

Wir bitten um ein differenziertes Urteil: Möglichkeiten zur Reduktion von Nondifferentiation bei Rating-Skalen

Eva Böbel, Beatrice Gerber-Braun, Martin J. Binser, Matthias Spörrle & Friedrich Försterling

Zusammenfassung

Nondifferentiation ist ein generelles Phänomen bei der Einschätzung von Meinungsgegenständen mittels Rating-Skalen: Zwischen den Meinungsgegenständen wird nicht so deutlich unterschieden, wie aufgrund der Skalen möglich. In zwei experimentellen Studien zu unterschiedlichen Meinungsgegenständen werden in der Literatur diskutierte Verfahren zur Reduzierung der Nondifferentiation zusammen mit einem hierfür bisher nicht verwendeten Verfahren (konstante Summe, KS) hinsichtlich ihres Einflusses auf Nondifferentiation verglichen. Eine erste Studie (N = 139) belegt eine signifikant erhöhte Differenzierung bei KS im Vergleich zu einem verbalen Einschub. Für eine zweite Studie werden in einer Voruntersuchung (N = 138) stark positiv und stark negativ eingeschätzte Meinungsgegenstände ausgewählt, die in der anschließenden Hauptuntersuchung (N = 200) ohne Messwiederholung bewertet werden. Hierbei werden insgesamt 3 verschiedene Verfahren zur Verbesserung der Differenzierung verwendet und einem rein verbalen Einschub gegenübergestellt. Auch hierbei erweist sich die KS generell als das Verfahren, das die höchste Differenzierung bewirkt.

Einleitung

Ein weit verbreitetes Verfahren zur Erfassung von Antworten ist die Rating-Technik, bei der Probanden auf einer mehrstufigen Skala die individuelle Ausprägung ihrer Antwort angeben. Neben den Vorteilen der Rating-Skalen, wie z.B. die schnelle Durchführbarkeit und leichte Auswertbarkeit, wird in der Literatur ein wichtiger methodischer Nachteil diskutiert (z.B. Munson & McIntyre, 1979): Werden Probanden um mehrfache Urteile bzw. Einstufungen gebeten, differenzieren sie häufig nicht in so deutlicher Weise in ihren Urteilen, wie dies aufgrund der Skalenvorgaben möglich wäre. Dies erklärt sich u.a. aus generellen Zustimmung- und Ablehnungstendenzen, sowie aus Satisficing-Tendenzen (Krosnick, 1999) und wird als Nondifferentiation bezeichnet (Krosnick & Alwin, 1988).

In der Literatur finden sich nur wenige Arbeiten, die sich mit einer Reduzierung der Nondifferentiation auseinandersetzen: So konnten McCarty und Shrum (2000) in ihrer Studie zeigen, dass der Einschub alternativer Skalierungen vor den Rating-Skalen zu einer Reduzierung der Nondifferentiation führt. In den hier beschriebenen Studien werden alle bisher in der Literatur diskutierten Verfahren zusammen mit einer hierfür bisher nicht verwendeten Skalierung (konstante Summe) hinsichtlich ihres Einflusses auf Nondifferentiation untereinander verglichen. Bei diesem Verfahren werden die Probanden gebeten vor der Einschätzung der Meinungsgegenstände auf Rating-Skalen eine fest vorgegebene Gewichtungssumme auf die Meinungsgegenstände aufzuteilen.

In einer ersten Studie wird der Einschub der konstanten Summe vor den mittels Rating-Skalen einzuschätzenden Meinungsgegenständen mit der Erfassung der Meinung ohne Vorschaltung einer alternativen Skalierung sowie mit der Vorgabe eines verbalen Einschubs verglichen.

Für eine zweite Studie werden in einer Voruntersuchung Meinungsgegenstände ausgewählt, die eine Ballung der Skalenwerte am oberen bzw. unteren Skalenende verursachen. In der Hauptuntersuchung werden diese dann ohne und mit Vorschaltung alternativer Skalierungen (konstante Summe, Ranking-Skala, Most-least Methode) bewertet.

Methode

Studie 1:

Es wurden 139 Personen (49.6 % Frauen, Altersspanne 15 bis 78, $M = 33.2$, $SD = 13.2$) zur gesellschaftlichen Verantwortung der Psychologie als Wissenschaft befragt. Die drei Fragebogenversionen unterschieden sich hinsichtlich des Vorhandenseins einer vorgeschalteten alternativen Skalierung vor den zur Einschätzung der Meinungsgegenständen verwendeten Rating-Skalen (konstante Summe (KS) vs. verbaler Instruktion (VI) vs. keiner Vorschaltung). Somit liegt der Untersuchung ein einfaktorielles Design mit 3 Faktorstufen ohne Messwiederholung zugrunde.

AVs: Varianz (s^2), Spannweite (sw) und Summe der Absolutbeträge aller Skalenwertdifferenzen (sad).

Studie 2:

Voruntersuchung: Es wurden 138 Personen (72.5 % Frauen, Altersspanne 16 bis 69, $M = 29.5$, $SD = 12.7$) zum Thema „Gentechnik und Lebensmittel“ befragt.

Hauptuntersuchung:

200 Personen (85 % Frauen, Altersspanne 16 bis 65, $M = 25.4$, $SD = 7.2$) wurden zu den, aus der Voruntersuchung ausgewählten, stark positiv bzw. stark negativ eingeschätzten Meinungsgegenständen befragt. Die 10 Fragebogenversionen unterschieden sich hinsichtlich des Vorhandenseins einer vorgeschalteten alternativen Skalierung vor den zur Einschätzung der Meinungsgegenständen verwendeten Rating-Skalen (konstante Summe (KS) vs. Ranking-Skala (R) vs. Most-least-Methode (ML) vs. verbaler Instruktion (VI) vs. keine Vorschaltung) und hinsichtlich der Einschätzungstendenz der Items (zustimmend vs. ablehnend). Somit liegt der Untersuchung ein 2 x 5 faktorielles Design zugrunde.

AVs: Varianz (s^2), Spannweite (sw) und Summe der Absolutbeträge aller Skalenwertdifferenzen (sad).

Literatur:

- Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology*, 50, 537-567.
- Krosnick, J. A. & Alwin, D. F. (1988). A test of the form-resistant correlation hypothesis: Ratings, rankings, and the measurement of values. *Public Opinion Quarterly*, 52, 526-38.
- McCarty, J. A. & Shrum, L. J. (2000). The measurement of personal values in survey research: A test of alternative rating procedures. *Public Opinion Quarterly*, 64, 271-298.
- Munson, J. M. & McIntyre, S. H. (1979). Developing practical procedures for the measurement of personal values in cross-cultural marketing. *Journal of Marketing Research*, 16, 48-52.

Ergebnisse I

- Zur Datenanalyse werden hierarchische multiple lineare Regressionen verwendet. Die unabhängige Variable „vorgeschaltete alternative Skalierung“ wird in Dummy-Variablen zerlegt.
- In den dargestellten Regressionsgleichungen werden nur die Variablen berichtet, deren beta-Koeffizienten als signifikant in der Regressionsgleichung verbleiben: In der ersten Studie ist dies nur KS, in der zweiten Studie sind dies KS, ML und R.
- In der zweiten Studie ist der Einfluss der unabhängigen Variablen Einschätzungstendenz auf die drei abhängigen Variablen (s^2 , sw, sad) mit $t(197/198) > 3.50$, $p < .001$ signifikant.

Studie 1:

a) $s^2 = 3.79 + 2.62 \cdot KS$

$F(2, 135) = 2.65$, $p = .074$
 $t(135) = 2.10$, $p < .05$, $R^2 = .04$

b) $sw = 3.24 + 1.38 \cdot KS$

$F(2, 135) = 3.29$, $p < .05$
 $t(135) = 2.32$, $p < .05$, $R^2 = .22$

c) $sad = 11.12 + 4.19 \cdot KS$

$F(135) = 2.85$, $p = .061$
 $t(135) = 2.11$, $p < .05$, $R^2 = .20$

Studie 2:

a) $s^2 = 3.30 + 2.52 \cdot KS$

$F(4, 195) = 2.79$, $p < .05$
 $t(194) = 2.50$, $p < .05$, $R^2 = .05$

b) $sw = 3.05 + 1.23 \cdot ML + 1.55 \cdot KS$

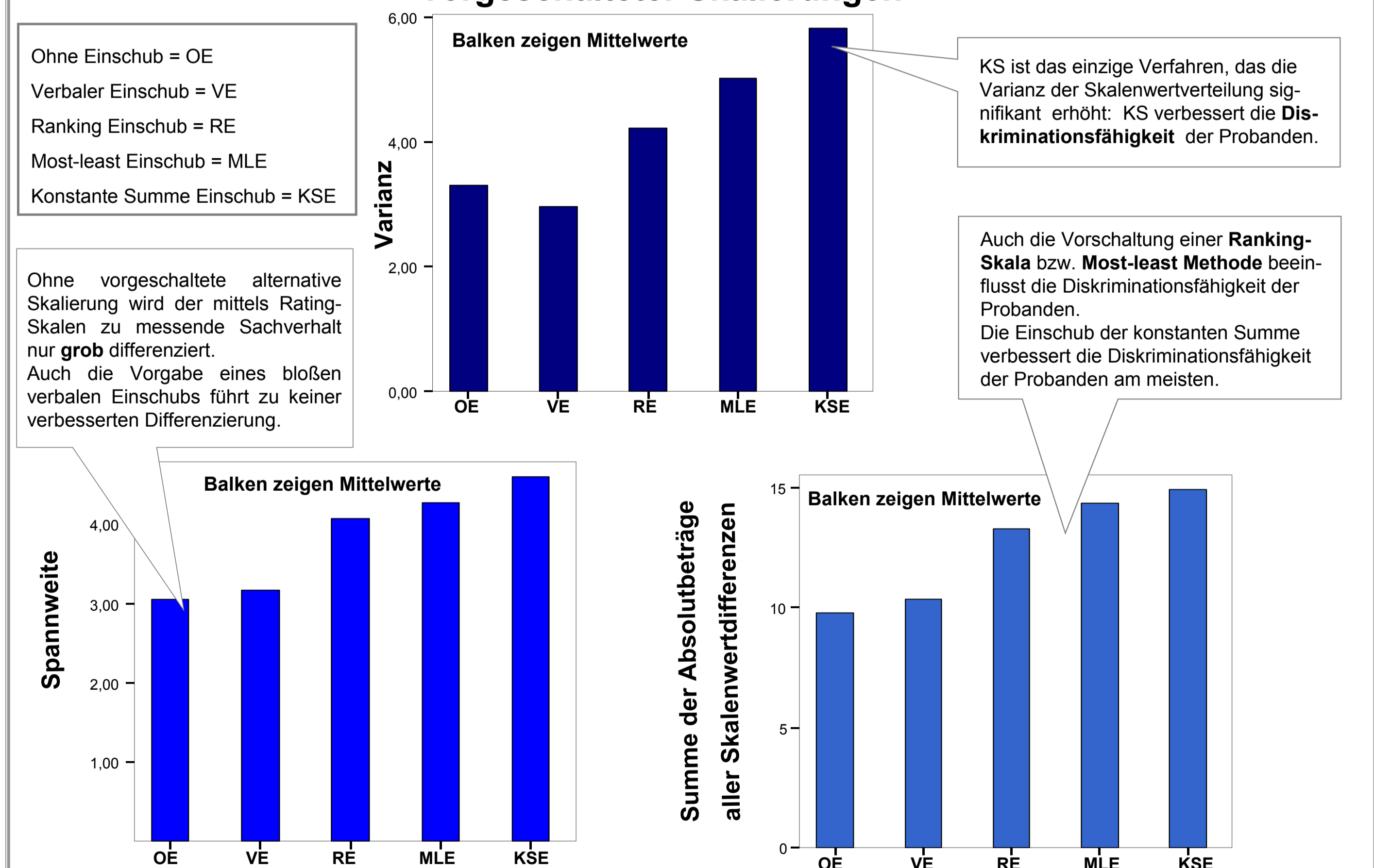
$F(4, 195) = 3.49$, $p < .05$
 $t_{ML}(194) = 2.53$, $p < .05$

c) $sad = 9.8 + 3.48 \cdot R + 4.56 \cdot ML + 5.13 \cdot KS$

$t_{KS}(194) = 2.98$, $p < .05$, $R^2 = .07$
 $F(4, 195) = 3.86$, $p < .05$
 $t_R(194) = 2.07$, $p < .05$,
 $t_{ML}(194) = 2.70$, $p < .05$
 $t_{KS}(194) = 3.05$, $p < .05$, $R^2 = .07$

Ergebnisse II

Verschiedene Streuungsmaße der mittels Rating-Skalen eingeschätzten Meinungsgegenstände in Abhängigkeit unterschiedlicher vorgeschalteter Skalierungen



Kontaktadressen eva.boebel@campus.lmu.de spoerrle@psy.uni-muenchen.de

Fazit

- Die Vorschaltung der Methode der konstanten Summe erweist sich in beiden Studien als das Verfahren, das die Diskriminationsfähigkeit der Probanden am meisten verbessert.
- Ohne Vorschaltung alternativer Skalierungen werden die mittels Rating-Skalen einzuschätzenden Meinungsgegenstände nur grob differenziert. Die Vorgabe einer bloßen verbalen Instruktion führt zu keiner verbesserten Differenzierung.