



# **THEO-WEB**

**Zeitschrift für Religionspädagogik**

**Academic Journal of Religious Education**

**16. Jahrgang 2017, Heft 2**

**ISSN 1863-0502**

**Thema: „Migration, Religion und Bildung.**

**Wege zu einer migrationssensiblen Religionspädagogik“**

Hermisson, S. (2017). Schülereinstellungen zu religiöser und weltanschaulicher Diversität – Entwicklung eines Erhebungsinstruments auf der Grundlage der Item-Response-Theorie. *Theo-Web*, 16(2), 75–92.

DOI: <https://doi.org/10.23770/two25>



Dieses Werk ist unter einer Creative Commons Lizenz vom Typ Namensnennung-Nichtkommerziell 4.0 International zugänglich. Um eine Kopie dieser Lizenz einzusehen, konsultieren Sie <http://creativecommons.org/licenses/by-nc/4.0/> oder wenden Sie sich brieflich an Creative Commons, Postfach 1866, Mountain View, California, 94042, USA.

# **Schülereinstellungen zu religiöser und weltanschaulicher Diversität – Entwicklung eines Erhebungsinstrumentes auf der Grundlage der Item-Response-Theorie**

VON  
Sabine Hermisson

## *Abstract*

*Der Beitrag beschreibt die Entwicklung eines Testinstruments, um Schülereinstellungen zu religiöser Diversität zu erheben. Die Konstrukt- und Testentwicklung folgt dem Ansatz des Construct Modelings. Das Erhebungsinstrument fokussiert auf Einstellungen gegenüber Muslimen, Christen, Juden und Nichtreligiösen unter Verwendung einer Skala unterschiedlicher Grade sozialer Proximität. In einer Pilotstudie wurden die Antworten von 114 Befragten erhoben und mittel der Item Response Theorie analysiert. Diskutiert werden die Kalibrierung des Testinstruments, Evidenz für seine Reliabilität und Validität sowie vorläufige Ergebnisse.*

*Stichworte: Religiöse Diversität – Einstellungen – Schüler/innen – Construct Modeling – Item Response Theory*

*Keywords: religious diversity – attitudes – students – Construct Modeling – Item Response Theory*

Krieg, Terror und wirtschaftliche Not veranlassen Millionen von Menschen in Afrika und dem Nahen Osten, ihre Heimat zu verlassen und in Nachbarstaaten, aber auch in Europa Zuflucht zu suchen. Die aktuellen Migrationsbewegungen stellen Europa vor neue Herausforderungen, insbesondere mit Blick auf den Umgang mit kultureller und religiöser Diversität. Während in Deutschland und Österreich im Jahr 2015 Geflüchtete, insbesondere aus Syrien, vielfach mit großer Offenheit willkommen geheißen wurden, brachte die ‚Flüchtlingskrise‘ auch antimuslimische Ressentiments an die Oberfläche sowie die Sorge über antijüdische Vorurteile unter den Migrant/innen.

Für die Integration jugendlicher Geflüchteter kommt den Schulen eine Schlüsselrolle zu. In diesem Zusammenhang gewinnen interkulturelles und interreligiöses Lernen an Bedeutung und fordern religionspädagogische Forschung, Reflexion und Praxis heraus. Ein Desiderat sind verlässliche empirische Daten zu den Einstellungen von Schülerinnen und Schülern in Bezug auf religiöse Diversität. Mit welcher Offenheit und mit welchen Vorbehalten begegnen sie Menschen unterschiedlicher religiöser und weltanschaulicher Prägung? Was akzeptieren sie, und wo gibt es Hürden? Dieser Beitrag beschreibt die Entwicklung eines Testinstruments, das darauf abzielt, dies zu erheben. Der fundierte Einblick in die Einstellungen der Schüler/innen stellt eine Grundlage für gezielte religionspädagogische Interventionen dar.

## **1 Begriffsklärungen, Fokus und Theorierahmen**

### ***1.1 Begriffsklärungen und Fokus der Studie***

Ziel des Testinstruments ist es, Schülereinstellungen zu religiöser Diversität reliabel und valide zu erfassen. Dafür ist zunächst zu klären wie der Terminus ‚Einstellungen zu religiöser Diversität‘ in diesem Projekt verstanden wird.

*Einstellungen* sind ein zentrales Forschungsgebiet der Sozialpsychologie, die diese als latente Konstrukte versteht, also als Sachverhalte, die nicht unmittelbar beobachtbar sind. Etabliert hat sich ein mehrdimensionales Konstrukt von Einstellungen, das die drei Dimensionen Kognition, Affektion und Verhalten unterscheidet (Zanna & Rempel, 1988; Eagly & Chaiken, 1993). Die hier beschriebene Studie fokussiert primär auf die *affektive* Dimension, der im Zusammenhang mit der Frage, wie Schüler/innen Menschen anderer Religiosität begegnen, eine wichtige Rolle zukommt. Hinzu kommt, dass beim Thema ‚religiöse Diversität‘ der Faktor soziale Erwünschtheit hoch ist. Um ihn begrenzt zu halten, sollten nicht sozial erwünschte Handlungen erhoben werden, sondern vielmehr die Affekte der Schüler/innen.

‚*Religiöse Diversität*‘ umfasst prinzipiell alle Weltreligionen und darüber hinaus auch nicht-institutionalisierte Formen von Religiosität. Dieses Projekt fokussiert auf Einstellungen gegenüber Muslimen, Christen, Juden sowie Nichtreligiösen. Dieser Fokus wurde gewählt, da diese vier Gruppen im europäischen Kontext wichtige Akteure, aber auch in der medialen Diskussion präsent sind. ‚Religiöse Diversität‘ ist daher im Folgenden im umfassenden, auch nichtreligiöse Weltanschauungen einschließenden Sinn zu verstehen.

Was den *Inhaltsbereich* von Einstellungen gegenüber religiöser Diversität betrifft, so liegt auf der Hand, dass das interessierende Konstrukt eine Vielfalt möglicher Aspekte umfassen kann. Dies ergibt sich zum einen aus der oben beschriebenen Mehrdimensionalität von Einstellungen und zum anderen aus der Vielfalt an relevanten Inhaltsbereichen. Dieses Projekt konzentriert sich auf die Akzeptanz sozialer Nähe, also die Frage: Mit Menschen welcher religiöser und weltanschaulicher Orientierung sind Schüler/innen bereit zu interagieren – sei es als Mitbürger, als Klassenkameradin oder als Freund oder Freundin? Was empfinden sie als schwierig, und wo sehen sie Grenzen? Im Kontext von Migration sind dies zentrale Fragestellungen und ein klassischer Topos, sodass an frühere Studien angeknüpft werden konnte.<sup>1</sup>

## **1.2 Construct Modeling nach Wilson**

Das methodische Vorgehen der Konstrukt- und Testentwicklung folgt dem von Wilson (2005) vorgeschlagenen Construct Modeling. Das Construct Modeling ist ein messtheoretisch fundierter Ansatz, der am *Berkeley Evaluation and Assessment Research (BEAR) Center* entwickelt wurde, um latente Konstrukte, insbesondere im Kontext der Bildungswissenschaften, zu erheben. Es umfasst vier ineinander verschränkte und sich gegenseitig bedingende Bausteine: *Construct Map*, *Item Design*, *Outcome Space* und *Measurement Model*. Wilsons Ansatz entstand im Kontext von Arbeiten mit der Item Response Theorie und findet vor allem in entsprechenden Studien Verwendung (u.a. Duckor, Draney & Wilson, 2009; Brown & Wilson, 2011; Morell u.a., 2017). Prinzipiell ist das theoretische Konzept aber auch für qualitative Studien gültig.

---

<sup>1</sup>

Siehe dazu unten die Abschnitte 2.1 sowie 2.2.

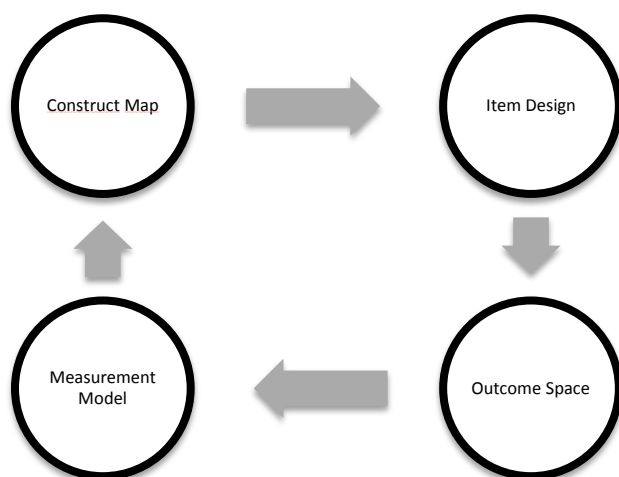


Abb.1 Construct Modeling nach Wilson (2005)

*Construct Map:* Ausgangspunkt des Construct Modeling ist eine Construct Map, die das zu erhebende latente Konstrukt beschreibt und in seinen unterschiedlichen Niveaus präzisiert.

*Item Design:* Auf der Grundlage der Construct Map werden Items entwickelt, die darauf abzielen, empirische Manifestationen des theoretischen Konstrukts zu erheben. Dies kann Testitems mit vorgegebenen Antwortoptionen (beispielsweise Likert-Skalen) umfassen, aber auch offene Fragen in qualitativen Interviews, Essays etc.

*Outcome Space:* Um von den Antworten der Befragten auf die jeweiligen Niveaus der Construct Map zurückschließen zu können, wird zunächst der Outcome Space festgelegt. Im Falle quantitativer Erhebungen werden zudem Punktwerte (Scores) zugeordnet.

*Measurement Model:* Schließlich kommen psychometrische Modelle aus der klassischen Testtheorie oder der Item Response Theorie zum Einsatz, um die Scores der Befragten auf die Niveaus der Construct Map zurückzubeziehen.

Das von Wilson vorgeschlagene Construct Modeling ist ein zirkuläres Verfahren und wird in der Entwicklung eines Testinstruments oft mehrfach durchlaufen. Wird etwa die anfangs theoretisch formulierte Construct Map durch die empirischen Daten, die im Verlauf der Konstrukt- und Testentwicklung gewonnen werden, nicht bestätigt, ist sie entsprechend zu korrigieren. Dies war auch bei der Entwicklung des Erhebungsinstruments der Fall, die im Folgenden beschrieben wird.

## 2 Theoretische Entwicklung des Erhebungsinstruments

### 2.1 Construct Map

Da das hier beschriebene Projekt auf die Akzeptanz sozialer Nähe fokussiert, galt es in einem ersten Schritt, unterschiedliche qualitative Niveaus der Akzeptanz sozialer Nähe zu systematisieren. Entwickelt wurde zunächst die folgende Construct Map, die sechs Stufen unterscheidet. Diese reichen von geringerer Akzeptanz zu höherer Akzeptanz respektive von Sozialbeziehungen, die ‚einfacher‘ zu akzeptieren sind, hin zu solchen, die ‚schwieriger‘ zu akzeptieren sind:

Tab. 1 Construct Map für die Akzeptanz religiöser Diversität

Niveau	Construct Map
6	Akzeptanz religiöser Diversität auf der Ebene <b>partnerschaftlicher Beziehungen</b>
5	Akzeptanz religiöser Diversität auf der Ebene <b>familiärer Beziehungen (z.B. Stiefschwester, Schwager)</b>
4	Akzeptanz religiöser Diversität auf der Ebene von <b>Freundschaftsbeziehungen</b>
3	Akzeptanz religiöser Diversität auf der Ebene <b>alltäglicher Sozialkontakte (z.B. Nachbar, Klassenkameradin)</b>
2	Akzeptanz religiöser Diversität auf der Ebene von <b>Mitbürgerschaft</b>
1	<b>Keine</b> Akzeptanz religiöser Diversität auf <b>jedwelcher Ebene sozialer Beziehungen</b>

Die hier entwickelte Construct Map ist eine Variante der in der Soziologie etablierten Bogardus-Skala oder Sozialdistanz-Skala, die bereits 1924 von dem amerikanischen Soziologen E.S. Bogardus konzipiert wurde, um Einstellungen gegenüber Einwanderern unterschiedlicher ethnischer Gruppen in den USA zu erheben (Wark & Galliher, 2007). Die Bogardus-Skala wurde für die Befragung erwachsener Männer entworfen und spiegelt die gesellschaftlichen und familiären Strukturen Anfang des 20. Jahrhunderts wider.<sup>2</sup> Im Unterschied dazu geht die in dieser Studie entwickelte Construct Map von postmodernen Familienstrukturen einschließlich Patchworkfamilien aus und hat vorrangig, wenn auch nicht notwendigerweise ausschließlich, Jugendliche im Blick. Die Bogardus-Skala setzt eine entscheidende Annahme voraus. Sie setzt voraus, dass die Bejahung eines engeren Grads an Intimität impliziert, dass alle vorausgehenden, d.h. weniger engen sozialen Beziehungen ebenfalls bejaht werden.<sup>3</sup> Dies wurde auch in der hier entwickelten Construct Map vorausgesetzt.

In der späteren empirischen Überprüfung bestätigte sich diese Annahme weitgehend, jedoch mit einer Ausnahme. Während ein Teil der Befragten die Akzeptanz von Freundschaftsbeziehungen als ‚einfacher‘ empfanden als von familiären Beziehungen, vertraten andere die umgekehrte Auffassung. Damit ließ sich weder die Hierarchisierung Familie über Freunde noch umgekehrt durchgängig empirisch bestätigen. Die empirisch revidierte Construct Map löste dieses Problem, indem sie zwischen Menschen, die entweder familiäre Beziehungen oder Freundschaften akzeptieren (‚einfacher‘), und solchen, die beides akzeptieren (‚schwieriger‘), differenzierte:

<sup>2</sup> Befragt wurden erwachsene Männer, für die vorausgesetzt wurde, dass sie Mitglieder in Clubs sind und über die Einheirat in ihre Familien (mit-)bestimmen. Die Befragten wurden gebeten anzugeben, bis zu welchem Grad an sozialer Nähe sie bereit seien, mit einem typischen Mitglied einer bestimmten ethnischen Gruppierung wie Iren, Deutschen oder Polen Beziehungen einzugehen. Dazu wurden ihnen die folgenden Abstufungen vorgeschlagen: citizenship in my country, employment in my occupation, to my street as neighbors, to my club as personal chums, to close kinship by marriage (Wark & Galliher, 2007).

<sup>3</sup> Damit ist die Bogardus-Skala ein Idealtyp einer nach dem Sozialforscher Louis Guttman benannten sog. Guttman-Skala. Unabhängig vom jeweiligen thematischen Kontext sind Guttman-Skalen dadurch charakterisiert, dass die Bejahung oder – in kognitiven Testkontexten – das Erreichen eines jeden höheren Niveaus alle niedrigeren umfasst (Guttman, 1944).

Tab. 2 Revision der Construct Map

Niveau	Hypothetisch entwickelte Construct Map	Empirisch revidierte Construct Map
6	Akzeptanz religiöser Diversität auf der Ebene <b>partnerschaftlicher Beziehungen</b>	Akzeptanz religiöser Diversität auf der Ebene <b>partnerschaftlicher Beziehungen</b>
5	Akzeptanz religiöser Diversität auf der Ebene <b>familiärer Beziehungen</b>	Akzeptanz religiöser Diversität auf der Ebene <b>familiärer und freundschaftlicher Beziehungen</b>
4	Akzeptanz religiöser Diversität auf der Ebene freundschaftlicher Beziehungen	Akzeptanz religiöser Diversität auf der Ebene <b>familiärer oder freundschaftlicher Beziehungen</b>
3	Akzeptanz religiöser Diversität auf der Ebene <b>alltäglicher Sozialkontakte</b>	Akzeptanz religiöser Diversität auf der Ebene <b>alltäglicher Sozialkontakte</b>
2	Akzeptanz religiöser Diversität auf der Ebene von <b>Mitbürgerschaft</b>	Akzeptanz religiöser Diversität auf der Ebene von <b>Mitbürgerschaft</b>
1	<b>Keine</b> Akzeptanz religiöser Diversität auf <b>jedwelcher Ebene sozialer Beziehungen</b>	<b>Keine</b> Akzeptanz religiöser Diversität auf <b>jedwelcher Ebene sozialer Beziehungen</b>

## 2.2 Item Design

Im Kontext religiöser Diversität wurden bereits in früheren Erhebungen Items zur Akzeptanz sozialer Nähe verwendet, sodass hier auf Bestehendem aufgebaut werden konnte. Der Religionsmonitor *Muslims in Europa* (2017) umfasst ein Item, das danach fragt, welche Personengruppen (Muslime, Ausländer, Christen, Juden, Homosexuelle etc.) die Befragten nicht gerne als Nachbarn hätten.<sup>4</sup> Die in Großbritannien durchgeführte Erhebung von Francis, Penny und McKenna (2016) verwendet Likert-Items mit Statements wie „I would be happy about a close relative marrying someone from a different faith“.<sup>5</sup>

In einer qualitativen Vorstudie im Verlauf der Entwicklung des hier beschriebenen Erhebungsinstruments wurden Interviews mit Schüler/innen geführt. Ihnen wurden die Statements des Fragebogens von Francis, Penny und McKenna (2016) vorgelegt mit der Bitte, zu diesen Stellung zu nehmen. Dabei erwies sich eindrücklich, dass eine maximale Konkretheit der zu entwickelnden Items entscheidend für die Reliabilität und Validität des Fragebogens sein würde. Dies spiegelt sich exemplarisch in der folgenden Interviewpassage mit einer 14-jährigen Schülerin wider:

<sup>4</sup> Das Item lautet präzise: „Ich werde Ihnen eine Reihe verschiedener Personengruppen vorlesen. Bitte sagen Sie mir jeweils, welche Sie nicht gerne als Nachbarn hätten bzw. ob es Ihnen egal ist.“ Gefragt wurde in Bezug auf Flüchtlinge, Muslime, Leute mit vielen Kindern, Ausländer/Gastarbeiter, Homosexuelle, Menschen, die eine andere Sprache sprechen, Juden, Menschen anderer Hautfarbe, Atheisten sowie Christen.

<sup>5</sup> Mit den Antwortoptionen: strongly agree, agree, not certain, disagree, disagree strongly.

I: “All religious groups must have equal rights.” What do you think?

R: I’m not sure. If “religious groups” means religions like Islam or Christianity or so, I would say “strongly agree”. But if it also means like the Flying Spaghetti Monster, then I’d rather choose “not certain”. [...]

I: “I would be happy about a close relative marrying someone from a different faith.” How do you feel about this?

R: If my sister wants to marry a Muslim and he is open, I’d choose “strongly agree”. But if he wants her to wear a head scarf, I don’t know. I wouldn’t say “I’m happy”. So, I’d say “agree”.

An diesem Interviewauszug wird deutlich: Wie die Schülerin das Item beantwortet, hängt nicht nur von ihren prinzipiellen Einstellungen zu religiöser Diversität ab, sondern auch davon, wie sie die Frage interpretiert, insbesondere wie sie die Formulierungen „religion“ und „someone from a different faith“ füllt (im Interview: Weltreligionen oder Flying Spaghetti Monster respektive offener oder restriktiver Muslim).

Der Beobachtung, dass die Konkretion der Items entscheidend für die Reliabilität und Validität des Fragebogens sein würde, wurde in der weiteren Itementwicklung Rechnung getragen. Im Vergleich zum Item des Religionsmonitors, das danach fragt, ob Muslime als Nachbarn akzeptiert werden, entfalten die hier entwickelten Items sowohl (a) den Terminus „Muslim“ als auch (b) den Terminus „Nachbar“.

(a) Für die vier Gruppen (Muslime, Christen, Juden und Nichtreligiöse) wurde ein Spektrum von je 12 (für Nichtreligiöse 11) konkreten Profilen entwickelt. Für diese Profile ist festzuhalten:

- Sie sind von unterschiedlichem ‚Schweregrad‘ und reichen von nicht praktizierenden Mitgliedern einer bestimmten Religionsgemeinschaft bis hin zu solchen, die auf der Grundlage ihrer Überzeugung religiöse Freiheit ablehnen.
- Innerhalb jeder Gruppe wurden die Profile nach ihrem vermuteten ‚Schweregrad‘ geordnet.
- Die vier Itemgruppen sind so parallel wie möglich konstruiert. Oft unterscheiden sich christliche, muslimische etc. Items nur durch die spezifische Religionszugehörigkeit der Person.<sup>6</sup> Aber auch bei Items, bei denen der inhaltliche Bezug religionspezifische Variationen erforderte, wurde maximale Parallelität intendiert.<sup>7</sup>

---

<sup>6</sup> So beispielsweise die vier parallelen Items 4 (A Muslim/Christian/Jew/Non-believer who in discussions brings up the topic of religion and worldview.)

<sup>7</sup> So beispielsweise die vier parallelen Items 6 (A Muslim who wears religious clothing in public (e.g. the female head scarf called hijab)./A Christian who wears religious clothing in public (e.g. a T-shirt with the message ‘Jesus saves’)./A Jew who wears religious clothing in public (e.g. the head covering for males called kippah)./A Non-believer who wears symbols or clothing displaying non-belief messages (e.g. a T-shirt with the message ‘Atheist’)

Tab. 3 Item Design: Profilspektrum

	Islam	Christentum	Judentum	Nichtreligiöse
<b>1</b>	A Muslim who is not practicing her/his religion	A Christian who is not practicing her/his religion	A Jew who is not practicing her/his religion	A Non-believer who doesn't bother about religion
<b>2</b>	A Muslim who practices his/her religion only in private	A Christian who practices his/her religion only in private	A Jew who practices his/her religion only in private	A Non-believer who keeps his worldviews private
<b>3</b>	A Muslim who sometimes shares publicly about his/her religious practice (e.g. praying, fasting during Ramadan)	A Christian who sometimes shares publicly about his/her religious practice (e.g. praying, fasting during Lent)	A Jew who sometimes shares publicly about his/her religious practice (e.g. praying, fasting on the Day of Atonement called Yom Kippur)	A Non-believer who sometimes shares publicly about his/her worldviews
<b>4</b>	A Muslim who in discussions brings up the topic of religion and worldview	A Christian who in discussions brings up the topic of religion and worldview	A Jew who in discussions brings up the topic of religion and worldview	A Non-believer in discussions brings up the topic of religion and worldview
<b>5</b>	A Muslim who invites me to celebrate the end of Ramadan with his/her family	A Christian who invites me to celebrate Christmas with his/her family	A Jew who invites me to celebrate Passover with his/her family	A Non-believer who invites me to celebrate a non-believers' naming ceremony with his/her family
<b>6</b>	A Muslim who wears religious clothing in public (e.g. the female head scarf called hijab)	A Christian who wears religious clothing in public (e.g. a T-shirt with the message 'Jesus saves')	A Jew who wears religious clothing in public (e.g. the head covering for males called kippah)	A Non-believer who wears symbols or clothing displaying non-belief messages (e.g. a T-shirt with the message 'Atheist')
<b>7</b>	A Muslim who interrupts our shared activities (e.g. having a meal together) for a religious activity (e.g. prayer)	A Christian who interrupts our shared activities (e.g. having a meal together) for a religious activity (e.g. prayer)	A Jew who interrupts our shared activities (e.g. having a meal together) for a religious activity (e.g. prayer)	-
<b>8</b>	A Muslim who in political and ethical debates (e.g. human rights, war and peace) argues from his/her standpoint as a Muslim	A Christian who in political and ethical debates (e.g. human rights, war and peace) argues from his/her standpoint as a Christian	A Jew who in political and ethical debates (e.g. human rights, war and peace) argues from his/her standpoint as a Jew	A Non-believer who in political and ethical debates (e.g. human rights, war and peace) argues from his/her standpoint as a Non-believer
<b>9</b>	A Muslim who believes that (his/her version of) Islam is the only true religion	A Christian who believes that (his/her version of) Christianity is the only true religion	A Jew who believes that (his/her version of) Judaism is the only true religion	A Non-believer who believes that (his/her version of) Non-belief is the only sensible worldview
<b>10</b>	A Muslim who repeatedly argues that my faith or worldview is wrong	A Christian who repeatedly argues that my faith or worldview is wrong	A Jew who repeatedly argues that my faith or worldview is wrong	A Non-believer who repeatedly argues that my faith or worldview is wrong
<b>11</b>	A Muslim who tries to convert me to (his/her version of) Islam	A Christian who tries to convert me to (his/her version of) Christianity	A Jew who tries to convert me to (his/her version of) Judaism	A Non-believer who tries to convert me to (his/her version of) Non-belief
<b>12</b>	A Muslim who on the basis of her/his faith rejects values such as religious freedom	A Christian who on the basis of her/his faith rejects values such as religious freedom	A Jew who on the basis of her/his faith rejects values such as religious freedom	A Non-believer who on the basis of her/his worldview rejects values such as religious freedom

(b) Neben nachbarschaftlichen Beziehungen wie im Religionsmonitor wurden auch die weiteren in der Construct Map systematisierten soziale Beziehungen erhoben: mitbürgerschaftliche, familiäre, freundschaftliche sowie partnerschaftliche Beziehungen.



### 2.3 Outcome Space

Im hier beschriebenen Projekt fallen Construct Map und Outcome Space weitgehend zusammen. Auch das Scoring der Items profitiert von dem oben beschriebenen systematisch strukturierten und empirisch revidierten Aufbau der Construct Map. Den jeweiligen Niveaus der Construct Map wurden Scores von 0 bis 5 zugeordnet.

Tab. 4 Outcome Space und Scoring

Niveau	Empirisch revidierte Construct Map	Outcome Space	Score
6	Akzeptanz religiöser Diversität auf der Ebene <b>partnerschaftlicher Beziehungen</b>	<b>(future) spouse or partner</b>	5
5	Akzeptanz religiöser Diversität auf der Ebene <b>familiärer und freundschaftlicher Beziehungen</b>	Wahl der beiden Optionen: <b>immediate family member</b> <i>und</i> <b>close friend</b>	4
4	Akzeptanz religiöser Diversität auf der Ebene <b>familiärer oder freundschaftlicher Beziehungen</b>	Wahl nur einer der beiden Optionen: <b>immediate family member</b> <b>und</b> <b>close friend</b>	3
3	Akzeptanz religiöser Diversität auf der Ebene <b>alltäglicher Sozialkontakte</b>	<b>neighbor or classmate</b>	2
2	Akzeptanz religiöser Diversität auf der Ebene von <b>Mitbürgerschaft</b>	<b>fellow citizen</b>	1
1	<b>Keine</b> Akzeptanz religiöser Diversität auf <b>jedwelcher Ebene sozialer Beziehungen</b>	<b>none of the above</b>	0

### 2.4 Der Fragebogen

Die vier Profilgruppen Muslime, Christen, Juden und Nichtreligiöse wurden den Befragten nicht parallel, sondern nacheinander vorgelegt. Für jedes einzelne Profil wurden die Befragten gebeten anzugeben, ob sie sich mit der Vorstellung, einen Menschen des betreffenden religiösen oder weltanschaulichen Profils als Mitbürger/in, Nachbar/in etc. zu haben, wohlfühlten. Die präzise Formulierung lautete: „I feel comfortable having the person as a ...“ mit den Antwortoptionen fellow citizen/neighbor or classmate/close friend/immediate family member (e.g. step sister, brother-in-law)/(future) spouse or partner/none of the above. Die Eingangsformulierung „I feel comfortable“ wurde gewählt, um die affektive Dimension der Schülereinstellungen zu ermitteln. Wie sich das im Fragebogen konkret gestaltet, zeigt der Auszug in Abbildung 2.

Each set of items below presents a different profile of a **Christian**.  
For each profile, please check **all** the answers that fit for you.

You can check 0 to 5 boxes for each profile.

**Or** you can check "none of the above".

1. I feel comfortable having a **Christian** who is **not practicing his/her religion** (e.g. doesn't pray, doesn't go to church) as a

- fellow citizen
- neighbor or classmate
- close friend
- immediate family member (e.g. stepsister, brother-in-law)
- (future) spouse or partner
- none of the above.

Abb. 2 Auszug aus dem Fragebogen

Darüber hinaus wurden als externe Variablen erhoben: Alter, Gender, Religionszugehörigkeit sowie Häufigkeit des Kontakts mit Muslimen, Christen, Juden sowie Nichtreligiösen.

In der Modellierung der Daten galt es, die dichte und komplexe Struktur des Erhebungsinstruments in ein angemessenes statistisches Modell zu übersetzen. Ausschlaggebend für die endgültige Entscheidung für ein bestimmtes Modell ist jedoch der Vergleich einer Reihe von Modellen auf den besten Modell-Fit hin. Dazu bedurfte es empirischer Daten.

### **3 Evaluierung des Erhebungsinstruments aufgrund der empirischen Daten**

#### ***3.1 Pilotdaten, Measurement Model und Item Response Theorie***

Es wurde eine Pilotstudie durchgeführt mit 114 Befragten unterschiedlichen Alters ab 13 Jahren. Die Stichprobe umfasste: Muslim/innen (n=4), Christ/innen (n=48), Angehörige anderer Religionen (n=8) sowie Religionslose (n=54). Die Daten wurden mittels eines Online-Fragebogens erhoben. Ziel der Pilotierung war zum einen, die Construct Map empirisch zu überprüfen und den Fragebogen zu testen, insbesondere in Bezug auf seine Kalibrierung, Reliabilität und Validität. Zum anderen ging es darum, auf der Grundlage der Pilotdaten zwar keine validen Ergebnisse zu erzielen, aber doch Hypothesen zu generieren, die in späteren umfassenderen Erhebungen zu testen sind.

Wie in Abschnitt 1.2 diskutiert, geht es bei dem 4. Baustein des Construct Modelings, dem Measurement Model, darum, mit Hilfe eines psychometrischen Modells, sei es aus der Klassischen Testtheorie oder der Item Response Theorie, die Scores der Befragten (hier: für 47 Items, je mit Scores von 0 bis 5, also max. 235 Punkte) auf die unterschiedlichen Niveaus der Construct Map zurückzubeziehen.

Die Klassische Testtheorie, auch ‚True Score Testtheorie‘, geht dabei von der Annahme aus, dass sich das erzielte Ergebnis einer Testperson zusammensetzt aus ihrem ‚wahren‘ Ergebnis (‚true score‘) und einem Messfehler, der zu berücksichtigen ist. Im Unterschied dazu zielt die Item Response Theorie, auch ‚Probabilistische Testtheorie‘, darauf ab, die Fähigkeit einer Testperson (person ability) in Abhängigkeit von

der Schwierigkeit der einzelnen Items (item difficulty) zu bestimmen. Ihr einfachstes Modell ist das Rasch-Modell, das in der folgenden Gleichung notiert wird:

(1.)

---

Dabei bezieht sich  $\delta_i$  auf die Schwierigkeit eines Items (item difficulty) und  $\theta$  auf die Fähigkeit einer Person (person ability). Im Kontext dieser Erhebung ist freilich der Terminus ‚Fähigkeit‘ im generischen Sinn zu verstehen und meint die Akzeptanzbereitschaft einer Person in Bezug auf soziale Beziehungen. Wäre die Schwierigkeit eines bestimmten Items  $\delta_i$  und die Fähigkeit, also Akzeptanzbereitschaft, einer Person  $\theta$  bekannt, dann gäbe die Formel (1.) die Wahrscheinlichkeit dafür an, dass diese Person für ein konkretes Item (Profil) eine bestimmte soziale Nähe (Nachbar, Freund etc.) bejaht. Aber beides ist nicht bekannt; jedoch kann diese Wahrscheinlichkeit, wenn empirische Daten vorliegen, aus den Antworten auf die Items geschätzt werden (rechte Seite der Formel). Damit können aus dem Formelzusammenhang auch die Werte für  $\delta_i$  und  $\theta$  geschätzt werden (linke Seite der Formel). Je größer die Stichprobe ist, desto genauer ist diese Schätzung.

Aufbauend auf dem Rasch-Modell liegt eine Reihe weiterer Modelle vor, die gegebenenfalls die Struktur des Testinstruments besser berücksichtigten. In dieser Studie betrifft dies beispielsweise die polytome Struktur der Items mit 6 Antwortmöglichkeiten und die Parallelität der vier Item-Gruppen. Welches Modell die jeweiligen Daten am besten modelliert, wird beurteilt, indem der Fit, d.h. die Anpassungsgüte einer Reihe von Modellen verglichen wird (De Boeck & Wilson, 2004). Modelliert wurden die Daten in dieser Studie schließlich mit einem konsekutiven eindimensionalen Rating Scale-Modell, einer Weiterentwicklung des Rasch-Modells (Andrich, 1978). Im Modell wurde die Religionszugehörigkeit der Befragten mittels latenter Regression berücksichtigt.<sup>8</sup> Die Analyse der Daten erfolgte mit Hilfe des Softwareprogramm ConQuest (Wu, Adams, Wilson & Heldane, 2007).

Aufgrund des verwendeten konsekutiven Modells wurde jede der vier Skalen einzeln in Bezug auf ihre Kalibrierung und Reliabilität evaluiert. Die folgende Darstellung berichtet umfassend die Ergebnisse für die Skala zur Akzeptanz von Juden und Jüdinnen. Aufgrund der Parallelität der vier Skalen fallen die Ergebnisse zu den übrigen Skalen ähnlich aus.

### **3.2 Testkalibrierung**

Die unten abgebildete Wright Map gibt auf der y-Achse die ‚person ability‘ auf einer sog. Logit-Skala an. Hohe Werte bedeuten in dieser Studie hohe Akzeptanz für soziale Beziehungen mit Juden und Jüdinnen. Die Kurve ganz links weist die aus den Pilotdaten geschätzte Akzeptanzbereitschaft der 114 Befragten aus. Alle übrigen Angaben erschließen, als wie ‚schwer‘ jedes der 12 jüdischen Items (Profile) empfunden wurde (item difficulty). Die unterste Zeile listet auf der x-Achse die 12 Items auf. Die vertika-

---

<sup>8</sup> Das konsekutive eindimensionale Modell, das die vier Itemgruppen (muslimisch, christlich, jüdisch und nichtreligiös) als individuelle Skalen modelliert, wies einen besseren Modell-Fit auf als ein zusammengesetztes eindimensionales Modell, das quer über die vier Itemgruppen modelliert. Dies galt sowohl für Modelle ohne latente Regression als auch für Modelle mit latenter Regression und sowohl in Bezug auf die Devianz als auch den AIC. Für die Modelle ohne Regression: 11964 vs. 12375 Devianz sowie 12097 vs. 12479 AIC, für die Modelle mit Regression: 12031 vs. 12442 Devianz sowie 12188 vs. 12552 AIC.

len Zahlenreihen über den einzelnen Items geben die Schwellen (*Thurstonian Thresholds*) in der Akzeptanzbereitschaft an, ab denen es wahrscheinlicher ist, dass eine befragte Person bereit ist, die nächstengere Beziehung einzugehen, als sie nicht einzugehen (Wilson, 2005). Beispielsweise gibt das in der Grafik markierte Item 5 die Schwellen für die Offenheit an, einen Juden, der Kippah trägt, als Mitbürger (5.1), als Nachbarn oder Klassenkameraden (5.2), als Freund oder Familienmitglied (5.3) etc. zu akzeptieren.

Für die Kalibrierung des Testinstruments sind einige Beobachtungen, die an der Wright Map ablesbar sind, von besonderem Interesse:

- Insgesamt zeigt sich, dass die einzelnen Items, wie vermutet, als zunehmend ‚schwieriger‘ zu akzeptieren empfunden wurden. Dazu gibt es jedoch zwei Ausnahmen:
- Item 5 wurde als leichter wahrgenommen als Item 4 und als ähnlich leicht wie Item 3. Da es somit redundant ist, kommt seine Streichung in Betracht.
- Item 8 wurde als leichter beurteilt als Item 7. Daher gab es eine relativ große Kalibrierungslücke zwischen Item 7 und Item 9. Diese sollte durch ein neu zu entwickelndes Item geschlossen werden.
- Relativ gleichmäßig verteilt sind für alle 12 Items die Schwellen 1 bis 4 (also von *none of the above* bis zu *close friend* und *family member*). Dagegen ist der Abstand zur 5. Schwelle (*(future) spouse or partner*) erheblich größer. Dafür kann es Gründe im Sinne des Konstrukts geben (dass Menschen weniger offen sind für Partnerschaften mit Juden und Jüdinnen oder generell mit Menschen anderer Religionszugehörigkeit). Aber auch andere Gründe sind denkbar (beispielsweise, dass manche Befragte generell nicht an Partnerschaften interessiert sind). Dies sollte, um die Validität des Fragebogens zu gewährleisten, ergründet werden.

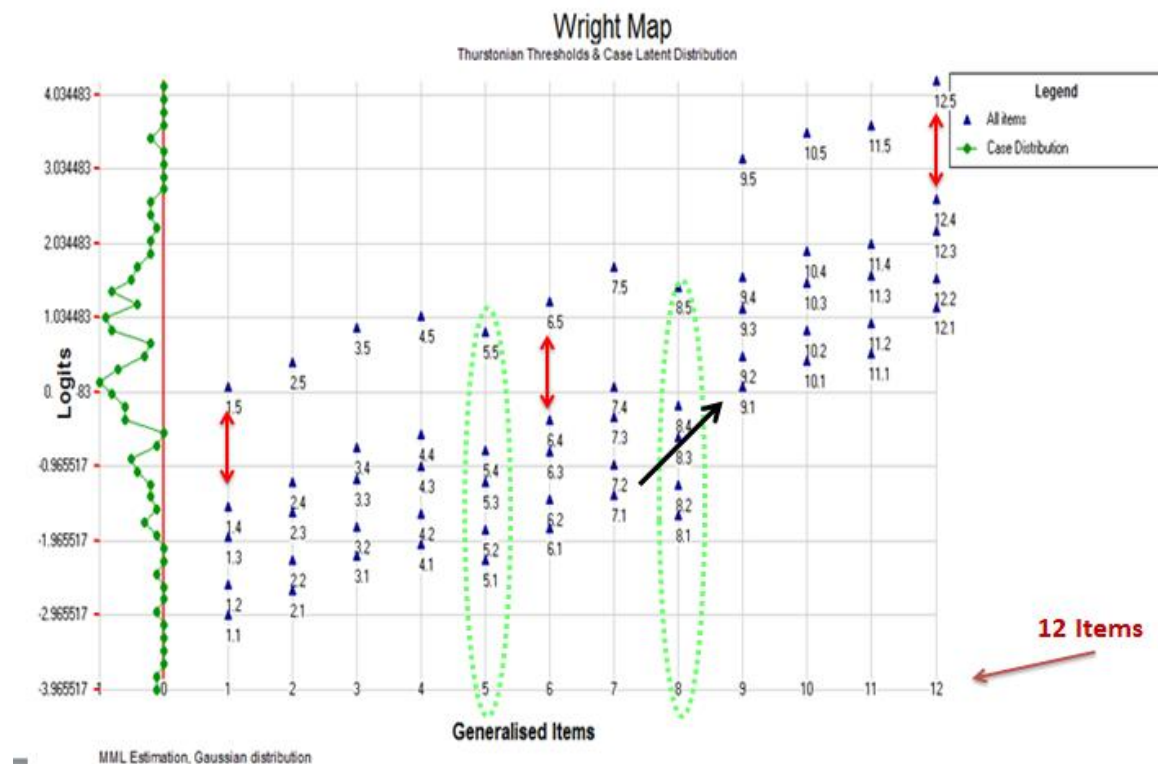


Abb. 3 Wright Map zur Akzeptanz von Jüdinnen und Juden

Mit Blick auf die Kalibrierung des Erhebungsinstruments wurden darüber hinaus die Passungsverhältnisse der einzelnen Items (*Item Fit*) und aller Befragten (*Respondent Fit*) überprüft, beide mit guten Ergebnissen.<sup>9</sup>

### **3.3 Reliabilität und Validität**

Es wurden drei Kriterien verwendet, um die Reliabilität zu evaluieren. Cronbachs Alpha, der die innere Konsistenz der Skala überprüft, beträgt 0,93. Die Split-half-Reliabilität wurde mit 0,96 berechnet. Der *standard error of measurement* liegt um den Logit-Wert 0 mit der höchsten Sensitivität von ca. -3 Logits bis +3 Logits.<sup>10</sup> Alle drei Kriterien weisen darauf hin, dass die Reliabilität des Fragebogens bereits in der vorliegenden Form hoch ist.

Die Validität wurden mit zwei Think-alouds und vier Ausgangsinterviews überprüft. Dabei ergab sich, dass die einzelnen Items insgesamt gut verstanden und in ihrem intendierten Sinn beantwortet wurden.<sup>11</sup> Dies weist auf eine hohe Validität des Fragebogens hin.

Darüber hinaus wurde die Validität mittels der Korrelation mit einer externen Variable überprüft: mit den Antworten der hier Befragten auf fünf Likert-Items der Studie von Francis, Penny und McKenna (2016), die ebenfalls Schülereinstellungen zu religiöser Diversität erhob.<sup>12</sup> Die berechnete Korrelation zwischen Werten für die Akzeptanz gegenüber Religiosität, die mit dem hier entwickelten Fragebogen geschätzt wurden, und der externen Variable betrug 0,48, was akzeptabel ist, aber auch auf Unterschiede zwischen beiden Testinstrumenten hinweist.<sup>13</sup>

## **4 Vorläufige Ergebnisse und Hypothesen**

Im Folgenden werden einige Ergebnisse sowohl zu den Befragten (person ability) als auch zu den Items (item difficulty) diskutiert. Dabei ist zu berücksichtigen, dass es sich lediglich um vorläufige Ergebnisse auf der Grundlage von Pilotdaten handelt, aus denen keine weitreichenden Schlüsse zu ziehen sind.

---

<sup>9</sup> Die Überprüfung der Item Fit-Statistiken ergab, dass alle Items in einem guten Varianzbereich liegen und keine ernsthaften Misfits zu verzeichnen sind. Was den Respondent Fit betrifft, so wurden Befragte als misfitting beurteilt bei einem infit mean square >1,33 und t-Wert >1,96. Legt man diese Kriterien zugrunde, ist bei 6 der 114 Befragten ein überraschendes Antwortmuster zu verzeichnen, davon 1 Befragte der Altersgruppe 13–24 (von 11) und 5 Befragte der Altersgruppe ≥ 35 (von 80).

<sup>10</sup> Siehe dazu die Abbildung im Anhang.

<sup>11</sup> Deutlich wurde allerdings, warum Item 5 (A Jew who invites me to celebrate Passover with his/her family) als leichter bewertet wurde als erwartet, und Item 7 (A Jew who interrupts our shared activities (e.g. having a meal together) for a religious activity (e.g. prayer)) als schwerer. Ausschlaggebend scheinen hier die positive Konnotation des Ausdrucks „invite“ bzw. die negative Konnotation des Ausdrucks „interrupt“ zu sein, die die übrigen Inhalte des Items in den Hintergrund treten ließen.

<sup>12</sup> Die Items lauteten: We must respect all religions./All religious groups must have equal rights./I would be happy to go out with someone from a different faith./I would be happy about a close relative marrying someone from a different faith./People from different religious backgrounds make where I live an interesting place.

<sup>13</sup> In den Think-alouds und Ausgangsinterviews mit Schüler/innen wurde deutlich, warum diese Korrelation nicht höher ist: Schüler/innen wogen bei der Aussage „We must respect all religions“ ab, wie die Formulierungen „all religions“ und „someone from a different faith“ zu interpretieren sei, und antworteten in Abhängigkeit ihrer jeweiligen Interpretation. Siehe dazu auch den Interviewausschnitt im Abschnitt 2.2.

#### 4.1 Akzeptanzbereitschaft der Befragten (person ability)

Tabelle 5 listet vorläufige Ergebnisse für die christlichen Befragten (n=48) und nicht-religiösen Befragten (n=54) zu den vier separaten Skalen auf.

Die *Zeilen* beziehen sich auf die *Probandengruppen* und weisen deren Mittelwerte für die Offenheit gegenüber religiöser Diversität aus. Die *Spalten* beziehen sich auf die vier *Skalen* – Akzeptanz von Diversität gegenüber Muslimen, Christen, Juden sowie Nichtreligiösen – und geben deren jeweiligen Mittelwerte regrediert auf die Religionszugehörigkeit der Befragten an.

Tab. 5 Mittelwerte der christlichen und nichtreligiösen Befragten

	Akzeptanz ggb. Muslimen	Akzeptanz ggb. Christen	Akzeptanz ggb. Juden	Akzeptanz ggb. Nichtreligiösen
<b>Mittelwerte christl. Befragte (n=48)</b>	0.233 (0.278) <sup>14</sup>	0.854 (0.214)	0.549 (0.238)	0.174 (0.209)
<b>Mittelwerte nichtrelig. Befragte (n=54)</b>	0.005 (0.178)	0.051 (0.137)	-0.051 (0.153)	0.997 (0.142)

Drei Beobachtungen seien festgehalten:

- Sowohl die christlichen als auch die nichtreligiösen Befragten unterschieden sich in ihrer Akzeptanzbereitschaft deutlich von Skala zu Skala. Obwohl die einzelnen Items parallel konstruiert sind, machten die Befragten in Bezug auf soziale Beziehungen Unterschiede zwischen Muslimen, Christen, Juden und Nichtreligiösen.<sup>15</sup>
- Die Akzeptanzbereitschaft war jeweils am höchsten gegenüber der eigenen Gruppe.
- Am geringsten war die Akzeptanzbereitschaft der christlichen Befragten gegenüber Nichtreligiösen. Die Akzeptanzbereitschaft der nichtreligiösen Befragten war in Bezug auf alle drei religiösen Gruppen eher gering.

Aus diesen Beobachtungen wurden die folgenden beiden Hypothesen abgeleitet:

*Hypothese 1:* Menschen unterscheiden zwischen Eigenreligion und Fremdreligion und sind eher bereit, Diversität in der eigenen Religion zu akzeptieren als gegenüber Fremdreligionen.

*Hypothese 2:* Die größte Trennlinie verläuft nicht zwischen dem Islam und den anderen Religionen, sondern zwischen den beiden Gruppen Religiöse und Nichtreligiöse.

Beides lässt sich in der Formulierung zusammenfassen: Religion macht einen Unterschied.

<sup>14</sup> Die Zahlen in Klammern weisen für jeden Wert den Standardmessfehler aus.

<sup>15</sup> Unberücksichtigt blieb hier eine mögliche Unschärfe in der Konsistenz der Kalibrierung der vier Skalen, die mittels Delta Dimensional Alignment zu überprüfen ist (Schwartz & Ayers, 2011).

#### 4.2 Schwierigkeit der Items (item difficulty)

Auch zu den Items seien zwei Beobachtungen festgehalten, die der in Abbildung 4 dargestellten Wright Map, hier zur Akzeptanz von Musliminnen und Muslimen, zu entnehmen sind:

- Wie bereits oben beobachtet, zeigte sich auch hier, dass die Schwelle für Partnerschaften hoch war, in Bezug auf Musliminnen und Muslimen in besonderem Maß.
- Dagegen stellte das Kopftuch (Item 6) nur eine mittlere Hürde für soziale Beziehungen dar. Es wurde als deutlich ‚leichter‘ gewertet als beispielsweise die Auffassung, der Islam (oder das Christentum etc.) sei die einzig wahre Religion (Item 9).

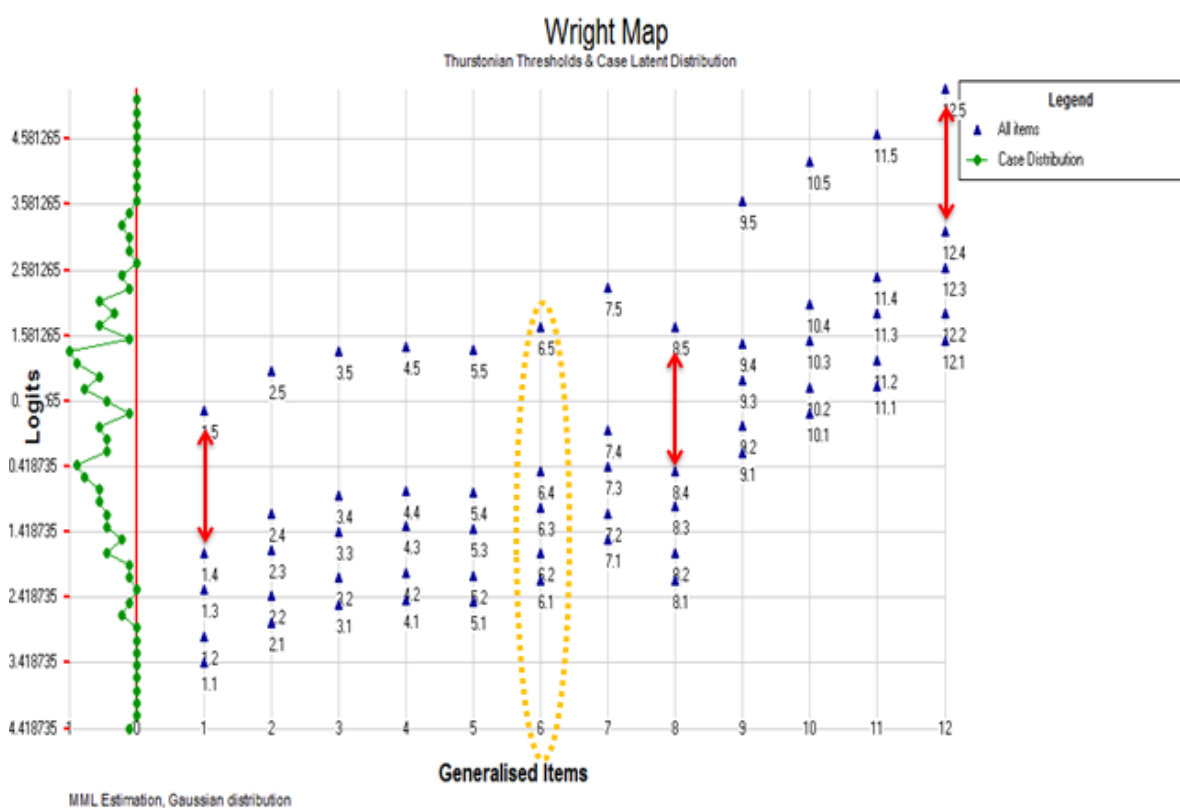


Abb. 4 Wright Map zur Akzeptanz von Musliminnen und Muslimen

Auch aus diesen Befunden wurden zwei Hypothesen abgeleitet:

*Hypothese 3:* Die Schwelle für Partnerschaften mit Angehörigen von Fremdreigionen ist hoch.

*Hypothese 4:* Das Kopftuch stellt nur eine mittlere Hürde für soziale Beziehungen dar, deutlich geringer als die offensiv vertretene Meinung, der Islam sei die einzig wahre Religion. Es ist schwieriger zu akzeptieren als die Kippah, aber leichter als christliche und atheistische ‚Bekenntnis-T-Shirts‘.<sup>16</sup>

<sup>16</sup>

Einmal mit der Aufschrift „Jesus saves“, einmal mit der Aufschrift „Atheist“.

Beide Beobachtungen zum Kopftuch lassen sich in der Formulierung zusammenfassen: Bekenntniseifer ist eine größere Hürde für soziale Akzeptanz als ein zwar sichtbares, aber nicht als offensiv wahrgenommenes Zeichen (Kleidungsstück).

## 5 Diskussion und Ausblick

Mit dem hier entwickelten Fragebogen liegt ein dichtes und komplexes Erhebungsinstrument vor, das Antworten auf eine Vielzahl spezifischer Forschungsfragen im Zusammenhang mit religiöser Diversität erlaubt. Dies umfasst die folgenden Fragen:

*Einzelne Items einer Skala:* Was empfinden die Befragten als ‚leicht‘, was als ‚schwer‘ zu akzeptieren?<sup>17</sup>

*Parallele Items der vier Skalen:* Für welche Religion oder Weltanschauung wird ein bestimmtes Verhalten oder Charakteristikum eher, für welche weniger akzeptiert?

*Soziale Nähe und Distanz:* Welche Grade an sozialer Nähe akzeptieren die Befragten bereitwillig, welche nur zurückhaltend?<sup>18</sup>

*Die vier Skalen:* Wie stehen die Befragten zu den vier Gruppen Muslime, Christen, Juden und Nichtreligiöse?<sup>19</sup>

*Externe Variable Religionszugehörigkeit:* Welche Unterschiede bestehen für die fünf Befragtengruppen Muslime, Christen, Juden, Angehörige anderer Religionsgemeinschaften und Nichtreligiöse in Bezug auf ihre Akzeptanz religiöser Diversität?<sup>20</sup>

*Externe Variablen Alter, Gender und Häufigkeit des Kontakts:* Sind jüngere Befragte aufgeschlossener gegenüber religiöser Diversität als ältere? Unterscheiden sich Männer und Frauen in Bezug auf ihre Bereitschaft, mit Menschen unterschiedlicher religiöser Profile soziale Beziehungen einzugehen? Gibt es geschlechtsspezifische Unterschiede in der Akzeptanz von Muslimen, Christen, Juden sowie Nichtreligiösen? Wirkt sich – wie Allports Kontakthypothese postuliert – der persönliche Kontakt positiv auf die Bereitschaft aus, soziale Beziehungen einzugehen (Allport, 1954)?

Die empirische Überprüfung des Instruments deutet darauf hin, dass dieses auch in seiner bisherigen Fassung geeignet ist, reliabel und valide Einstellungen zu religiöser Diversität zu erheben. Dennoch legen die Beobachtungen, die hier exemplarisch für die jüdische Skala berichteten wurden, sowie die Ergebnisse zu den drei übrigen Skalen eine Reihe von Überarbeitung nahe. Dies betrifft:

- die Streichung von Items, die im Vergleich zu anderen keine neue Information bringen (beispielsweise Item 5 „celebrate“), sowie von Items, deren Validität nicht voll gewährleistet ist (beispielsweise Item 5 „celebrate“, Item 7 „prayer“),<sup>21</sup> und
- die Entwicklung eines neuen Items, um die Lücke im Schwierigkeitsgrad zwischen Item 7 und Item 9 zu schließen.

---

<sup>17</sup> Siehe dazu die Wright Maps Abb. 3 und 4.

<sup>18</sup> Siehe dazu die Wright Maps Abb. 3 und 4.

<sup>19</sup> Siehe dazu Tabelle 5.

<sup>20</sup> Siehe dazu Tabelle 5.

<sup>21</sup> Dabei wurde darauf geachtet, dass die Reliabilität des Fragebogens nach wie vor hoch ist.



Dagegen empfiehlt es sich, die Frage nach der Akzeptanz partnerschaftlicher Beziehungen beizubehalten, da dies ein relevantes Ergebnis sein könnte. Die Gründe für die hier beobachtete hohe Schwelle sollten jedoch näher untersucht werden.<sup>22</sup>

Wenngleich die Pilotierung nicht ausschließlich mit Schüler/innen erfolgt ist, sondern auch ältere Befragte umfasste, weisen die folgenden Beobachtungen darauf hin, dass das Testinstrument für die Verwendung mit Schüler/innen gut geeignet ist: Zum einen ergab die Überprüfung des *Respondent Fits* ein gutes Ergebnis für jüngere Befragte. Das weist darauf hin, dass der Fragebogen auch in der Altersgruppe der 14- bis 18jährigen gut verstanden wurde. Zum anderen wurden die zwei Think-alouds sowie zwei der vier Ausgangsinterviews mit Schüler/innen der 8. bis 10. Klasse geführt und ergaben, dass von ihnen die Fragen so verstanden wurden, wie sie intendiert waren. Dennoch sollte die Eignung des Fragebogens für Schüler/innen ab der 8. Klasse in einem zweiten Pretest umfassend evaluiert werden.

## Danksagung

Mein Dank gilt Mark Wilson (UC Berkeley) für seinen kontinuierlichen fachlichen Rat in der Entwicklung des Erhebungsinstruments sowie Perman Gochyyev (UC Berkeley) für weiterführende Diskussionen zur Modellierung der Daten.

## Literaturverzeichnis

- Allport, G.W. (1954). *The Nature of Prejudice*. Cambridge: Addison-Wesley.
- Andrich, D. (1978). A Rating Formulation for Ordered Response Categories. *Psychometrika*, 43(4), 561–573.
- Bertelsmann-Stiftung (Hrsg.) (2017). *Religionsmonitor. Muslime in Europa – integriert, aber nicht akzeptiert?* Gütersloh: Bertelsmann.
- Brown, N. & Wilson, M. (2011). A Model of Cognition. The Missing Cornerstone of Assessment. *Educational Psychology Rev*, 23, 221–234.
- Duckor, B., Draney, K. & Wilson, W. (2009). Measuring Measuring. Toward a Theory of Proficiency with the Constructing Measures Framework. *Journal of Applied Measurement*, 10(3), 296–319.
- Eagly, A. & Chaiken, S. (1993). *The Psychology of Attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich College.

---

<sup>22</sup> Je nachdem kann diese höchste Schwelle in künftigen Analysen berücksichtigt oder ausgelassen werden.

- Francis, L., Penny, G. & McKenna, U. (2016). Does RE Work and Contribute to the Common Good in England? In E. Arweck (Hrsg.), *Young People's Attitudes to Religious Diversity* (S. 153–169). Abingdon: Routledge.
- Guttman, L. (1944). A Basis for Scaling Qualitative Data. *American Sociological Review*, 9, 139–150.
- Morell, L. u.a. (2017). A Construct Modeling Approach to Develop a Learning Progression of how Students Understand the Structure of Matter. *Journal of Research in Science Teaching*, 54(8), 1024–1048.
- Schwartz, R. & Ayers, E. (2011). *Delta Dimensional Alignment*. Comparing Performances Across Dimensions of the Learning Progression for Assessing Data Modeling and Statistical Reasoning. Unpublished manuscript, University of California, Berkeley.
- Wark, C. & Galliher, F. (2007). Emory Bogardus and the Origins of the Social Distance Scale. *The American Sociologist*, 38(4), 383–395.
- Wilson, M. (2005). *Constructing Measures. An Item Response Modeling Approach*. Mahwah, NJ: Erlbaum.
- Zanna, M. & Rempel, J. (1988). Attitudes. A New Look at an Old Concept. In D. Bartal & A. Kruglanski (Hrsgs.), *The Social Psychology of Knowledge* (S. 316–334). Cambridge: Cambridge University Press.

## 6 Anhang

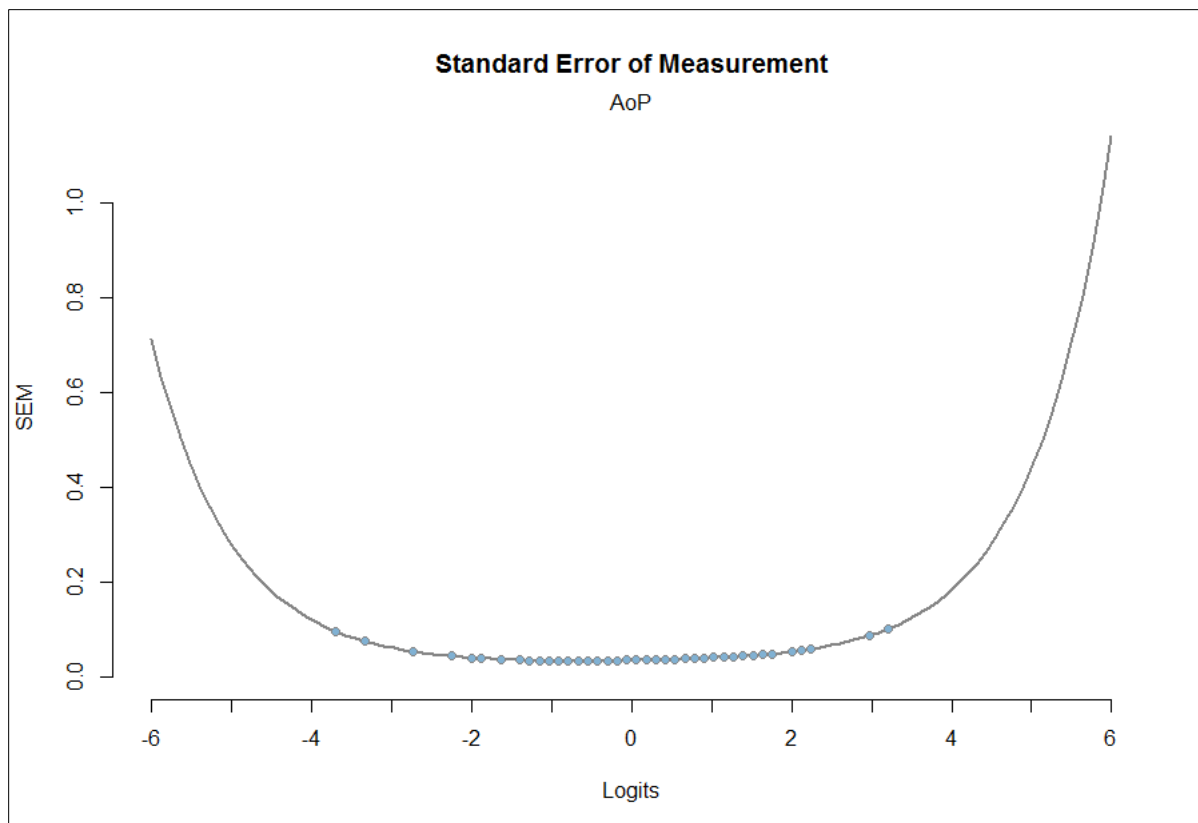


Abb. 5 Standard error of measurement

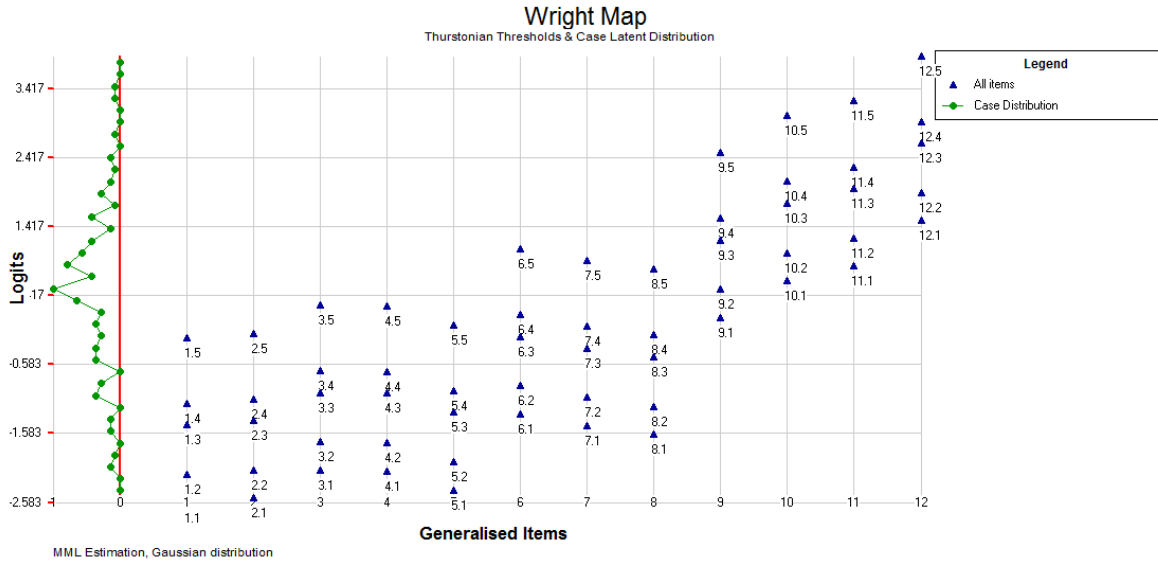


Abb. 6 Wright Map zur Akzeptanz von Christ/innen

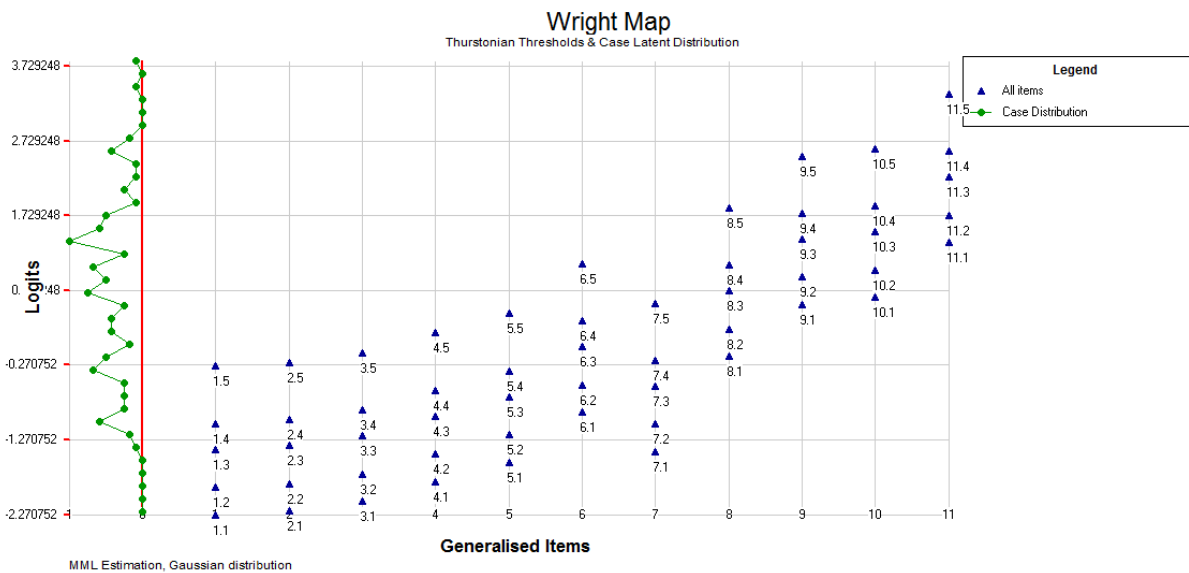


Abb. 7 Wright Map zur Akzeptanz von Nichtreligiösen

*Dr. Sabine Hermisson, Institut für Religionspädagogik der Evangelisch-Theologischen Fakultät der Universität Wien*