Behinderung vorliegt oder nicht. Die Reihenfolge, in der die einzelnen Lebensbereiche durch die Befragten bearbeitet wurden, wurde randomisiert, um eine mögliche Beeinflussung des Antwortverhaltens (Priming-Effekte), die eine festgelegte Abfolge der Lebensbereiche womöglich mit sich bringen könnte, bestmöglich zu verhindern.

20.3 Haupterhebung

Die Haupterhebung des Surveys erfolgte anhand eines deutschlandweiten Panels, das repräsentativ die Bevölkerung Deutschlands abbildet. Die Befragung wurde in zwei Wellen durchgeführt, in denen – nach Bereinigung des Datensatzes (unter anderem, um sogenannte Erwünschtheitsantworten herauszufiltern) – eine Anzahl von 3695 Teilnehmenden erreicht werden konnte. Die Stichproben wurden zusätzlich zur Quotierung gewichtet (ausgehend von den Merkmalen Alter, Geschlecht und Religion). Die Stichprobe setzt sich zu 49,85 % aus Frauen und zu 49,82 % aus Männern zusammen. 0,32 % der Befragten machten hinsichtlich ihres Geschlechts keine Angabe. Das Durchschnittsalter der Teilnehmenden beläuft sich auf 48,81 Jahre (SD 16,045). Befragt wurden Personen zwischen 18 und 95 Jahren. Nach eigenen Angaben haben 11,64 % der befragten Personen einen Migrationshintergrund. Zum Zeitpunkt der Befragung waren mehr als zwei Drittel der Personen hauptbeschäftigt angestellt (43,09 %) oder im Ruhestand (25,85 %). Etwa die Hälfte (47,01 %) gab an, regelmäßig Kontakt zu Menschen mit Behinderung (nicht nur geistige Behinderung) zu haben. Von diesen wiederum gaben 67,13 % an, regelmäßig Kontakt zu Menschen mit geistiger Behinderung zu haben (dies entspricht 31,56 % der gesamten Stichprobe). Nach eigenen Angaben haben 12,29 % der Befragten selbst eine Behinderung. 1,19 % der Teilnehmenden gaben an, selbst eine geistige Behinderung zu haben.

21. Auswertung: Clusteranalyse

Die Frage nach der Einstellung der Gesamtbevölkerung zu Inklusion wurde über die Durchführung einer Clusteranalyse operationalisiert, die sich im Wesentlichen dadurch kennzeichnet, dass innerhalb des Datensatzes nach signifikanten Beantwortungsmustern (Clustern) gesucht wird, die in der Folge klassifiziert und beschrieben werden. Die Grundlage für die Clus-

teranalyse bildete eine weitere Faktorenanalyse, mittels derer die innere Struktur des Datensatzes bestimmt wurde. Diese bestätigte das Ergebnis der gleichförmigen Strukturierung der Lebensbereiche Wohnen, Arbeit und Freizeit, wie sie bereits im Zuge des Pre-Tests festgestellt wurde. Die Analyse offenbarte jeweils für die vergleichenden Items von acht der generierten Kategorien, die die Gegenüberstellung und den Vergleich der verschiedenen Lebensbereiche gewährleisten, eine konsistente Zweifaktorenstruktur. Auf dem ersten Faktor, der sich als ›Befürwortung von Inklusion‹ bestimmen lässt, luden in jedem Lebensbereich jeweils die gleichen fünf Items. Die übrigen drei Items hingegen luden auf dem zweiten Faktor, der sich als ›Ablehnung von Inklusion‹ fassen lässt. Die übrigen vergleichenden Items ließen sich nicht klar zuordnen (unter anderem aufgrund von Doppelladungen) und wurden insofern aus der Analyse ausgeklammert. Dies betraf die Items der Überkategorien »Inklusion als Aufgabe«, »Ausgrenzung«, »Veränderungen«, »Schulungsbedarfe« und »Überforderung«. Für die Clusteranalyse blieben insofern die Items folgender Überkategorien in den Lebensbereichen Wohnen, Arbeit und Freizeit bestehen:

Tabelle 41: Itemkategorien (2)

	Kategorien
1	Unterstützung
2	Engagement
3	Erfordernisse
4	AdressatInnenauswirkung
5	Umsetzung
6	Geteilte Lebenspraxis
7	Protektion
8	Finanzierung

Um die Aussagen der verschiedenen Faktoren im Hinblick auf die Einstellung(en) zu Inklusion miteinander vergleichen zu können, musste die Likert-Skala der Items, die auf dem inklusionsablehnenden Faktor luden, umgepolt werden, damit in diesen Fällen – wie auch bei Items des inklusionsbefürwortenden Faktors – ein hoher Zustimmungsgrad der Items einen hohen Zustimmungsgrad zu Inklusion bedeutet und umgekehrt. Beispielhaft veranschaulicht heißt das: Der fiktive Zustimmungsgrad einer Person von 80 % in Bezug auf die inklusionsablehnende Aussage »*Wenn an meinem Arbeitsplatz Menschen mit geistiger Behinderung integriert würden, würde ich lieber meinen Arbeitsplatz wechseln*« bedeutet, dass die Person eine eher negative Einstellung zu Inklusion hat. Im Zuge der Umpolung werden aus den 80 % nun 20 %. Dieser Wert stellt infolgedessen nicht mehr den Zustimmungsgrad zur Aussage dar, sondern, und darin liegt ein zentrales Ergebnis der Analyse, basiert vielmehr auf einer inneren Kohärenz des Datensatzes, die als Zustimmungsgrad zu Inklusion gefasst werden kann. Sind im Folgenden Zustimmungsgrade abgebildet, so stellen diese nicht den Zustimmungsgrad zur Aussage dar, sondern immer den Zustimmungsgrad zu Inklusion. Auf dieser Grundlage wurde anhand des Datensatzes eine Clusteranalyse entlang des Ward-Verfahrens[4] durchgeführt. Zur Identifizierung der geeignetsten Clusterlösung wurden jeweils pro Clusterlösung eine Einfaktorielle

4 Die Clusteranalyse, die zu den explorativen Verfahren der multivariaten Datenanalyse gehört, fasst unterschiedliche Verfahren zur Gruppenbildung zusammen. Dabei unterscheiden sich die Verfahren hinsichtlich des Proximitätsmaßes und der Wahl des Gruppierungsverfahrens (Backhaus et al. 2018). Das agglomerative Ward-Verfahren gehört zu den hierarchischen Clusterverfahren, die nicht von einer gegebenen Gruppierung der Objekte und einer festgelegten Anzahl an Clustern ausgehen, sondern die zu gruppierenden Objekte schrittweise zu immer größeren Clustern zusammenfassen (Backhaus et al. 2018). Die Anwendung des Verfahrens setzt metrische Merkmale voraus, die hier gegeben sind. Das Ward-Verfahren vereinigt diejenigen Objekte, die die Fehlerquadratsumme in einer Gruppe am wenigsten erhöhen und bildet somit möglichst intern homogene Cluster. Dadurch kann das Ziel erreicht werden, möglichst homogene Gruppen zu identifizieren, die in ihrer Einstellung zu Inklusion gleich oder sehr ähnlich sind.

Varianzanalyse (ANOVA)[5] sowie der Tukey-Test[6] durchgeführt. Im Vergleich stellte sich die Vier-Cluster-Lösung als beste Clusterlösung heraus, da sich dort alle vier Cluster signifikant (p<=.00) voneinander unterscheiden – sowohl bei der Einfaktoriellen Varianzanalyse als auch beim Tukey-Test. Des Weiteren verweisen die Werte des Eta2 bei allen sechs Faktoren auf eine große Effektstärke (nach Cohen 1988, S. 79ff). Dies bedeutet, dass ein großer Prozentsatz (mind. 58,5 %) der beobachteten Variationen der Zustimmungswerte auf die Clusterzugehörigkeit zurückgeführt werden kann. Die identifizierten Cluster sind damit das Ergebnis statistischer Berechnungen, die in der Folge zu interpretieren beziehungsweise inhaltlich zu bestimmen sind. Die vier Cluster lassen sich insofern auch als eine Form der Typisierung verstehen, die entlang des Antwortverhaltens der TeilnehmerInnen erstellt wurde und die statistisch fundiert ist.

5 Die Einfaktorielle Varianzanalyse beschreibt im Wesentlichen ein Verfahren, bei dem über einen Mittelwertvergleich der Frage nachgegangen wird, ob sich signifikante – also statistisch aussagekräftige – Unterschiede zwischen mehreren unabhängigen Gruppen in Bezug auf eine abhängige Variable finden lassen (Diekmann 2016, S. 694ff). Im hiesigen Zusammenhang ging es folglich um die Frage, ob sich in dem verfügbaren Datensatz tatsächlich belastbar unterschiedliche Typen (Cluster) in Bezug auf die »Einstellung zu Inklusion« (abhängige Variable) finden und beschreiben lassen.

6 Mit dem signifikanten Ergebnis der Einfaktoriellen Varianzanalyse wird zunächst lediglich die Aussage getroffen, dass es einen Unterschied zwischen den Gruppen gibt, allerdings nicht, worin dieser liegt. Um dies zu identifizieren, muss ein Post-hoc-Test herangezogen werden. Die paarweisen Vergleiche des Post-hoc-Tests suchen durch sogenannte multiple Mittelwertvergleiche nach signifikanten Unterschieden zwischen den Gruppen. Eine besondere Eigenschaft des Tukey-Tests ist dabei, dass das Fehlerniveau konstant nahe 5 % gehalten wird. Die Anwendung des Tukey-Tests ist zu empfehlen, wenn Varianzhomogenität gegeben und die Gruppengröße gleich ist. Die Varianzhomogenität wurde in der hiesigen Auswertung zuvor mit dem Levene-Test bestätigt. Allerdings handelt es sich bei den generierten Clustern um ungleiche Gruppengrößen. Das Statistikprogramm SPSS berechnet in diesem Fall automatisch den Tukey-Kramer-Test, der speziell für ungleiche Gruppengrößen entwickelt wurde, jedoch in der Interpretation identisch ist.