gelernt hat. Fairerweise müsste man also G8 Absolventen nur in dem testen, was sie nicht gelernt haben. Die einfache Logik widerspricht also politischen Sprachregelungen, die da lauten „Die vergleichbare Leistungsfähigkeit von Abiturientinnen und Abiturienten des achtjährigen wie des neunjährigen gymnasialen Bildungsganges ist im „doppelten Abiturjahrgang“ im Jahr 2013 wie auch in mehreren wissenschaftlichen Untersuchungen hinlänglich bestätigt worden und in der öffentlichen Debatte mittlerweile unstrittig.“<sup>4</sup> Das hieße also: ein Jahr mehr Schule fördert die Leistungsfähigkeit nicht. Das widerspricht sämtlichen wissenschaftlichen Erkenntnissen der Lernzeitforschung. Natürlich kann man in einem Jahr noch mehr lernen: zum Beispiel Grundkenntnisse in einer weiteren Sprache erwerben, neue Fächer erschließen, die akademische Allgemeinbildung verbessern.

---

<sup>4</sup> aus einem Brief eines Ministeriums

Verräterisch ist die Vokabel „Leistungsfähigkeit“ die auf „Kompetenzorientierung“ verweist. Bei der Kompetenzorientierung kommt es ja ohnehin weniger darauf an, **was man kann und weiß**, sondern ob man fähig ist. Kompetenz lässt sich im Prinzip schon im vierten Schuljahr feststellen, sogar noch früher: im Kindergarten. Die Intelligenzquotienten bei kleinen Kindern korrelieren systematisch hoch signifikant und relevant mit dem Intelligenzquotienten im Alter von 18 Jahren. Also kann ein Kompetenztest lange Schuljahre ersparen? Wenn ein Jahr nichts ausmacht - warum dann nicht G7 oder G6? Intelligenz ist übrigens die Kompetenz, auf die es vor allen Dingen bei der Bewältigung von Schule im Kompetenzmodus ankommt. Wer Wissen und Bildung zusammen denkt, muss zu dem Schluss kommen, dass man auf Teile des Wissenserwerbs und des Stoffs verzichten muss, wenn man die gymnasiale Zeit verkürzt. Das ist eine rein logische, sachlogisch begründete Aussage und keine politische Meinung für G8 oder G9. Das ist gewissermaßen Fakt - auch für jene Ländern, die schon lange G8 haben.<sup>5</sup>

5.. Die **Forschungen zu temporalen Mustern**<sup>6</sup> Diese Forschungen haben gezeigt, dass Kinder, Jugendliche und Erwachsene über Wünsche nach „idealen Reihenfolgen von Tätigkeiten“ verfügen. Der Lebenslauf wird unterteilt von normativen Übergängen und Einschnitten, die als langfristige Erwartungen von Generation zu Generation weitergegeben werden. Eine Veränderung dieser Normen ist nur langsam und schleichend möglich. G8 verändert diese zeitliche Strukturierung des Lebenslaufs und sorgt für eine erneute Verunsicherung. Die Untersuchungen zu temporalen Mustern haben übrigens auch gezeigt, dass Eltern und Kinder spontan den schulischen Ganzttag nicht wünschen, sondern lieber der Regel folgen „erst die Arbeit und dann das Spiel“ und damit ebenfalls für eine historisch gewachsene Zeitstrukturierung votieren. Morgens Schule - nachmittags Freizeit.

6.. Der amerikanische Psychologe **John Bissel Carroll (1916-2003)** hat einen besonders originellen und empirisch belegten Ansatz zur Bedeutung des Zeitbedarfs für das Lernen

---

<sup>5</sup> Dollase, R. 2016 Ein Jahr Schule macht ein Jahr mehr Qualität. G8 Bildungspolitik: mit Widersprüchen ins argumentative Abseits, in: Ladenthin, V., Nostadt, A., Krautz, J. (Hg.) Weniger ist weniger. Bonn: Verlag für Kultur und Wissenschaft,

<sup>6</sup> Dollase, R., Hammerich, K., Tokarski, W., 2000, Temporale Muster. Die ideale Reihenfolge von Tätigkeiten. Opladen: Leske und Budrich.

entworfen. Nach diesem kann man alle Schülereigenschaften und Systemeigenschaften des Schul- und Bildungssystems in Zeitbedarf umwandeln. Ein Schüler hat keine „geringere Fähigkeit“ - wie man sich auszudrücken beliebt - sondern der Schüler muss länger lernen, um dasselbe Ziel zu erreichen. Wird jemand durch sein soziales Herkunftsmilieu beeinträchtigt, so braucht der Schüler länger Zeit, um dasselbe Ziel wie ein Kind aus der Bildungsschicht zu erreichen. Haben Schüler schlechte Lehrer, so brauchen die Schüler länger Zeit zum Lernen, um dasselbe Ziel wie andere mit guten Lehrern zu erreichen. Zahllose weitere Variablen sind von dem Autor entwickelt worden (zum Beispiel perseverance) um alle sogenannten Lern- und Bildungsnachteile in Zeitbedarf umzuwandeln. Daraus lernt man, dass die **Zeitregelung in Schulen ein Diskriminierungsfaktor** sein kann. Verhindert man zusätzliche Lernzeiten, zeitraubende Lernwege, nachholende Lernzeiten, dann wird die Zeit zu einem neuen Diskriminierungsfaktor, der Bildungsgerechtigkeit verhindert. Bezogen auf die Situation in Nordrhein-Westfalen bzw. einigen anderen Bundesländern - wie das Saarland - die die Institutionsform mit einer bestimmten Zeitmenge verknüpft haben, d.h. G8 im Gymnasium und G9 auf der Gesamtschule, besteht folgendes Problem: Eltern wählen eine bestimmte Institutionsform von Schule, nämlich das Gymnasium. Diese Entscheidung wird mit einer sehr restriktiven Zeitzuordnung, d.h. also mit einer Diskriminierungsoption verbunden. Die Wahl der Institution Gymnasium wird quasi mit einer Strafoption, nämlich Verkürzung der Schulzeit, verbunden. Damit soll die Attraktivität der Gesamtschule (G9) vergrößert werden. Diese Schulform gehört aber nicht zur Wahl der Eltern. Es bildet sich das Gefühl der Reaktanz aus, d.h. aus der Freiheitseinschränkung der Schulwahl heraus entwickelt sich Widerstand und Trotz. (Vergleiche Reaktanztheorie nach Wicklund)<sup>7</sup>.

---

<sup>7</sup> Wicklund, Robert A. 1974, Freedom and reactance. Oxford, England: Lawrence Erlbaum Freedom and reactance.(1974).



## 4. Bedienungsanleitung für die Rezeption von Meinungsumfragen - Häufig gestellte Fragen

Der Verfasser führt seit den späten sechziger Jahren (1968, Mitarbeit im Projekt „Strukturanalyse der Schule“ Leitung: Rene König) regelmäßig bis heute größere empirische Umfragen in unterschiedlichen Themengebieten durch, rund 40 Projekte mit einem Befragtenvolumen von ca. 50 000 Personen (ohne die jetzige). Aus dieser Erfahrung haben sich Regeln herausgebildet, bzw. häufig gestellte Fragen, die es notwendig machen, einige typische Spezifikationen der empirischen Meinungsumfragen hervorzuheben. Hin und wieder sind auch kritische Anmerkungen von Befragten bzw. von wenig erfahrenen „Experten“ erhoben worden, die freundlicher Weise natürlich beantwortet werden müssen. Die Kritik der „Experten“ wird meist von einem eher dogmatischen und bürokratischen Verständnis von Forschung getragen, denn von wirklich praktischer Erfahrung mit der Durchführung von Meinungsumfragen.

Man mag sich angesichts des Titels „Bedienungsanleitung für die Rezeption von Meinungsumfragen“ wundern, warum so etwas notwendig ist. Meinungsumfragen sind schließlich einfach dargestellt: man nennt alle Fragen, man nennt die Antworten dazu und wandelt diese eventuell in Prozentsätze um, ehe man einige Grafiken dazu anfertigt. So ist das dann auch im Abschlußbericht geschehen.

Leider ist es nicht ganz so einfach. Das liegt aber auch daran, dass Meinungsumfragen schon bei der Planung der Kritik ausgesetzt sind. Manche glauben gar, wenn sie den Fragebogen kennen, schon mit einem finalen Urteil eine Untersuchung insgesamt als „unwissenschaftlich“, „unbrauchbar“ und „völlig überflüssig“ darstellen zu können.

Warum reagieren einige Menschen so allergisch auf Meinungsumfragen? Vor allen Dingen: warum reagieren Menschen, die mal ein bisschen über empirische Forschung gelesen oder gehört haben, in ähnlicher Weise wie völlige Laien?

Das liegt ziemlich eindeutig daran, dass Meinungsumfragen für jeden, der eine Meinung hat, immer eine Herausforderung bzw. ein Unsicherheit erzeugendes Forschungsinstrumentarium sind. Nichts fürchten Menschen so sehr wie die Tatsache, dass ihre Meinungen widerlegt werden. Nahezu jeder Mensch in unserer Gesellschaft möchte ein finaler Besserwisser sein, d.h. genau wissen, wie die Realität beschaffen ist und was man tun muss, um sie in die gewünschte Richtung zu verändern.

Diese Sicherheit bezogen auf die eigene Meinung ist bei Politikern wie Wissenschaftlern und auch Laien immer falsch. Sie ist falsch darum, weil wir die gesamte Realität alleine überhaupt nicht genau kennenlernen können, weil wir alle eine Privatmeinung haben, die selbstverständlich, wenn man alle Informationen auch für die Zukunft hätte, rein hypothetisch, völlig falsch sein könnte.

Weil wir alle nur urteilen und meinen aufgrund einer unvollständigen Realitätserkenntnis müssten wir also entsprechend vorsichtig, tolerant und offen mit den Meinungen anderer umgehen. Das tun die meisten nicht, sondern sie gehen mit anderer Leute's Meinung rabiat um: sie denunzieren, sie halten sie selbstverständlich für Unsinn, für belanglos, dumm, hirnrissig und falsch etc. Sie beschimpfen Andersdenkende als „Pack“ und verweigern die Diskussion mit ihnen.

Aus der Erfahrung mit vielen Studien weiß man: Etwa ein halbes Prozent der Befragten hat immer schon angesichts des Fragebogens „schwere Bauchschmerzen“ und fühlt sich zur Kritik bis hin zur Anzeige bei einer Staatsanwaltschaft wegen Volksverhetzung bemüßigt. 0,5 % von 1000 Menschen sind fünf Personen. Nur, 0,5 % von 50.000 Menschen sind schon 250 Menschen, also „ganz viele“. Solche Zahlen verleiten dann zu dem Fehlurteil - „die Umfrage war von Anfang an in der Kritik“. Diese Situation liegt in der gegenwärtigen Umfrage, über die hier berichtet wird, vor. Erfreulicherweise waren es in dieser Umfrage prozentual deutlich weniger, die Kritik am Fragebogen offen geäußert haben - z.B. auch durch Anmerkungen auf den Fragebögen (ca. 50). Das erscheint vermutlich an Dyskalkulie leidenden ZeitgenossInnen auch als „ganz viele“.

Auch wenn die Kritik prozentual zu vernachlässigen ist, gibt es immer und bei jeder Umfrage ein Rumoren bei viel mehr Befragten: „Ich kann meine Meinung nicht richtig äußern“, „Die Fragen sind tendenziös“, „Das ist keine richtige wissenschaftliche Untersuchung“, „Die Ergebnisse können nur falsch sein“ etc. - alles banal und nie zu verhindern. Diese Urteile zeigen nur eines: wir sind alle verschieden, haben unterschiedliche Meinungen, haben ein unterschiedliches Sprachverständnis. Und das ist gut so.

Ein weiteres zentrales, nicht überall bekanntes Problem: Meinungsumfragen können keine Wirkungsaussagen liefern. Das ist banal und bedarf keinerlei Hervorhebung mehr. Meinungsumfragen können auch nicht falsch sein – auch das ist banal. Sie können methodisch unglücklich angelegt sein, wenn man beispielsweise nur G8 Anhänger fragen würde, was sie von G8 oder G9 halten würden oder umgekehrt nur G9 Anhänger. Die Fragen, die man stellt, muss man akzeptieren. Jeder Mensch hat das Recht, Fragen zu stellen, zu denen er eine Antwort haben möchte.

In dieser Umfrage hat die Landeselternschaft der Gymnasien in NRW e.V. zusammen mit dem Verfasser und 15 weiteren Personen überlegt, welche Fragen interessant sind. Der abschließende Fragebogen ist vom Verfasser akzeptiert worden, wenngleich einige Vorschläge, die die 15 weiteren Personen gemacht haben, nicht berücksichtigt werden konnten, weil die Personen sich untereinander nicht einig waren oder Gegenteiliges vorgeschlagen haben.

Nun kommt es bei jedem Fragebogen vor, dass Leute, die den Fragebogen zum ersten Mal sehen folgendes sagen:

1. „Also ich hätte einen ganz anderen Fragebogen konstruiert.“
2. „Ich finde, die Fragen überhaupt nicht richtig und sie sind nicht zielführend.“
3. „Die Fragen drängen die Antworten alle in eine bestimmte Richtung.“
4. „Das müsste man viel differenzierter erfragen.“
5. „Es fehlen eine ganze Menge von wichtigen Fragen.“

Mit solchen Anmerkungen - so paradox es klingt - kann man nichts anfangen. Diese Kritik gilt immer und für jede Umfrage. Stellt man einen mit der Formulierung von Fragen zufrieden, gibt es den anderen, der damit überhaupt nicht zufrieden ist.

Im Jahre 1984 habe ich in einem Buch „Grenzen der Erziehung“<sup>8</sup> das Kaffeekannen Beispiel als Metapher für sinnlose Kritik vorgestellt: mit der Kaffeekanne kann man keinen Nagel in die Wand schlagen, die Kaffeekanne kann nicht mit mehreren Personen gleichzeitig angefasst werden, sie müsste mindestens drei Henkel haben, damit man auch gemeinsam Kaffee einschenken kann usw. Das bedeutet ziemlich hart gesagt: die oben genannten Kritikpunkte sind kabarettistische Kritikpunkte, sie sind banal und sie treffen den methodischen Kern keiner durchgeführten Meinungsumfrage.

Fazit:

1. Jeder Mensch hätte einen anderen Fragebogen konstruiert – wenn man 100 Menschen beauftragt, einen Fragebogen zur G8 /G9 Problematik zu konstruieren, so wird man 100 verschiedene Fragebögen erhalten.
2. Jeder Mensch findet andere Fragen im Zusammenhang mit der Fragestellung zielführend und wichtig. Um zu einem einsetzbaren Fragebogen zu kommen, muss man also Kompromisse eingehen.
3. In der Fremdenfeindlichkeitsforschung ist der Verfasser in der Tat bei der Staatsanwaltschaft Wuppertal angezeigt worden, weil die Schüler in einer Umfrage die Menschen unterschiedlicher Nationen mit Schulnoten von 1 = sehr gut bis 6 =ungenügend beurteilen sollten. Der Anzeiger meinte, wenn alle Noten als Antworten erlaubt wären, wäre das eine Aufforderung schlechte Noten für Menschen unterschiedlicher Nationen zu geben. Natürlich hat die Staatsanwaltschaft noch nicht einmal eine Ermittlung eingeleitet, weil ja jeder sowohl eine „1“ als auch eine „6“ geben könnte. Auch in unserem Fragebogen kam die politische Kritik, dass Antwortalternativen darin enthalten seien, die doch kein Mensch wolle. O. k., dann war der Fragebogen genau richtig und jeder hätte mit einer Ablehnung antworten können. Entscheidend ist also, dass alle Arten von Zustimmung oder Ablehnung möglich sind – das ist das Argument gegen den Suggestionvorwurf.
4. Man kann Antworten in einer Umfrage immer differenzierter erfragen als es geschehen ist. Wenn man fünf Antwortmöglichkeiten (1 bis 5) hat, dann könnte man auch 7 vorgeben, wenn man 7 hat, könnte man auch 20 nehmen (ist ja noch differenzierter), wenn man 20 hat, könnte man auch 50 vorgeben usw. So führen sich solche Argumente selber ad

---

<sup>8</sup> Dollase, R. 1984, Grenzen der Erziehung. Düsseldorf: Schwann

absurdum – sie sind nicht praktikabel. Sie werden von Menschen vertreten, die keine wirkliche Erfahrung mit großen Umfragen haben.

5. In jedem Fragebogen fehlen immer ganz viele Fragen – das liegt daran, dass der Fragebogen nur einen begrenzten Umfang und eine begrenzte Fragestellung haben kann. Niemand beantwortet einen Fragebogen von über 40 Seiten. Was man erfragt, muss also immer auch mit Rücksicht auf den Umfang und die Dauer der Beantwortung eines Fragebogens geschehen. Bei der Auswahl der Fragen beschränkt man sich auf das, was der Auftraggeber gerne wissen möchte.

Fragebogen - Studien können als Meinungsumfragen – wie alle Meinungsumfragen auf der ganzen Welt – nur für das in Anspruch genommen werden, für das sie konstruiert werden. Wichtig ist Transparenz: man teilt mit, wie die Fragen aussehen und man teilt mit, welche Ergebnisse mit welcher Stichprobe erhalten wurden. Man sorgt dafür, dass jeder Leser weiß, wie man an die Daten gekommen ist.

Der Verfasser weiß, dass man mit vielen Meinungsumfragen nicht einverstanden ist. In einem solchen Fall muss man selbst eine eigene Umfrage machen. Wissenschaftler sind da viel toleranter: Sie beziehen sich, wenn sie eine neue Studie machen, auf die, die früher gemacht worden sind und zeigen, wo sie sich unterscheiden, damit der Leser genau weiß, an welcher Stelle diese Meinungsumfrage als Alleinstellungsmerkmal Bedeutung bekommt.

Die im folgenden thematisierten Probleme von Meinungsumfragen sind keineswegs ein Spezifikum dieser Meinungsumfrage sondern sie gelten für alle Meinungsumfragen und sie gelten vor allen Dingen auch für Kritik an Meinungsumfragen.

**I. Was heißt „wissenschaftlich“?** Große Unklarheit besteht darüber, was eigentlich wissenschaftlich heißt, wenn es um Meinungsumfragen geht. Wie weiter oben schon ausgeführt, kann „wissenschaftlich“ bedeuten, dass die Fragestellung aus der wissenschaftlichen Diskussion abgeleitet wird. Das ist aber kein „Muss“ sondern man kann es so machen. Es ist auch möglich, dass Praktiker Fragestellungen entwickeln, eine Meinungsumfrage selber durchführen und die Auswertung einem statistikkundigen „Wissenschaftler“ übertragen. „Wissenschaftlich“ kann also auch heißen, ein

Wissenschaftler oder eine Wissenschaftlerin führte diese Umfrage durch. Die dritte Bedeutung des Begriffes „wissenschaftlich“ könnte sein, dass sie mit wissenschaftlichen Methoden durchgeführt und ausgewertet wurde. Allerdings ist auch dies etwas oberflächlich, weil diese wissenschaftlichen Methoden natürlich auch von Laien, wenn sie diese denn verstehen, durchgeführt werden können. Für die Wissenschaft insgesamt sind alle Informationen wichtig: es gibt keine Informationen erster oder zweiter Klasse sondern Meinungsumfragen können ebenso wie Experimente, wie neuropsychologisch-apparative Untersuchungen zur Verbesserung unserer Kenntnisse über einen Themenbereich beitragen. Bei einer Meinungsumfrage ist eigentlich nur die Zweckgerechtigkeit der Umfrage wichtig, die in diesem Fall dadurch gegeben ist, dass ein Verband eine Mitgliederbefragung durchführen will, so dass sichergestellt werden muss, dass der Verband auch tatsächlich die Fragen stellt, die er beantwortet haben möchte. Mehr nicht. Alle anderen Ansinnen sind nicht sachgerecht.

**2. Toleranz gegenüber Methoden, Stichproben und Operationalisierungen.** Empirische Untersuchungen haben immer wieder das Image, dass sie final richtig eine Frage klären könnten. Manche Ergebnisse von Meinungsumfragen gelten auch als Diskussionskiller – was soll man schon gegen Fakten ins Felde führen? Man kann sie bestenfalls akzeptieren und überlegen, wie man Meinungen ändern kann. Fatalerweise ziehen empirische Studien auch Forscher mit stark ausgeprägtem Machtmotiv und mit einem „hierarchischen Selbst Interesse“ an, d.h., diese führen empirische Untersuchungen durch, um als finale Rechthaber dazustehen. Die Geschichte der Meinungsumfragen hat gezeigt, dass solche Attitüden ebenfalls nicht sachgerecht sind. Jede Meinungsumfrage kann auch anders mit anderen Fragen an anderen Stichproben zu anderen Zeitpunkten und mit anderen Zwecken durchgeführt werden. Die Wissenschaft ist hier deutlich toleranter und demokratischer, wenn man so will: transparente und intersubjektiv überprüfbare Informationen – egal wie – sind für den Fortschritt der Wissenschaft willkommen.

**3. Was ist Meinung? Ist sie unwichtig?** Gegen manche Meinungsumfragen zu G8 und G9 Problemen wird angeführt, dass ja nur „subjektive Meinungen“ der Befragten registriert werden und keine objektiven Fakten. Das ist eine banale Feststellung. Selbstverständlich erfasst man in einer Meinungsumfrage nur die Meinungen vieler Subjekte. Man stellt ein

Meinungsbild her, d.h. die unterschiedlichen subjektiven Ansichten über ein Thema. In einer solchen Umfrage ist es selbstverständlich erlaubt, auch Menschen, die keine Erfahrung mit G8 haben, nach ihrer Meinung zu G8 oder G9 zu fragen. Man kann die Meinung unserer Mitmenschen nicht dadurch desavouieren, indem man ihnen sagt, du hast ja keine Erfahrung mit dem Problem, also ist deine Meinung wertlos. Das ist einer Demokratie unwürdig. Man stelle sich einmal vor, dass man bei der Wahl von Parteien alle diejenigen ausschließt, die von Schulpolitik und von Schule keine Ahnung haben oder keine Kinder. Dürfen diese sich über schulpolitische Maßnahmen keine Meinung bilden? Wenn man zum Beispiel in einer Untersuchung findet, dass alle Abiturienten, die G8 absolviert haben, im Nachhinein das Ganze als nicht so schlimm wahrgenommen haben, dann ist das ein interessanter Punkt im komplexen Meinungsbild, aber das macht die Befragung von Grundschulleitern zu G8 und G9 nicht überflüssig. Meinungen kann man überdies auch ändern und Meinungen stimmen nicht mit der faktischen Realität zwangsweise überein – auch das sind absolute Banalitäten die nicht mehr thematisiert werden. Richtig ist, dass Meinungen alle unsere Entscheidungen, die ja immer in der Zukunft liegen, im Alltag und in jeder Stunde und Sekunde determinieren. Deswegen werden sie untersucht. Wir haben keine andere Wahl, als unser gesellschaftliches Leben nach Meinungen über die Zukunft zu entscheiden.

**4. Vielfalt der erlaubten Fragen und Operationalisierungen.** Um Meinungen zu erfragen, muss man Fragen konkretisieren. Diesen Konkretisierungsvorgang nennt man Operationalisierung, d.h. „messbar Machung“. Nun gibt es viele Menschen, die glauben, dass man eindeutig festlegen kann, welche Operationalisierung zur Erfassung etwa der Meinung über G8 und G9 richtig ist. Wenn man 20 Leute bittet, die richtigen Fragen zur G8/G9 Problematik zu stellen, so erhält man 20 verschiedene Frageformulierungen. Im Anschluss daran ergeben sich in Forschungsgruppen manchmal jahrelange Diskussionen darüber, welche Operationalisierung die richtige oder die falsche ist, welche wertvoll ist und welche wertlos ist. Das sind Anfängerdebatten, denn bereits als Erstsemester lernt man, dass jede Operationalisierung Vor- und Nachteile hat, eine richtige oder falsche Operationalisierung kann es überhaupt nicht geben, sondern: unterschiedliche Operationalisierungen zum selben Sachverhalt erlauben eine bessere Schätzung des wahren Sachverhaltes. Anfänger der empirischen Meinungsforschung verfallen häufig auf den Fehler, die Fragen so komplex und kompliziert zu stellen und vor allen Dingen auf germanistische Perfektion hin zu

konstruieren, wo es doch eigentlich notwendig wäre, die Fragen im „Idiom der Befragten“ anzubieten d.h. sie so zu formulieren, wie der Stand des Sprachgebrauchs gerade in der betroffenen Gruppe ist.

In der methodischen Erforschung von Meinungsumfragen hat man ein interessantes Phänomen gefunden. Man kann zum Beispiel mit hunderten von Begriffen die Persönlichkeit eines Menschen erfragen – die empirische Untersuchung ergibt bei der Auswertung (Faktorenanalysen) eigentlich nur fünf Faktoren (Extraversion/Introversion, Gewissenhaftigkeit, Verträglichkeit, Stabilität/Labilität, Offenheit für Erfahrungen), d.h. die sprachliche Oberfläche ist deutlich differenzierter und vielfältiger als die psychischen Gegenstände der sogenannten „latenten Variablen“<sup>9</sup>. Man kann sich eine Kugel vorstellen und auf der Oberfläche dieser Kugel sind alle möglichen Fragen und Begriffe zu einem Gegenstand versammelt, die alle eigentlich nur einen zentralen Punkt haben, den Mittelpunkt dieser Kugel. Ein Verband wie die Landeselternschaft der Gymnasien in NRW e.V. hat natürlich an den manifesten Fragen der Kugeloberfläche in einigen Fällen ein nicht hinterfragbares Interesse, sofern es sich um ganz konkrete Vorschläge handelt, wie etwa die Verbreitung des Ganztags (gebunden oder offen). Hinter der Entscheidung für einen Ganztag in gebundener oder offener Form können dann allerdings wieder dieselben latenten Variablen stehen, zum Beispiel „Konservativität“ oder „Progressivität“ oder „Entschleunigung“.

**5. Mitarbeit der Betroffenen bei der Fragebogenkonstruktion.** Warum werden bei der Konstruktion des Fragebogens die Betroffenen zur Mitarbeit aufgefordert? Diese Frage beantwortet sich durch die vorangegangenen Erklärungen von selbst. Weil es darauf ankommt, die Erwartungen und den Erkenntnishorizont der Betroffenen möglichst genau abzubilden. Hierzu hilft nicht der Bezug zur Germanistik der Frageformulierung (also philologische Kompetenz), sondern nur der Bezug zur Praxis. Am besten wird diese Praxis aber eingebunden, indem man die Betroffenen bittet, den Fragebogen im Vorfeld zu korrigieren, zu verändern und eigene Fragen zu entwerfen. Der in dieser Studie eingesetzte Fragebogen ist also noch nicht einmal zur Hälfte auf den Verfasser zurückzuführen. Deshalb wurde auch nicht nach parteipolitischer Orientierung, nach Berufstätigkeit beider Eltern,

---

<sup>9</sup> Borkenau, P. & Ostendorf, F. (1993). NEO-Fünf-Faktoren-Inventar (NEO-FFI) nach Costa und McCrae (S. 5–10, 27–28). Hogrefe.



nach dem Alter der Eltern etc. gefragt, weil es dem Auftraggeber nur um eine Befragung zu G8 und G9 und einigen darüber hinaus strittigen Fragen zur zeitlichen Organisation von Schule etc. ging. Fragen nach der Berufstätigkeit oder gar der politischen Orientierung hätten die Vertrauensgrundlage im Verband gestört. Daran muss sich eine praxisnahe Meinungsbefragung dann orientieren. Wissenschaftlich wären solche Informationen - wie Parteizugehörigkeit oder elterliche Berufe- natürlich interessant.

**6. Stichproben und Repräsentativität von Stichproben.** Die Öffentlichkeit hat mittlerweile gelernt, dass eine repräsentative Stichprobe so ungefähr 1000 Menschen umfasst - oder? Nein – das ist zum Beispiel eine falsche Auffassung von Repräsentativität. Repräsentativität hat zumindest zweierlei gedankliche Elemente: 1. die Größe der Stichprobe ist in erster Linie für die Fehlermarge verantwortlich (das Konfidenzintervall). Je kleiner die Stichprobe, desto größer die Fehlermarge der Untersuchung. 2.,,Repräsentativ“ heißt eigentlich stellvertretend, d.h. die Stichprobe muss in ihrer Zusammensetzung ein verkleinertes Abbild der Grundgesamtheit sein, deren Meinungsbild sie schätzen möchte. Eine Stichprobe von zwei Personen kann, bezogen auf das Geschlecht, durchaus repräsentativ für die bundesdeutsche Bevölkerung sein, wenn man zum Beispiel einen Mann und eine Frau mittleren Alters per Zufall ausgewählt habe. Die Meinung dieser zwei Menschen ist allerdings mehr oder weniger wertlos, weil man sie nicht verallgemeinern kann. Auch wenn man diese Stichprobe nur aus Menschen bestehen lasse, die eine besondere Position haben, zum Beispiel Direktoren von Gymnasien, dann ist sie kein repräsentatives Abbild der Elternschaft der Gymnasien, sondern sie wäre evtl. repräsentativ für die Direktorenschaft- was jeder verstehen kann. Wenn ein Politiker sagt, er hätte sich mit fünf Direktoren unterhalten und alle fünf hätten ihm gesagt ,dass sie an der Umfrage „G8 und mehr“ nicht teilgenommen haben, und daraus auf den Unwert der Studie schließen möchte, so ist das ebenfalls Indiz für ein fehlendes Verständnis von Stichproben. Zu der letztgenannten Meinung: Es gibt seit gefühlten 50 Jahren eine Registrierung der Einschaltquoten – viele von uns sind niemals an diesen Untersuchungen beteiligt gewesen. „Stichprobe“ heißt, es wird nur ein Teil der Betroffenen ausgewählt. Die postalische Umfrage in dieser Studie ist eine Zufallsstichprobe in zweierlei Hinsicht: es wurden zufällige Klassenstufen und zufällige Eltern bzw. Erziehungsberechtigte ausgewählt. Müssen Stichproben immer repräsentativ sein? Auch das nicht – wenn eine Meinung so banal ist

und von allen Menschen vertreten wird, dann würde eine beliebige Menge von Menschen von irgendwoher reichen, um auf das Meinungsbild der Gesamtheit zu schließen. Nur in wenigen Bereichen ist die Meinung der Gesamtheit schon so weit einhellig, dass man dieses Verfahren ernsthaft empfehlen könnte.

**7. „Die Gesamtheit ist gesetzmäßig“ oder „der Einzelfall ist gesetzmäßig“?** Die amerikanische, deutschstämmige Wissenschaftlerin Miriam Lewin hat diese beiden Prinzipien (galileisches und aristotelisches Prinzip genannt) in ihren methodischen Publikationen herauskristallisiert.<sup>10</sup> Meinungsumfragen werden danach beurteilt, ob es ihnen gelingt, das Meinungsbild einer Gesamtheit korrekt wiederzugeben. Es geht nicht darum, ob der Fragebogen jedem einzelnen und seiner kognitiven Repräsentation des Problembereiches gerecht wird, sondern ob es dem Meinungsbild in der Gesamtheit entspricht. Es geht um Mittelwertstabilität (Reliabilität und Validität) und nicht um Einzelfallstabilität. Hierzu liegen vom Verfasser frühere Untersuchungen vor, die zeigen, dass die Mittelwertstabilität von Fragebogendaten sehr hoch ist - obwohl gleichzeitig die Einzelfallstabilität nur mittelmäßig bzw. durchschnittlich ist!<sup>11</sup> Einige Mittelwerte konnten sogar nach einem erheblichen Zwischenzeitraum **16 Jahre später** an anderen Stichproben repliziert werden.<sup>12</sup> Das ist ein wesentlicher Unterschied zwischen Tests und Meinungsumfragen. Wer einen Test konstruiert, der in der individuellen Diagnose eingesetzt werden soll, muss auch auf eine hohe Einzelfallstabilität, Reliabilität und Validität achten. Viele Laien haben die Befürchtung, dass bei Fragebögen eine zufällige Ankreuzung stattfindet. Das mag wohl sein, dass Menschen, denen zwei Antwortmöglichkeiten vorgegeben werden (zum Beispiel: „mein Kind ist belastet“ und „mein Kind ist nicht belastet“) zwischen den beiden hin und her schwanken und beim Ausfüllen des Fragebogens zufällig die linke oder zufällig die rechte Alternative ankreuzen, aber nicht wirklich damit zufrieden sind. Das macht nichts, denn für jeden, der sich per Zufall für die linke Alternative entscheidet, gibt es per Zufall auch jemanden, der sich zufällig für die rechte Alternative entscheidet. Also ergibt sich kein wirklicher Unterschied im Mittelwert. Wer einem Einzelfall gerecht werden möchte, muss diesen Einzelfall extra untersuchen, sich länger mit diesem unterhalten, damit er genau weiß,

---

<sup>10</sup> Lewin, M. 1986 Psychologische Forschung im Umriß. Berlin: Springer

<sup>11</sup> Dollase, R., Hammerich, K., Tokarski, W., 2000, Temporale Muster. Opladen: Leske und Budrich

<sup>12</sup> Schmücker, F. 1993, Das Jazzkonzertpublikum. Münster: Lit Verlag

welche Meinung dieser Einzelfall hat. Hierfür sind üblicherweise qualitative Studien nötig. Die brauchen aber wiederum so viel Zeit, dass man nur wenige Fälle untersuchen kann. Die darf man dann noch nicht einmal verallgemeinern. Von Einzelfallstudien auf die Allgemeinheit zu schließen ist falsch - genauso wie es falsch ist, von Durchschnittsaussagen über eine große Grundgesamtheit auf den Einzelfall zu schließen (sog. „ökologischer Fehlschluss“).

**8. Es machen nicht alle mit – Missing Values.** Die Erfahrung zeigt, dass es schwer ist, Menschen zur Mitarbeit bei einer Umfrage zu bewegen. Bei den großen Meinungsforschungsinstituten, die mit telefonischen Interviews arbeiten, kann man schon mal 100 Telefonate tätigen und fünf davon nehmen an einer Umfrage teil. Der Verfasser hat Studien mit postalischem Rücklauf von Fragebögen durchgeführt, die lediglich 20 % bis 30 % der Befragten ausgefüllt zurückgeschickt haben. Untersuchungen haben generell gezeigt, dass diejenigen, die bei einer Umfrage mitmachen, eher zu den wohlwollenden und interessierten Mitbürgern gehören und dass sie keinesfalls in den Fragen wesentlich andere Antworten geben als diejenigen, die eine Umfrage nicht mitmachen.

Aber damit ist das Problem noch nicht zu Ende: bei denjenigen, die an einer Befragung teilnehmen, stellt man immer wieder fest, dass sie Teile des Fragebogens oder einzelne Fragen nicht beantworten. Deswegen hat man bei der Auswertung – und das wird, weil es so banal ist, im weiteren auch nicht erklärt – ein unterschiedliches N (von number = Anzahl der Befragten). Wenn man zwei Fragen miteinander vergleicht (z.B. bei einer Kreuztabulierung oder einer Korrelation), dann kann die Zahl der Befragten noch mal sinken, weil bei der ersten Frage andere Menschen nicht geantwortet haben wie bei der zweiten Frage. Beides kann sich im Extremfall addieren.

**9. Signifikanz und Relevanz.** Ist ein Unterschied von 79,5 % und 81,2 % schon ein bedeutsamer Unterschied? Ja, wenn man nur die Signifikanz prüft und die Stichprobe sehr groß ist, dann kann ein so minimaler kleiner Prozentsatz Unterschied schon signifikant sein und könnte interpretiert werden als ein Unterschied, der nicht durch Zufall zustande gekommen ist. Eine solche Argumentationsweise war früher üblicher, zumal viele nicht wussten, was ein Signifikanztest wirklich ist (zufallskritische Absicherung einer Hypothese und kein Effektmaß) und ebenfalls nicht wussten, dass, je größer die Stichprobe ist, je kleiner die Unterschiede sind, die signifikant werden. Der Unterschied von 79,5 % zu

81,2 % ist möglicherweise signifikant, das hängt von der Größe der Stichprobe ab, aber er ist praktisch nicht relevant. In dieser Studie, in der es häufig um Stichprobengrößen im fünfstelligen Bereich geht, werden solche Signifikanzteste gar nicht erst angestellt und mitgeteilt, weil dann nur sehr kleine und wenig relevante Unterschiede signifikant werden und bei der Interpretation in die Irre führen. Welche Unterschiede man für relevant hält oder halten kann, wurde bei den einzelnen Ergebnissen dargestellt.

**10. Notwendigkeit pragmatischer Entscheidungen.** Man kann zu jedem Thema einer Meinungsumfrage tausende von unterschiedlichen Fragebögen konstruieren, das muss immer wiederholt werden, weil es sich offenbar so schlecht einprägt. Auch zur G8/G9 Problematik. Die Kritik, dass man eine Frage auch anders hätte formulieren können, ist im Vorfeld schon häufig genug als belanglos bewertet worden. Das ist richtig, aber auch nicht zu verhindern. Wenn man eine Kritik realisiert, den Fragebogen also ändert, produziert man gleich wieder andere Kritiken, die genau diese geänderte Formulierung für nicht passend halten. Ein regressus ad infinitum. Aus pragmatischen Gründen muss man, wenn man sich auf ein Verfahren der Konstruktion des Fragebogens geeinigt hat, damit abfinden, dass er so geworden ist, wie er ist. Damit nicht genug. Immer wieder wird moniert, dass man doch hätte diese und jene Auswertung vornehmen können, müssen, sollen. Richtig: auch hier sind pragmatische Entscheidungen nötig – denn es sind Tausende von möglichen Auswertungen denkbar. Beispiel: man kann alle Fragen nach Jungen und Mädchen auswerten, dann alle Fragen nach jungen Jungen und jungen Mädchen, dann alle Fragen nach jungen Jungen und jungen Mädchen, je nachdem ob Vater oder Mutter den Fragebogen ausgefüllt hat. Und dann könnte man noch differenzieren danach, ob die jungen Mädchen und jungen Jungen, wenn der Vater oder die Mutter den Fragebogen ausgefüllt hat, gut in der Schule sind und dann damit vergleichen, ob der Vater einer Tochter oder eines Sohnes öfter zu G8 oder G9 neigt. etc. Damit wäre die Differenzierung aber noch nicht am Ende. Anfänger werden mit der Auswertung einer solchen Umfrage nie fertig, weil sie nicht verstehen, dass man zu einem pragmatischen Kompromiss auch bei der Vielzahl der möglichen Auswertungen gezwungen ist. Standard ist heute, dass man eine „Grundauszählung“ macht, das heißt, die Antworten auf jede einzelne Frage darstellt und dass man die Ergebnisse nach relevanten Unterscheidungskriterien differenziert. Das wird hier in Bezug auf G8 und G9 Präferenzen

gemacht. Die Auswertungsmöglichkeiten sind bei einer solchen Datei im Prinzip unendlich groß – deshalb müssen pragmatische Kompromisse der Begrenzung gemacht werden.

**11. Selektionsverführung der Kritik.** Weil im Prinzip eine unendliche Menge von Fragen zu ein und derselben Problematik gestellt werden könnten und weil im Prinzip eine unendliche Menge von Auswertungsoptionen denkbar wäre, gibt es eine Verführung für jene, die eine Untersuchung abwerten wollen. Wir brauchen ja nur zu schreiben „hier hätte er eine andere Frage stellen müssen“ und da „hätte er eine andere Auswertung machen können“ um Laien (zum Beispiel auch Journalisten oder Politiker) zu beeindrucken. Die würden dann selbstbewußt auftreten und sagen „Ja, ich habe die Untersuchung von zwei ausgewiesenen Fachleuten prüfen lassen und sie haben festgestellt, dass die Untersuchung eine Reihe von Mängeln hat. Bei einigen Fragen ergibt sich deutliche Kritik. Auch alle wichtigen Auswertungsverfahren sind nicht durchgeführt worden.“ Eine solche Kritik ist naiver Unsinn. Eine Untersuchung macht das, was sie getan hat. Wenn man etwas anderes machen möchte, sollte man das tun. Angesichts der Tatsache, dass hier unendlich viele Möglichkeiten der Frageformulierung und der Auswertungsstrategien denkbar sind, sind solche Pauschalkritiken der Ausweis von Naivität.

**12. Kausalität.** Es gibt Menschen, die glauben, dass man mithilfe von Meinungsumfragen die Ursachen für eine bestimmte Meinung herausbekommen kann. Das ist nicht ganz richtig und wurde weiter oben schon erwähnt. Kausale Analysen setzen zwingend Experimente voraus, die auch im Humanbereich gemacht werden können. In einer normalen Meinungsumfrage kann man über mögliche Ursachen spekulieren, d.h. diskutieren und Überlegungen anstellen, wie die Daten zu erklären sind. Man kann einige dieser Überlegungen auch anhand von Daten eher ausschließen, wie man im weiteren Verlauf dieses Berichtes sehen wird. Eine finale Ursachenfeststellung ist auch dann nicht möglich, wenn man etwa fragt „Was ist die Ursache dafür, dass Sie für G8 sind?“ und dann sieben Gründe vorgibt. Auch diese dann genannten Meinungen der Befragten sind Meinungen und zwar Meinungen über Ursachen. Mehr nicht.

**13. Die Unabhängigkeit von Untersuchungen vom Durchführenden.** Empirische Meinungsumfragen dienen der transparenten Kritik an vorgefassten Meinungen über das

Meinungsbild der Bevölkerung. Oder zur Überprüfung wissenschaftlicher Hypothesen über die Verbreitung von Meinungen. Wenn jemand glaubt, dass die G8 Präferenz bei gymnasialen Eltern die Mehrheit hat und er führt anschließend eine Untersuchung durch, in der dieses geprüft wird, so ist das sehr loblich, weil man in der Berichtslegung über die Untersuchung sehen kann, ob er ein Vorurteil in der Fragestellung bzw. Auswertung angewendet hat oder ob er die Untersuchung fair ausgewertet hat. Insofern ist irrelevant, dass der Verfasser sich vor Durchführung dieser Untersuchung in zwei Artikeln gegen die G8 Lösung gewendet hat, weil der falsche Eindruck entsteht, dass ein Jahr weniger Schule keine Folgen für die Bildung von Menschen hätte (s.o.).

Ist der Untersucher unabhängig? Ja - er ist parteilos, allerdings G10 Befürworter. War die G10 Befürwortung vorher bekannt? Ja - Die gezeigten Folien am 16.11.15 (der Herbst Mitgliederversammlung) sprechen eine eindeutige Sprache: das persönliche G10 Schicksal (einmal sitzengeblieben) wurde geschildert, ein Fairness Angebot deswegen unterbreitet, d.h. auch Gegner von G9 wurden explizit an der Fragebogenkonstruktion beteiligt. (Wo gibt es das sonst noch?)

Hat der Untersucher die Studie beeinflusst? Nein. Der Fragebogen ist unter Mitarbeit von 15 anderen Personen entstanden, nur ein geringer Teil der Fragen stammt vom Verfasser. Onlinefragebögen werden automatisch ausgewertet, andere werden wie in der Wissenschaft üblich, transparent und wiederholbar im Abschlußbericht dargestellt.

### **Ein Dauerthema: Operationalisierungen der „Belastung“ - Möglichkeiten**

Zwei FAQ's sollen hier beantwortet werden:

1. Wenn die Eltern sagen, ihr Kind sei belastet, dann muss man doch erst „Belastung“ definieren. Wer weiß schon, was Eltern darunter verstehen?
2. Wie differenziert müssen Antwortmöglichkeiten sein?

**Zur ersten Frage:** Nein - „Belastung“ wird nicht definiert. Wörter aus der Alltagssprache bleiben bei solchen Umfragen undefiniert - ihre Bedeutung ergibt sich nach der Erhebung durch die empirischen Korrelationen mit vergleichbaren Begriffen (sog. face oder construct validity). Eine Definition würde auch wieder Worte benutzen, die dann ebenfalls alle definiert werden müssten. Also: man käme aus dem Dilemma unterschiedlichen Sprachverständnisses nicht heraus. Dass der eine den Begriff zu streng auslegt, ist normal

### Darstellung I: Grundraten einer Frage nach der Belastetheit des Kindes mit 5 Antwortalternativen vs. einer dichotomen Antwortvorgabe

zu niedrig	0,7/0,6
niedrig	7,9/4,6
ok	48,6/36,0
hoch	34,5/43,7
zu hoch	8,3/15,0

#### Grundraten postalisch/online

nicht unangemessen belastet	42,4/45,9
belastet	57,6/54,1

und führt zu dem oben schon erläuterten Fehlerausgleich bei Mittelwerten, weil es genauso viele gibt, die ihn zu „lasch“ auslegen.

**Zur zweiten Frage:** Im folgenden Beispiel soll zur Verdeutlichung der unterschiedlichen Operationalisierungen der Vergleich der Antworten zu einer Belastungsfrage mit fünf Antwortvorgaben und einer Belastungsfrage mit zwei Antwortmöglichkeiten dargestellt werden. In der Darstellung I sind zunächst die Grundraten für die postalische und die Onlinebefragung in Prozent dargestellt.

In den beiden folgenden Darstellungen – Darstellung 2 und Darstellung 3 – ist der Vergleich zwischen der Dichotomie und der vierstufigen Antwortskala dargestellt. Das nennt man Kreuztabulierung (crosstabs bei SPSS).

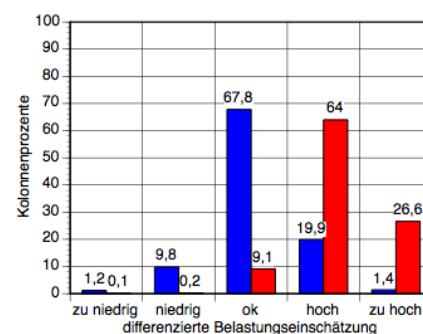
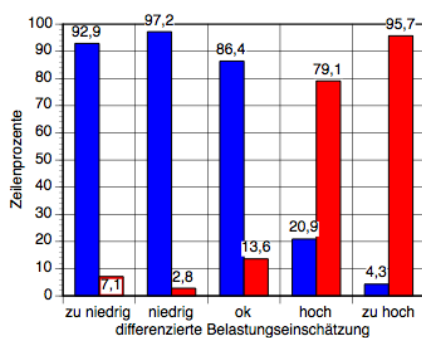
Man sieht, dass bei der Onlinebefragung nur ein kleiner Prozentsatz bei der Mittelkategorie fehlplatziert ist und bei der postalischen Umfrage waren in der Mittelkategorie und auch in

den beiden unteren Kategorien „zu niedrig“ und „niedrig“ einige Eltern, die bei der dichotomen Antwortalternative „belastet“ angekreuzt haben, aber bei der mehrstufigen nicht.

Wie sieht dasselbe bei der postalischen Umfrage, der Zufallsumfrage also, aus? Man erkennt, dass die Fehlplatzierungen etwas häufiger sind als bei der Onlinebefragung.

Onlinebefragungen lassen sich im Nachhinein - nach Ende des Ausfüllens - nicht korrigieren (einige haben auch dies bemängelt), weil die Seiten nicht zurückgeblättert werden können.

## Darstellung 2: N=29 603 (online) - Vergleich einer differenzierten mit einer dichotomen Antwortvorgabe zur Belastung



■ Es ist nicht unangemessen belastet  
■ Es ist belastet

■ Es ist nicht unangemessen belastet  
■ Es ist belastet

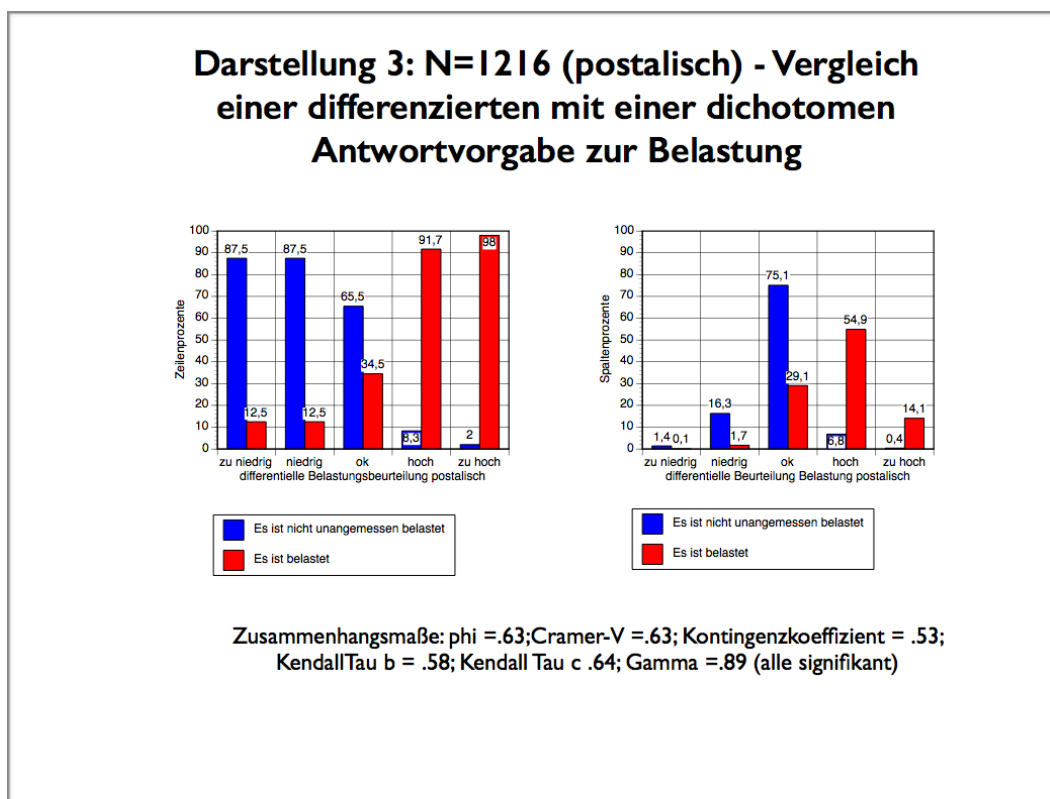
Zusammenhangsmaße:  $\phi = .71$ ; Cramer-V =  $.71$ ; Kontingenzkoeffizient =  $.58$ ;  
KendallTau b =  $.64$ ; Kendall Tau c =  $.74$ ; Gamma =  $.92$  (alle signifikant)

Obwohl die postalische Umfrage als Zufallsstichprobe das höhere Vertrauen genießt, ist sie für nachträgliche Änderungen durch die Befragten offen. Das kann auch zu solchen Unterschieden führen.



Dennoch: Sowohl online wie auch postalisch sind alle in SPSS verfügbaren Zusammenhangsmaße bei Kreuztabulierungen (s. Darstellung 2) hoch und signifikant. Dh. es gibt empirisch keinen Grund, für die eine oder andere Operationalisierung zu sein.

Die Onlinebefragung ist also konsistenter als die postalische Befragung. Dennoch kann man nicht sagen, dass die Eltern, die bei einer fünfstufigen Antwortskala zunächst antworten, die Belastung ist „zu niedrig“ bzw. „niedrig“ und später meinen, das „Kind ist belastet“ einen Fehler gemacht haben. Bei der zweiten Antwortalternative (Dichotomie) folgten Belastungsfaktoren, von denen man dann einige ankreuzen konnte. Dabei waren dann auch einige, die vielleicht von den Eltern nicht direkt mit der „schulischen“ Belastung verbunden waren, sondern mit der Belastung durch Klassenkameraden, Lehrer oder sonstigen Unvermeidbarkeiten.



Von Bedeutsamkeit für die Interpretation ist aber die Tatsache, dass alle Zusammenhangsmaße hoch und signifikant waren. Und zwar wurden sowohl jene für nominale wie ordinale Daten unterhalb der Darstellung 2 und 3 abgedruckt. Man darf also daraus schließen, dass unterschiedliche Operationalisierungen – und diese beiden waren besonders unterschiedlich wegen der zusätzlichen Benennung von Belastungsfaktoren –

trotzdem zu vergleichbaren Ergebnissen führen. Dies nur als Beispiel zur Illustration einer der oben genannten Befürchtungen.

## **5. Diskussion: Beitrag der Ergebnisse der Umfrage „G8 und mehr“ zu bildungspolitischen, wissenschaftlichen und theoretischen Fragen**

Eine wissenschaftliche Untersuchung wird üblicherweise mit einem Kapitel „Diskussion“ beendet. In einem solchen Abschnitt hat der Autor die Möglichkeit, interpretative, wertende und handlungsrelevante Ideen zu seinen Ergebnissen äußern. Diese Äußerung von „Meinungen“ darf allerdings nicht unabhängig von den gefundenen Ergebnissen sein. Sie müssen sich mit den Hauptergebnissen vertragen, man sagt auch, „kompatibel“ sein. Auch wenn im Abschlussbericht betont wurde, dass man mit Meinungsumfragen keine Kausalitäten final feststellen kann, so sind doch aufgrund der Ergebnisse manche Ideen zur Verursachung der Ergebnisse ausgeschlossen. Eine ähnliche gedankliche Figur liegt den statistischen Strukturgleichungsmodellen bzw. Pfadanalysen zu Grunde, die unter der Auflage, mit den empirischen Interkorrelationen kompatibel zu sein, errechnet werden.

In der Diskussion werden die Ergebnisse der Studie nicht mehr erneut zusammengefasst. Deswegen ist zu Beginn dieses Kommentarbandes die Zusammenfassung der Ergebnisse und Methoden vorangestellt. Die folgenden Ausführungen können nur verstanden werden, wenn man die Ergebnisse präsent hat.

Zentrale Ergebnisse – das ist eine Gewichtung und gehört in die Diskussion hinein – sind neben der deutlichen Präferenz für G9, vor allen Dingen die Befunde zur Verbindung der G8/G9 Entscheidung mit der differenziellen temporalen Belastbarkeit der Schüler und Schülerinnen aus der Sicht der Eltern, sowie die strikte Ablehnung sowohl von G8 wie G9 Befürwortern gegen jede Form der gebundenen Ganztagsregelungen. Diese Ergebnisse sind im Zusammenhang mit der eindeutigen Ablehnung der Gesamtschule der eigentliche wissenschaftliche Beitrag dieser Umfrage zur Diskussion der Psychologie von bildungspolitischen Entscheidungen.

## 5.1. Die Verlässlichkeit der Ergebnisse

Die erste Diskussion betrifft die Verlässlichkeit der Ergebnisse. Die Hoffnung jener, die stur bei G8 bleiben wollen, ist ja immer, dass man einem pro G9 Befund aus Meinungsbefragungen methodische Zweifel anhängen kann. Sind diese methodischen Zweifel einmal in der medialen Diskussion, kann man stets locker darauf hinweisen dass „die Ergebnisse fragwürdig“ seien. Ein solches Argumentationsspiel von Talkshows und medialen Kommentaren kann natürlich einen Wissenschaftler nicht überzeugen.

Aus diesem Grunde ist die vorliegende Umfrage von Anfang an mit besonderer methodischer Sorgfalt angelegt und durchgeführt worden. Die Mitarbeit der Betroffenen bei der Fragebogengestaltung, die multimethodale Erhebung durch zwei Online- und eine Zufallsstichprobe mit postalischer Verteilung, die Größen der Stichproben, die Erhebung der G9 Präferenz gegen Distraktoren und die Befragung mehrerer Personengruppen zum Gegenstand der G8 und G9 Präferenz, sind die wesentlichen methodischen Vorkehrungen, um die Verlässlichkeit dieser Umfrage zu stützen.

Zugleich sind die Ergebnisse durch vorauslaufende, parallele und mittlerweile auch nachherige Umfragen gestützt worden. Nach dieser Umfrage, die im Februar und März des Jahres 2016 durchgeführt wurde, gab es in Baden-Württemberg eine Umfrage der SPD, eine Umfrage der FDP bei den Kreisvorsitzenden in NRW im Sommer 2016, sowie zahllose kleinere Studien bzw. Registrierungen des Wahlverhaltens von Eltern und Schülern (zum Beispiel in Hessen), die das Hauptergebnis der sehr großen Akzeptanz von G9 bestätigen. Aus diesem Grunde kann man sich auf die Ergebnisse dieser Studie natürlich auch verlassen.

Die große Zahl der teilnehmenden Eltern an den drei Umfragen geht einher mit einem statistisch geringen Konfidenzintervall, d.h. die Fehlermarge, in der der „wahre Wert“ der Grundgesamtheit liegt, wird enger.

Von besonderem Interesse bei der Auswertung der Daten war der Vergleich zwischen einer postalischen Zufallsstichprobe und den beiden Onlinebefragungen. Gemeinhin wird pauschal davon gesprochen, dass Onlineumfragen nicht verlässlich seien. Die Ergebnisse dieser Studie zeigen ein anderes Bild.

Einerseits sind Antworten, die mit den zentralen und in der Öffentlichkeit umstrittenen Fragen von G8 und G9 zusammenhängen, durchaus unterschiedlich, bei der Größe der Stichproben selbstverständlich und banalerweise auch signifikant, aber in den numerischen Dimensionen keineswegs so unterschiedlich, dass sich beispielsweise ein korrelatives Meinungsbild von unterschiedlicher Qualität zeichnen ließe. Kurz: auch in den umstrittenen Fragen liefern Onlineumfragen dimensional kaum andere Ergebnisse als postalische Zufallsumfragen.

Erstaunlicherweise, so muss man aufgrund der pauschalen Unterstellungen über den Wert von Onlineumfragen konstatieren, sind die Ergebnisse beider Umfragen nahezu identisch, sobald es sich um nicht aktuell umstrittene, sondern überdauernde Einstellungen, etwa zur Präferenz des Gymnasiums und zur Ablehnung gebundener Ganztagsregelungen, handelt.

Onlinebefragungen und postalische Zufallsbefragungen haben einen entscheidenden Vorteil gegenüber telefonischen oder face to face Befragungen: in beiden Fällen ist der Antwortende ohne informelle Kontrolle durch einen Interviewpartner. Auch bei telefonischen Befragungen gibt es auf Seiten des Befragten Tendenzen zum „Impression Management“, d.h. zum hinterlassen eines bestimmten „erwünschten Eindrucks“ beim Interviewpartner. Das kann zu beeinflussten Prozentsätzen führen.

Berühmtes Beispiel aus dem Jahre 2016: die Meinungsforschungsinstitute haben die Resonanz der AFD bei den Landtagswahlen deutlich unterschätzt – sie ist offenbar von weniger Menschen im Telefoninterview genannt worden als sie dann nachher gewählt wurde. (Ähnliches kann bei den Fehlprognosen der Institute zum Brexit geschehen sein.) Möglicherweise hängt das damit zusammen, dass etwa zum Zeitpunkt der Wahl in einigen Bundesländern die Partei zu stark in der öffentlichen Kritik stand. Einige Anhänger der Partei trauen sich dann nicht, diese Parteipräferenz im Gespräch zuzugeben.

In der zweiten Onlinebefragung haben nur 127 Direktoren teilgenommen. Warum so wenige? Zunächst einmal ist der allgemeine Forschungsbefund zu zitieren, dass höhere

Statusgruppen seltener an Umfragen teilnehmen als die weniger einflussreichen Statusgruppen. Mit anderen Worten: Manager zu befragen ist schwieriger als das „gemeine Volk“ bitten, mitzumachen.

Zum anderen ist denkbar, dass die allgemeine Einstellung „an solchen Umfragen nehmen wir erst gar nicht Teil, weil das nur die Gegenseite stärkt“ oder „weil das den Eindruck erweckt, es wäre an G8 noch was zu ändern“ (Originalwortlaut von Gesprächen mit Direktoren) – zu einer Teilnahmeabstinenz geführt hat. Möglicherweise ist das eine falsche Einschätzung der Bedeutung von Meinungsumfragen.

## 5.2. Darf Politik getroffene Entscheidungen revidieren?

Am 20.8.2016 hat ein Kommentator im WDR die anstehenden Entwicklungen der G8/G9 Problematik kommentiert und am Ende der amtierenden Ministerin, die vor ihrem Amtsantritt gegen G8 war, aufgefordert, konsequent zu sein, andernfalls „würde ihre Glaubwürdigkeit“ auf dem Spiel stehen. Ein solcher Ratschlag ist absurd. Und das aus mehreren Gründen. Die Ministerin war Anfang der 2000er Jahre selber gegen G8, hat sich also dem Druck des Ministeriums und der Regel, dass die Entscheidungen der Vorgängerregierung nicht sofort geändert werden sollen, weil das eine unübersehbare Instabilität und zusätzliche Arbeit erfordern würde, gebeugt. Wenn sie jetzt zu ihren eigenen Überzeugungen zurückkehren sollte, so wäre das ein ehrenwerter Vorgang, mit dem eine einmal getroffene Maßnahme aufgrund der geringen Akzeptanz in der Bevölkerung und bei den Betroffenen rückgängig gemacht würde.

Die potentiell mögliche Revision aller Entscheidungen ist das Prinzip jeder erfolgreichen Qualitätssteuerung. Qualitätssteuerung besteht aus einem mehrschrittigen Prozess, der da lautet „planen – umsetzen – überprüfen – handeln“. Im Englischen: PDCA Zyklen oder auch Deming - Kreis genannt.“ (Plan – Do – Check – Act).

Als Kritiker der bürokratisch – administrativen Qualitätsteuerung im Bildungssystem wundert den Verfasser dennoch, dass ein wesentliches Kennzeichen aller Qualitätsverbesserung bei der G8/G9 Problematik nicht beachtet wird: wenn eine Maßnahme derartig stabil von der Bevölkerung und den Betroffenen abgelehnt wird, wenn sie erkennbar zu Entwicklungen führt, die keiner gewollt hat (z.B. Diskriminierung der

langsamen Lerner), dann ist es völlig normal und keineswegs problematisch und kein Makel der Glaubwürdigkeit von Politikern, wenn sie diese Maßnahme wieder rückgängig machen. Nur dann, wenn eine Maßnahme, die nicht den gewünschten Erfolg hat, auch wieder rückgängig gemacht werden kann, stellt sich ein Verbesserungseffekt ein. Die Ablehnung von G8 ist ein Mißerfolg der Einführung von G8, weil die Akzeptanz fehlt. Die Rückgängigmachung dieser Entscheidung wäre ein Akt der Qualitätsverbesserung.

Nun hat das Ministerium mit dem „runden Tisch“ (Mai 2014) und der wissenschaftlichen Begleitung einen Prozess der Veränderung von G8 eingeleitet, hält sich scheinbar also an den PDCA Zirkel. Diese wissenschaftliche Begleitung scheint sich wenig um die Akzeptanz als Teil des Bewertungskriteriums zu scheren. Man versucht hingegen, objektiv nachzuweisen, dass G8 positive Effekte habe. Man will G8 nicht generell auf den Prüfstand stellen, sondern bestenfalls Modifikationen an G8 anbringen. Das verstößt gegen das Prinzip aller Verfahren der Qualitätssteuerung. Am Ende muss auch die Abschaffung von G8 möglich sein. Natürlich nur, wenn die Ergebnisse entsprechend ungünstig ausfallen. Und das tun sie ja, weil das wichtigste Kriterium - die Akzeptanz - auch nach 8 Jahren nicht erreicht wird. Oder aber - man bemüht den argumentativen Trick und erklärt die fehlende Akzeptanz als unwichtig, weil das Volk zu dumm ist, die Vorteile zu erkennen.

Ein negativer Ausgang aller wissenschaftlichen Begleitung zugunsten von G8 ist allerdings bei der Bewertung von Reformmodellen eher unwahrscheinlich, da die Bevölkerung insgesamt noch Ahnung vom G9 Abitur hat und weiß, das damit Millionen von Menschen in unserem Lande ein erfolgreiches Leben führen konnten. Es ist für ein Reformmodell kaum möglich, eine derartig transhistorische Validität nachzuweisen.

Die Logik sämtlicher Qualitätssteuerungen geht im übrigen davon aus, dass man bei Einführung einer Maßnahme und sei sie noch so gut geplant, nicht davon ausgehen kann, dass sie problemlos funktioniert. Qualitätssteuerung geht also von Unsicherheit aus, die vor der Realisierung vorhanden ist. Erst die Bewährung in der Praxis zeigt, ob die Planung eine tatsächliche Verbesserung der Realität darstellt. Reformen sind immer Eingriffe ins Unbekannte mit unbekanntem Ausgang.

In der modernen Sozialpsychologie spielt die Unsicherheit von Meinungen, Plänen und Vorstellungen von der Zukunft, eine viel größere Rolle als früher (s.o.). Man sieht die Chancen einer Akzeptanz aktueller Unsicherheit gegenüber der Zukunft (Veränderung), aber auch die Gefahren. In einigen Varianten dieser neuen „uncertainty“ Psychologie (Hogg und Blaylock, s.o.) wird der Zusammenhang zwischen politischem Extremismus und der Unsicherheit über die Zukunft deutlich gemacht. Wir wissen alle nicht, was die

Flüchtlingskrise noch bringen wird, allerdings lösen wir die Unsicherheit, die wir alle logisch und sachlich haben, entweder durch Optimismus oder durch Pessimismus. Beide Grundstimmungen führen, wie jeder weiß, zu sehr unterschiedlichem politischen Wahlverhalten. Optimismus wie auch Pessimismus sind allerdings nicht sachlogisch begründbar - es sind Glaubenshaltungen. Das ist bei Einführung von G8 ebenso gewesen - eine reine Glaubenshaltung war die Einführung, zumal man sich nicht auf interdisziplinäre Prognosen über den Wert einer Verkürzung der Gymnasialzeit gestützt hat. Lediglich der „bildungsökonomische Komplex“ (mdl. Mitteilung Nida Rümelin, 2016) und das „Auslandsargument“ („die neuen Bundesländer und das Ausland machen es ja auch vor“), das niemals von Kultur zu Kultur transferiert werden darf, standen Pate bei der Entscheidung.

Ein häufig genanntes Argument gegen „die Rolle rückwärts zu G9“ ist der Arbeitsanfall, der mit einer neuerlichen Reform verbunden wäre. Wenn dem so ist, so geht es nun darum, ein möglichst einfaches und wenig aufwändiges Verfahren zu finden, wie man zu G9 zurückkehren kann. Das ist selbstverständlich möglich.

In Diskussionen hört man häufig, dass die DirektorInnen der Gymnasien ziemlich einhellig für G9 seien, weil sie den Rückbau zu G8 nur als sehr dornenreich, aufwändig und demzufolge als zusätzliche Belastung erleben würden. Deswegen würde bei einer Rückkehr zu G9 damit zu rechnen sein, dass die G8 Befürworter unter den Direktoren und Direktorinnen stärkeren Widerstand leisten. Aber hier muss man fragen, für wen Politik gemacht wird – für Eltern und SchülerInnen oder für die Schulleiter von Gymnasien?

### **5.3. Zeitdiskriminierung - ein neues Phänomen**

Die Forschungsarbeiten des Amerikaners Carroll, die weiter oben diskutiert worden sind, zeigen, dass man die unterschiedlichen Lern- und Leistungsniveaus auch in Zeitbedarf umrechnen kann - z.B. wer in der Schule schlecht ist, braucht mehr Zeit, um besser zu werden. Dieses ist in zahlreichen empirischen Untersuchungen im Umkreis der Theorie von Carroll belegt worden.



Mit diesem Forschungszweig wird deutlich, dass die Lern- und Bildungszeit zu einer Diskriminierungsvariable werden kann. Wer schneller lernt, ist besser. In der Intelligenzforschung unterscheidet man auch die „speed“ Komponente der Intelligenz von der „Power“ oder „Niveau“ Komponente, letztere misst keine Dauer für eine vorgegebene Aufgabe. Die empirischen Ergebnisse der vorliegenden Meinungsumfragen zu G8/G9 stützen diese Ansicht. Mit der Verkürzung der gymnasialen Zeit wird eine neue Diskriminierungsvariable eingeführt, d.h. eine neue Möglichkeit, die Bildungsgerechtigkeit zu verhindern. Die Verkürzung der Zeit schafft Bildungsungerechtigkeit. G8 wird von jenen abgelehnt, die dadurch „belastet“ werden und im Durchschnitt eher weniger erfolgreiche SchülerInnen sind.

Das Ergebnis ist nahe liegend. Man hätte es auch erwarten können und man hätte es auch bei der politischen Entscheidung berücksichtigen können, dass man mit der Zeit bis zum Abitur eine neue Diskriminierungsvariable einführt. Unter Umständen war es und ist es auch so gewollt, um dem Gymnasium einen exklusiven Rang zu sichern.

Diskriminierungsvariable heißt: es ist immer zu erwarten, dass einige Menschen weniger Zeit brauchen, um dasselbe zu erreichen und andere länger. Und nun wird durch die Zeitverknappung das G8 Gymnasium nur für die weniger belasteten SchülerInnen reserviert, d.h. für jene, die auch in weniger Zeit mehr lernen können.

Man könnte überspitzt formulieren, dass das G8 Gymnasium für die Streber reserviert ist. Und nicht nur für diese, sondern für alle Entwicklungsverläufe ohne Brüche und ohne „ups“ and „downs“. In der Bevölkerung und in der Entwicklungspsychologie ist hingegen bekannt, dass die Entwicklung eines jungen Menschen nicht einfach und linear verläuft, sondern in vielen Fällen mal akzeleriert, mal retardiert. Junge Menschen mit unregelmäßigen Entwicklungswegen werden durch G8, werden aber auch durch die Abschaffung des Sitzenbleibens benachteiligt. Es gibt eine deutliche Evidenz dafür, dass im gegliederten Schulsystem Deutschlands, das Sitzenbleiben, bei grundsätzlicher Eignung für das Gymnasium, in der späteren Karriere der ehemaligen Sitzenbleiber keine negativen Folgen zeitigt, eher sogar positive. Einen „Pubertätsdurchhänger“ zu haben, ist kein Grund zur Annahme, dass jemand nicht für die allgemeine Hochschulreife geeignet ist. Sitzenbleiber in der Pubertät konnten auch den Nobelpreis gewinnen (z.B. Thomas Mann, Albert Einstein).

Die angloamerikanischen Untersuchungen weisen für das Sitzenbleiben nur innerhalb von Gesamtschulsystemen (also im gemeinsamen Lernen) nach, dass diese einen negativen Effekt haben. Kein Wunder: dort ist man stets dem Vergleich mit den Besten ausgesetzt, verliert also immer. Dieser Nachweis fehlt empirisch für ein gegliedertes Schulsystem. Deswegen sind die vorhandenen Untersuchungen zum Sitzenbleiben nicht auf die deutsche Schuldiskussion übertragbar.

#### **5.4. Sind Gesamtschulen und gebundene Ganztagschulen noch zeitgemäß?**

Ein besonderes Ergebnis dieser Umfrage ist die ziemlich eindeutige Ablehnung von Gesamtschule und gebundenen Ganztagschulen. Zur Erklärung kann man Vorurteile gegen gymnasiale Eltern bemühen. Sie praktizieren angeblich noch ein traditionelles Familienmodell bzw. sie können es sich finanziell leisten, dass nicht beide Ehepartner erwerbstätig sind, weshalb sich ein Elternteil stets um die Kinder nachmittags kümmern könne. Außerdem favorisieren sie das Gymnasium, um sich von anderen abzusetzen, quasi außer reinem Distinktionsbedürfnis heraus.

Abgesehen davon, dass jede Pauschalisierung falsch ist, entspricht dieses Bild eher der Realität im 19. oder 20. Jahrhundert. Die Ablehnung von gebundenem Ganztags und gleichzeitig auch von der Gesamtschule, kann unmittelbar mit der Modernisierung der Gesellschaft und mit der Flexibilisierung der Arbeitszeiten zusammenhängen.

Beginnen wir mit der in dieser Untersuchung festgestellten eindeutigen Ablehnung von gebundenen Ganztagsregelungen. Nur wer davon ausgeht, dass moderne Arbeitsverhältnisse sich in das Muster von „nine to five“ fügen, kann die Hoffnung haben, dass eine Ganztagschule wesentliche Zeitprobleme der Eltern löst. Gebundene Ganztagsregelungen zerstören offenbar die flexibilisierte familiäre Zeitkultur und führen zu den festgestellten Ablehnungen.

Warum? In den Forschungen zu den temporalen Mustern (s.o.) wurde mehrfach festgestellt, dass Wunscharbeitszeiten nicht Gleitzeitarbeit, also jeden Tag ein bißchen früher kommen oder etwas eher gehen können, bedeutet, sondern, dass sich eine stärkere Blockung von

Arbeitstagen und Freizeittagen für viele Menschen als günstiger erweist. Freitag, Samstag und Sonntag ganz frei haben oder Sonntag, Montag und Dienstag - das ist gewünscht. Bei Befragungen zur Notwendigkeit der öffentlichen Betreuung bei Kleinkindern wurde etwas ähnliches festgestellt: 23 % der Eltern benötigen nur bis 20 Stunden lang Betreuung - 46 % bis 30 Stunden und über 15 % mehr als 60 Stunden. An manchen Tagen ist Betreuung bis 22:00 Uhr nötig, vor allen Dingen auch samstags und sonntags für Eltern, die am Wochenende zu Hause oder unterwegs oder in Unternehmen arbeiten müssen. Die Zeitflexibilisierung der modernen Gesellschaft, die auch durch Teilzeitarbeit beschleunigt wird, macht völlig neue Überlegungen zur Zeitregelung in Schulen notwendig. Welche das im einzelnen sind, ist noch nicht genügend erforscht. Die benötigten neuen Zeitregelungen erfordern sehr viel organisatorische Fantasie und völlig neue Modelle der Beschulung.

Wenn diese Überlegungen stimmen, dann ist es kein Wunder, dass sowohl von G8 wie auch G9 Befürwortern alle gebundenen und nicht flexiblen Lösung für den Ganzttag deutlich abgelehnt werden. Gebundene Zeitregelungen von Schulen stören die modernen, flexibilisierten Zeitstrukturen der Familien.

Eine ähnliche Antiquiertheitsdiagnose könnte man der Gesamtschule und allen ähnlichen Schulformen stellen. Möglicherweise ist die Gesamtschule nur eine Lösung für eine sehr traditionell organisierte Gesellschaft gewesen, d.h. wie eine Gesellschaft von früher. Sie hat heute natürlich noch ihre Aufgabe, bedient aber in einer pluralisierten Gesellschaft nur spezifische Zielgruppen. Und das ist gut so.

Aber - die Umfrage zeigt, dass die Gesamtschule keineswegs das Wunschmodell aller Eltern ist. Und das kann auf ganz nüchternen Überlegungen basieren.

Mit zunehmender Entwertung akademischer Abschlüsse durch die Akademikerschwemme – im Sommer 2016 wurde in der IAB Studie festgestellt, dass sich manche akademischen Berufe finanziell nicht mehr lohnen – werden die auf den Facharbeitermarkt bezogenen Schulabschlüsse aufgrund der Marktnachfrage wesentlich attraktiver werden. Oder: in einer pluralisierten, heterogener werdenden, Gesellschaft setzt sich auch eine größere Heterogenität der Schulformen durch (d.h. es gibt immer mehr Spezialisierungen oder eine vielfältigere Schullandschaft). Getreu dem Grundsatz „Only variety can destroy

variety“ (Ashby's Law). Oder aber, es muss zu einer Entkopplung zwischen Bildungsabschluss und Einkommen (auch gedanklich) kommen. In einem solchen Fall werden anspruchsvolle Bildungsgänge vermutlich überwiegend von jenen gewählt werden, die sich für Bildung interessieren, allerdings nicht mit dem Hintergedanken, dass sie sich dadurch später finanziell verbessern. Denen bringt dann ein Modell der Gesamtschule nichts besonderes, weil die bislang leitende Philosophie „Aufstieg durch Bildung - Bildungsgerechtigkeit“ ungültig wird, sondern eher „Bildung als Erweiterung und Bereicherung von Arbeit, Beruf und Leben“ sinnvoll wird.

Diese Überlegungen mögen jenen fremdartig erscheinen, die noch immer dem herkömmlichen Modell von Bildungsgerechtigkeit hinterher laufen. Wenn jemand lernt, wie er den Maurer-Beruf optimal ausüben kann, so ist das ebenfalls Bildung und wenn der Markt dem Maurer mehr Geld zahlt als dem Doktor phil., der sich mit der „Gebärde des Weinens in der mittelländischen Literatur“ auseinandergesetzt hat, dann führt ein neuer Bildungsbegriff sowohl zu einem neuen Verständnis von Bildungsgerechtigkeit als auch möglicherweise zur Ablehnung von Schulformen, die eigentlich nur zur Überwindung der Ungerechtigkeit im frühen Industriezeitalter geeignet gewesen waren.

An den Ergebnissen dieser vorliegenden Umfrage ist, bei der Frage nach den Übergangsempfehlungen und der Begründung für den Besuch des Gymnasiums, kein Dünkel erkennbar gewesen. Wenn der Grund „wegen der sozialen Situation“ (ein Gymnasium gegenüber der Gesamtschule gewählt wird) von über 40 % der Befragten angekreuzt wird, so muss dies kein Dünkel sein, sondern kann auch ein Ausdruck der sozialen und räumlichen Aufsplitterung unserer Gesellschaft sein. Dies kann in einer modernen und zukünftigen Gesellschaft allerdings auch eine rein kulturelle Trennung (taste cultures, Milieus) nach dem Prinzip „gleich zu gleich gesellt sich gern“ sein.

**Fazit:** Es lohnt sich, aufgrund der Ergebnisse dieser Umfrage, den Zeitbedarf moderner Familien neu zu erforschen, um ein schulisches Angebot zu bieten, das diesen modernen Anforderungen gerecht wird. Gleichzeitig ist zu überlegen, ob die bisherige Hierarchisierung der Schulformen nicht durch einen egalitären Nebeneinander ersetzt werden kann, und dass vor allem eine Vielfalt von Schulformen der Pluralität moderner

Gesellschaften besser entspricht. In dieser Vielfalt haben selbstredend Gesamtschulen ebenso Platz wie Gymnasien. Aber auch Haupt- und Realschulen.

### **5.5. Die lebensalterlich frühe Entstehung der G8/G9 Meinung bei Eltern**

Ein auffälliges Ergebnis dieser Umfrage „G8 und mehr“ war, dass die Präferenz für G8 oder G9 bereits im frühen Lebensalter bzw. im Grundschulalter der zukünftigen AbiturientInnen fällt. Dafür spricht die große Beteiligung von Grundschulleitern und die ebenfalls stärkere Beteiligung von Eltern der Unter- und Mittelstufe der Gymnasien. Und natürlich deren höhere Präferenz für G9, die in den unteren Klassen des Kindes größer ausfällt als in den oberen.

Die Entstehung der G9 Präferenz im frühen Kindesalter ist weder pathologisch noch einer bildungsbürgerlichen Gesellschaft und ihrer Bildungshysterie geschuldet, sondern es ist ein ganz normaler entwicklungspsychologischer und pädagogischer Vorgang, dass man sich bei den kleinsten Kindern schon fragt, was denn später aus ihnen werden könnte (Beispiel: früh abgeschlossene Ausbildungsversicherungen).

Bedauerlicherweise gab es bei der Einführung von G8 keine interdisziplinäre Beteiligung an der politischen Entscheidung. Für lebensalterrelevante Entscheidungen sind entwicklungspsychologische Kompetenzen ebenso gefragt wie fachdidaktische. Die Bedeutung des lebenslangen Lernens, die Bedeutung der Kompetenzorientierung für die zukünftige Qualität der Menschen sowohl für den Arbeitsmarkt als auch für ihr Leben insgesamt, wurde nicht ausreichend bedacht. Insbesondere die fachdidaktische Kompetenz hätte stärker berücksichtigt werden müssen, weil nur diese für die Verkürzung bzw. Verlängerung von Bildungszeiten relevant ist. Es sind eben nicht nur BildungsökonomInnen für eine solche Entscheidung zu hören, deren Beitrag lediglich einen Nebenaspekt bildungspolitischer Entscheidungen abbilden kann. Deren Versuche, aufgrund bildungsökonomischer Zusammenhänge inhaltliche Schlussfolgerungen zu ziehen, lassen höflicherweise nur den Schluss auf Kompetenzüberschreitung zu.

## 5.6. Die Rolle der Kompetenzorientierung in der G8/G9 Problematik

Die Kompetenzorientierung nordrhein-westfälischer Schulen trifft alle Schulformen. Sie ist einer Übereinkunft der Kultusministerkonferenz geschuldet. Es gäbe auch noch andere didaktische Ansätze, zum Beispiel der früher sehr prominente wissenschaftspropädeutische Ansatz oder der Situationsansatz. Man hätte in empirischen Studien leicht klären können, welcher dieser (und anderer) didaktischen Ansätze günstigere Ergebnisse liefern würde. Das hat man nicht getan. Man wäre dann auf eine komplexe Verschränkung des gewählten didaktischen Ansatzes mit dem Lebensalter gekommen - fraglich, ob die G8 Entscheidung damals dann gefallen wäre.

Die Kompetenzorientierung hat allerdings den Freibrief und das grüne Licht für eine Schulzeitverkürzung vorbereitet. Wie weiter oben schon erläutert, verlangt der kompetenzorientierter Ansatz weniger Wissen, das man sich mit weiterem Zeitbedarf einverleiben müsste. Er beschränkt sich auf die Vermittlung der grundlegenden Fähigkeit, nachdenken und Probleme lösen zu können, wenn die Wissenstatbestände gleich in der Aufgabenstellung mit genannt werden. Diese Fähigkeit nennt man auch „Intelligenz“ und es wundert nicht, dass die PISA-Teste alle eine substantiell hohe signifikante und relevante Korrelation mit Intelligenzquotienten aufweisen (um einen Wert von  $r = 0.80$  herum). Das ist einerseits ein Beweis, dass die Kompetenzorientierung mehr oder weniger so etwas wie Intelligenz ist, andererseits aber auch der Beweis, dass es auf Wissen, Erfahrung und Bildung nicht ankommt. Hier wird, wenn man es stark überzeichnet, Intelligenztraining mit Schule und Bildung verwechselt. Intelligenz und Kompetenz ist nicht akademische Allgemeinbildung.

Das muss nicht schlecht sein – aber es verändert Schule und Gymnasien in einer Form, die kaum noch Akzeptanz bei den Eltern von gymnasialen Kindern findet. In der Frage zur Lösung der G8/G9 Problematik war die Antwortoption „die Rückkehr zu einem Gymnasium wie früher“ enthalten. Dieser Antwort haben sich trotz der altmodischen Formulierung immer noch die Mehrheit der Befragten angeschlossen.

.

Das „wie früher“ – so eine andere Kritik, müsse genau datiert werden. Das hat der Verfasser abgelehnt, weil in den Vorgesprächen deutlich wurde, dass dieses „wie früher“ wegen der laufenden Reformen mit einer Jahreszahl hätte verbunden werden müssen oder mit einer „Erlassgeschichte“ und mit zahllosen Veränderungen, so dass die Umfrage hätte abgetrennt werden müssen, um das Ideal des Gymnasiums „von früher“ (zum Beispiel: vor der reformierten Oberstufe) differenziert zu erfassen. Dies hätte eine erhebliche Verlängerung des Fragebogens bedeutet.

## **5.7. Ein Gesamtfazit: Die Reformfalle der differentiellen Schulzeit bei Gymnasien und Gesamtschulen**

Die Vermischung der Schulzeitfrage mit der Institutionsfrage (Gymnasium G8 - Gesamtschule G9) war und ist das politische Problem: potentielle und reale Gymnasialeltern wollen ihr Kind auf das Gymnasium und nicht auf die Gesamtschule schicken. Und nun wird durch die Zeitverknappung das G8-Gymnasium nur für die weniger belasteten SchülerInnen reserviert, d.h. für jene, die auch in weniger Zeit viel lernen können.

Wenn diese Absicht offen und deutlich kommuniziert wird, so ist das eine politische Maßnahme, die offen diskutiert werden kann. Wenn es nur eine klammheimliche, bauernschlaue und trickreiche Maßnahme war, dann hätte man damit rechnen können, ja müssen, dass eine solche Regelung umstritten ist und Widerstand erzeugt. Die Absicht - mit einer differentiellen Schuldauer die Gesamtschule zu stärken - ist von den Befragten klar erkannt worden. Und möglicherweise regen sich deshalb so viele auf. Vielleicht hätte G8 mehr Chancen, wenn auch die Gesamtschulen G8 einführen müssten. Das müsste aber erst empirisch untersucht werden.

Die Alternative „Gesamtschule mit 9 Jahren“ - scheint zu einer Trotzreaktion zu führen, weil Eltern nicht einsehen, dass sie diese Schulform wählen müssen, wenn ihr Kind zum Lernen länger brauchen sollte.

Die Zeit, die man für den Erwerb von Bildung braucht, ist individuell unterschiedlich. Sie variiert auch über einen diskontinuierlichen Entwicklungsweg der Kinder: mal gibt es Phasen schneller Veränderungen - mal Phasen der Stagnation. In dieser Situation ist Zeit im Bildungssystem ein Diskriminierungsfaktor. Wer keinen linearen, ansteigenden Entwicklungsweg nimmt, wer länger braucht, wird durch eine Verkürzung der Gymnasialzeit diskriminiert.

Zeitregelungen im Schulsystem gehen offenbar immer noch von einem „nine to five“ job der Eltern aus. Viele wundern sich über die Ablehnung von verpflichtenden Ganztagsangeboten - die offenbar nicht aus Konservativität geschieht - sondern weil die mittlerweile schnell fortschreitende Flexibilisierung der Arbeitswelt keine verpflichtenden Eckpfeiler benötigt. Der gebundene Ganzttag kann die zeitflexible Arbeits- und Familienwelt derangieren.



## **6. Anmerkung zur Auswertung und Darstellung der Ergebnisse im Abschlußbericht:**

Das verwendete Rechenprogramm ist meist SPSS. Explorativ wurden auch Routinen aus Aabel benutzt. Die grafischen Darstellungen wurden mit Deltagraph erzeugt.

Weitere Auswertungen der Daten sind in Vorbereitung.

Die Fa. Unipark hat den SPSS Datensatz zu den beiden Onlinebefragungen automatisch geliefert.

Die per Hand eingegebenen Daten der Zufallsstichprobe wurden alle einer mehrfachen Plausibilitätsprüfung unterzogen.

Signifikanztestungen sind meistens unterblieben, da wegen der Größe der Stichproben auch kleine und kleinste Unterschiede signifikant werden. Es darf daran erinnert werden, dass Signifikanzen kein Ersatz für Effektstärken sind, sondern lediglich die Wahrscheinlichkeit der Thesen auf Nichtzufälligkeit testen.

Die Diagramme in den Darstellungen des Abschlußberichtes sind aus Gründen der Anschaulichkeit verändert worden: es sind nicht nur die Mittelwerte angegeben worden (das wären ja bloß Punkte), sondern die Mittelwerte sind auch noch miteinander durch einen Linienzug verbunden worden. Das darf nicht missverstanden werden als eine inhaltliche Verbindung zwischen den einzelnen Mittelwerten, sondern die Verbindungen dienen nur der Verdeutlichung von Mittelwertsunterschieden in unterschiedlichen Stichproben (z.B. G8 vs. G9 Anhänger bzw. zwischen verschiedenen Antwortvorgaben). So sieht man die Unterschiede schneller. Eine weitere Veränderung besteht darin, dass die Reihenfolge der Items nach Größe des Mittelwertes sortiert worden ist (oben die guten Noten - unten die schlechten bzw. gelegentlich umgekehrt) und nicht in derselben Reihenfolge wie im Fragebogen. Hätte man die Mittelwerte in derselben Reihenfolge wie im

Fragebogen dargestellt, würden die Linien im Zickzack verlaufen und die Übersichtlichkeit der Darstellung würde leiden. Deshalb wurde auf die originale Abfolge der Items im Fragebogen bei der Darstellung der Ergebnisse verzichtet. Die Darstellung der Mittelwerte stellt also eine Vereinfachung dar, die mehr Übersichtlichkeit über die Ergebnisse liefert.

Steinhagen, im August 2016

Anschrift des Verfassers

rdollase@gmx.de