



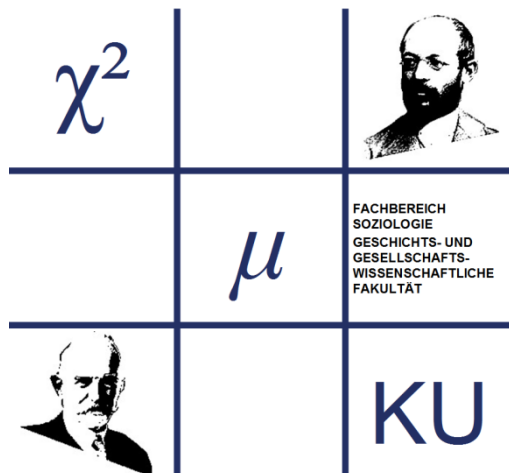
KATHOLISCHE UNIVERSITÄT
EICHSTÄTT-INGOLSTADT

EICHSTÄTTER BEITRÄGE ZUR SOZIOLOGIE

Metrische Eigenschaften des Semantischen Differentials zur Messung von
Selbst- und Fremdbeschreibungen

Stefanie Eifler, Natalja Menold und Sara Pinkas

Nr. 7
November 2015



Soziologische Forschungsberichte,
herausgegeben von Stefanie Eifler, Robert Schmidt und Joost van Loon

Eichstätter Beiträge zur Soziologie

Die Eichstätter Beiträge zur Soziologie erscheinen in unregelmäßiger Reihenfolge mehrmals im Jahr und können unter der angegebenen Adresse angefordert werden. Für die Inhalte sind allein die jeweiligen Autoren verantwortlich.

Redaktion: Dipl.-Soz. Ramona Kay, M. A.

Kontakt: Katholische Universität Eichstätt-Ingolstadt

Kapuzinergasse 2
85072 Eichstätt
+49 8421 93 21168
ramona.kay@ku.de

<http://www.ku.de/ggf/soziologie/schriftenreihe-eichstaetter-beitraege-zur-soziologie/>

Metrische Eigenschaften des Semantischen Differentials zur Messung von Selbst- und Fremdbeschreibungen

Stefanie Eifler, Natalja Menold und Sara Pinkas

Korrespondenzanschrift:

Prof. Dr. Stefanie Eifler
Catholic University of Eichstätt-Ingolstadt
Chair for Sociology and Empirical Social Research
Kapuzinergasse 2
85072 Eichstätt
Germany
stefanie.eifler@ku.de

Abstract:

In order to apply Semantic Differential Scales (SDS), linearity, symmetry and equidistance of scale points must be met. Violations can be explained by the presence of so-called positivity bias or by the reading behavior of respondents. The metric properties of SDS related to the two explanations were analyzed on the basis of a randomized between-subjects experiment (split-ballot) which was part of a web survey of German residents (internet users, n=537), comparing classical, unipolar and meaningless SDS. The metric properties of SDS are investigated by the variability and central tendencies of the data as well as by means of Multiple Correspondence Analysis (MCA). The results reveal that a) the extent of positivity bias was rather negligibly in the bipolar SDS, b) bipolar SDS allow for linear ordinal measurement, c) the symmetry of the data is rather observed in the case of bipolar SDS but not in the case of meaningless SDS.

Um Semantische Differentiale (SD) anwenden zu können, müssen die Linearität, Symmetrie und Äquidistanz der Skalenpunkte der einzelnen Skalen des Semantischen Differentials (SSD) vorausgesetzt werden. Verletzungen dieser Voraussetzungen können durch den sogenannten Positivity Bias oder die Leserichtung der Befragten erklärt werden. Die metrischen Eigenschaften der SSD im Zusammenhang mit den beiden Erklärungen wurden auf der Grundlage eines randomisierten Between-Subjects Experiment (Split-Ballot), das Teil einer Umfrage von deutschen Internetbenutzern (n = 537) war, analysiert. Dabei werden klassische, unipolare und unsinnige SSD verglichen. Die metrischen Eigenschaften der SSD werden sowohl anhand von Streuungsmaßen und Lagemaßen als auch mittels Multipler Korrespondenzanalyse (MCA) untersucht. Die Ergebnisse zeigen, dass a) das Ausmaß an Positivity Bias in den bipolaren SSD vernachlässigbar ist, b) bipolare SSD lineares und ordinales Messen erlauben, c) die Symmetrie der Daten eher im Falle bipolarer SSD aber nicht bei unsinnigen SSD beobachtet wird.

Keywords:

Semantic Differential Scales, Bipolar Rating Scales, Linearity, Equidistance, Symmetry, Multiple Correspondence Analysis

1 Einleitung

Seitdem Osgood, Suci und Tannenbaum (1957) das Semantische Differential (SD) entwickelt haben, gehört es zu den klassischen Methoden der Datenerhebung. In aktuellen Studien werden SDs häufig zur Erfassung von Konnotationen verwendet, die mit alltäglichen Begriffen wie Zirkus, Nebel, Donner, Tafel oder Ampel (Skrandies, 2011) assoziiert werden, oder zur Erfassung von sinnlichen Reizen, wie beispielsweise das Berühren von Pflanzen (Koga und Iwasaki, 2013). SDs werden meistens zur Analyse von Urteilen über Personen oder Personengruppen herangezogen, wie zum Beispiel über Homosexualität unter Männern und Frauen (Breen und Karpinski, 2013) oder über Arme, Reiche, Ältere und Teenager (Kervyn, Fiske und Yzerbyt, 2013). Einen Überblick über bisherige Anwendungen des SD gibt Heise (2010). Ein SD besteht aus einer beliebigen Anzahl an Skalen (Skalen des Semantischen Differentials, SSD), wobei die Endpunkte der Skalen jeweils durch bipolare Adjektive (gut-schlecht) verankert werden (Osgood et al., 1957). Die Befragten werden gebeten, spezifische Begriffe und Objekte anhand dieser Skalen zu bewerten. Die Urteile der Befragten werden analysiert, indem die Mittelwerte der Skalen bestimmt und in Form von Polaritätsprofilen dargestellt werden (Heise, 2010).

Obwohl das SD als eine klassische Methode der Datenerhebung gesehen wird, wurden bisher nur relativ wenige methodische Studien zum SD veröffentlicht. Die meisten von ihnen entstanden in den 1960er und 1990er Jahren und behandeln zwei Themen. Ein erster Forschungszweig bezieht sich auf die Dimensionalität des SDs. In den klassischen Studien von Osgood (1964) und Osgood, May und Miron (1975) konnte mittels Hauptkomponentenanalyse gezeigt werden, dass die Skalen des SD durch die drei Faktoren Evaluation, Potency und Activity repräsentiert werden. Diese drei Faktoren stellen die zentralen Dimensionen des SDs dar und werden als EPA-Struktur bezeichnet. Diese Struktur hat sich in kulturübergreifenden Studien, die im Anschluss durchgeführt wurden, als stabil erwiesen (Heise, 1977, 2010; Himmelfarb, 1993). Ein zweiter Forschungszweig bezieht sich auf die Analyse der metrischen Eigenschaften des SDs. Die Verwendung der SSD beruht auf bestimmten Annahmen über ihre metrischen Eigenschaften, nämlich Linearität, Symmetrie um den Skalenmittelpunkt und Äquidistanz der Antwortkategorien der SSD (Osgood et al., 1957). Hinsichtlich dieser metrischen Eigenschaften wurden bisher nur wenige Studien veröffentlicht. Während einige dieser Studien in den 1960er und 1970er Jahren durchgeführt wurden, wurden andere in den 1990er Jahren realisiert. Sowohl Himmelfarb (1993) als auch kürzlich Heise (2010) kommen in ihren systematischen Literaturreviews zum SD zu dem Schluss, dass die empirische Untersuchung der metrischen Eigenschaften des SDs zu widersprüchlichen Ergebnissen führte.

Das Ziel der vorliegenden Studie besteht daher darin, die metrischen Eigenschaften des SDs zu untersuchen, um die Gründe für die inkonsistenten und widersprüchlichen Ergebnisse

zu identifizieren. Zu diesem Zweck werden zunächst empirische Studien herangezogen, die den Effekt der verbalen Polarität von Ratingskalen auf die Qualität der mit diesen Skalen erhobenen Daten untersuchen (eine Zusammenfassung findet sich bei Menold und Raykov, 2015). Es wird dabei zwischen unipolaren und bipolaren Skalen unterschieden. Eine unipolare Ratingskala beinhaltet nur eine semantische Dimension, zum Beispiel Zustimmung (stimme nicht zu – stimme zu) oder Wichtigkeit (nicht wichtig – wichtig). Eine bipolare Skala hingegen umfasst zwei semantische Dimensionen, zum Beispiel Ablehnung und Zustimmung (lehne ab – stimme zu) oder Aktivität und Passivität (aktiv – passiv). Dementsprechend sind die klassischen SSD verbal bipolar verankert. Wir erwarten, dass zwei verschiedenen Arten der verbalen Verankerung der SSD zu Unterschieden im Antwortverhalten der StudienteilnehmerInnen führt. Diese Vermutung wird auf der Grundlage vorhandener Forschungsliteratur entwickelt und zu empirisch prüfbareren Hypothesen konkretisiert (Abschnitt 2). Im folgenden Abschnitt wird die Methode beschrieben (Abschnitt 3). In einem Split-Ballot-Experiment wurden verschiedene verbale Verankerungen von SSD realisiert. Die Befragten werden gebeten, anhand dieser Skalen sich und eine andere Person zu beschreiben. Im darauf folgenden Abschnitt werden die Ergebnisse der Studie berichtet (Abschnitt 4). Zusätzlich zu deskriptiven Maßen wird die Struktur jeder einzelnen Antwortskala durch eine Multiple Korrespondenzanalyse analysiert und visualisiert. Schließlich werden die Ergebnisse interpretiert, und es werden weiterführende Überlegungen entwickelt (Abschnitt 5).

2 Qualität der Skalen des Semantischen Differentials

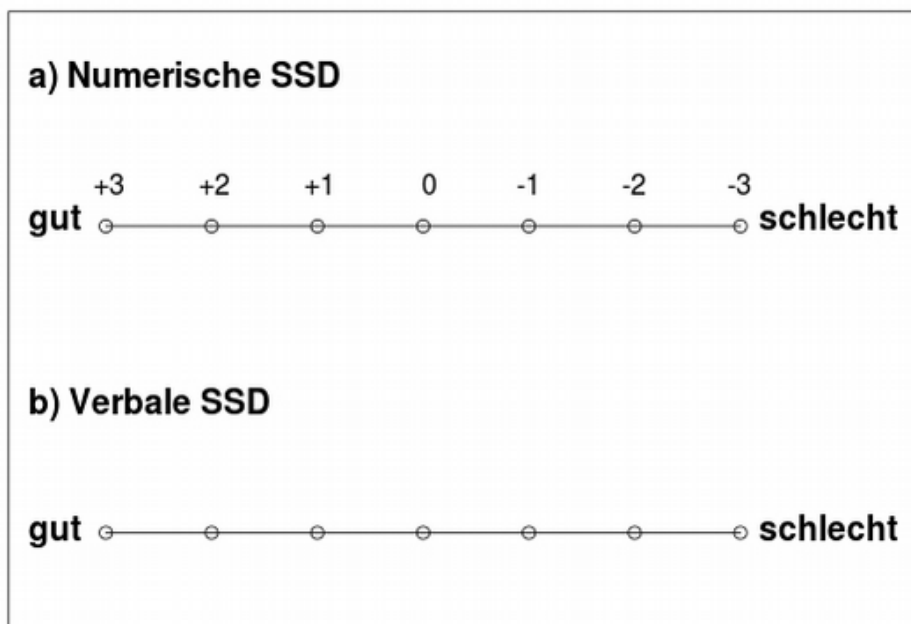
Die metrischen Eigenschaften der einzelnen bipolaren Skalen Semantischer Differentiale stehen im Mittelpunkt der vorliegenden Studie. In einem ersten Schritt werden die metrischen Eigenschaften der SSD beschrieben, im Hinblick auf die verfügbare Literatur diskutiert und in den Kontext der bisher durchgeführten Forschung zur Qualität von Erhebungsdaten eingeordnet (Abschnitt 2.1). Auf dieser Grundlage werden in einem zweiten Schritt die Forschungsfragen und Hypothesen der vorliegenden Studie abgeleitet (Abschnitt 2.2).

2.1 Die metrischen Eigenschaften der Skalen des Semantischen Differentials und ihre empirische Analyse

Mit einem SD kann die konnotative Bedeutung von Konzepten oder Objekten quantitativ erfasst werden (Osgood et al., 1957). Die Anwendung Semantischer Differentiale zur Erhebung der Konnotation bestimmter Begriffe findet ihre Begründung in der psychophysischen Wahrnehmungsforschung. Stevens' (1975) Methode der Category Estimation folgend werden verbale Urteile über die konnotative Bedeutung von Konzepten oder Objekten entlang einer vorher definierten Skala mit verbal verankerten Endpunkten (z. B. gut-schlecht) gemessen. Die Methode ist

durch das Modell des Semantischen Raums begründet. Der Semantische Raum wird als linearer euklidischer Raum angenommen. Es wird weiterhin angenommen, dass jede SSD eine lineare Funktion durch den Ursprung des Semantischen Raums bildet. Die Annahme der Linearität zwischen den Antwortkategorien, die durch ein Adjektivpaar verankert sind, spiegelt die Idee wieder, dass diese Adjektivpaare Antonyme sind und echte Gegensätze im Semantischen Raum darstellen. Linearität impliziert, dass die Antwortkategorien entlang jeder SSD als geordnet angenommen werden. Die Annahme, dass die Skalenpunkte symmetrisch um den Nullpunkt der Skala verteilt sind bedeutet auch, dass jede SSD einen Nullpunkt in ihrer Mitte hat, welcher die Skala in zwei symmetrische Teile zerlegt. Idealerweise sind der Nullpunkt der Skala und der subjektiv wahrgenommene Mittelpunkt identisch. Die Annahme der Äquidistanz der Skalenpunkte bezieht sich auf die Idee, dass die Befragten die Abstände zwischen den Antwortkategorien der Skala als gleich groß wahrnehmen. SSD zu beantworten unterscheidet sich nicht von der Beantwortung von Ratingskalen, die entwickelt wurden, um andere Einstellungen zu messen. SSD werden manchmal nur mit verbaler Verankerung und manchmal mit zusätzlich numerisch verankerten Skalenpunkten präsentiert (Abbildung 1).

Abbildung 1: Numerische und verbale Form der Skalen des Semantischen Differentials



Die vorliegende Forschungsliteratur zu den metrischen Eigenschaften des SD wurde hauptsächlich in den 1960er, 1970er und 1990er Jahren veröffentlicht. Es liegt keine aktuelle Forschung zu diesem Thema vor. In systematischen Reviews verweisen Heise (1977, 2010) und Himmel-
 farb (1993) auf inkonsistente und widersprüchliche Ergebnisse in den Studien, in denen Lineari-

tät, Symmetrie und Äquidistanz der Antwortkategorien der SSD untersucht werden. Manche Autoren wie Anderson (1986), Baxter (1962), Bentler (1969), und Mann, Phillips und Thompson (1979) finden empirische Unterstützung für die Annahme der Symmetrie der einzelnen SSD. Andere wie Attneave (1959), Ross und Leavy (1960), und Schriesheim, Cogliner, Newmark und Lowensohn (1994) finden keine Evidenz dafür, dass die Annahme der Symmetrie erfüllt wird. Insbesondere konnte Messick (1957) mittels der Methode der sukzessiven Intervalle zeigen, dass die wahrgenommenen Kategorien auf der linken Seite von neun SSD größer ausfallen als die Kategorien auf der rechten Seite der entsprechenden Skalen. Somit war der wahrgenommene Mittelpunkt jeder SSD leicht auf die rechte Seite mit der negativen Bedeutung verschoben. Zusätzlich platzierten die Befragten mehr Urteile auf der linken Seite mit der positiven Bedeutung. Entsprechend fanden Schriesheim et al. (1994), dass von 18 analysierten SSD nur sechs Skalen Symmetrie aufwiesen. Hinsichtlich der Äquidistanz schließen Osgood et al. (1957, S. 146) und Schriesheim et al. (1994), dass die Annahme der Äquidistanz der Antwortkategorien jeder SSD zutreffend ist. Jedoch kann nach Fuchs (1974), Messick (1957) und Schriesheim et al. (1994) die bloße Annahme, dass Äquidistanz gegeben ist, nicht gerechtfertigt werden. Insbesondere fand Messick (1957), dass die Kategoriegrenzen von neun SSD nicht den Positionen entsprechen, die aufgrund gleicher Intervallabstände zu erwarten wären, sondern nur lediglich ähnlich platziert sind. Zwar wird die Linearität der Antwortkategorien aufgrund sachlogischer Überlegungen angenommen, jedoch konnte sie unserem Wissen nach bislang nicht empirisch unterstützt werden (Heise, 1977, 2010; Himmelfarb, 1993).

Vor dem Hintergrund der Ergebnisse bisheriger Studien zeichnet sich durch die empirischen Analysen kein klares Bild der metrischen Eigenschaften der SSD ab. Zusätzlich fehlt es, trotz des weit verbreiteten Gebrauchs des SD in verschiedenen Zweigen aktueller empirischer Forschung (Abschnitt 1), an Studien, die die metrischen Eigenschaften der SSD untersuchen. Drei Überlegungen können angeführt werden, um die Verletzungen der Annahmen der metrischen Eigenschaften der SSD zu erklären.

Erstens ist es fragwürdig, ob die verbalen Verankerungen der SSD tatsächlich Antonyme sind und von den Befragten also solche wahrgenommen werden. Verhagen, van den Hoof und Meents (2015) bemerken, dass es sich in der Praxis als schwierig gestaltet, exakte Gegenteile zu finden, was einen Effekt auf Linearität und Symmetrie der SSD ausüben könnte. Hinsichtlich der semantischen Bipolarität schlagen die Autoren eine Methode für die Entwicklung der SSD vor, welche semantisch bipolare Adjektive aufzufinden erlaubt.

Zweitens sind die Verteilungen der Antworten von Befragten auf SSD häufig rechtsschief (Messick, 1957; Schriesheim et al., 1994). In den Anwendungen des SD wird meistens das positiv konnotierte Adjektiv, gegebenenfalls auch das positive numerische Label auf die linke Seite der Skala platziert. Entsprechend ist eine SSD wie „gut-schlecht“ numerisch von +3 bis -3 in

Leserichtung verankert (Abbildung 1). Die bei SSD beobachtete Verteilungsschiefe könnte mit der Leserichtung zusammenhängen, welche zur Folge hat, dass die Befragten der linken Seite der Ratingskala mehr Aufmerksamkeit als der rechten Seite geben. Menold, Kaczmirek, Lenzner und Neusar (2014) unterstützen diese Interpretation aufgrund der Ergebnisse einer Eye-Tracking-Studie.

Drittens gibt es ähnliche Probleme im Feld der Einstellungsforschung. Ergebnisse aus Studien zur Qualität von Umfragedaten können dabei helfen, die Forschung zu metrischen Eigenschaften der SSD weiter zu verfeinern. In der Einstellungsforschung werden sowohl bipolare („stimme voll zu – lehne voll ab“) als auch unipolare Ratingskalen („stimme überhaupt nicht zu – stimme voll zu“) verwendet. Nach Schaeffer und Presser (2003) wurde häufig ein übermäßiger Gebrauch der positiven Seite der Skala im Falle bipolarer Ratingskalen beobachtet. Dieser sogenannte Positivity Bias setzt sich aus der zusätzlichen Verwendung numerischer Labels in bipolaren Skalen, z.B. von -3 bis +3 oder von -5 bis +5, im Vergleich mit bipolaren Skalen, z.B. von 0 bis 7 oder 0 bis 10, zusammen (Schwarz, Knäuper, Hippler, Noelle-Neumann und Clark, 1991; Schaeffer und Presser, 2003). Als Konsequenz daraus wurden negative numerische Labels in der Einstellungsforschung vermieden. Aktuelle Studien zu der verbalen Bipolarität von Ratingskalen lieferten gemischte Ergebnisse. Während Saris und Gallhofer (2007) keinen Effekt auf die Polarität der Messergebnisse fanden, berichtet Krebs (2012) höhere Alpha-Koeffizienten für bipolare Skalen. Sie befasste sich auch mit dem Effekt von verbalem oder numerischem positiven oder negativen Labeling, was den Alpha-Koeffizient nicht beeinflusste. Mittels alternativer Methoden zur Schätzung der Reliabilität, fanden Menold und Raykov (2015) eine höhere Reliabilität bei unipolaren Skalen. Diese gemischten Ergebnisse deuten darauf hin, dass sich die verbale Polarität auf das Antwortverhalten und die Messqualität auswirken könnte.

Zusammenfassend können also verschiedene Faktoren zu den widersprüchlichen Ergebnissen der bisherigen methodischen Studien zu den metrischen Eigenschaften von SSD geführt haben. All diese Gründe hängen entweder indirekt oder direkt mit den verschiedenen Arten der verbalen und numerischen Verankerung der Ratingskalen zusammen. In der vorliegenden Studie werden allerdings nur die Effekte verbaler Verankerung bei Ratingskalen untersucht. Im nächsten Abschnitt werden die Forschungsfragen entwickelt, die mit den möglichen Effekten der verbalen Verankerung von SSD zusammenhängen.

2.2 Forschungsfragen und Hypothesen

Wenn ein SD im Rahmen einer Umfrage verwendet wird, besteht die Aufgabe der Befragten darin, Urteile über die interessierenden Begriffe oder Personen oder Personengruppen auf den entsprechenden verbal bipolar verankerten Skalen abzugeben. Dabei wird angenommen, dass

die verbalen Verankerungen der Skala Antonyme sind, mit anderen Worten also echte sprachliche Gegensätze darstellen. Nach Tourangeaus (1984, Tourangeau, Rips und Rasinski, 2000) Modell des Antwortprozesses muss der Befragte zuerst den Inhalt der Frage verstehen. Sodann muss der Befragte Informationen aus seinem Gedächtnis abrufen, ein Urteil bilden und letztendlich seine Antwort in die vorgegebenen Antwortkategorien einordnen. Wird dieses Modell auf die Verwendung des SD angewandt, besteht die Herausforderung für die Befragten darin, die verbalen Verankerungen der einzelnen SSD als echte Antonyme zu erkennen, und ihre Antworten auf dem gedachten Kontinuum zu platzieren.

Im Hinblick auf dieses Kontinuum und die entsprechenden Ratingskalen formulieren Osgood et al. (1957) sehr strenge Annahmen. Sie nehmen an, dass die Urteile der Befragten über alle Punkte der siebenstufigen Ratingskala verteilt sind. Weiterhin wird postuliert, dass die Skalenpunkte symmetrisch um den Mittelpunkt der Skala liegen und äquidistant sind. Ergebnisse von Studien zur Qualität von Erhebungsdaten stehen jedoch mit dieser Annahme im Widerspruch. Erstens muss in diesem Zusammenhang genannt werden, dass Äquidistanz der Skalenpunkte in den Sozialwissenschaften kaum vorkommt (Viellemann und Wilkinson, 1994). Zweitens lieferten viele Studien Evidenz dafür, dass der positive Pol der Ratingskalen bei verbal bipolar verankerten Skalen bevorzugt wird (Abschnitt 2.1). Darüber hinaus konnten eye-tracking-Studien zeigen, dass aufgrund der Leserichtung von links nach rechts die Befragten dem linken Pol einer Skala mehr Aufmerksamkeit schenken als dem rechten Pol (Abschnitt 2.1). In Anbetracht dieser Ergebnisse wird eine solche Verzerrung auch bei den einzelnen Skalen eines SDs erwartet. Effekte der Leserichtung können nicht nur durch Vergleich von SSD mit verbal bipolaren Verankerungen (gut-schlecht) und SSD mit verbal unipolaren Verankerungen (gut-nicht gut) untersucht werden, sondern auch durch Vergleich der verbal bipolar verankerten Skalen (gut-schlecht) mit unsinnigen Skalen. Dabei werden solche SSD als unsinnig bezeichnet, bei denen die rechte Verankerung der SSD aus einem zufällig ausgewählten und inhaltlich unangemessenen Adjektiv besteht (gut-träge). Falls Befragte den rechten Anker der unsinnigen SSD tatsächlich nicht beachten, würden sie nicht bemerken, dass der linke und rechte Anker nicht zusammenpassen.

Unter Berücksichtigung dieser theoretischen Überlegungen können zwei Hypothesen abgeleitet und empirisch analysiert getestet werden. Basierend auf der Idee des SDs sollten Skalen mit bipolaren Verankerungen zu Antworten führen, die gleichmäßig über den Wertebereich verteilt sind. Es wird angenommen, dass die Befragten bei der Urteilsbildung sowohl den linken als auch den rechten Pol einer SSD berücksichtigen. Anstelle der sehr strengen Annahme der Äquidistanz wird die Annahme einer ordinalen Anordnung der einzelnen Skalenpunkte von SSD untersucht. Es sollte also möglich sein, die Antwortkategorien entlang einer zumindest ordinalen (und im besten Fall metrischen) Skala anzuordnen. Diese Ideen werden anhand eines

systematischen Vergleichs zwischen verschiedenen Arten der verbalen Verankerung von SSD untersucht. Konkret werden Skalen mit verbal bipolarer Verankerung mit Skalen mit verbal unipolarer Verankerung und Skalen mit verbal unsinniger Verankerung verglichen.

Es würde für das SD sprechen, wenn die Antworten keinen Positivity Bias aufweisen. Das wäre beim Vergleich der Skalen mit verbal unipolarer und unsinniger Verankerung der Fall, wenn allgemein eine größere Streuung der Antworten sowie eine größere Anzahl an Urteilen auf dem rechten Pol und eine stärkere zentrale Tendenz zum rechten Pol auftreten würden. Um die Effekte der verbalen Verankerung auf die Urteile der Befragten unabhängig von den Effekten numerischer Verankerung analysieren zu können, wurden die SSD lediglich mit rein verbalen Verankerungen präsentiert (Abbildung 1, Abschnitt 2.1 und 3.2). Entsprechend wurden die folgenden zwei Hypothesen formuliert:

Hypothese 1: Diese Hypothese bezieht sich auf den Vergleich zwischen der bipolaren und unipolaren Version des Semantischen Differentials.

1.1 Die Antworten auf die bipolare Version jeder SSD weisen mehr Variation über die Antwortkategorien hinweg auf als die der unipolaren Version.

1.2 Es liegen in der bipolaren Version mehr Urteile auf der rechten Seite, und es gibt eine stärkere zentrale Tendenz in Richtung des rechten Pols als bei der unipolaren Version.

1.3 Die Antworten auf die bipolare und unipolare Version der SSD ermöglichen beide eine ordinale Anordnung der Antwortkategorien.

Hypothese 2: Diese Hypothese bezieht sich auf den Vergleich zwischen der bipolaren und unsinnigen Version des Semantischen Differentials.

2.1 Die Antworten auf die bipolare Version jeder SSD weisen mehr Variation über die Antwortkategorien hinweg auf als die der unsinnigen Version.

2.2 Es liegen in der bipolaren Version mehr Urteile auf der rechten Seite, und es gibt eine stärkere zentrale Tendenz in Richtung des rechten Pols als bei der unsinnigen Version.

2.3 Die Antworten auf die bipolare Version der SSD ermöglichen eine ordinale Anordnung der Antwortkategorien.

3. Methode

3.1 Experimentelles Design

Die vorliegende empirische Studie wurde als Split-Ballot Experiment auf der Basis eines einfaktoriellen Between-Subjects Designs durchgeführt. Der Faktor umfasst drei Stufen, wobei jede Stufe einer Art der verbalen Verankerung eines Semantischen Differentials entspricht. Es wurden jeweils Selbst- und Fremdbeschreibungen erfasst. Der Arbeit von Osgood et al. (1957) wurden Adjektive für die drei Dimensionen des Semantischen Differentials, Evaluation, Potency und Activity, entnommen. Es wurden nur kontrastierende Adjektive, die sich auf die Bewertung von Personen beziehen, ausgewählt und in die deutsche Sprache übertragen. In der ersten Versuchsbedingung wurde den Probanden das klassische Semantische Differential von Osgood et al. (1957) mit bipolaren verankerten Skalen (Experimentelle Bedingung 1: verbal bipolar verankerte SSD) präsentiert. Den Probanden der zweiten Versuchsbedingung wurde ein Semantisches Differential präsentiert, dessen Skalen verbal unipolar verankert waren. Die linke Seite umfasste die Adjektive des klassischen SD, während die rechte Seite Verneinungen dieser Adjektive mit „nicht“ umfasste (Experimentelle Bedingung 2: verbal unipolar verankerte SSD). In der dritten Versuchsbedingung wurde den Probanden ein Semantisches Differential präsentiert, dessen Skalen auf der rechten Seite mit unpassenden bzw. unsinnigen Adjektiven verankert wurden. Es wurde die klassische Form des SD verwendet, jedoch waren die Adjektive auf der rechten Seite in einer anderen Reihenfolge, sodass die einzelnen Skalen unsinnige Adjektivpaare bildeten, da auf der rechten Seite keine echten Gegensätze zu der linken Seite standen (Dritte Versuchsbedingung: unsinnige SSD). Abbildung 2 zeigt die Realisierung des SD in den drei Versuchsbedingungen.

Abbildung 2: Experimentelle Bedingungen – Verbale Verankerung der Skalen des Semantischen Differentials (SSD)*

Experimentelle Bedingung 1 Bipolare SSD	Experimentelle Bedingung 2 Unipolare SSD	Experimentelle Bedingung 3 Unsinnige SSD
<i>Evaluation</i>		
gut – schlecht	gut – nicht gut	gut – träge
optimistisch – pessimistisch	optimistisch – nicht optimistisch	optimistisch – gelassen
positiv – negativ	positiv – nicht positiv	positiv – mild
vollkommen – unvollkommen	vollkommen – nicht vollkommen	vollkommen – nachgiebig
zeitgemäß – altmodisch	zeitgemäß – nicht zeitgemäß	zeitgemäß – schwach
<i>Potency</i>		
hart – weich	hart – nicht hart	hart – passiv
tiefgründig – oberflächlich	tiefgründig – nicht tiefgründig	tiefgründig – feminin
maskulin – feminin	maskulin – nicht maskulin	maskulin – oberflächlich
streng – mild	streng – nicht streng	streng – negativ
stark – schwach	stark – nicht stark	stark – altmodisch
hartnäckig – nachgiebig	hartnäckig – nicht hartnäckig	hartnäckig – unvollkommen
<i>Activity</i>		
aktiv – passiv	aktiv – nicht aktiv	aktiv – weich
reizbar – gelassen	reizbar – nicht reizbar	reizbar – pessimistisch
energisch – träge	energisch – nicht energisch	energisch – schlecht

* Herkunft der bipolaren Adjektive des Semantischen Differentials: Osgood, Suci und Tannenbaum, 1957.

3.2 Erhebungsinstrument und unabhängige Variablen

In jeder Versuchsbedingung wurden die Probanden gebeten, eine Fremd- und eine Selbstbeschreibung vorzunehmen. Zunächst wurden die Befragten gebeten, eine andere Person mit Hilfe der angegebenen Adjektive zu beschreiben, wobei hier der deutsche Bundespräsident Joachim Gauck, also ein bekannter deutscher Politiker, als Beispiel gewählt wurde (Wörtliche Formulierung: „Bitte beurteilen Sie den Bundespräsidenten Herrn Joachim Gauck anhand der folgenden Eigenschaftswörter!“). Danach wurden die Probanden gebeten, sich selbst zu beschreiben, und zwar auch anhand der angegebenen Adjektive (Wörtliche Formulierung: Bitte beurteilen Sie nun sich selbst anhand dieser Eigenschaftswörter!). Abbildung 3 zeigt die Darbietungsform der SSD beispielhaft anhand des Adjektivpaars „gut-schlecht“.

Abbildung 3: Darbietungsform der Skalen des Semantischen Differentials



Die SSD wurden mit sieben Antwortoptionen präsentiert. Um den Positivity Bias, der mit der Verwendung negativer Zahlen zusammenhängt, und mögliche Verzerrungen im Antwortverhalten aufgrund von unabhängigen Effekten durch numerische Labels (Teil 2 der vorliegenden Studie) zu kontrollieren, wurden nur die Pole des SD verbalisiert, während die dazwischen liegenden Antwortoptionen nicht numerisch gelabelt wurden (Abbildung 3).

3.3 Stichprobe

Die Daten wurden im Sommer 2012 am Gesis Leibniz-Institut für Sozialwissenschaften im Rahmen eines probabilistischen Online Access Panels an mindestens 18-jährigen Personen mit deutscher Staatsangehörigkeit erhoben. Das Panel erfasste Personen, welche das Internet für private Zwecke nutzen. Details zum Stichprobendesign, zur Anwerbung von Studienteilnehmer/innen und Stichprobenmerkmalen können Struminskaya, Kaczmirek, Schaurer und Bandilla (2014) entnommen werden.

Die Daten wurden im Juli und August 2012 erhoben. Insgesamt wurden 644 Personen eingeladen, an der Umfrage teilzunehmen. Von diesen antworteten $n = 537$, womit sich eine Rücklaufquote von 82,91% ergab. Die Versuchspersonen wurden zufällig einer der drei Experimentalbedingungen zugeordnet. Der ersten Experimentalbedingung (bipolare SSD) wurden insgesamt $n = 175$ Versuchspersonen zugeordnet, der zweiten Experimentalbedingung (unipolare SSD) $n = 188$ Versuchspersonen, und der dritten Experimentalbedingung (unsinnige SSD) $n = 174$ Versuchspersonen. Die Stichprobe umfasste 53,79 % Frauen und 46,21 % Männer. Das mittlere Alter der Versuchspersonen betrug 43,2 Jahre mit einer Standardabweichung von 14,84 Jahren. Von den Versuchspersonen hatten 21,91 % keinen oder einen Hauptschulabschluss, 14,74 % hatten die Mittlere Reife, und 63,35 % hatten als Schulabschluss die Hochschulreife angegeben. Ungefähr 66,60 % der Versuchspersonen waren berufstätig. Um eine Verzerrung der Ergebnisse aufgrund wesentlicher Stichprobenmerkmale zu vermeiden, wurde die Zusammensetzung der drei Experimentalgruppen anhand soziodemographischer Merkmale verglichen. Wie der Tabelle A-1 entnommen werden kann, ergaben sich hinsichtlich Geschlecht, Alter, Bildung und Erwerbstätigkeit keine signifikanten Unterschiede zwischen den drei Experimentalgruppen.

3.4 Datenanalyse

Die Datenanalyse der beiden oben formulierten Untersuchungshypothesen erfolgte in mehreren, aufeinander aufbauenden Schritten. Die erste Hypothese bezieht sich auf den Vergleich der verbal bipolar verankerten SSD mit den verbal unipolar verankerten SSD, die zweite Hypothese

bezieht sich auf den Vergleich der verbal bipolar verankerten SSD mit den verbal unsinnig verankerten SSD. Die einzelnen Analyseschritte sind jeweils für beide Hypothesen analog:

Für die Analyse der Hypothesen 1.1 und 2.1 wird das Ausmaß der Variation der einzelnen Antworten über die sieben Antwortkategorien jeder SSD für Selbst- und Fremdbild bestimmt und über die beiden Arten der verbalen Verankerung verglichen. Hierzu wird für jede Skala ein Streuungsmaß für nominale Daten, und zwar der von Gibbs (1975) vorgeschlagene Index qualitativer Variation (IQV) M2 bestimmt:

$$IQV = \frac{K}{K-1} (1 - \sum_{i=1}^K p_i^2) \quad (1)$$

Gleichung 1 gibt die Formel für den Index qualitativer Variation (IQV) an, wobei K der Anzahl der Antwortkategorien entspricht, p_i die relative Häufigkeit jeder Antwortkategorie darstellt, und der Faktor der Standardisierung auf den Wertebereich von 0 bis 1 dient (Gibbs, 1975, S. 472). Der IQV nimmt den Wert 0 an, wenn nur eine Antwortkategorie einer Skala besetzt ist. Er erreicht den Wert 1, wenn alle Antworten gleichmäßig über alle Antwortkategorien verteilt sind, wenn mit anderen Worten jede Kategorie gleich häufig gewählt wird. Zusätzlich wird der Quartilabstand ($QA = Q75 - Q25$) benutzt, um das Ausmaß der Streuung der Antworten über den Wertebereich jeder SSD zu untersuchen.

Für die Analyse der Hypothesen 1.2 und 2.2 werden der Median und die Rangsummen für jede SSD für Fremd- und Selbstbeschreibungen bestimmt. Die vermutete stärkere zentrale Tendenz zum rechten Pol und die höheren Werte von verbal bipolar verankerten SSD im Vergleich zu verbal unipolar oder unsinnig verankerten SSD wird anhand eines Mediantests und anhand eines Kruskal-Wallis-H-Tests der Rangsummen untersucht (Kruskal und Wallis, 1952).

Die Hypothesen 1.3 und 2.3 werden anhand multipler Korrespondenzanalysen (MCA) untersucht. Die MCA basiert auf der Annahme, dass es zwischen den Antwortkategorien der SSD keine Beziehungen gibt. Das Ziel dieser Analyse besteht darin, die Struktur der Beziehungen zwischen den Antwortkategorien der SSD zu analysieren und in einem mehrdimensionalen Raum zu visualisieren. Im vorliegenden Zusammenhang geht es insbesondere darum, die Struktur der Beziehungen zwischen den Antwortkategorien der SSD zwischen den drei experimentellen Bedingungen systematisch miteinander zu vergleichen, nicht jedoch um einen Modelltest im engeren Sinne. Es wird daher für jede experimentelle Bedingung gesondert und getrennt nach Selbst- und Fremdbeschreibungen und nach den drei Dimensionen Evaluation, Potency und Activity jeweils eine MCA gerechnet. Grundlage jeder MCA ist eine Indikatormatrix, deren Zeilen durch die Befragten und deren Spalten durch die Antwortkategorien der jeweiligen SSD gebildet werden. Die Einträge der Indikatormatrix bestehen nur aus den Werten 0 und 1, wobei eine 0 bedeutet, dass die betreffende Kategorie nicht beobachtet wurde, und eine 1 bedeutet, dass

die betreffende Kategorie beobachtet wurde. Im Rahmen dieser Analysen wurden die einzelnen Antwortkategorien der verbal verankerten SSD von links nach rechts aufsteigend nummeriert (z.B. gut-schlecht: 1-7). Diese Nummerierung wird in der nachfolgenden Darstellung der Ergebnisse verwendet. Im Rahmen einer MCA wird das Verfahren der Simple Correspondence Analysis, das im Anschluss an eine Normalisierung der absoluten Häufigkeiten eine Singulärwertzerlegung mit anschließender Reskalierung der Koordinaten beinhaltet, auf eine Indikatormatrix generalisiert (Blasius und Greenacre, 1998; Greenacre, 1984; Greenacre und Blasius, 1994; Weller und Romney, 1990). Gleichung 2 zeigt die Grundstruktur der MCA in formalisierter Struktur, wobei U die Matrix der Zeilenvektoren, V^T die transponierte Matrix der Spaltenvektoren und D die Diagonalmatrix der Eigenwerte ist.

$$H = UDV^T \quad (2)$$

Im Rahmen der MCA werden die optimal scores für die Spaltenvariablen, mit anderen Worten die Antwortkategorien der jeweils einbezogenen SSD ermittelt und dargestellt. Es wird gefragt, ob sich die Antwortkategorien im Sinne einer optimalen Skala, die sich vom positiven Pol der jeweiligen SSD bis zum negativen Pol der jeweiligen SSD aufspannen lässt, anordnen lassen. Unter der Annahme, dass die einbezogenen Variablen nominal skaliert sind, wird gefragt, ob sich die einzelnen Antwortkategorien in eine ordinale Anordnung bringen lassen. In Bezug auf die sogenannte französische und niederländische Tradition der MCA wurde von Di Franco (2015) gezeigt, dass die Ergebnisse einer MCA mit den Ergebnissen einer CatPCA, die unter der Annahme nominaler Daten gerechnet wird, übereinstimmen. Um im vorliegenden Zusammenhang zu verhindern, dass geringe Anteilswerte zu überhöhten Ausprägungen der Koordinatenwerte führen, und um zu ermöglichen, dass alle Koordinatenwerte tatsächlich im Sinne der Struktur der Beziehungen zwischen den Antwortkategorien interpretiert werden können, wurden Skalen und Antwortkategorien ausgewählt, bei denen die relativen Häufigkeiten p der Antwortkategorien j $p_j \geq .05$ sind. In einem ersten Schritt wurden deshalb für alle experimentellen Bedingungen und getrennt nach Selbst- und Fremdbeschreibungen und nach den drei Dimensionen Evaluation, Potency und Activity vorbereitende multiple Korrespondenzanalysen gerechnet. Im Anschluss daran wurden alle Skalen des semantischen Differentials ausgewählt, in denen in allen experimentellen Bedingungen dieselben Antwortkategorien in ausreichender Stärke besetzt waren oder zu ausreichend besetzten Kategorien zusammengefasst werden konnten. Die Interpretation erfolgt auf Basis der zwei- und eindimensionalen Lösungen der MCA. Die erste Dimension zeigt die Zusammenhgangsstruktur der Antwortkategorien, während die zweite Dimension typischerweise eine methodeninduzierte kurvilineare Anordnung der Repräsentationen der einzelnen Antwortkategorien in Form einer Parabel zeigt (horseshoe-Effekt, Kendall, 1971;

Greenacre, 1984, S. 226-232). Wenn sich in der zweidimensionalen Darstellung dieser methodeninduzierte Effekt zeigt, wird nur die erste Dimension der MCA im Sinne unserer Fragestellung untersucht und interpretiert. Die standard coordinates werden im Hinblick auf ihre Anordnung auf der ersten Dimension untersucht. Dabei werden Abweichungen von der Ordinalität am Anteil solcher Abweichungen von der Anzahl der jeweils möglichen Abweichungen gemessen. Aufgrund einer visuellen Inspektion werden die Projektionen im Hinblick auf die Symmetrie und Äquidistanz beurteilt. Die in der Korrespondenzanalyse ermittelten standard coordinates werden für alle experimentellen Bedingungen und alle jeweils einbezogenen SSD in Tabelle A-4 dargestellt. Eine Interpretation der Anteile der ersten beiden Dimensionen an der gesamten Inertia, dem Chi-Quadrat-basierten Maß der Variation innerhalb der jeweils analysierten Matrix (Blasius und Greenacre, 1998; Greenacre, 1984; Greenacre und Blasius, 1994; Weller und Romney, 1990), erfolgt im vorliegenden Zusammenhang nicht, weil weder eine faktorenanalytische noch eine clusteranalytische Interpretation der MCA angestrebt wird. Auch ist es nicht möglich, die Häufigkeit der Nichtbeantwortung von Items als gesondertes Kriterium in die MCA einzubeziehen, da der Anteil fehlender Werte insbesondere in den ersten beiden experimentellen Bedingungen sowohl bei Selbst- als auch bei Fremdbeschreibungen sehr gering war (Tabelle A-2). Ebenso erfolgen keine zusätzlichen Analysen in Bezug auf die Merkmale Alter, Geschlecht, Bildung und Erwerbstätigkeit, da sich diese nicht signifikant voneinander unterscheiden (Tabelle A-1 und Abschnitt 3.3).

4 Ergebnisse

Zunächst wurden die univariaten Häufigkeitsverteilungen der einzelnen SSD (vgl. Tabelle A-2, Anhang) betrachtet. Aus der visuellen Überprüfung ergibt sich, dass die Befragten nur vier von fünf Antwortkategorien benutzt haben, um sich selbst und andere zu beschreiben. In den meisten Fällen sind die Antwortkategorien ausreichend besetzt, das heißt, dass die relativen Häufigkeiten p_j der Antwortkategorie j größer oder gleich 0.05 sind ($p_j \geq .05$). Mit Ausnahme der Skalen „vollkommen“, „streng“ und „reizbar“ liegen diese vier der fünf Antwortkategorien überwiegend auf der linken Seite der entsprechenden SSD. Dies gilt sowohl für die verbal bipolar, unipolar als auch unsinnig verankerten SSD. Es zeigt sich weiterhin, dass bei den verbal bipolar und unipolar verankerten SSD ein sehr geringer Anteil fehlender Werte ($p_j \leq .05$), bei den verbal unsinnig verankerten SSD ein geringer Anteil fehlender Werte ($p_j \leq .10$) zu beobachten ist.

4.1 Vergleich der verbal bipolar und unipolar verankerten Skalen des Semantischen Differentials

Im Folgenden werden die Ergebnisse der empirischen Analyse der ersten Untersuchungshypothese berichtet. Zur empirischen Analyse des ersten Teils der Untersuchungshypothese wurde

für verbal bipolar und unipolar verankerte Skalen die Variation der Antworten über die einzelnen Antwortkategorien anhand des Index qualitativer Variation (IQV) und anhand des Quartilabstandes (QA) berechnet und verglichen (vgl. Abschnitt 3.4). Die Maße der qualitativen Variation verdeutlichen, dass sowohl die Selbst- als auch die Fremdbeschreibungen anhand von verbal bipolar verankerten Skalen in den meisten Fällen ein geringeres Maß an Variation aufweisen als die Selbst- und Fremdbeschreibungen anhand von verbal unipolar verankerten Skalen. Lediglich für eine SSD, nämlich für „hartnäckig-nachgiebig/nicht hartnäckig“, zeigt sich im Falle der Selbstbeschreibung die erwartete stärkere Variation der Antworten über die einzelnen Antwortkategorien (vergleiche Tabelle A-3). Wie aus Tabelle A-3 weiterhin hervorgeht, zeigt der QA sowohl bei Selbst- als auch bei Fremdbeschreibungen kaum Unterschiede zwischen verbal bipolar und unipolar verankerten SSD. In den meisten Fällen beträgt der QA = 2. Es gibt keine nennenswerten Unterschiede zwischen den zwei Arten der verbalen Verankerung in Bezug auf den QA. Hypothese 1.1 findet folglich keine empirische Unterstützung, da die verbal unipolare Verankerung mit einer höheren Variabilität der Werte zusammenhängt.

Zur empirischen Analyse des zweiten Teils der Untersuchungshypothese wurde für verbal bipolar und unipolar verankerte Skalen ein Maß für die zentrale Tendenz der Antworten über die einzelnen Antwortkategorien berechnet. Dieses gibt an, wie die Antworten über die einzelnen Antwortkategorien verteilt sind.

Tabelle 1: Unterschiede in den zentralen Tendenzen der Skalen des Semantischen Differentials (Hypothese 1)

Selbstbeschreibungen

	Experimentelle Bedingung 1 Bipolare SSD		Experimentelle Bedingung 2 Unipolare SSD		Mediantest	Kruskal-Wallis- Test
	Median	Rangsumme	Median	Rangsumme	X ² (p-Wert)	X ² (p-Wert)
Gut	2	27074.50	3	32265.50	2.69 (0.10)	1.91 (0.17)
Optimistisch	2	28887.50	2	33593.50	1.67 (0.20)	1.69 (0.19)
Positiv	2	28494.50	2	32930.50	2.66 (0.10)	1.67 (0.20)
Vollkommen	4	25803.00	5	35973.00	21.11 (0.00)	17.08 (0.0001)
Zeitgemäß	3	28898.50	3	33229.50	0.02 (0.89)	1.02 (0.31)
Hart	4	29783.00	4	32345.00	0.09 (0.76)	0.06 (0.81)
Tief	2	27878.00	3	34250.00	4.65 (0.03)	4.53 (0.03)
Maskulin	4	28840.50	4	32584.50	0.001 (0.94)	0.77 (0.38)
Streng	4	30026.50	4	32454.50	0.89 (0.35)	0.005 (0.95)
Stark	3	30109.50	3	32371.50	0.008 (0.93)	0.00 (0.98)
Hartnäckig	3	32880.50	3	29247.50	13.78 (0.00)	10.90 (0.001)
Aktiv	3	30673.50	3	30751.50	2.44 (0.12)	1.72 (0.19)
Reizbar	4	31476.50	4	31004.50	0.27 (0.60)	2.16 (0.14)
Energisch	3	29665.50	3	32110.50	0.21 (0.65)	0.007 (0.93)

Fremdbeschreibungen

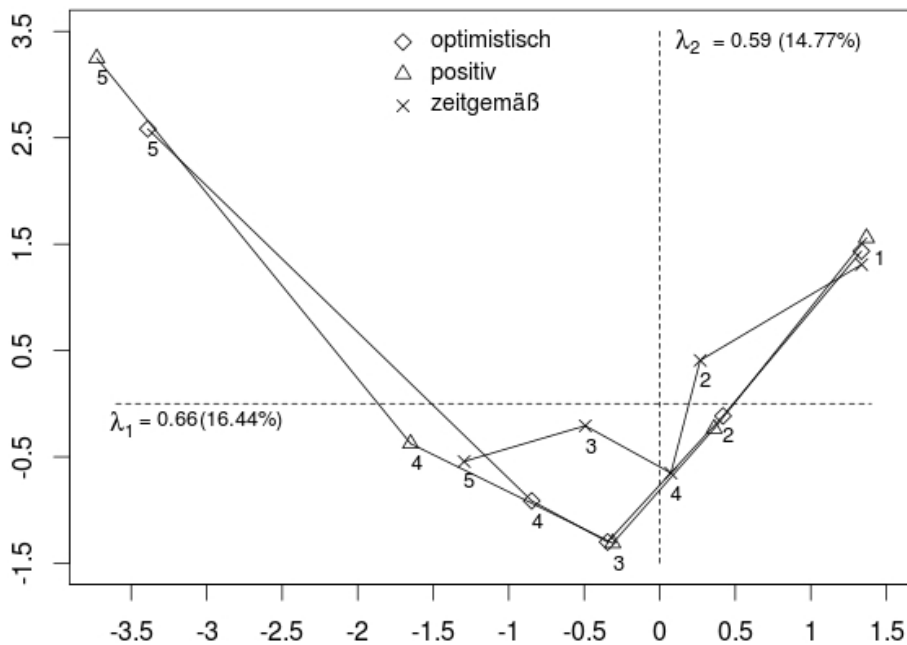
	Experimentelle Bedingung 1 Bipolare SSD		Experimentelle Bedingung 2 Unipolare SSD		Mediantest	Kruskal-Wallis- Test
	Median	Rangsumme	Median	Rangsumme	X ² (p-Wert)	X ² (p-Wert)
Gut	2	29367.50	3	32760.50	1.14 (0.29)	0.77 (0.38)
Optimistisch	3	29927.50	3	32553.50	0.20 (0.65)	0.31 (0.58)
Positiv	2	29552.00	2	32929.00	3.03 (0.08)	0.60 (0.44)
Vollkommen	4	27461.50	4	35373.50	15.47 (0.00)	11.55 (0.0007)
Zeitgemäß	3	29517.50	4	32963.50	0.85 (0.36)	0.98 (0.32)
Hart	4	28274.50	4	33501.50	4.11 (0.04)	4.87 (0.03)
Tief	2	29951.00	3	31825.00	0.61 (0.43)	0.03 (0.87)
Maskulin	3	25817.00	4	36664.00	19.37 (0.00)	25.23 (0.0001)
Streng	4	30038.50	4	32442.50	0.37 (0.54)	0.20 (0.66)
Stark	3	29207.50	3	33273.50	2.40 (0.12)	1.76 (0.18)
Hartnäckig	3	30923.50	3	30852.50	0.84 (0.36)	0.81 (0.37)
Aktiv	2.5	29954.00	3	32174.00	0.00 (1.00)	0.06 (0.81)
Reizbar	5	32478.00	4	30003.00	5.10 (0.02)	4.86 (0.03)
Energisch	3	29846.50	3	32634.50	0.005 (0.94)	0.42 (0.52)

Auch die Werte des Medians und die Rangsummen stimmen überwiegend zwischen beiden experimentellen Bedingungen überein, wie aus Tabelle 1 hervorgeht. Im Falle von Selbstbeschreibungen zeigen die Ergebnisse von Mediantest und Kruskal-Wallis-H-Test, dass lediglich bei den Skalen „vollkommen-unvollkommen/nicht vollkommen“ und „heavy-light/not heavy“ der Median und die Rangsummen der verbal bipolar verankerten Skala jeweils kleiner als der Median und die Rangsummen der verbal unipolar verankerten Skala sind. Für die Skala „hartnäckig-nachgiebig/nicht hartnäckig“ zeigt sich allerdings, dass der Median und die Rangsummen der verbal bipolar verankerten Skala größer sind als der Median und die Rangsummen der verbal unipolar verankerten Skala.

Für Fremdbeschreibungen führen Median- und Kruskal-Wallis-H-Test zu dem Ergebnis, dass bei drei Skalen („vollkommen-unvollkommen/nicht vollkommen“, „hart-weich/nicht hart“, „maskulin-feminin/nicht maskulin“) sowohl der Median als auch die Rangsummen bei den verbal bipolar verankerten SSD kleiner ist als bei den verbal unipolar verankerten SSD. Nur für die Skala „reizbar-gelassen/nicht reizbar“ ergibt sich aufgrund des Mediantests und des Kruskal-Wallis-H-Tests, dass der Median und die Rangsummen bei der verbal bipolar verankerten Skala gegenüber der verbal unipolar verankerten Skala signifikant in Richtung des rechten Pols der Skala verschoben sind. Hypothese 1.2 findet folglich keine empirische Unterstützung. Die zentrale Tendenz der SSD stimmt bei verbal bipolar und bei verbal unipolar verankerten Skalen im überwiegenden Teil aller Fälle überein. Wenn Unterschiede vorhanden sind, dann sind die Werte im Falle der unipolaren Skalen in der Regel höher.

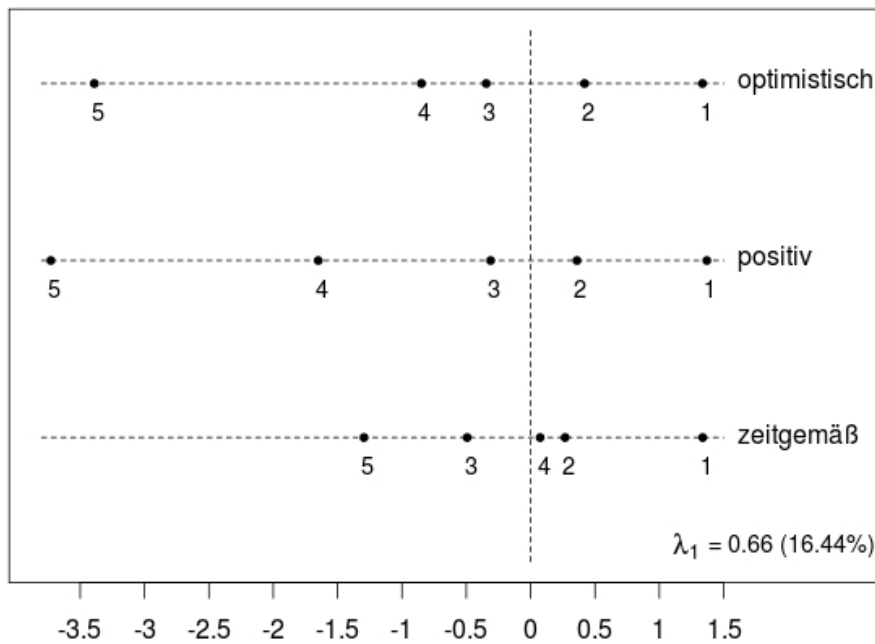
Hypothese 1.3 wurde anhand multipler Korrespondenzanalysen untersucht. Das Vorgehen wird am Beispiel der Analyse von Selbstbeschreibungen entlang der Dimension Evaluation ausführlich dargestellt. In die multiple Korrespondenzanalyse der Dimension Evaluation konnten aufgrund der vorbereitenden multiplen Korrespondenzanalysen (vgl. Abschnitt 3.4) jeweils die ersten fünf Antwortkategorien (ausgehend vom linken Pol der Skala also die Antwortkategorien 1, 2, 3, 4 und 5) der SSD „optimistisch-pessimistisch/nicht optimistisch/gelassen“, „positiv-negativ/nicht positiv/mild“, und „zeitgemäß-altmodisch/ nicht zeitgemäß/schwach“ einbezogen werden. Tabelle A-4 sind die standard coordinates dieser Antwortkategorien auf den ersten beiden Dimensionen der multiplen Korrespondenzanalyse zu entnehmen. Auf der Grundlage der standard coordinates der verbal bipolar verankerten SSD wird eine zweidimensionale graphische Darstellung erstellt, eine sogenannte ‚map‘ (vgl. Abbildung 4). Wie aus Abbildung 4 hervorgeht, zeigt sich für alle einbezogenen SSD („optimistisch-pessimistisch“, „positiv-negativ“, „zeitgemäß-altmodisch“) ein deutlicher, methodeninduzierter horseshoe-Effekt auf der zweiten Dimension, weshalb nur die erste Dimension interpretiert wird.

Abbildung 4: Zweidimensionaler MCA-Plot für die Evaluations-Dimension – Selbstbeschreibungen



Der eindimensionalen Darstellung der in der Korrespondenzanalyse ermittelten Koordinaten (Abbildung 5) kann entnommen werden, dass auf der ersten Dimension ein Kontrast zwischen den beiden ersten Antwortkategorien am linken Pol der einzelnen SSD gebildet wird. Vor allem bei den beiden Skalen „optimistisch-pessimistisch“ und „positiv-negativ“ wird sehr stark zwischen den ersten vier Antwortkategorien am linken Pol und der fünften Antwortkategorie differenziert. Nur bei der Skala „zeitgemäß-altmodisch“ gibt es darüber hinaus Abweichungen von der Annahme einer ordinalen Ordnung der Antwortkategorien: Die Reihenfolge lautet 1, 2, 4, 3, 5; entsprechend gibt es eine Abweichung, nämlich die Reihenfolge 4, 3. Gemessen an der Anzahl aller möglichen Paare bei fünf Antwortkategorien, nämlich = 10, beträgt daher der Anteil der abweichenden Paare . Die Antwortskalen weisen also zu einem überwiegenden Anteil die erwartete ordinale Ordnung auf.

Abbildung 5: Eindimensionaler MCA-Plot für die Evaluations-Dimension - Selbstbeschreibungen



Die Ergebnisse aller weiteren multiplen Korrespondenzanalysen werden im Folgenden nur zusammenfassend berichtet. Hierzu wird die Ergebnisdarstellung in Tabelle A-4 herangezogen. In Tabelle 2 wird vermerkt, ob sich in der zweidimensionalen Betrachtung ein methodeninduzierter horseshoe-Effekt zeigt und wie hoch ggf. der Anteil von Abweichungen von der ordinalen Anordnung der Antwortkategorien für jede einzelne SSD ist, welcher der eindimensionalen Darstellung entnommen werden kann.

Tabelle 2: Ergebnisse der Multiplen Korrespondenzanalyse

Selbstbeschreibungen

	Experimentelle Bedingung 1 Bipolare SSD		Experimentelle Bedingung 2 Unipolare SSD		Experimentelle Bedingung 3 Unsinnige SSD	
	Horseshoe	Abweichung von der Ordinalität	Horseshoe	Abweichung von der Ordinalität	Horseshoe	Abweichung von der Ordinalität
Dimension Evaluation						
Optimistisch	+	0	+	0	+	0
Positiv	+	0	+	0	+	0.10
Zeitgemäß	+	0.10	+	0.10	+	0
Dimension Potency						
Hart	+	0	+	0	+	0.17
Maskulin	+	0.17	+	0	+	0
Hartnäckig	+	0	+	0	+	0
Dimension Activity						
Aktiv	+	0	+	0	+	0
Energisch	+	0	+	0	+	0.17

Fremdbeschreibungen

	Experimentelle Bedingung 1 Bipolare SSD		Experimentelle Bedingung 2 Unipolare SSD		Experimentelle Bedingung 3 Unsinnige SSD	
	Horseshoe	Abweichung von der Ordinalität	Horseshoe	Abweichung von der Ordinalität	Horseshoe	Abweichung von der Ordinalität
Dimension Evaluation						
Gut	+	0	+	0	+	0
Optimistisch	+	0	+	0	+	0
Positiv	+	0	+	0	+	0
Zeitgemäß	+	0	+	0	+	0
Dimension Potency						
Tiefgründig	+	0	+	0	+	0
Maskulin	+	0	+	0	+	0
Stark	+	0	+	0	+	0
Hartnäckig	+	0	+	0	+	0
Dimension Activity						
Aktiv	+	0	-	0.20	+	0
Energisch	+	0.10	-	0.20	+	0

Im Hinblick auf Selbstbeschreibungen zeigt sich für die Dimension Evaluation in allen experimentellen Bedingung ein horseshoe-Effekt (vgl. Tabelle 2). Die erste Dimension kontrastiert im

Fälle der verbal unipolar verankerten Ratingskalen den rechten Pol der einzelnen SSD stärker als in der ersten experimentellen Bedingung (vgl. Tabelle A-4). Nur für die Skala „zeitgemäß-altmodisch“ zeigt sich eine Abweichung von der ordinalen Ordnung der Antwortkategorien, mit hin ein Anteil von 0.10 (vgl. Tabelle 2).

Hinsichtlich der Dimension Potency weisen die Ergebnisse der multiplen Korrespondenzanalyse darauf hin, dass in allen Fällen ein horseshoe-Effekt gegeben ist (vgl. Tabelle 2). Bei den verbal bipolar verankerten Ratingskalen werden auf der ersten Dimension die ersten beiden und die letzten beiden Skalenpunkte kontrastiert. Die Abstände zwischen den einzelnen Skalenpunkten sind insbesondere für die Skala „hartnäckig-nachgiebig“ annähernd äquidistant (Tabelle A-4). Nur für die Skala „maskulin-feminin“ findet sich eine Abweichung von der ordinalen Anordnung der einzelnen Antwortkategorien, was einem Anteil von 0.17 entspricht (Tabelle 2). Bei den verbal unipolar verankerten Ratingskalen findet sich demgegenüber eine stärkere Kontrastierung des linken Pols der Skala, der bei zwei Skalen auf der ersten Dimension gegen alle anderen Skalenpunkte abgebildet wird (Tabelle A-4). Es findet sich in dieser Bedingung keine Abweichung von der ordinalen Anordnung der einzelnen Skalenpunkte. Dieses Ergebnis spricht also eher für die Idee des Semantischen Differentials.

Hinsichtlich der Dimension Activity führt die multiple Korrespondenzanalyse zu dem Ergebnis, dass in allen Fällen ein horseshoe-Effekt zu beobachten ist (Tabelle 2). Für die verbal bipolar verankerten Ratingskalen zeigt sich auf der ersten Dimension eine Kontrastierung zwischen den ersten beiden und den letzten beiden Skalenpunkten. Bei beiden SSD zeigt sich eine starke Differenzierung zwischen den beiden Antwortkategorien am linken Pol der Skala (Tabelle A-4), und es finden sich keine Abweichungen von der ordinalen Anordnung der einzelnen Skalenpunkte (Tabelle 2). Demgegenüber findet sich für die verbal unipolar verankerten Ratingskalen eine sehr starke Kontrastierung des linken Pols der Skala gegenüber allen anderen Skalenpunkten (Tabelle A-4). Es findet sich auch in dieser Bedingung keine Abweichung von der ordinalen Anordnung der einzelnen Skalenpunkte.

Bei Fremdbeschreibungen zeigt sich für die Dimension Evaluation in beiden experimentellen Bedingungen – also bei verbal bipolar und verbal unipolar verankerten SSD – ein horseshoe-Effekt bei allen einbezogenen Skalen (Tabelle 2). Die Antwortkategorien aller Skalen sind in beiden experimentellen Bedingungen relativ gleichmäßig über die erste Dimension verteilt. Es finden sich weder bei verbal bipolar noch bei verbal unipolar verankerten SSD Hinweise auf eine stärkere Kontrastierung des linken Pols der Skala (Tabelle A-4). Außerdem weisen die Antwortkategorien aller Skalen in beiden experimentellen Bedingungen ausnahmslos eine ordinale Anordnung auf (Tabelle 2).

Für die Dimension Potency zeigt sich, dass sowohl bei den verbal bipolar als auch den verbal unipolar verankerten Skalen ein horseshoe-Effekt beobachtet werden kann (Tabelle 2). In

der ersten experimentellen Bedingung werden auf der ersten Dimension die ersten drei Antwortkategorien am linken Pol gegen die letzte Antwortkategorie kontrastiert. Die Abstände zwischen den Antwortkategorien weisen allerdings auf eine nur mäßige Differenzierung hin (Tabelle 4). Demgegenüber findet sich für die verbal unipolar verankerten Skalen eine deutliche Differenzierung zwischen der ersten Antwortkategorie und allen anderen Antwortkategorien auf der ersten Dimension. Weder in der ersten noch in der zweiten experimentellen Bedingung lassen sich Abweichungen von der ordinalen Anordnung der einzelnen Skalenpunkte finden.

Im Hinblick auf die Dimension Activity zeigt sich, dass nur im Falle der verbal bipolar verankerten Skalen ein horseshoe-Effekt gegeben ist (Tabelle 2). Auf der ersten Dimension zeigt sich für die verbal bipolar verankerten Skalen eine Differenzierung zwischen den ersten beiden und den letzten beiden Antwortkategorien. Bei der Skala „aktiv-passiv“ fallen die beiden rechten Antwortkategorien zusammen (Tabelle A-4), und bei der Skala „energisch-träge“ besteht eine Abweichung von der ordinalen Anordnung der einzelnen Skalenpunkte, mithin ein Anteil von 0.10 (Tabelle 2). Für die verbal unipolar verankerten Skalen finden sich dagegen Abweichungen von der ordinalen Anordnung der Skalenpunkte derart, dass die erste und die vierte Antwortkategorie auf der ersten Dimension gegen die anderen Antwortkategorien kontrastiert werden. Der Anteil der Abweichungen beträgt bei beiden Skalen 0.20. Zudem differenziert die multiple Korrespondenzanalyse stark zwischen den einzelnen Antwortkategorien, und ein horseshoe-Effekt lässt sich nicht beobachten (Tabelle A-4). Die Ergebnisse der multiplen Korrespondenzanalyse weisen also darauf hin, dass bei Fremdbeschreibungen auf der Dimension Activity mit verbal bipolar verankerten Skalen die erwarteten Beziehungen zwischen den einzelnen Antwortkategorien bestehen, bei verbal unipolar verankerten Skalen jedoch nicht.

4.2 Vergleich der verbal bipolar und unsinnig verankerten Skalen des Semantischen Differentials

Im Folgenden werden die Ergebnisse der empirischen Analyse der zweiten Untersuchungshypothese berichtet. Auch hier wurde der erste Teil der Hypothese – der Vergleich der Variation der Antworten über die einzelnen Antwortkategorien zwischen verbal bipolar und verbal unsinnig verankerten Skalen – anhand des Index der qualitativen Variation (IQV) und anhand des Quartilabstandes (QA) untersucht (Abschnitt 3.4). Die Maße der qualitativen Variation verdeutlichen, dass bei Selbstbeschreibungen annähernd gleich viele verbal bipolar verankerte SSD ein höheres und ein geringeres Maß an Variation im Vergleich zu verbal unsinnig verankerten SSD aufweisen. Bei Fremdbeschreibungen ist die qualitative Variation in jeweils annähernd der Hälfte aller Skalen zwischen verbal bipolar und verbal unsinnig verankerten SSD vergleichbar oder bei verbal bipolar verankerten SSD kleiner als bei verbal unsinnig verankerten SSD. Lediglich für eine SSD, nämlich für „reizbar-gelassen/pessimistisch“, zeigt sich im Falle der Fremdbeschrei-

bung die erwartete stärkere Variation der Antworten über die einzelnen Antwortkategorien (Tabelle A-3). Wie aus Tabelle A-3 weiterhin hervorgeht, ist im Falle der Selbstbeschreibungen bei nur fünf SSD der QA bei verbal bipolar verankerten Skalen größer als bei verbal unsinnig verankerten Skalen. Bei annähernd der Hälfte der SSD ist der QA in beiden experimentellen Bedingungen identisch. Im Falle der Fremdbeschreibungen ist der QA nur bei zwei SSD bei verbal bipolar verankerten Skalen größer als bei verbal unsinnig verankerten Skalen. Bei den meisten SSD stimmt der QA in beiden experimentellen Bedingungen überein. Hypothese 2.1 findet folglich keine empirische Unterstützung.

Tabelle 3: Unterschiede in den zentralen Tendenzen der Skalen des Semantischen Differentials (Hypothese 2)

Selbstbeschreibungen

	Experimentelle Bedingung 1 Bipolare SSD		Experimentelle Bedingung 2 Unipolare SSD		Mediantest	Kruskal-Wallis-Test
	Median	Rangsumme	Median	Rangsumme	X ² (p-Wert)	X ² (p-Wert)
Gut	2	26140.00	3	26510.00	1.47 (0.23)	0.40 (0.53)
Optimistisch	2	25732.50	3	30547.50	7.75 (0.005)	10.90 (0.001)
Positiv	2	26441.50	2	29169.50	4.98 (0.03)	4.52 (0.03)
Vollkommen	4	28760.00	4	25855.00	0.20 (0.65)	1.35 (0.25)
Zeitgemäß	3	28259.00	3	27019.00	0.73 (0.39)	0.02 (0.89)
Hart	4	31711.50	4	22573.50	20.70 (0.00)	19.16 (0.0001)
Tief	2	26275.50	3	29335.50	10.46 (0.001)	5.28 (0.02)
Maskulin	4	31076.00	3	20605.00	44.27 (0.00)	22.63 (0.0001)
Streng	4	32292.50	3	21335.50	37.25 (0.00)	28.81 (0.0001)
Stark	3	30104.50	3	24841.50	4.82 (0.03)	5.04 (0.02)
Hartnäckig	3	29028.50	3	25586.50	2.93 (0.09)	1.57 (0.21)
Aktiv	3	28434.00	3	26844.00	0.07 (0.80)	0.30 (0.58)
Reizbar	4	31369.50	4	22915.50	40.23 (0.00)	15.79 (0.0001)
Energisch	3	30015.50	3	24930.50	8.96 (0.003)	5.56 (0.02)

(Fortsetzung Tabelle 3)

Fremdbeschreibungen

	Experimentelle Bedingung 1 Bipolare SSD		Experimentelle Bedingung 2 Unipolare SSD		Mediantest	Kruskal-Wallis- Test
	Median	Rangsumme	Median	Rangsumme	X ² (p-Wert)	X ² (p-Wert)
Gut	2	28309.50	2	26968.50	0.09 (0.76)	0.04 (0.85)
Optimistisch	3	24198.00	4	31413.00	24.97 (0.00)	27.71 (0.0001)
Positiv	2	25646.50	3	28309.50	11.54 (0.001)	8.87 (0.003)
Vollkommen	4	28541.00	4	25415.00	2.91 (0.09)	0.09 (0.76)
Zeitgemäß	3	29547.50	3	24737.50	0.53 (0.47)	1.93 (0.16)
Hart	4	29662.00	4	24294.00	3.14 (0.08)	2.87 (0.09)
Tief	2	28365.50	3	25590.50	0.11 (0.74)	0.08 (0.78)
Maskulin	3	27951.50	3	26004.50	0.01 (0.93)	0.17 (0.68)
Streng	4	31162.00	4	22466.00	10.62 (0.001)	13.39 (0.0003)
Stark	3	27439.50	3	27175.50	0.37 (0.54)	1.50 (0.22)
Hartnäckig	3	28597.50	3	24703.50	1.16 (0.28)	0.61 (0.44)
Aktiv	2.5	27984.00	3	26631.00	0.02 (0.88)	0.14 (0.71)
Reizbar	5	34898.00	4	18403.00	68.44 (0.00)	72.09 (0.0001)
Energisch	3	29344.00	3	25934.00	2.81 (0.09)	0.70 (0.40)

Auch der zweite Teil der zweiten Untersuchungshypothese wurde anhand eines Vergleichs der zentralen Tendenz der Antworten für verbal bipolar und verbal unsinnig verankerte SSD getestet. Im Falle von Selbstbeschreibungen führen die Ergebnisse von Mediantest und Kruskal-Wallis-H-Test zu dem Ergebnis, dass bei nahezu der Hälfte aller SSD die erwarteten Unterschiede zwischen den Maßen der zentralen Tendenz und den Rangsummen auftreten, nämlich bei den Skalen „hart-weich/passiv“, maskulin-feminin/oberflächlich“, „streng-mild/negativ“, „stark-schwach/altmodisch“, „aktiv-passiv/weich“ und „reizbar-gelassen/pessimistisch“ (Tabelle 3). Allerdings gibt es auch drei Skalen, in denen Mediane und Rangsummen der verbal bipolar verankerten Skalen kleiner sind als die der verbal unsinnig verankerten Skalen, nämlich die Skalen „optimistisch-pessimistisch/gelassen“, „positiv-negativ/mild“ und „tiefgründig-oberflächlich/feminin“ (Tabelle 3). Für Fremdbeschreibungen führen die Analysen zu dem Ergebnis, dass die erwarteten Unterschiede zwischen verbal bipolar und verbal unsinnig verankerten Skalen in nur zwei Fällen auftreten, und zwar bei „streng-mild/negativ“ und bei „reizbar-gelassen/pessimistisch“. Bei ebenso vielen Skalen zeigen sich im Falle der verbal bipolar verankerten Skalen ein jeweils kleinerer Median und kleinere Rangsummen, nämlich bei „optimistisch-pessimistisch/gelassen“ und bei „positiv-negativ/mild“ (Tabelle 3). Alles in allem weisen die Ergebnisse der Mediantests und der Kruskal-Wallis-Tests darauf hin, dass Hypothese 2.2 nur mit Einschränkung empirische Unterstützung findet.

Hypothese 2.3 wurde anhand multipler Korrespondenzanalysen untersucht. Dabei kam das bereits im Zuge der Analyse von Hypothese 1.3 beschriebene Vorgehen analog zur Anwendung (vgl. Abschnitte 3.4, 4.1). Die Ergebnisse werden auf der Grundlage der Darstellung in Tabelle A-4 interpretiert. Tabelle 2 in Abschnitt 4.1 zeigt, ob sich in der zweidimensionalen Betrachtung ein horseshoe-Effekt zeigt und wie hoch gegebenenfalls der Anteil von Abweichungen von der ordinalen Anordnung der Antwortkategorien für jede einzelne SSD ist. Im Hinblick auf Selbstbeschreibungen zeigt sich ergänzend zu den bereits berichteten Ergebnissen für die erste experimentelle Bedingung, dass sich für die verbal unsinnig verankerten Skalen in allen Fällen ein horseshoe-Effekt zeigt (Tabelle 2). Außerdem zeigt sich, dass im Falle der verbal unsinnig verankerten Skalen auf der ersten Dimension ein sehr starker Kontrast zwischen der jeweils ersten Antwortkategorie und allen anderen Antwortkategorien gebildet wird (Tabelle A-4). Bei einer Skala, nämlich der Skala „positiv-mild“, findet sich eine Abweichung von der Annahme der Ordinalität, also ein Anteil von 0.10. Hinsichtlich der Dimension Potency wurde bereits gezeigt, dass bei den verbal bipolar verankerten Ratingskalen auf der ersten Dimension die ersten beiden und die letzten beiden Skalenpunkte kontrastiert werden (Abschnitt 4.1).

Die Ergebnisse der multiplen Korrespondenzanalyse weisen darauf hin, dass bei den verbal unsinnig verankerten Skalen in allen Fällen ein horseshoe-Effekt gegeben ist (Tabelle 2). Auf der ersten Dimension kontrastiert die fünfte Antwortkategorie sehr stark gegen alle anderen einbezogenen Antwortkategorien (Tabelle A-4). Nur bei der Skala „hart-passiv“ lässt sich eine Abweichung von der ordinalen Anordnung der einzelnen Antwortkategorien beobachten, mithin ein Anteil von 0.17. Für die Dimension Activity hatte die multiple Korrespondenzanalyse bereits gezeigt, dass bei den verbal bipolar verankerten SSD auf der ersten Dimension eine Kontrastierung zwischen den beiden ersten und den beiden letzten Skalenpunkten besteht, und dass die Antwortkategorien am linken Pol der Skala stark differenziert sind. Demgegenüber zeigt sich für die verbal unsinnig verankerten SSD in allen Fällen ein horseshoe-Effekt (Tabelle 2), und eine sehr starke Kontrastierung der ersten Antwortkategorie gegenüber allen anderen Antwortkategorien (Tabelle A-4). Darüber hinaus findet sich eine Abweichung von der ordinalen Anordnung der einzelnen Skalenpunkte bei der Skala „energisch-schlecht“ (Tabelle 2).

Bei Fremdbeschreibungen zeigt sich für die Dimension Evaluation, dass – anders als im Falle der verbal bipolar verankerten SSD – zwar auch in allen Fällen ein horseshoe-Effekt beobachtet werden kann, die erste Antwortkategorie aber auf der ersten Dimension sehr stark gegen alle anderen Antwortkategorien kontrastiert (Tabelle A-4). Es finden sich darüber hinaus keine Abweichungen von der ordinalen Anordnung der einzelnen Antwortkategorien (Tabelle 2).

Für die Dimension Potency hatte sich bereits gezeigt, dass im Falle der verbal bipolar verankerten SSD auf der ersten Dimension die ersten drei Antwortkategorien am linken Pol gegen die letzte Antwortkategorie abgebildet werden. Im Unterschied dazu findet sich zwar auch

im Falle der verbal unsinnig verankerten SSD in allen Fällen ein horseshoe-Effekt, es wird aber auf der ersten Dimension durchgängig sehr deutlich die erste Antwortkategorie gegen alle anderen Antwortkategorien kontrastiert (Tabelle A-4). Abweichungen von der ordinalen Anordnung der einzelnen Skalenpunkte bestehen nicht (Tabelle 2).

Im Hinblick auf die Dimension Activity zeigt sich, dass bei beiden verbal unsinnig verankerten Skalen ein horseshoe-Effekt zu beobachten ist (Tabelle 2). Auf der ersten Dimension zeigt sich für die verbal unsinnig verankerten Skalen ähnlich wie für die verbal bipolar verankerten Skalen eine Differenzierung zwischen den ersten beiden und den letzten beiden Antwortkategorien, dabei bildet die dritte Antwortkategorie einen Mittelpunkt (Tabelle A-4). Abweichungen von der ordinalen Anordnung der Antwortkategorien bestehen nicht (Tabelle 2). Im Vergleich der verbal bipolar mit den verbal unsinnig verankerten Skalen zeigen sich bei den verbal unsinnig verankerten Skalen in den meisten Fällen starke Kontraste zwischen der ersten Antwortkategorie am linken Pol der SSD und allen anderen Antwortkategorien. Dies spricht grundsätzlich für die Idee des Semantischen Differentials, da bei den verbal bipolar verankerten Skalen die Antwortkategorien stärker auf der ersten Dimension variieren. Es finden sich fast keine Abweichungen von der ordinalen Anordnung der einzelnen Antwortkategorien, und zwar weder bei den verbal bipolar noch bei den verbal unsinnig verankerten SSD. Die Ergebnisse der multiplen Korrespondenzanalysen sprechen daher nur mit Einschränkung für Hypothese 2.3.

5 Diskussion

Das Ziel der Studie war die Analyse der metrischen Eigenschaften verschiedener Skalen des semantischen Differenzials (SSD). Die grundlegende Idee war, dass Linearität, Symmetrie und Äquidistanz von der Art der verbalen Verankerung der Skalen durch Adjektivpaare abhängen. Falls die Annahme der Äquidistanz verletzt wird, muss zumindest Ordinalskalenniveau gegeben sein: Linearität und Ordinalskalenniveau der Daten sind wichtige Voraussetzungen für allgemein übliche Methoden der Datenanalyse (Lineare Regression oder Varianzanalyse). Zunächst wurden verschiedene Erklärungen der häufig auftretenden Rechtsschiefe in den SSD-Daten beurteilt (Schriesheim et al., 1994). In der vorliegenden Studie wurde ein experimentelles Design gewählt, das einen Vergleich verschiedener Erklärungen ermöglicht. Dadurch konnten die Ergebnisse aktueller Studien berücksichtigt werden, welche sich mit Bipolarität von Ratingskalen befassen (Krebs, 2012; Menold und Raykov, 2015; Schaeffer und Presser, 2003). Da Bipolarität mit gleicher Verschiebung in den Daten (sogenannter Positivity Bias/Positivitätsverzerrung) zusammenhängt, wurden klassische bipolare SSD mit der unipolaren Version verglichen, um das Ausmaß des Positivity Bias in den klassischen SSD zu beurteilen. Eine weitere Erklärung der Rechtsschiefe in den SSD war das Leseverhalten der Befragten, da sie möglicherweise nur die

linke Seite beachten, während sie die rechte Seite aufgrund ihrer Leserichtung außer Acht lassen. Um diese Annahme zu überprüfen, wurde die Bedingung der bedeutungslosen SSD realisiert.

Die Hypothesen wurden sowohl im Hinblick auf den Effekt getestet, den die verbale Verankerung eines Semantischen Differenzials auf die Variabilität und auf zentrale Tendenzen der Daten hat, als auch hinsichtlich Linearität, Symmetrie und Äquidistanz. Da die Annahme der Äquidistanz in den Sozialwissenschaften kaum erfüllt wird (Vielleman und Wilkinson, 1994), wurde der Fokus auf die Frage gelegt, ob SSD Ordinalskalenniveau erlauben, was auch Linearität der Daten impliziert.

Die Hypothesen wurden im Rahmen einer Onlineumfrage getestet, in welcher die Befragten ihr Selbstbild und das Bild, welches sie über bestimmte andere Menschen haben, beurteilten. Dabei wurden verbal unterschiedlich verankerte Formen der SSD verwendet. Die Ergebnisse sind für das Selbst- und Fremdbild ähnlich. Die Ergebnisse zeigen zunächst, dass die Variabilität der Daten bei unipolaren Formen der SSD höher ist als bei klassischen bipolaren SSD. Die Rangsummen waren manchmal bei der unipolaren Realisierung der SSD größer als bei den klassischen SSD. Das bedeutet, dass negative Dimensionen der SSD bei unipolar verankerten SSD häufiger genutzt wurden. Der Positivity Bias war deutlich, jedoch nur zu einem geringen Ausmaß in den klassischen SSD zu sehen, da Unterschiede in den Werten der zentralen Tendenzen nur für ein paar wenige SSD beobachtet wurden. Hinsichtlich eines Vergleichs zwischen klassischen und bedeutungslosen SSD, ergaben sich die größten Verschiebungen nach der rechten oder manchmal auch nach der linken Seite bei den bedeutungslosen SSD. Die Unterschiede in der Variabilität und die Verschiebung der Daten in Richtung eines Pols zeigen, dass Befragte sich der bedeutungslosen Seite der SSD bewusst waren. Nur eine kleine Anzahl Befragter gab in der bedeutungslosen Experimentalgruppe überhaupt keine Antwort, während die meisten Befragten nur eine Seite der bedeutungslosen SSD für ihre Antwort nutzten. Die Ergebnisse der MCA, welche sich durch Testung des dritten Teils der zwei Hypothesen ergaben, lieferte Evidenz dafür, dass die Annahme der Äquidistanz zwischen den Kategorien in einem semantischen Differenzial, durch die Daten nicht unterstützt werden kann. Dieses Ergebnis war unabhängig von der verbalen Verankerung der Skalen. Jedoch unterstützen die Ergebnisse der Daten größtenteils die Annahme des Ordinalskalenniveaus. Hier ergab sich nur ein kleiner Unterschied (für zwei SSD, Potency Dimension) zwischen der klassischen und der unipolaren Realisierung. Im Vergleich zwischen bipolarer und bedeutungsloser Experimentalgruppe konnten bei der bedeutungslosen Gruppe kleine Verletzungen der Annahme des Ordinalskalenniveaus beobachtet werden. Da Ordinalität auch Linearität impliziert, kann daraus geschlossen werden, dass insbesondere bipolare SSD die Annahme der Linearität und des Ordinalskalenniveaus erlauben. Hinsichtlich der Symmetrie der Daten, ergab sich diese eher bei den klassischen SSD

und war insbesondere bei den bedeutungslosen SSD nicht gegeben. Jedoch kann dieser Schluss nur unter großer Vorsicht gezogen werden, da der Range der SSD sich auf fünf oder vier Werte beschränkte, um das Problem weniger oder keiner Beobachtungen in den Kategorien zu vermeiden. Das bedeutet, dass die Befragten nicht den vollen Range der sieben Kategorien in den SSD ausschöpften, so dass es angemessener ist nur fünf Kategorien anzubieten. Diese Annahme zu überprüfen, bedarf weiterer Forschung. In Bezug auf die Erklärung der ausgeprägten Rechtsschiefe der Daten bei SSD, zeigten die Ergebnisse, dass diese bei bipolaren SSD nicht vollständig durch den Positivity Bias erklärt werden kann, da es signifikante Unterschiede zwischen den Werten der zentralen Tendenzen zwischen klassischen bipolaren SSD und der unipolaren Alternative nur für ein paar SSD gab. In Bezug auf die Linearität, Ordinalität und Symmetrie des bipolaren SSD im Vergleich zu den unipolaren und bedeutungslosen Realisierungen, lassen die Ergebnisse vermuten, dass die Befragten die Adjektivpaare im Falle klassischer SSD als Antonyme sehen. Die Erklärung, dass die Leserichtung von links nach rechts zu geringerer Wahrnehmung der rechten Seite der SSD führt, konnte durch die Daten nicht gestützt werden, weil die Befragten sich der bedeutungslosen rechten Seite bewusst waren.

Aus den Ergebnissen kann geschlossen werden, dass die klassischen SSD die Annahme der Linearität, Ordinalität und Symmetrie am ehesten erlauben und daher der unipolaren Realisierung vorgezogen werden sollten. Letzteres scheint für die Befragten ein eher ungewöhnliches Format zu sein. Da die unipolaren Ratingskalen eine größere Variabilität der Daten ergaben, kann die Entwicklung alternativer unipolarer Messinstrumente in Form von Item-spezifischen Fragen in Betracht gezogen werden (Saris, Revilla, Krosnick und Shaeffer, 2010). Ob solche Instrumente messtheoretisch besser wären als klassische SSD, muss sich in zukünftiger Forschung zeigen. Die Ergebnisse der bedeutungslosen SSD zeigen deutlich, dass Forscher/innen darauf achten sollten, semantische Gegenteile bei der Entwicklung von SSD zu finden.

Die Stärke der verwendeten Methode liegt darin, dass ein randomisiertes Between-Subjects-Design verwendet wurde, welches kausale Erklärungen der Effekte zwischen den Experimentalgruppen erlaubt. Außerdem ergaben sich ähnliche Ergebnisse für verschiedene Konzepte, nämlich Selbst- und Fremdbild, was eine Generalisierung der Ergebnisse über solche Bewertungsobjekte erlaubt.

Jedoch sollten auch Einschränkungen der Ergebnisse diskutiert werden. Die Stichprobe umfasste nur deutsche Internetnutzer, schloss also einige Menschen mit niedrigem Bildungsstand, ältere Menschen und Menschen, die Onlineumfragen meiden, aus. Zusätzlich lassen sich die Ergebnisse nur auf Onlineumfragen übertragen und sollten mit anderen und gemischten Formen der Datenerhebung repliziert werden. In der vorliegenden Studie wurden verbal verankerte Skalen mit sieben Kategorien verwendet. Wie O'Muirheartaigh, Gaskell und Wright (1995)

zeigten, können verbal und numerisch verankerte Skalen unabhängige Effekte auf das Antwortverhalten haben. Um derartige Verzerrungen zu vermeiden, wurden verschiedene Formen der verbalen Verankerung in den SSD analysiert. Weitere Studien sollten diese Beschränkung überwinden und die unabhängigen und gemeinsamen Effekte verbaler und numerischer Labels bei gleichzeitiger Variation der Kategorien in den SSD untersuchen.

Literatur

- Anderson, A. B. (1968). A Nonmetric Multidimensional Scaling Analysis of Structural Assumptions in the Semantic Differential. *Dissertation Abstracts International*, 28 (10-A), 4281.
- Attneave, F. (1959). *Applications of Information Theory to Psychology: A Summary of Basic Concepts, Methods, and Results*. New York: Henry Holt and Company, Inc.
- Baxter, J. C. (1962). Mediated generalization as a function of semantic differential performance. *The American Journal of Psychology*, 75(1), 66-76.
- Bentler, P. (1969). Semantic Space Is (Approximately) Bipolar. *Journal of Psychology*, (71), 33-40.
- Blasius, J. & Greenacre, M. (1998). *Visualization of categorical data*. New York: Academic Press.
- Breen, A. A. & Karpinski, A. (2013). Implicit and Explicit Attitudes Toward Gay Males and Lesbians Among Heterosexual Males and Females. *Journal of Social Psychology*, 153(3), 351-374.
- Di Franco, G. (2015). Multiple correspondence analysis: one or only several techniques. *Quality & Quantity*, first online, 1-17.
- Fuchs, A. (1974). A study of metric problems in the semantic differential technique. *Archiv Für Psychologie*, 126(2-4), 114-124.
- Gibbs, J. P. & Poston, D. L. (1975). The Division of Labor: Conceptualization and Related Measures. *Social Forces*, 53(3), 468-476.
- Greenacre, M. J. (1984). *Theory and applications of correspondence analysis*. London: Academic Press.
- Greenacre, M. J. & Blasius, J. (Eds.) (1994). *Correspondence analysis in the social sciences: Recent developments and applications*. London: Academic Press.
- Heise, D. R. (1977). Social action as the control of affect. *Behavioral Science*, 22(3), 163-177.
- Heise, D. R. (2010). *Surveying Cultures. Discovering Shared Conceptions and Sentiments*. Hoboken: Wiley.
- Himmelfarb, S. (1993). The measurement of attitudes. In A. H. Eagly & S. Chaiken (Eds.), *The psychology of attitudes* (pp. 23-87). Fort Worth, TX: Harcourt Brace Jovanovich.
- Kendall, D. G. (1971). Seriation from abundance matrices. In C. R. Hodson, D. G. Kendall, & P. Táutu (Eds.), *Mathematics in the Archaeological and Historical Sciences* (pp. 215-252). Edinburgh: Edinburgh University Press.
- Kervyn, N., Fiske, S. T. & Yzerbyt, V. Y. (2013). Integrating the stereotype content model (warmth and competence) and the Osgood semantic differential (evaluation, potency, and activity). *European Journal of Social Psychology*, 43(7), 673-681.
- Koga, K. & Iwasaki, Y. (2013). Psychological and physiological effect in humans of touching plant foliage - using the semantic differential method and cerebral activity as indicators. *Journal of Physiological Anthropology*, 32(1), 7.

- Krebs, D. (2012). The impact of response format on attitude measurement. In S. Salzborn, E. Davidov, & J. Reinecke (Eds.), *Methods, Theories, and Empirical Applications in the Social Sciences* (pp. 105-113). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kruskal, W. H. & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47, 583-621.
- Mann, I. T., Phillips, J. L. & Thompson, E. G. (1979). An Examination of Methodological Issues Relevant to the Use and Interpretation of the Semantic Differential. *Applied Psychological Measurement*, 3(2), 213-229.
- Menold, N. & Raykov, T. (2015). Can Reliability of Multiple Component Measuring Instruments Depend on Response Option Presentation Mode? *Educational and Psychological Measurement*. Doi: 10.1177/0013164415593602.
- Menold, N., Kaczmirek, L., Lenzner, T., & Neusar, A. (2014). How do respondents attend to verbal labels in rating scales? *Field Methods*, 26(1), 21-39.
- Messick, S. J. (1957). Metric properties of the semantic differential. *Educational and Psychological Measurement*, 17, 200-206.
- O'Muircheartaigh, C. A., Gaskell, G. D., & Wright, D. B. (1995). Weighing Anchors: Verbal and Numeric Labels for Response Scales. *Journal of Official Statistics*, 11(3), 295-307.
- Osgood, C. E. (1964). Semantic Differential Technique in the Comparative Study of Cultures. *American Anthropologist*, 66(3), 171-200.
- Osgood, C. E., May, W. H. & Miron, M. S. (1975). *Cross-Cultural Universals of Affective Meaning*. Urbana: University of Illinois Press.
- Osgood, C. E., Suci, G. J. & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.
- Ross, B. M., & Levy, N. (1960). A comparison of adjectival antonyms by simple card-pattern formation. *The Journal of Psychology: Interdisciplinary and Applied*, 49, 133-137.
- Saris, W. E., & Gallhofer, I. N. (2007 a). Estimation of the effects of measurement characteristics on the quality of survey questions. *Survey Research Methods*, 1, 31-46.
- Saris, W. E., Revilla, M. A., Krosnick, J. A., & Shaeffer, E. M. (2010). Comparing questions with agree/ disagree response options to questions with item-specific response options. *Survey Research Methods*, 4(1), 61-79.
- Schaeffer, N. C., & Presser, S. (2003). The science of asking questions. *Annual Review of Sociology*, 29, 65–88. doi:10.1146/annurev.soc.29.110702.110112
- Schriesheim, C. A., Coglisier, C. C., Newmark, R. I. & Lowensohn, S. H. (1994). The Equal-Interval Nature of Semantic Differential Scales: An Empirical Investigation Using Fiedler's Least Preferred Coworker (LPC) Scale and Magnitude Estimation and Case III Scaling Procedures. *Educational and Psychological Measurement*, 54(2), 253-62.
- Schwarz, N., Knäuper, B., Hippler, H.-J., Noelle-Neumann, E., & Clark, L. (1991). Rating Scales: Numeric Values may change the meaning of scale labels. *Public Opinion Quarterly*, 55, 570–582.

- Skrandies, W. (2011). The structure of semantic meaning: A developmental study. *Japanese Psychological Research*, 53(1), 65-76.
- Stevens, S. S. (1975). *Psychophysics. Introduction to its perceptual, neural, and social prospects*. New York: John Wiley & Sons.
- Tourangeau, R. (1984). Cognitive science and survey methods. In T. Jabine, M. Straf, J. Tanur, & R. Tourangeau (Eds.), *Cognitive aspects of survey methodology: Building a bridge between disciplines* (pp. 73- 100). Washington, DC: National Academy Press.
- Tourangeau, R., Rips, L. J. & Rasinski, K. (2000). *The Psychology of Survey Response*. Cambridge: Cambridge University Press.
- Verhagen, T., van Den Hooff, B., & Meents, S. (2015). Toward a Better Use of the Semantic Differential in IS Research: An Integrative Framework of Suggested Action. *Journal of The Association For Information Systems*, 16(2), 108-143.
- Vielleman, P. & Wilkinson, L. (1994). Nominal, ordinal, interval and ratio typologies are misleading. In I. Borg & P. Ph. Mohler (Eds.), *Trends and Perspectives in Empirical Social Research* (pp. 161 – 176). Berlin: Walter de Gruyter & Co.
- Weller, S. C. & Romney, A. K. (1990). *Metric scaling. Correspondence analysis*. Newbury Park: Sage.

Anhang

Tabelle A-1: Stichprobenbeschreibung nach experimentellen Bedingungen

	Experimentelle Bedingung 1 Bipolare SSD	Experimentelle Bedingung 2 Unipolare SSD	Experimentelle Bedingung 3 Unsinnige SSD	Total*
Geschlecht				
männlich	90# 52.9%+	101 54.3%	93 54.1%	284 53.8%
weiblich	80 47.1%	85 45.7%	79 45.9%	244 46.2%
$\chi^2_{(df=2)} = 0.07$				
Alter				
Mittelwert	43.04	42.25	43.84	43.02
SD	14.79	15.33	14.38	14.84
$F_{(df1=2; df2=529)} = 0.51; \text{Partielles } \eta^2 = .002$				
Bildung				
elementar	33 20.63%	37 20.79%	40 24.39%	110 21.91%
mittel	23 14.38%	28 15.73%	23 14.02%	74 14.74%
hoch	104 65.00%	113 63.48%	101 61.59%	318 63.35%
$\chi^2_{(df=4)} = 1.02$				
Erwerbstätigkeit				
ja	110 63.95%	124 67.03%	119 68.79%	353 66.60%
nein	62 36.05%	61 33.97%	54 31.21%	177 33.40%
$\chi^2_{(df=2)} = 0.93$				

*Abweichung von insgesamt N = 537 aufgrund fehlender Daten;

Absolute Häufigkeiten, + relative Häufigkeiten

Tabelle A-2: Randverteilungen der Skalen des Semantischen Differentials (SSD)

Selbstbeschreibungen

Dimension Evaluation	1	2	3	4	5	6	7	FW*
Bipolare SSD+ (N = 175)								
gut – schlecht	10,86%	41,71%	25,14%	14,29%	1,71%	0,00%	0,00%	6,29%
optimistisch – pessimistisch	17,71%	38,86%	23,43%	9,71%	2,29%	4,57%	0,57%	2,86%
positiv – negativ	17,14%	44,00%	21,14%	9,14%	4,57%	0,57%	0,00%	3,43%
vollkommen – unvollkommen	0,00%	3,43%	17,71%	49,14%	17,14%	7,43%	1,14%	4,00%
zeitgemäß – altmodisch	6,29%	37,14%	21,14%	24,00%	5,71%	1,71%	0,57%	3,43%
Unipolare SSD (N = 188)								
gut – nicht gut	12,77%	31,91%	29,26%	19,68%	1,06%	1,06%	0,00%	4,26%
optimistisch – nicht optimistisch	17,55%	31,91%	24,47%	10,11%	9,04%	3,72%	0,53%	2,66%
positiv – nicht positiv	16,49%	35,64%	28,19%	7,98%	6,38%	1,60%	0,00%	3,72%
vollkommen – nicht vollkommen	0,53%	3,19%	14,89%	28,72%	23,40%	15,96%	10,64%	2,66%
zeitgemäß – nicht zeitgemäß	9,04%	28,19%	26,60%	18,09%	9,57%	5,32%	0,53%	2,66%
Unsinnige SSD (N = 174)								
gut – träge	17,24%	27,59%	27,59%	13,22%	4,02%	2,30%	0,00%	8,05%
optimistisch – gelassen	10,92%	29,31%	22,99%	20,69%	3,45%	5,17%	2,30%	5,17%
positiv – mild	13,79%	33,91%	27,01%	12,64%	3,45%	2,87%	0,57%	5,75%
vollkommen – nachgiebig	1,72%	7,47%	22,99%	33,33%	15,52%	9,20%	2,87%	6,90%
zeitgemäß – schwach	6,90%	31,61%	28,74%	20,11%	4,60%	1,72%	0,00%	6,32%

* Fehlende Werte

+ Skalen des Semantischen Differenzials

Dimension Potency	1	2	3	4	5	6	7	MV
Bipolare SSD (N = 175)								
hart – weich	0,57%	7,43%	24,00%	25,14%	21,71%	12,57%	5,71%	2,86%
tiefgründig – oberflächlich	9,14%	39,43%	29,14%	14,29%	2,86%	1,71%	0,00%	3,43%
maskulin – feminin	4,00%	18,86%	13,14%	18,29%	8,00%	24,00%	10,29%	3,43%
streng – mild	1,14%	13,14%	17,71%	33,71%	16,00%	13,14%	2,29%	2,86%
stark – schwach	4,57%	22,86%	30,86%	29,14%	6,86%	1,71%	1,14%	2,86%
hartnäckig – nachgiebig	7,43%	27,43%	21,14%	24,00%	8,57%	7,43%	0,57%	3,43%
Unipolare SSD (N = 188)								
hart – nicht hart	2,13%	7,45%	18,09%	31,38%	16,49%	16,49%	4,79%	3,19%
tiefgründig – nicht tiefgründig	9,57%	30,32%	28,19%	17,55%	6,91%	4,26%	0,53%	2,66%
maskulin – nicht maskulin	2,13%	16,49%	19,15%	15,43%	9,57%	12,23%	21,28%	3,72%
streng – nicht streng	2,66%	14,89%	20,21%	22,87%	17,02%	12,77%	6,91%	2,66%
stark – nicht stark	5,85%	25,53%	27,13%	22,87%	11,17%	4,26%	0,53%	2,66%
hartnäckig – nicht hartnäckig	10,64%	34,57%	29,79%	15,43%	4,79%	2,13%	0,00%	2,66%
Unsinnige SSD (N = 174)								
hart – passiv	4,02%	10,34%	27,01%	33,91%	13,79%	1,72%	0,57%	8,62%
tiefgründig – feminin	12,07%	27,01%	21,26%	20,69%	4,60%	5,17%	3,45%	5,75%
maskulin – oberflächlich	4,02%	19,54%	22,41%	32,76%	6,32%	2,30%	0,00%	12,64%
streng – negativ	3,45%	13,79%	33,91%	34,48%	3,45%	1,15%	0,00%	9,77%
stark – altmodisch	8,62%	25,29%	32,76%	17,82%	4,02%	3,45%	0,57%	7,47%
hartnäckig – unvollkommen	9,77%	24,71%	28,16%	16,67%	9,20%	3,45%	0,57%	7,47%

(Fortsetzung Tabelle A-2)

Dimension Activity	1	2	3	4	5	6	7	MV
Bipolare SSD (N = 175)								
aktiv – passiv	8,00%	35,43%	28,00%	14,29%	6,86%	1,71%	1,71%	4,00%
reizbar – gelassen	2,29%	15,43%	18,86%	24,00%	13,71%	15,43%	7,43%	2,86%
energisch – träge	5,71%	29,71%	30,86%	23,43%	4,57%	1,71%	0,57%	3,43%
Unipolare SSD (N = 188)								
aktiv – nicht aktiv	12,23%	34,04%	32,98%	10,11%	5,32%	2,13%	0,00%	3,19%
reizbar – nicht reizbar	7,45%	15,96%	19,68%	20,74%	13,30%	16,49%	3,72%	2,66%
energisch – nicht energisch	6,91%	26,60%	35,64%	17,55%	3,72%	6,38%	0,00%	3,19%
Unsinnige SSD (N = 174)								
aktiv – weich	15,52%	27,01%	23,93%	17,24%	6,90%	1,72%	0,00%	5,75%
reizbar – pessimistisch	5,75%	13,79%	21,26%	43,68%	4,60%	1,72%	0,57%	8,62%
energisch – schlecht	9,77%	29,31%	38,51%	12,07%	2,30%	1,15%	0,00%	6,90%

Fremdbeschreibungen

Dimension Evaluation	1	2	3	4	5	6	7	MV
Bipolar (N = 175)								
gut – schlecht	19,43%	32,57%	20,00%	21,14%	2,29%	1,14%	1,14%	2,29%
optimistisch – pessimistisch	16,00%	32,00%	26,29%	20,57%	1,71%	1,14%	0,57%	1,71%
positiv – negativ	19,43%	40,57%	17,14%	16,00%	1,71%	1,71%	1,14%	2,29%
vollkommen – unvollkommen	1,71%	4,00%	26,86%	51,43%	6,86%	4,57%	2,86%	1,71%
zeitgemäß – altmodisch	6,29%	22,29%	25,14%	28,57%	8,57%	5,14%	2,29%	1,71%
Unipolar (N = 188)								
gut – nicht gut	19,15%	26,06%	23,40%	20,21%	2,13%	2,66%	2,66%	3,72%
optimistisch – nicht optimistisch	14,89%	29,79%	25,53%	21,81%	2,66%	0,53%	1,06%	3,72%
positiv – nicht positiv	22,34%	27,66%	25,00%	15,96%	2,13%	2,13%	1,60%	3,19%
vollkommen – nicht vollkommen	1,06%	7,45%	14,36%	42,02%	11,70%	11,70%	8,51%	3,19%
zeitgemäß – nicht zeitgemäß	7,98%	18,09%	21,28%	27,13%	13,83%	5,32%	2,66%	3,72%
Unsinnig (N = 174)								
gut – träge	21,84%	25,29%	18,97%	18,39%	5,17%	2,30%	0,57%	7,47%
optimistisch – gelassen	11,49%	16,09%	17,24%	19,54%	13,79%	9,77%	4,60%	7,47%
positiv – mild	17,82%	20,11%	19,54%	19,54%	10,34%	2,30%	0,57%	9,77%
vollkommen – nachgiebig	3,44%	10,34%	19,54%	36,21%	13,79%	5,17%	1,15%	10,34%
zeitgemäß – schwach	11,49%	21,26%	20,69%	23,56%	8,05%	2,30%	2,87%	9,77%

Dimension Potency	1	2	3	4	5	6	7	MV
Bipolar (N = 175)								
hart – weich	0,00%	6,86%	21,14%	44,00%	19,43%	5,71%	1,14%	1,71%
tiefgründig – oberflächlich	10,86%	38,29%	17,71%	22,29%	3,43%	4,00%	1,14%	2,29%
maskulin – feminin	8,00%	24,57%	26,86%	34,29%	3,43%	1,14%	0,00%	1,71%
streng – mild	0,57%	8,57%	24,57%	42,86%	15,43%	6,29%	0,00%	1,71%
stark – schwach	7,43%	28,00%	26,86%	28,57%	3,43%	3,43%	0,57%	1,71%
hartnäckig – nachgiebig	8,00%	31,43%	20,57%	33,14%	3,43%	1,14%	0,00%	2,29%
Unipolar (N = 188)								
hart – nicht hart	1,60%	3,19%	17,02%	37,77%	22,34%	10,11%	3,19%	4,79%
tiefgründig – nicht tiefgründig	17,02%	28,19%	15,96%	23,94%	5,85%	2,13%	2,66%	4,26%
maskulin – nicht maskulin	4,26%	13,83%	18,09%	39,89%	10,64%	4,79%	4,79%	3,72%
streng – nicht streng	1,60%	10,64%	20,74%	38,83%	11,17%	10,11%	3,19%	3,72%
stark – nicht stark	11,17%	21,81%	19,68%	26,60%	11,70%	2,13%	3,19%	3,72%
hartnäckig – nicht hartnäckig	13,30%	27,13%	23,40%	25,53%	4,26%	0,53%	1,60%	4,26%
Unsinnig (N = 174)								
hart – passiv	1,72%	7,47%	21,84%	42,53%	8,05%	5,75%	2,30%	10,34%
tiefgründig – feminin	14,94%	24,71%	24,14%	23,56%	1,15%	0,57%	1,15%	9,77%
maskulin – oberflächlich	7,47%	20,11%	27,59%	25,86%	5,17%	2,30%	1,15%	10,34%
streng – negativ	2,87%	12,64%	28,16%	37,93%	5,75%	1,15%	0,57%	10,92%
stark – altmodisch	8,05%	19,54%	26,44%	23,56%	5,75%	4,60%	2,87%	9,20%
hartnäckig – unvollkommen	12,07%	24,71%	23,56%	21,84%	2,87%	2,30%	1,72%	10,92%

(Fortsetzung Tabelle A-2)

Dimension Activity	1	2	3	4	5	6	7	MV
Bipolar (N = 175)								
aktiv – passiv	17,71%	31,43%	22,86%	22,29%	2,29%	1,14%	0,00%	2,29%
reizbar – gelassen	0,00%	1,71%	7,43%	33,14%	17,14%	30,29%	8,57%	1,71%
energisch – träge	8,57%	25,14%	24,57%	34,29%	5,14%	0,57%	0,00%	1,71%
Unipolar (N = 188)								
aktiv – nicht aktiv	17,55%	30,32%	22,87%	18,09%	5,32%	1,06%	1,06%	3,72%
reizbar – nicht reizbar	2,13%	3,19%	7,45%	38,30%	18,62%	18,62%	7,98%	3,72%
energisch – nicht energisch	9,04%	21,81%	26,06%	27,66%	6,91%	2,13%	2,66%	3,72%
Unsinnig (N = 174)								
aktiv – weich	17,24%	27,59%	21,26%	17,24%	5,75%	1,72%	0,57%	8,62%
reizbar – pessimistisch	0,57%	4,60%	16,67%	55,75%	9,20%	1,72%	0,00%	11,49%
energisch – schlecht	7,47%	26,44%	29,31%	20,11%	5,17%	1,72%	1,72	8,05%

Tabelle A-3: Zentrale Tendenzen und Variation der Skalen des Semantischen Differentials (SSD)

Selbstbeschreibungen

	Experimentelle Bedingung 1 Bipolare SSD		Experimentelle Bedingung 2 Unipolare SSD		Experimentelle Bedingung 3 Unsinnige SSD	
	IQR	IQV	IQR	IQV	IQR	IQV
Dimension Evaluation						
Gut	1	0.81	1	0.86	1	0.89
Optimistisch	1	0.86	1	0.91	2	0.91
Positiv	1	0.82	1	0.86	1	0.87
Vollkommen	1	0.78	2	0.92	2	0.89
Zeitgemäß	2	0.86	2	0.92	2	0.86
Dimension Potency						
Hart	2	0.93	2	0.93	1	0.86
Tiefgründig	1	0.83	2	0.90	2	0.93
Maskulin	3	0.96	3	0.97	2	0.86
Streng	2	0.91	2	0.96	1	0.80
Stark	2	0.87	2	0.91	2	0.88
Hartnäckig	2	0.92	2	0.86	2	0.91
Dimension Activity						
Aktiv	2	0.87	1	0.85	2	0.91
Reizbar	2	0.96	2	0.97	1	0.80
Energisch	2	0.86	2	0.87	1	0.82

(Fortsetzung Tabelle A-3)

Fremdbeschreibungen

	Experimentelle Bedingung 1 Bipolare SSD		Experimentelle Bedingung 2 Unipolare SSD		Experimentelle Bedingung 3 Unsinnige SSD	
	IQR	IQV	IQR	IQV	IQR	IQV
Dimension Evaluation						
Gut	2	0.89	2	0.91	2	0.92
Optimistisch	1	0.88	2	0.88	3	0.98
Positiv	1	0.85	1	0.90	2	0.94
Vollkommen	1	0.75	1	0.87	1	0.87
Zeitgemäß	2	0.91	2	0.94	2	0.93
Dimension Potency						
Hart	2	0.82	1	0.87	1	0.81
Tiefgründig	2	0.87	2	0.92	2	0.88
Maskulin	2	0.86	1	0.88	2	0.89
Streng	1	0.83	2	0.88	1	0.81
Stark	2	0.88	2	0.93	2	0.92
Hartnäckig	2	0.85	2	0.90	2	0.90
Dimension Activity						
Aktiv	2	0.88	2	0.90	2	0.91
Reizbar	2	0.87	2	0.88	0	0.65
Energisch	2	0.86	2	0.91	2	0.88

Tabelle A-4: Ergebnisse der Multiplen Korrespondenzanalyse nach experimentellen Bedingungen

Selbstbeschreibungen

		Experimentelle Bedingung 1 Bipolare SSD		Experimentelle Bedingung 2 Unipolare SSD		Experimentelle Bedingung 3 Unsinnige SSD	
		Dim. 1	Dim. 2	Dim. 1	Dim. 2	Dim. 1	Dim. 2
Dimension Evaluation							
Optimistisch	1	1.335	1.434	1.137	1.981	3.058	0.239
	2	0.419	-0.115	0.853	-0.987	-0.228	-1.464
	3	-0.345	-1.300	-0.440	-0.812	-0.359	0.316
	4	-0.848	-0.913	-1.323	0.877	-0.389	0.876
	5	-3.390	2.584	-1.699	0.481	-0.617	1.802
Positiv	1	1.369	1.559	1.123	2.255	2.699	0.376
	2	0.360	-0.234	0.852	-0.849	-0.144	-1.310
	3	-0.311	-1.311	-0.691	-0.577	-0.621	0.222
	4	-1.650	-0.375	-1.458	0.502	-0.502	1.695
	5	-3.728	3.249	-2.099	0.759	-0.826	1.905
Zeitgemäß	1	1.337	1.308	0.547	1.654	3.087	0.862
	2	0.268	0.407	0.653	-0.525	0.195	-0.171
	3	-0.493	-0.208	0.519	0.039	-0.286	-0.637
	4	0.074	-0.652	-0.680	-0.432	-0.601	0.325
	5	-1.295	-0.543	-1.654	0.489	-0.649	2.125
Anteil gesamte Inertia		16.44%	14.77%	18.10%	14.89%	16,85%	13,75%
Dimension Potency							
Hart	2	1.610	2.214	2.751	1.386	0.670	1.595
	3	1.195	-0.631	1.114	-0.211	0.911	0.094
	4	0.196	-0.439	-0.129	-1.238	-0.027	-0.791
	5	-1.257	0.368	-1.187	1.291	-2.096	0.501
Maskulin	2	1.014	1.117	1.302	1.006	1.172	1.508
	3	1.188	-0.512	-0.055	-0.156	0.330	-0.751
	4	-0.108	-1.926	-0.263	-1.756	-0.372	-0.807
	5	-1.012	0.625	-0.697	0.567	-2.508	2.360
Hartnäckig	2	1.118	1.121	1.398	0.233	0.823	0.921
	3	0.323	-1.474	-0.560	-1.044	0.553	-0.399
	4	-0.673	-0.400	-1.111	0.777	-0.891	-1.476
	5	-1.655	0.719	-1.299	1.756	-1.885	1.375
Anteil gesamte Inertia		19.17%	14.06%	16.57%	15.61%	21.49%	15.06%
Dimension Activity							
Aktiv	1	2.007	1.870	2.345	0.638	2.023	0.757
	2	0.592	-0.595	-0.092	-1.194	0.104	-1.304
	3	-0.713	-0.576	-0.573	0.461	-0.765	0.095
	4	-1.366	1.591	-0.876	1.816	-0.752	1.317
Energisch	1	1.559	2.812	3.183	1.130	2.502	1.293
	2	1.088	-0.393	0.246	-1.245	0.288	-1.140
	3	-0.562	-0.762	-0.508	0.083	-0.622	0.172
	4	-1.188	1.029	-0.629	1.513	-1.064	1.891
Anteil gesamte Inertia		25,84%	21.09%	26.61%	23.61%	26.59%	21.90%

(Fortsetzung Tabelle A-4)

Fremdbeschreibungen

		Experimentelle Bedingung 1 Bipolare SSD		Experimentelle Bedingung 2 Unipolare SSD		Experimentelle Bedingung 3 Unsinnige SSD	
		Dim. 1	Dim. 2	Dim. 1	Dim. 2	Dim. 1	Dim. 2
Dimension Evaluation							
Gut	1	1.336	1.434	1.545	1.308	1,647	0.667
	2	0.522	-0.870	0.591	-1.156	0.057	-1.250
	3	-0.186	-0.913	-0.410	-0.888	-0.576	-0.680
	4	-1.560	0.745	-1.293	0.930	-1.023	1.238
Optimistisch	1	1.462	1.748	1.683	1.603	2.030	1.000
	2	0.603	-0.783	0.523	-0.914	0.571	-1.160
	3	-0.325	-1.030	-0.332	-0.786	-0.440	-0.950
	4	-1.465	1.007	-1.287	0.937	-0.786	0.736
Positiv	1	1.448	1.640	1.465	1.112	1.905	0.847
	2	0.407	-0.966	0.413	-1.256	0.232	-1.664
	3	-0.483	-0.946	-0.512	-0.711	-0.549	-0.492
	4	-1.781	1.119	-1.453	1.237	-0.888	1.033
Zeitgemäß	1	1.620	2.275	1.784	2.215	2.228	1.056
	2	0.873	-0.348	0.989	-0.420	0.373	-0.797
	3	0.315	-0.528	0.438	-0.853	-0.260	-1.039
	4	-0.850	0.146	-0.823	0.180	-0.851	0.870
Anteil gesamte Inertia		24.24%	19.11%	25.43%	18.26%	25.59%	17.24%
Dimension Potency							
Tiefgründig	1	1.377	2.190	1.968	1.081	2.057	1.082
	2	0.749	-0.421	0.130	-1.551	0.166	-1.575
	3	0.184	-1.332	-0.346	-0.241	-0.363	-0.263
	4	-1.520	0.517	-1.064	0.919	-0.996	0.887
Maskulin	1	0.850	1.647	2.877	1.781	2.456	1.876
	2	0.750	-0.106	0.727	-0.826	0.813	-0.843
	3	0.380	-0.628	0.451	-0.726	0.530	-0.784
	4	-0.901	0.160	-0.526	0.322	-0.518	0.292
Stark	1	1.920	3.724	2.441	1.726	2.478	1.765
	2	0.939	-0.466	0.631	-1.415	0.681	-1.458
	3	0.326	-1.207	-0.131	-0.941	-0.088	-1.006
	4	-1.365	0.475	-0.927	0.814	-0.916	0.762
Hartnäckig	1	1.644	3.055	2.324	1.281	2.383	1.326
	2	0.887	-0.741	0.210	-1.306	0.180	-1.227
	3	0.184	-0.616	-0.211	-0.463	-0.166	-0.564
	4	-1.189	0.306	-1.001	0.968	-0.984	0.926
Anteil gesamte Inertia		19.47%	14.29%	21.34%	16.54%	20.91%	16.60%

(Fortsetzung Tabelle A-4)

Dimension Activity							
Aktiv	1	1.137	1.502	1.754	0.677	1.220	1.589
	2	0.820	-0.697	-0.452	0.826	0.617	-0.722
	3	-0.540	-1.000	-0.540	0.051	-0.117	-0.790
	4	-1.361	0.885	0.523	-1.866	-1.150	-0.133
	5	-1.412	-0.684	-2.054	-0.791	-2.058	1.611
Energisch	1	1.042	1.954	2.158	0.798	1.549	2.649
	2	1.319	0.028	0.102	1.119	0.836	-0.120
	3	-0.003	-1.465	-0.717	0.423	0.074	-0.845
	4	-1.010	0.571	0.501	-1.395	-1.021	-0.068
	5	-1.402	-0.137	-1.838	-0.530	-2.283	1.747
Anteil gesamte Inertia		20.95%	17.83%	18.78%	18.25%	21.45%	19.06%

Bisher erschienene Eichstätter Beiträge zur Soziologie

Nr. 1: Knut Petzold, Thomas Brunner & Carlos Watzka: Determinanten der Zustimmung zur Frauenquote. Ergebnisse einer Bevölkerungsumfrage (12/2013)

Nr. 2: Joost van Loon & Laura Unsöld: The Work and the Net: a Critical Reflection on Facebook-Research Methods and Optical Mediation (2/2014)

Nr. 3: Stefanie Eifler: Projekt „Zusammenleben in der Stadt“ - Methodendokumentation (3/2014)

Nr. 4: Basil Wiese: Intercultural Atmospheres – The Affective Quality of Gift Situations (9/2014)

Nr. 5: Florian Mayr: Gibt es Krisen, und wenn ja wie viele? – Theoretisch-konzeptionelle Überlegungen zu einer Soziologie der Krise (12/2014)

Nr. 6: Ramona Kay: Delinquente Opfer und viktimisierte Täter? Eine Analyse des Offending-Victimization-Overlap mit Hilfe der International Self-Report Delinquency Study 2 (ISRD-2) (Working Paper) (5/2015)

Nr. 7: Stefanie Eifler, Natalja Menold & Sara Pinkas: Metrische Eigenschaften des Semantischen Differentials zur Messung von Selbst- und Fremdbeschreibungen (11/2015)