SSOAR

Open Access Repository
www.ssoar.info

Online- und Offlinevergleich visueller Analogskalen mit 4- und 8-stufig skalierten Likert-Skalen bei einem Fragebogen zum Verhalten in sozialen Gruppen

Funke, Frederik

Veröffentlichungsversion / Published Version
Sammelwerksbeitrag / collection article

Empfohlene Zitierung / Suggested Citation:
Funke, F. (2006). Online- und Offlinevergleich visueller Analogskalen mit 4- und 8-stufig skalierten Likert-Skalen bei einem Fragebogen zum Verhalten in sozialen Gruppen. In K.-S. Rehberg (Hrsg.), *Soziale Ungleichheit, kulturelle Unterschiede: Verhandlungen des 32. Kongresses der Deutschen Gesellschaft für Soziologie in München. Teilbd. 1 und 2* (S. 4826-4838). Frankfurt am Main: Campus Verl. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-141683

Nutzungsbedingungen:
Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:
This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.
By using this particular document, you accept the above-stated conditions of use.

gesis Leibniz-Institut für Sozialwissenschaften

Mitglied der Leibniz-Gemeinschaft

Online- und Offlinevergleich Visueller Analogskalen mit 4- und 8-stufig skalierten Likert-Skalen bei einem Fragebogen zum Verhalten in sozialen Gruppen

Frederik Funke

1. Einleitung

Bei Befragungen können Befragte ihre Antwort meist nicht beliebig differenziert zum Ausdruck bringen, sondern müssen ihre Einstellung in ein Kategoriensystem einordnen. Kategorialskalen können mit einer doppelten Korsettmetapher beschrieben werden: Die Begrenzung der Auswahlmöglichkeiten kann einerseits beengen, so dass der Befragte das Gefühl hat, sich nicht richtig entfalten zu können, seine Einstellung nicht exakt genug zum Ausdruck bringen zu können und mit dem Festlegen auf eine Kategorie nur eine unvollständige Beurteilung zu geben. Andererseits haben vorgegebene Kategorien auch eine stützende Wirkung, indem das Festlegen auf eine Kategorie durch eine überschaubare Anzahl deutlich unterschiedlicher Alternativen erleichtert wird.

Bei Visuellen Analogskalen (VAS) entfällt dies, da auf einem Kontinuum geantwortet wird. Möchte man Kategorialdaten mit Daten aus VAS vergleichen, so stellt sich die Frage, wie man die Rohwerte dieses Skalentyps angemessen kategorisiert. In dieser Untersuchung werden zwei Methoden miteinander verglichen. Die Methode der linearen Kategorisierung von VAS spiegelt die Annahme wider, dass die Abstände zwischen allen Kategorien gleich groß sind. Die Methode der reduzierten Extreme sieht nicht alle Kategorien als gleichabständig, äquidistant, an, so dass unterschiedliche Kategorien unterschiedliche Einstellungsintensitäten repräsentieren.

Um diese beiden Vorgehensweisen zu testen und die Abstände zwischen den einzelnen Kategorien bei Kategorialskalen beschreiben zu können, bedarf es eines Maßstabes, an dem kategoriale Skalen gemessen werden können. Visuelle Analogskalen sind dafür gut geeignet. Ihr wesentlicher Vorteil ist, dass sie als Kontinuum wahrgenommen werden und ihre Daten somit als intervallskaliert gelten. Zwei gleichgroße Abstände auf der Visuellen Analogskala werden von den Befragten immer als zwei gleich große Unterschiede interpretiert werden. Mit identischen verbalen Ankern ausgestattet, decken sie dieselbe Bandbreite an Einstellungen ab,

wie dies bei entsprechend verankerten Kategorialskalen der Fall ist. Da Visuelle Analogskalen eine sehr feine Untergliederung zwischen den Extremen erlauben, eignen sie sich zur präzisen Bestimmung der Breite von Kategorien bei Kategorialskalen.

Die Kategorienbreite wird in der vorliegenden Untersuchung sowohl bei vier-, als auch bei achtstufigen Kategorialskalen untersucht. Dieser doppelte Vergleich soll klären, ob sich die Einstellungsintensität der Rand- und Mittelkategorien auf ein und dieselbe Weise beschreiben lässt. Zusätzlich soll ein Vergleich von Offline- und Online-Design klären, ob die Beschreibung der Kategorienbreite durch Visuelle Analogskalen vom Erhebungsmodus unabhängig ist, oder ob die Skalen offline und online unterschiedlich wahrgenommen werden.

2. Visuelle Analogskalen (VAS)

Eine Visuelle Analogskala ist eine in der Regel 100 Millimeter lange horizontale Linie, die an den Extremen verbale Anker besitzt. Befragte markieren auf der Linie die Stelle, die ihre Einstellung am besten widerspiegelt. VAS sind zur Messung eindimensionaler Konstrukte geeignet und werden auch als räumlich eindimensional wahrgenommen (Brunier/Graydon 1996). VAS fallen unter die Gruppe der kontinuierlichen Antwortformate. Das bedeutet, dass dem Befragten keine feste Anzahl von Antwortmöglichkeiten vorgegeben ist, sondern dass sich die Antworten auf einem Kontinuum, einer lückenlos zusammenhängenden Abstufung, bewegen.

Mit Visuellen Analogskalen können sowohl bei Individuen als auch bei Gruppen kleinere Unterschiede im Antwortverhalten aufgezeigt werden, als dies mit Kategorialskalen möglich ist. Inwiefern solch kleine Änderungen in der Praxis relevant sind, ist eine Frage, die jeweils im theoretischen Kontext entschieden werden muss. Ein konkreter Vorteil Visueller Analogskalen ist, dass sich die Frage nach der Anzahl der Kategorien nicht stellt. Wenn digitale Datenerhebungsmöglichkeiten nicht angewandt werden können, fallen die Nachteile des manuellen Ausmessens der Visuellen Analogskalen erheblich ins Gewicht, so dass es starke Argumente für den Einsatz von VAS gibt, die den Aufwand des Auslesens aufwiegen. Der Befund, dass eine geringere schulische Bildung mit einem geringeren Verständnis des Gebrauchs Visueller Analogskalen einhergeht, spricht deutlich gegen deren Einsatz zur Untersuchung der Allgemeinbevölkerung und anderer Gruppen, deren Bildungsniveau niedrig oder unbekannt ist. Wenn zu befürchten ist, dass Effekte sozialer Erwünschtheit Ergebnisse stark verzerren, dann stellt der Einsatz von VAS eine gute Alternative zu Kategorialskalen dar.

Visuelle Analogskalen bieten einen kontinuierlichen und linearen Raum zur Bewertung von Items und lassen sich nahezu beliebig genau auslesen. Werden Visuelle Analogskalen mit identischen verbalen Ankern wie Kategorialskalen ausgestattet, dann steht in beiden Fällen dieselbe Bandbreite zur Einschätzung von Items zur Verfügung. Somit sind Visuelle Analogskalen grundsätzlich auch dafür geeignet, als Maßstab für die Bestimmung der Breite von Kategorialskalen zu dienen, ebenso wie sie – wie oben beschrieben – dafür benutzt werden können, um die Distanz verbaler Marken (oder grafischer Marken (Jäger 2004)) zu bestimmen. Es kann keine allgemeine Empfehlung für oder gegen den Gebrauch dieses Skalentyps gegeben werden. In erster Linie ist es eine Frage des Untersuchungszusammenhanges, aber auch der jeweiligen Forschungsinteressen, der finanziellen und zeitlichen Ressourcen und des persönlichen Geschmacks; »the most efficacious method for data collection remains inconclusive« (Flynn u.a. 2004: 49).

3. Untersuchungsdesign

Es wurden zwei Erhebungen durchgeführt. Eine im Offline-Design und eine weitere Untersuchung im Online-Design. Die Offline-Befragung wurde als Panelstudie mit zwei Erhebungen als Papier-und-Bleistift-Untersuchung umgesetzt. Bei der ersten Erhebungswelle standen zur Beantwortung der Items ausschließlich Visuelle Analogskalen zur Verfügung. Die Wiederholungsmessung erfolgte unter Einsatz von vier- und achtstufigen Kategorialskalen.

Zusätzlich zu der Befragung des Offline-Panels wurde ein Online-Experiment durchgeführt (Reips 2003), um zu prüfen, ob eventuell auftretende Effekte stabil sind, oder ob sie vom Erhebungsmodus abhängen. Die Onlinebefragung war eine Umsetzung des oben beschriebenen Fragebogens ins Computer-Assisted-Web-Interview-Design (CAWI), wobei die Erhebungsform von einem Panel- zu einem experimentellen Design abgeändert wurde. Es wurden also alle Teilnehmenden auf eine der drei Skalen randomisiert. Um Artefakte durch unzureichende Verteilung auf die Skalen ausschließen und Skaleneffekte bestimmen zu können, wurden nicht nur soziodemographische Faktoren analysiert. Es sollte auch prüfbar sein, ob sich die Untergruppen auf inhaltlicher Ebene voneinander unterscheiden oder ob der wahre Wert der Stichproben ohne Unterschiede ist. Dies lässt sich allerdings nur dann bestimmen, wenn den Befragten identische Items mit identischen Skalen vorgelegt wurden, Skaleneffekte also ausgeschlossen werden können. Um diese, auf den ersten Blick nicht miteinander zu vereinbarenden Forderungen nach Gleichheit und Verschiedenheit der Skalen erfüllen zu können, wurde der Fragebogen halbiert (Frage 1–8 und Frage 9–16). Den beiden Teilen wurden jeweils die drei zu untersu-

chenden Skalen zugewiesen, was zu neun experimentellen Gruppen führte. Bei der Planung und Programmierung des Onlinefragebogens wurden die 16 von Ulf-Dietrich Reips (2002) formulierten Standards für Internet-basierte Befragungen umgesetzt.

Dass sich Visuelle Analogskalen prinzipiell für Onlinebefragungen eignen, haben Mick P. Couper u.a. (2004) mehrfach herausgestellt. Für die vorliegende Untersuchung wurde eine von Timo Gnambs programmierte Visuelle Analogskala eingesetzt, die große Ähnlichkeit mit der VAS der Offlineuntersuchung hat. Im Ruhezustand sieht der Befragte eine nackte, an den Extremen verbal verankerte horizontale Linie. Sobald das erste Mal auf die Line geklickt wird, erscheint an der entsprechenden Stelle ein Kreuz, dass sich mit weiteren Klicken an einen anderen Ort auf der Linie bewegen lässt. Analog zum Offline-Design wurde die VAS so programmiert, dass sie sich (für den Nutzer nicht sichtbar) in 80 Einzelstücke unterteilte und folglich Werte zwischen 1 und 80 annehmen konnte. Die VAS wurde mit einem JavaScript ausgelesen.

4. Kategorisierung Visueller Analogskalen

Ein Ziel dieser Untersuchung ist die Bestimmung der Kategorienbreite bei Kategorialskalen, um die Bedeutung, die Kategorien für den Befragten haben, besser verstehen zu können und Änderungen im Antwortverhalten den richtigen Wert beizumessen. Visuelle Analogskalen sind aus drei Gründen zur Bestimmung der Kategorienbreite gut geeignet. Zum einen messen sie dieselbe Einstellungsbandbreite wie Kategorialskalen. Zum anderen lassen sie sich beliebig genau auslesen, so dass sich Antwortausprägungen mit Visuellen Analogskalen exakt bestimmen lassen. Und letztlich gelten die Daten Visueller Analogskalen als intervallskaliert, so dass sie als Maßstab für die Untersuchung kategorialer Skalen dienen können. Die Breite der Kategorien gilt genau dann als gut beschrieben, wenn sie zwischen Messungen eines identischen Konstrukts zu denselben Ergebnissen führt. Die Daten der VAS müssen, um Vergleiche anstellen zu können, zu Kategorien zusammengefasst werden. Die Art und Weise, nach der diese Kategorisierung geschieht, dient gleichzeitig als Beschreibung der Breite der Kategorialskalen. Konkret stellt sich also die Frage, wie die achtzig Einheiten umfassende VAS so zu vier beziehungsweise acht Kategorien zusammengefasst werden kann, dass eine möglichst hohe Übereinstimmung zwischen den beiden Skalen erzielt wird. Diesem Vorgehen liegt folgende Logik zugrunde: »If the data can be transformed to fit the model, the scale values can be derived from the transformed data, and the inverse transformation could be interpreted as the judgement function.« (Birnbaum 1978)

Im Folgenden werden zwei Verfahren vorgestellt, um die Werte Visueller Analogskalen zu Kategorien zusammenzufassen, die lineare Reduktion und das Modell der reduzierten Extreme.

4.1. Lineare Transformation

Fasst man immer gleich große Intervalle einer VAS zu je einer Kategorie zusammen, dann sind die Abstände zwischen diesen Kategorien gleich groß und das Niveau der Daten bleibt intervallskaliert. Bei dieser linearen Kategorienreduktion werden die Daten der VAS derart transformiert, als hätte man beim Auslesen der VAS nur ein grobes vier- beziehungsweise achtstufiges Raster statt der feinen 80fachen Untergliederung angelegt. Nach linearer Transformation können VAS als Maßstab zur Prüfung dienen, ob Befragte alle Kategorien bei Kategorialskalen für gleich große Einstellungsintervalle nutzen. Das Prinzip der linearen Transformation wird in Abbildung 1 verdeutlicht.

Wenn die lineare Transformation zu einer hohen Übereinstimmung mit den Kategorialskalen führt, dann bedeutet dies, dass die Kategorien als gleich breit und somit äquidistant wahrgenommen werden, und dass Kategorialskalen die Voraussetzungen für die Anwendung parametrischer Verfahren erfüllen. Besteht zwischen den Kategorienbelegungen bei Kategorialskalen und linear transformierten Visuellen Analogskalen eine hohe Übereinstimmung, bedeutet dies, dass die Breite der Kategorien der Kategorialskala durch die linear transformierte VAS gut beschrieben wird.

4.2. Modell der reduzierten Extreme

Eine Alternative zu der linearen Kategorienreduktion stellt das Modell der reduzierten Extreme dar, bei dem die mittleren Kategorien ebenfalls linear transformiert werden, die Intervalle der Randkategorien jedoch in ihrer Breite reduziert sind. Die Breite der Randkategorien beträgt nach diesem Modell nur zwei Drittel der Breite der inneren Kategorien. Die Herleitung für diese Verkleinerung der Extreme ist, dass die inneren Kategorien jeweils zwei angrenzende Kategorien haben, während die Randkategorien immer nur eine benachbarte Kategorie besitzen. Dabei wird folgende Wahrnehmung der Kategorien unterstellt: Die mittleren

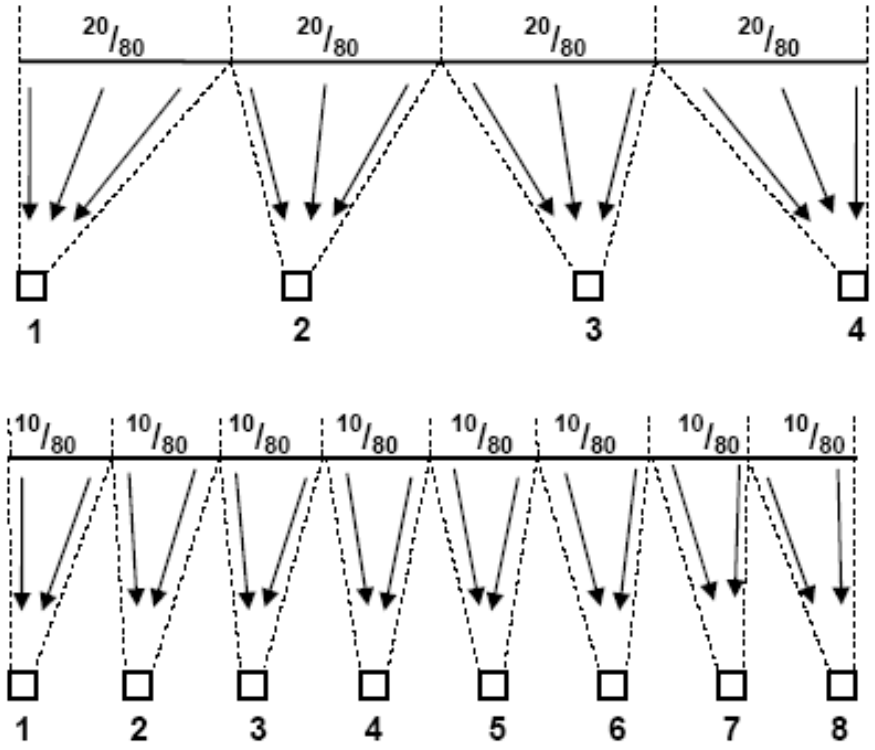

Abbildung 1: Lineare Transformation der VAS; A zu vier Kategorien, B zu acht Kategorien.

Kategorien lassen sich jeweils in drei Stufen unterscheiden: Vollkommene Übereinstimmung mit einer Kategorie, Übereinstimmung mit einer Kategorie mit Tendenz zur rechts angrenzenden Kategorie und Übereinstimmung mit Tendenz zur links angrenzenden Kategorie. Bei den Randkategorien gibt es lediglich zwei Stufen: Übereinstimmung und die Tendenz zu einer angrenzenden Kategorie. Nach dem Modell der reduzierten Extreme gibt es bei den mittleren Kategorien jeweils drei Fälle, bei denen eine Kategorie als zutreffend erachtet werden kann; bei den extremen Kategorien sind es nur zwei Fälle, wie in Abbildung 2 verdeutlicht wird.

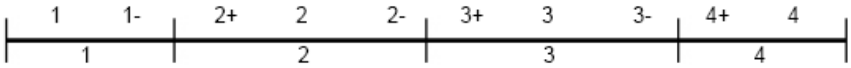

Abbildung 2: Abstufung von Kategorien am Beispiel der vierstufigen Kategorialskala.

Die Anwendung des Modells der reduzierten Extreme führt bei Transformation der VAS in vier Kategorien (wie Abbildung 3 A zeigt) dazu, dass die zu Randkategorien zusammengefassten Intervalle der VAS jeweils 16 der 80 Einheiten umfassen und die mittleren Kategorien je 24 Einheiten breit sind. Die Transformation zu acht Kategorien (Abbildung 3 B) führt nach dem Modell der reduzierten Extreme dazu, dass die Intervalle der Randkategorien jeweils sieben Einheiten und die Intervalle, die die inneren Kategorien bilden, jeweils elf Einheiten breit sind.

Falls die durch Transformation nach dem Modell der reduzierten Extreme entstandenen VAS Kategorien zu größerer Übereinstimmung mit den Daten aus Kategorialskalen führen, als es nach linearer Transformation der Fall ist, dann ist dies ein deutlicher Beleg dafür, dass Befragte die Randkategorien schmaler als die übrigen Kategorien wahrnehmen. Die Randkategorien decken somit einen geringeren Einstellungsbereich ab. Das bedeutet, dass bei Randkategorien in der Regel bereits ein geringerer Einstellungswechsel zur Belegung einer angrenzenden Kategorie führt, als dies im mittleren Skalenbereich der Fall ist.

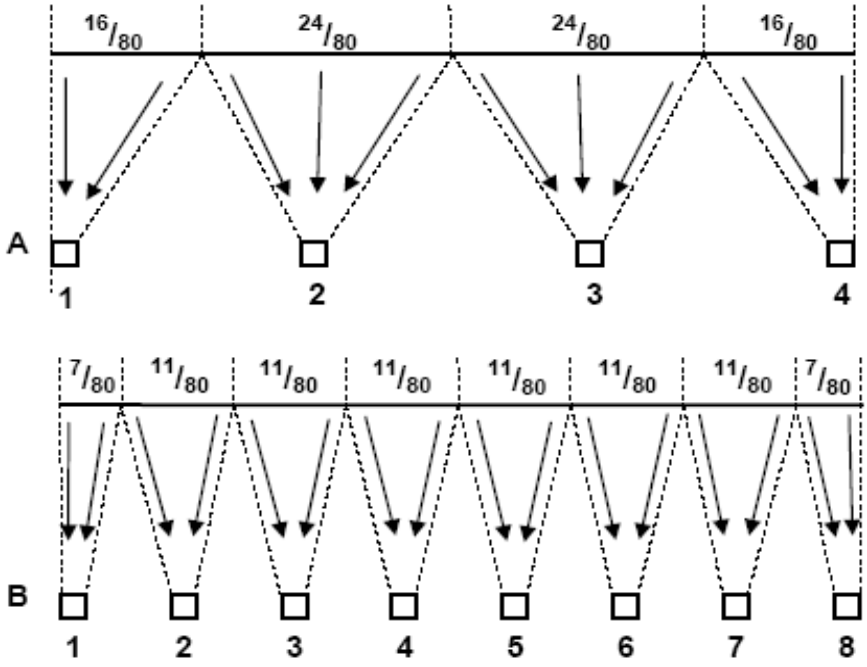

Abbildung 3: Transformation der VAS nach dem Modell der reduzierten Extreme; A zu vier Kategorien, B zu acht Kategorien.

5. Ergebnisse

5.1. Maße der zentralen Tendenz

Die Übereinstimmung der Mittelwerte bei den kontinuierlich wahrgenommenen Visuellen Analogskalen und den diskreten Kategorialskalen ist unabhängig vom Erhebungsmodus sehr hoch. Die Berechnung des arithmetischen Mittels bei Kategorialskalen ist sinnvoll, da sie zu Werten führt, die die Grundgesamtheit gut abbilden. Die Konvention, Daten aus Kategorialskalen als intervallskaliert zu betrachten, erweist sich also in beiden Designs als zweckmäßig. Offline-Design (Tabelle 1 und 2) und Online-Design (Tabelle 3) unterscheiden sich hinsichtlich der Mittelwertdifferenzen bei Visuellen Analogskalen und Kategorialskalen nicht voneinander.

	MIN	MAX	$\bar{x}$	s
ADDITIVER INDEX, 16 ITEMS				
KS-4	0	100	56,6	28,4
VAS, 80-stufig	0	100	56,4	24,7
DIMENSION FÜHRUNG, 6 ITEMS				
KS-4	0	100	51,2	26,5
VAS, 80-stufig	0	100	50,9	23,6
DIMENSION BEKANNTHEIT, 6 ITEMS				
KS-4	0	100	53,7	26,5
VAS, 80-stufig	0	100	53,0	24,9
DIMENSION HILFELEISTUNG, 4 ITEMS				
KS-4	0	100	69,2	25,9
VAS, 80-stufig	3	100	69,6	21,1

Tabelle 1: Vergleich von Mittelwerten und Standardabweichungen nach additivem Index und nach Dimensionen, bei vierstufigen Kategorialskalen und Visuellen Analogskalen im Offline-Design.

	MIN	MAX	$\bar{x}$	s
ADDITIVER INDEX, 16 ITEMS				
KS-8	0	100	58,2	26,2
VAS, 80-stufig	0	100	56,4	25,5
DIMENSION FÜHRUNG, 6 ITEMS				
KS-8	0	100	51,5	24,7
VAS, 80-stufig	0	100	51,2	24,4
DIMENSION BEKANNTHEIT, 6 ITEMS				
KS-8	0	100	55,5	26,9
VAS, 80-stufig	0	100	53,6	26,1
DIMENSION HILFELEISTUNG, 4 ITEMS				
KS-8	0	100	72,2	21,8
VAS, 80-stufig	4	100	68,2	22,3

Tabelle 2: Vergleich von Mittelwerten und Standardabweichungen, additiv und nach Dimensionen, bei achtstufigen Kategorialskalen und Visuellen Analogskalen im Offline-Design.

	MIN	MAX	$\bar{x}$	s
ADDITIVER INDEX, 16 ITEMS				
KS-4	0	100	58,9	32,1
VAS, 80-stufig	0	100	57,0	27,5
KS-8	0	100	60,4	28,3
DIMENSION FÜHRUNG, 6 ITEMS				
KS-4	0	100	55,5	31,2
VAS, 80-stufig	0	100	53,7	28,5
KS-8	0	100	57,8	29,1
DIMENSION BEKANNTHEIT, 6 ITEMS				
KS-4	0	100	61,7	34,4
VAS, 80-stufig	0	100	58,0	28,1
KS-8	0	100	61,5	28,2
DIMENSION HILFELEISTUNG, 4 ITEMS				
KS-4	0	100	60,3	29,9
VAS, 80-stufig	0	100	59,9	25,1
KS-8	0	100	62,7	26,9

Tabelle 3: Vergleich von Mittelwerten und Standardabweichungen, nach additivem Index und nach Dimensionen, bei vier- und achtstufigen Kategorialskalen und Visuellen Analogskalen im Online-Design.

5.2. Kategorienbelegungen

Ziel bei dem Vergleich der Kategorienbelegung ist es festzustellen, auf welche Art sich die Breite von Kategorialskalen adäquat beschreiben lässt. Zunächst werden Vergleiche zwischen vier- beziehungsweise achtstufigen Skalen und den beiden Transformationen der VAS durchgeführt. Dies geschieht sowohl mit den Offline- als auch mit den Onlinedaten. In einem weiteren Schritt werden die Ergebnisse des Offline-Designs mit denen des Online-Designs verglichen, um das Vorliegen von Modeeffekten zu prüfen. Je geringer die Differenzen zu den Kategorialskalen sind, desto besser ist die Beschreibung der Kategorienbreite bei Kategorialskalen.

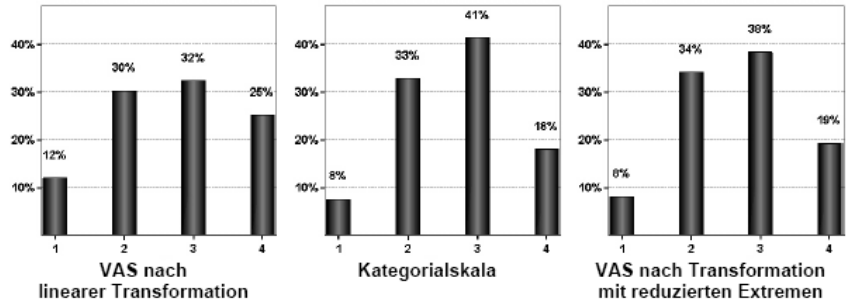

Abbildung 4: Vergleich Kategorienbelegung im Offline-Design: vierstufige Kategorialskala (Mitte) – VAS nach linearer Transformation (links) – VAS nach Transformation mit reduzierten Extremen.

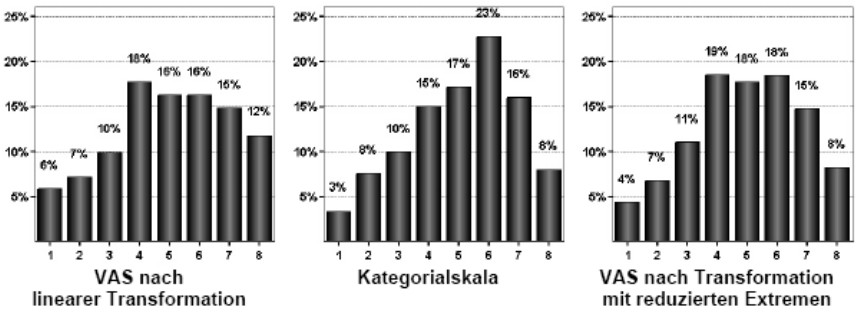

Abbildung 5: Vergleich Kategorienbelegung im Offline-Design: achtstufige Kategorialskala (Mitte) – VAS nach linearer Transformation (links) – VAS nach Transformation mit reduzierten Extremen.

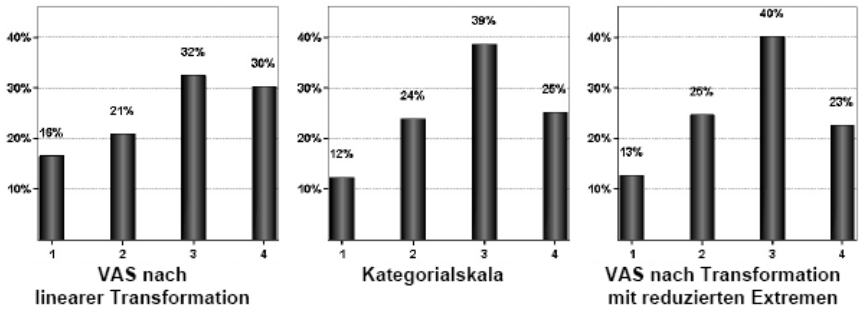

Abbildung 6: Vergleich Kategorienbelegung im Online-Design: vierstufige Kategorialskala (Mitte) – VAS nach linearer Transformation (links) – VAS nach Transformation mit reduzierten Extremen.

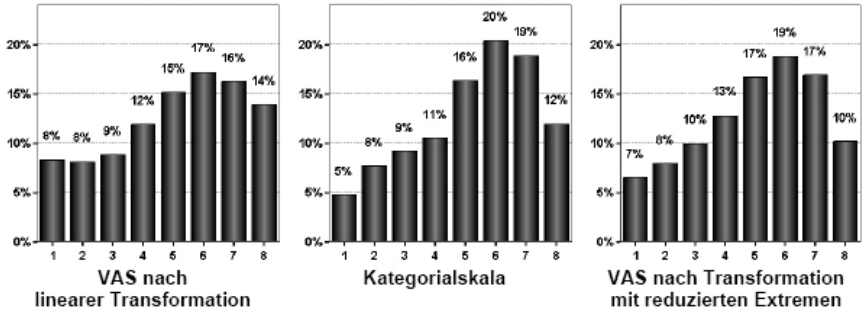

Abbildung 7: Vergleich Kategorienbelegung im Online-Design: achtstufige Kategorialskala (Mitte) – VAS nach linearer Transformation (links) – VAS nach Transformation mit reduzierten Extremen.

Bei der Betrachtung des Offline-Panels (Abbildung 4 und 5) ergibt sich ein ganz klares Bild. Sowohl bei den vier-, als auch bei den achtstufigen Kategorialskalen führt die Transformation mit reduzierten Extremen (mit der Ausnahme der Mittelkategorien einer Unterdimension bei der achtstufigen Skala) immer zu größeren Übereinstimmungen sowohl bei Gesamtbetrachtung als auch bei gesonderter Betrachtung der Mittel- und der Randkategorien. Im Online-Design (Abbildung 6 und 7) ist jeweils eine der drei Dimensionen bei Vergleichen mit den vier- und achtstufigen Kategorialskalen treffender durch lineare Transformation dargestellt. Sowohl bei der Betrachtung aller Items, als auch bei Betrachtung der beiden anderen

Dimensionen zeigt sich jedoch, dass das Modell der reduzierten Extreme zu größeren Übereinstimmungen mit den Kategorialskalen führt.

Die Befunde bei Offline- und Online-Design unterscheiden sich demnach grundsätzlich nicht voneinander und die Transformation mit reduzierten Extremen erweist sich unabhängig von dem Erhebungsmodus als zutreffender. Diese Unabhängigkeit von der Erhebungsart in Kombination mit unterschiedlich strukturierten Stichproben führt zu der Schlussfolgerung, dass die Randkategorien bei Kategorialskalen weniger breit wahrgenommen werden als die mittleren Kategorien. Das im Modell der reduzierten Extreme beschriebene Verhältnis zwischen den Kategorien beschreibt die Wahrnehmung der Kategorien besser, als dies das lineare Modell der gleich breiten Kategorien vermag.

6. Zusammenfassung

Die Untersuchung der Kategorienbreiten zeigt, dass die einzelnen Kategorien von Befragten als unterschiedlich große Intensitätsintervalle wahrgenommen werden und die Kategorien somit nicht äquidistant benutzt werden. Die wesentliche messtheoretische Voraussetzung für die Anwendung parametrischer Verfahren ist somit nicht gegeben. Allerdings führt die Bestimmung des arithmetischen Mittels bei Kategorialskalen zu den gleichen Ergebnissen wie bei den intervallskalierten Visuellen Analogskalen. Die Verletzung des Intervallniveaus der Daten hat bei der Berechnung des arithmetischen Mittels de facto keine Auswirkungen, so dass es zu sinnvoll interpretierbaren Ergebnissen führt, wenn man die Daten aus Kategorialskalen wie intervallskaliert behandelt. Die weitgehende Unabhängigkeit der Befunde in Offline- und Online-Design ist ein Beleg dafür, dass es sich um echte Effekte und nicht nur um Artefakte handelt. Befragte weisen den äußeren Kategorien einen schmaleren Einstellungsbereich zu als den mittleren Kategorien. Bei der Interpretation der Randkategorien ist diese Wahrnehmungsdiskrepanz zu berücksichtigen. Den Randkategorien ist ein größeres Gewicht beizumessen als den inneren Kategorien; ihre Belegung stellt somit eine gewichtigere Aussage dar.

Literatur

Aitken, R. C. B. (1969), »Measurement of Feeling Using Visual Analogue Scales«, *Proceedings of the Royal Society of Medicine*, Jg. 62, S. 989–993.

Birnbaum, Michael H. (1978), »Differences and Ratios in Psychological Measurement«, in: Castellan, N. John/Restle, Frank (Hg.), *Cognitive Theory – Volume 3*, New York.

Couper, Mick P. u.a. (2004), *Evaluating the Effectiveness of Visual Analog Scales: A Web Experiment*. Präsentation bei der »RC33 Sixth International Conference on Social Science Methodology«, Amsterdam, August 2004.

Brunier, Gillian/Graydon, Jane (1996), »A Comparison of Two Methods of Measuring Fatigue in Patients on Chronic Haemodialysis: Visual Analogue versus Likert Scale«, *International Journal of Nursing Studies*, Jg. 33, S. 338–348.

Flynn, Darren/Schaik, Paul van/Wersh, Anna van (2004), »A Comparison of Multi-Item Likert and Visual Analogue Scales for the Assessment of Transactionally Defined Coping Function«, *European Journal of Psychological Assessment*, Jg. 20, H. 1, S. 49–58.

Jäger, Ruth (2004), »Konstruktion einer Ratingskala mit Smilies als symbolische Marken«, *Diagnostica*, Jg. 50, H. 1, S. 31–38.

Reips, Ulf-Dietrich (2002), »Standards for Internet-based Experimenting«, *Experimental Psychology*, Jg. 49, H. 4, S. 243–256 (http://www.psychologie.unizh.ch/sowi/reips/papers/exppsy/ExPsyReipsReprint.pdf).

Reips, Ulf-Dietrich (2003), »Online-Erhebungen in der wissenschaftlichen Sozialforschung«, in: Informationszentrum Sozialwissenschaften (Hg.), *Online-Erhebungen – Sozialwissenschaftliche Tagungsberichte Band 7*, Bonn.