

Ja-Prozente bei 6-stufigen Antwortskalen

Andreas Kögel

Zusammenfassung

Im Gegensatz zu den gängigeren 5er-Skalen sind bei Antwortskalen mit 6 Antwortabstufungen zwei Varianten der Ermittlung von Ja-Prozenten möglich. Anhand einer Kundenbefragung werden die Ja-Prozente auf Basis der beiden stärker zustimmenden Items mit den Ja-Prozenten auf Basis aller 3 zustimmenden Items verglichen. Dabei stellt sich heraus, dass die Unterschiede zwischen beiden Varianten gering sind, dass aber die Variante mit 3 zustimmenden Items eine geringere Schwankung der Passung aufweist und insgesamt mit den Antwortverteilungen der Items etwas stärker korreliert, als die andere Variante. Die Korrelationen mit den Item-Mittelwerten sind für beide Varianten extrem hoch.

Wozu Ja-Prozente?

Mit sogenannte „Likert-Skalen“ können Befragte Ihre Einschätzungen zu bestimmten Sachverhalten durch vorgegebene Antwortkategorien abgeben. Standard sind 5 Antwortabstufungen, aber auch 6 Abstufungen werden häufig verwendet. Die beiden Varianten entscheiden sich darin, dass es für die 6er-Skala keine Mittelkategorie gibt. Eine eindeutige Klärung, ob eine der Variante „besser“ ist, gibt es nicht, da hier auch inhaltliche Aspekte (Thema, Zielgruppe etc.) eine Rolle spielen.

Ich kann bei meiner Arbeit immer wieder Neues dazulernen

trifft zu	trifft eher zu	teils/teils	trifft eher nicht zu	trifft nicht zu
<input type="checkbox"/> (5)	<input type="checkbox"/> (4)	<input type="checkbox"/> (3)	<input type="checkbox"/> (2)	<input type="checkbox"/> (1)

trifft voll zu	trifft zu	trifft eher zu	trifft eher nicht zu	trifft nicht zu	trifft überhaupt nicht zu
<input type="checkbox"/> (6)	<input type="checkbox"/> (5)	<input type="checkbox"/> (4)	<input type="checkbox"/> (3)	<input type="checkbox"/> (2)	<input type="checkbox"/> (1)

Eine Reihe mehrerer ähnlich strukturierter Items (Einzelfragen mit gleicher Anzahl von Antwortabstufungen) zu einem bestimmten Thema wird „Likert-Batterie“ genannt. Bei größeren Item-Batterien (z. B. 40 Einzelfragen bei einer Mitarbeiterbefragung) wird die Sichtung der Verteilungen anhand von Tabellen oder Grafiken schnell ermüdend, weshalb die einzelnen Verteilungen üblicherweise nochmals komprimiert und durch einzelne Maßzahlen repräsentiert werden. Naheliegend ist hier der **Mittelwert** der Antwortkategorien, der aber für viele Betrachter zunächst recht abstrakt ist, da der Mittelwert aus den Zahlencodes gebildet wird und der Betrachter einen Mittelwert entsprechend auf den Wertebereich der Antwortskala übertragen muss. Als Alternative können die Antwortkategorien kollabiert und die zustimmenden bzw. inhaltlich „positiven“ Items als **Ja-Prozente** zusammengefasst werden. Jede derartige Zusammenfassung bedeutet einen Informationsverlust – die genaue Verteilung der Antworten im Detail wird durch einen Einzelwert ersetzt, der diese Verteilung dann mehr oder weniger genau repräsentiert. Dafür können sich die Betrachter der Ergebnisse mehr Einzelwerte als Einzelverteilungen merken, weshalb die Einzelwerte besser geeignet sind, um sich einen Überblick über eine Vielzahl von Einzelfragen zu verschaffen.

Ja-Prozente in 5er- und 6er-Skalen

Gabler und Borg¹ weisen für 5er-Skalen nach, dass die Ja-Prozente hoch mit den Item-Mittelwerten korrelieren und daher als Alternative verwendet werden können; der Informationsverlust ist damit nicht größer als bei der Reduktion der Antwortverteilungen auf die Mittelwerte. Die Ja-Prozente werden für die gängigen 5er-Skalen aus den beiden eindeutig zustimmenden Kategorien (z. B. „trifft zu“ und „trifft eher zu“) gebildet, die Mitte bleibt außen vor. Wie beim Mittelwert ist der Informationsverlust am geringsten für Items, deren Antwortverteilung sich einer Normalverteilung annähert und eingipflig ist. Verteilungen mit hohen Werten an den Rändern (ein Teil der Befragten stimmt eher zu, ein anderer Teil lehnt ab, die Mitte ist kaum vertreten) werden schlechter repräsentiert.

In der dieser Untersuchung zugrundeliegenden Befragung² wurden allerdings für die relevanten Items 6er-Skalen verwendet, so dass hier 2 Varianten für die Ermittlung der Ja-Prozente denkbar sind.

Stimme voll und ganz zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme überhaupt nicht zu
<input type="checkbox"/> (6)	<input type="checkbox"/> (5)	<input type="checkbox"/> (4)	<input type="checkbox"/> (3)	<input type="checkbox"/> (2)	<input type="checkbox"/> (1)

Für die „härtere“ Variante werden die Kategorien (6) und (5) als Ja-Prozente verwendet (2 Zustimmungskategorien gegen 4 Nicht-Zustimmungskategorien, kurz 2/4), für die „weichere“ Variante werden die Kategorien (6) bis (4) als Ja-Prozente verwendet, so dass die Skala einfach anhand der Skalenmitte dichotomisiert wird (Variante 3/3). Die Ja-Prozente sind in der folgenden Untersuchung mit „1 = Ja%“ und „2 = übrige“ kodiert. Die Ja-Prozente wurden für alle Likert-Items der Beispielbefragung jeweils nach beiden Varianten ermittelt und anschließend untersucht.

Korrelation der Ja-Prozente mit den Items

Zunächst werden die beiden Varianten der Ja-Prozente mit dem entsprechenden Item korreliert (Rangkorrelation nach Spearman). In der folgenden Tabelle sind die Ja-Prozente, die Item-Mittelwerte sowie die Korrelationskoeffizienten aufgelistet:

¹ Gabler, Siegfried; Borg, Ingwer: Zustimmungssanteile und Mittelwerte von Likert-skalierten Items; in: ZUMA-Nachrichten 50, Mannheim 2002, S.7-25

² Eine Kundenbefragung mit knapp 170 Teilnehmern

Item	N	mean Item	Ja% 2/4	Ja% 3/3	r _s 2/4*Item	r _s 3/3*Item	r _s 2/4 – r _s 3/3	Schiefe
v3	151	4,43	56,95	77,48	-0,89	-0,75	-0,14	-0,757
v4	150	4,29	47,33	76,00	-0,892	-0,763	-0,13	-0,518
v5	150	3,52	22,67	51,33	-0,746	-0,891	0,15	-0,07
v7	141	4,68	61,70	84,40	-0,874	-0,652	-0,22	-0,843
v8	141	3,88	29,79	64,54	-0,82	-0,858	0,04	-0,226
v9	138	3,06	17,39	39,13	-0,673	-0,867	0,19	0,096
v10	141	4,26	46,81	80,14	-0,907	-0,725	-0,18	-0,77
v11	141	3,59	23,40	56,74	-0,756	-0,884	0,13	-0,232
v12	142	3,82	35,92	59,86	-0,846	-0,864	0,02	-0,253
v14	141	4,40	52,48	77,30	-0,891	-0,748	-0,14	-0,609
v15	140	3,69	24,29	59,29	-0,771	-0,883	0,11	-0,272
v16	139	2,88	15,11	33,09	-0,637	-0,836	0,2	0,216
v17	139	3,84	31,65	65,47	-0,834	-0,853	0,02	-0,384
v18	140	3,36	17,86	43,57	-0,686	-0,888	0,2	0,047
v19	141	3,57	31,21	55,32	-0,817	-0,877	0,06	-0,155
v21	129	4,51	64,34	82,17	-0,867	-0,693	-0,17	-1,201
v22	129	3,50	23,26	51,16	-0,751	-0,888	0,14	-0,151
v23	128	2,72	11,72	28,13	-0,575	-0,804	0,23	0,409
v24	128	3,97	41,41	67,97	-0,875	-0,829	-0,05	-0,593
v25	128	3,30	14,84	46,88	-0,635	-0,892	0,26	-0,047
v26	129	3,95	48,06	65,89	-0,886	-0,841	-0,05	-0,559
v28	122	4,39	55,74	81,15	-0,893	-0,703	-0,19	-1,035
v29	123	3,56	30,08	52,03	-0,813	-0,886	0,07	-0,241
v30	122	2,77	16,39	32,79	-0,658	-0,834	0,18	0,352
v31	123	3,81	37,40	63,41	-0,861	-0,857	0	-0,492
v32	121	3,29	17,36	43,80	-0,673	-0,882	0,21	0,077
v33	123	3,90	44,72	65,04	-0,881	-0,845	-0,04	-0,514
v35	125	4,94	74,40	90,40	-0,801	-0,541	-0,26	-1,203
v36	127	3,69	29,13	52,76	-0,812	-0,892	0,08	-0,066
v37	127	4,15	42,52	71,65	-0,879	-0,802	-0,08	-0,389
v38	126	4,21	49,21	70,63	-0,887	-0,808	-0,08	-0,559
v40	126	4,47	59,52	80,95	-0,884	-0,708	-0,18	-0,997
v41	125	3,62	24,80	53,60	-0,77	-0,889	0,12	-0,016
v42	126	4,15	46,83	71,43	-0,888	-0,804	-0,08	-0,577
v43	125	3,54	24,00	54,40	-0,759	-0,886	0,13	-0,196
v44	124	3,71	27,42	59,68	-0,794	-0,873	0,08	-0,184
v45	124	3,46	25,00	50,00	-0,765	-0,883	0,12	-0,032
v46	125	3,93	34,40	68,80	-0,85	-0,829	-0,02	-0,406
v47	126	3,48	22,22	51,59	-0,742	-0,892	0,15	-0,29
v48	124	3,17	8,87	40,32	-0,513	-0,885	0,37	-0,218
v69	121	4,02	36,36	70,25	-0,865	-0,822	-0,04	-0,43
v70	122	4,07	36,89	73,77	-0,87	-0,793	-0,08	-0,483
v71	122	3,92	37,70	59,84	-0,859	-0,868	0,01	-0,121
v72	121	2,72	7,44	21,49	-0,472	-0,739	0,27	0,365
v73	120	2,60	5,83	20,83	-0,419	-0,727	0,31	0,483
v74	121	3,36	20,66	40,50	-0,727	-0,882	0,16	0,093
v75	120	3,17	12,50	40,83	-0,591	-0,878	0,29	0,12
v76	121	2,24	4,96	17,36	-0,391	-0,682	0,29	0,679

Skala von 6 = stimme voll zu bis 1 = stimme überhaupt nicht zu

Rechnerische Skalenmitte = 3,5

Insgesamt 48 Items mit 6-stufigen Antwortskalen

Mehrheitlich korrelieren die Ja-Prozente der Variante 3/3 höher mit dem jeweiligen Item als die Ja-Prozente der Variante 2/4. Im Einzelnen gibt es aber durchaus Abweichungen, die zur Veranschaulichung als einfache Differenz der beiden Korrelationskoeffizienten abgebildet sind. Da die Werte der Koeffizienten immer negativ sind (aufgrund der Kodierung der Ja-Prozente), bedeutet eine negative Differenz (grün) eine höhere Korrelation der 2/4-Variante mit dem Item, eine positive Differenz (orange) steht für eine höhere Korrelation der 3/3-Variante mit dem Item.

Korrelation 2/4 > Korrelation 3/3?		Häufigkeit	Prozent
Gültig	Korr. 2/4 > 3/3	19	39,6
	Korr. 3/3 > 2/4	29	60,4
Gesamt		48	100,0

Bei den meisten Items korrelieren die Ja-Prozente der Variante 3/3 höher mit dem jeweiligen Item als die Ja-Prozente der Variante 2/4, wobei die Höhe der Differenzen uneinheitlich ist. Die stärkste Differenz besteht bei Item v48 („6. Das PMS ist in unserem Unternehmen ein großer Erfolgsfaktor“) mit $r_s = -0,513$ (2/4) zu $r_s = -0,885$.

Korrelation der Ja-Prozente mit den Itemmittelwerten

Um die Frage zu klären, welche Variante der Ja-Prozente besser den Item-Mittelwert repräsentiert, wurden die Ja-Prozente für jedes einzelne Item aggregiert und anschließend mit den Itemmittelwerten korreliert. Wieder schneidet die Variante 3/3 geringfügig besser ab als die Variante 2/4, der Unterschied ist aber unerheblich.

Korrelationen

			Item: Mittelwert	Item: Ja-Prozente (2/4)	Item: Ja-Prozente (3/3)
Spearman-Rho	Item: Mittelwert	Korrelationskoeffizient	1,000	,973**	,991**
		Sig. (2-seitig)	.	,000	,000
		N	48	48	48
	Item: Ja-Prozente (2/4)	Korrelationskoeffizient	,973**	1,000	,958**
		Sig. (2-seitig)	,000	.	,000
		N	48	48	48
	Item: Ja-Prozente (3/3)	Korrelationskoeffizient	,991**	,958**	1,000
		Sig. (2-seitig)	,000	,000	.
		N	48	48	48

** . Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Dieser Effekt war zu erwarten, da sich die Summe der Einzelverteilungen der Items über viele Items hinweg einer Normalverteilung annähert, ggf. linear verschoben um den durchschnittlichen Schwierigkeitsgrad der Items. Beide Varianten der Dichotomisierung nähern sich als Aggregat somit dem Itemmittel an. Maßgeblich für die Bewertung der Aussagekraft der Ja-Prozente ist die Korrelation mit den ursprünglichen Items. Hier schneidet die Variante 3/3 besser ab, weshalb ihr bei der reduzierten Darstellung der Antwortverteilungen der Einzelitems der Vorzug zu geben ist.

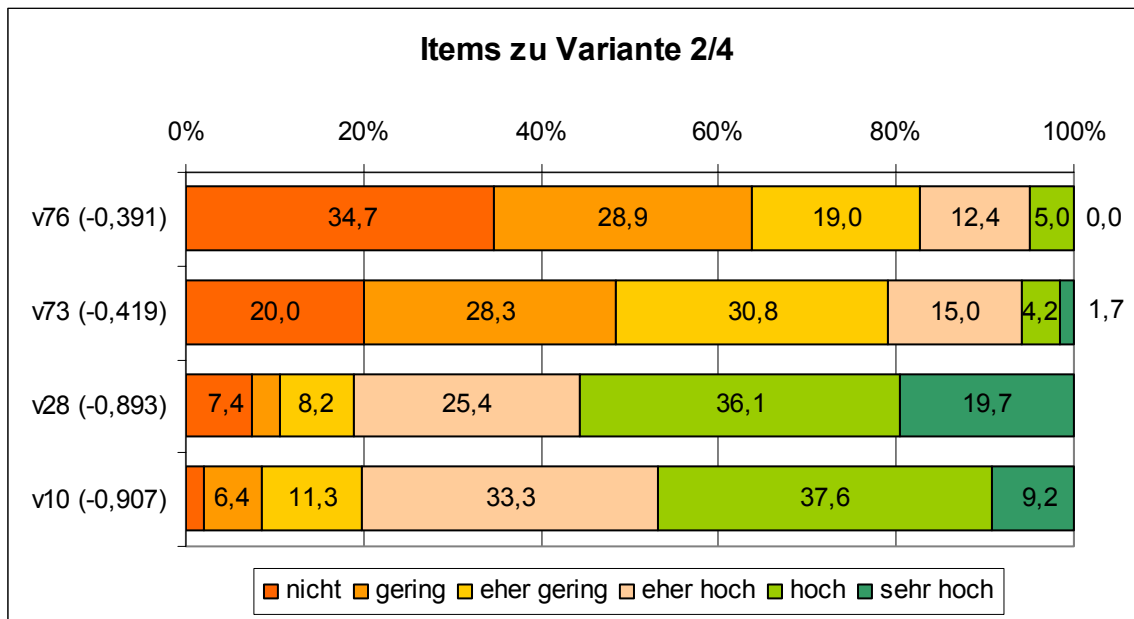
Einzelbetrachtung: Hohe und niedrige Korrelationen

Die Werte der Korrelationskoeffizienten der einzelnen Items mit deren Ja-Prozenten variieren für die Variante 2/4 zwischen $r_s = -0,391$ (v76) und $-0,907$ (v10) und für die Variante 3/3 zwischen $r_s = -0,541$ (v35) und $-0,892$ (v25, v36, v47). Sowohl die minimalen als auch die maximalen Werte treten in der Variante 2/4 auf.

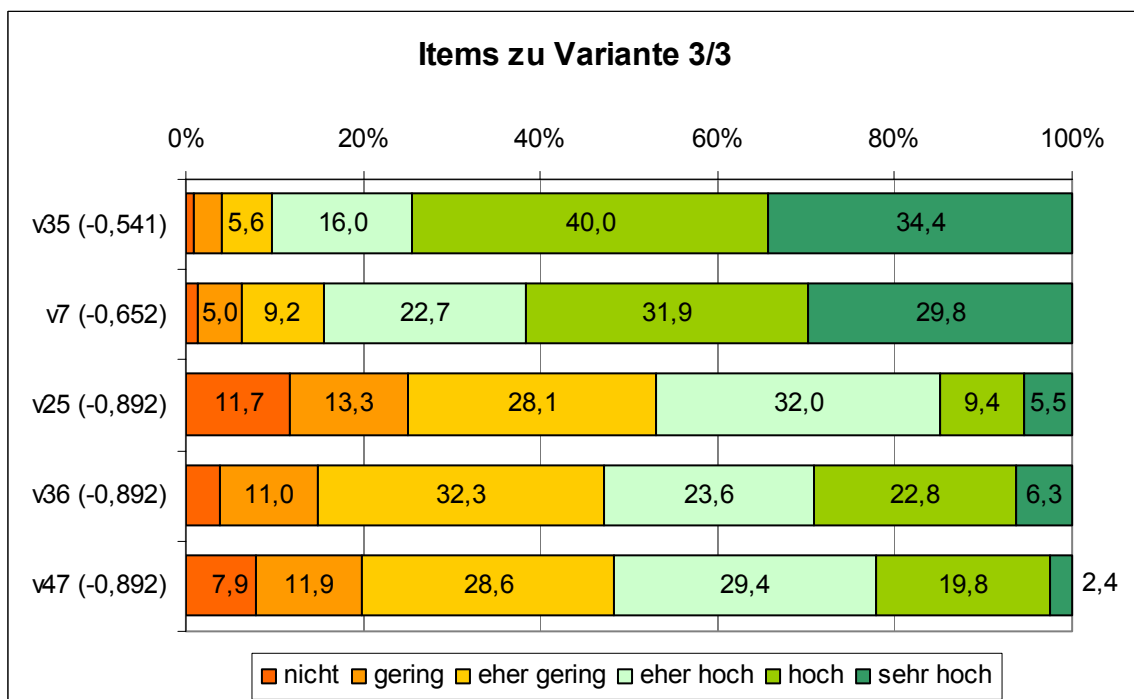
Übersicht	N	Minimum	Maximum	Mittelwert	Standardabweichung
Item: Mittelwert	48	2,24	4,94	3,6988	,57905
Item: Ja-Prozente (2/4)	48	4,96	74,40	32,3027	16,83520
Item: Ja-Prozente (3/3)	48	17,36	90,40	57,6074	17,84294
Korrelation Ja% 2/4 mit Item	48	-,91	-,39	-,7697	,13198
Korrelation Ja% 3/3 mit Item	48	-,89	-,54	-,8203	,07929
Differenz der Korrelationen	48	-,26	,37	,0506*	,15845
Gültige Werte (Listenweise)	48				

* der Mittelwert hat hier keine Aussagekraft, da sich positive und negative Werte neutralisieren. Hier müsste der Mittelwert der Beträge errechnet werden.

Nachfolgend eine grafische Darstellung der Häufigkeitsverteilungen in den Items mit den niedrigsten und höchsten Einzelkorrelationen:



Die Items mit den niedrigsten Korrelationen sind hier die Items mit den insgesamt niedrigsten Ja-Prozentsen, die Mittelwerte sind ebenfalls die geringsten. Die Items mit den höchsten Korrelationen weisen zwar demgegenüber hohe Ja-Prozente und hohe Mittelwerte auf, es gibt aber Items mit höheren Ja-Prozentsen bzw. Mittelwerten, die geringere Korrelationen aufweisen. Insgesamt werden Items mit hoher Zustimmung durch die Ja-Prozentsen der Variante 2/4 besser repräsentiert als Items mit geringer Zustimmung.



Bei den Ja-Prozenten der Variante 3/3 ist der Effekt genau umgekehrt. Hier weisen die Items mit den niedrigsten Korrelationen deutlich höhere Zustimmungswerte auf als die Items mit den höchsten Korrelationen. Bei den Items mit den höchsten Korrelationen befinden sich die Ja-Prozente nahe bei 50%, wobei es aber andere Items mit niedrigeren Korrelationen gibt, die sich näher bei 50% befinden. Die Items mit den niedrigsten Korrelationen weisen tendenziell stark schiefe Verteilungen auf: Die in der Grafik abgebildeten Items v35 und v7 in Richtung Zustimmung; aber auch das Item v76 mit der geringsten Zustimmung von 17,36% hat hier mit -0,682 die drittniedrigste Korrelation (auch wenn sie gegenüber dem Wert von -0,391 bei der 2/4-Variante deutlich höher ist). Ausschlaggebend ist für die Variante 3/3 also vermutlich nicht der absolute Schwierigkeitsgrad der Items, sondern ein mittlerer Schwierigkeitsgrad bzw. die Annäherung an eine Normalverteilung der Antwortkategorien. Diese Tendenz ist hier nicht so deutlich, da die meisten Items im Beispieldatensatz eine hohe Zustimmung aufweisen, also rechtssteil verteilt sind. Die wenigen linkssteil verteilten Items haben nur einen geringen Einfluss auf die Gesamteffekte.

Aufgrund dieser Befunde wird noch eine Korrelation zwischen den Werten für die Schiefe und den Korrelationskoeffizienten Ja%*Item durchgeführt:

Korrelationen

		Korrelation Ja% 2/4 mit Item	Korrelation Ja% 3/3 mit Item	Item: Standardab- weichung	Item: Schiefe
Korrelation Ja% 2/4 mit Item	Korrelation nach Pearson	1	-,067	-,225	,798**
	Signifikanz (2-seitig)		,652	,124	,000
	N	48	48	48	48
Korrelation Ja% 3/3 mit Item	Korrelation nach Pearson	-,067	1	-,221	-,464**
	Signifikanz (2-seitig)	,652	,130	,130	,001
	N	48	48	48	48
Item: Standardabweichung	Korrelation nach Pearson	-,225	-,221	1	,009
	Signifikanz (2-seitig)	,124	,130	,952	,952
	N	48	48	48	48
Item: Schiefe	Korrelation nach Pearson	,798**	-,464**	,009	1
	Signifikanz (2-seitig)	,000	,001	,952	,952
	N	48	48	48	48

** . Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Es wird deutlich, dass die Korrelationskoeffizienten für die Ja-Prozente der Variante 2/4 positiv mit den Werten für die Schiefe korrelieren ($r = 0,798$). Je stärker die Schiefe in Richtung einer hohen Zustimmung (Hohe Zustimmung = Schiefe stark negativ), desto stärker ist die Korrelation der Ja-Prozente mit dem jeweiligen Item (die Korrelationskoeffizienten Item*Ja% sind ja durchweg negativ).

Für die Ja-Prozente der Variante 3/3 verhält es sich auf den ersten Blick umgekehrt, allerdings mit einem geringeren Wert ($r = -0,464$). Hier muss berücksichtigt werden, dass die meisten Items – wie bereits erwähnt – schief in Richtung einer hohen Zustimmung sind. Es wäre zu erwarten, dass die Ja-Prozente der Variante 3/3 am stärksten mit dem Item korrelieren, wenn die Schiefe 0 ist. Daher wird nochmals eine Korrelation der Ja-Prozente mit den jeweiligen Beträgen der Schiefe durchgeführt. Wie erwartet ist der Wert für die Variante 3/3 deutlich höher, wenn statt den Werten für die Schiefe die Beträge verwendet werden:

Korrelationen

		Korrelation Ja% 2/4 mit Item	Korrelation Ja% 3/3 mit Item	Item: Betrag der Schiefe
Korrelation Ja% 2/4 mit Item	Korrelation nach Pearson	1	-,067	-,330*
	Signifikanz (2-seitig)		,652	,022
	N	48	48	48
Korrelation Ja% 3/3 mit Item	Korrelation nach Pearson	-,067	1	,888**
	Signifikanz (2-seitig)	,652		,000
	N	48	48	48
Item: Betrag der Schiefe	Korrelation nach Pearson	-,330*	,888**	1
	Signifikanz (2-seitig)	,022	,000	
	N	48	48	48

*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

** . Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Fazit

Im Vergleich zu der Repräsentation der Häufigkeitsverteilungen 6-stufiger Likertskalen durch das arithmetische Mittel besteht kein nennenswerter Unterschied zwischen den Ja-Prozenten der Varianten 2/4 und 3/3. Werden die Korrelationskoeffizienten (Rangkorrelation nach Spearman) zwischen Ja-Prozenten und Items herangezogen, schneidet die Variante 3/3 im hier vorliegenden Datensatz etwas besser ab. Die Höhe der Korrelation zwischen Ja-Prozenten und Item hängt deutlich mit der Schiefe der Einzelverteilungen zusammen, und zwar für beide Varianten der Ja-Prozent-Ermittlung unterschiedlich. Die Ja-Prozente der Variante 2/4 charakterisieren Items mit hoher Zustimmung (= niedrigem Schwierigkeitsgrad) besser, die Ja-Prozente der Variante 3/3 charakterisieren Items mit einem mittleren Schwierigkeitsgrad besser.

Auf Basis des vorliegenden Beispieldatensatzes – der mit einem leichten Übergewicht von Items mit hoher Zustimmung z. B. auch für Mitarbeiterbefragungen charakteristisch ist – wird empfohlen, im Zweifelsfall die Variante 3/3 zu verwenden, die hier neben einer geringfügig besseren Gesamtpassung vor allem eine geringere Schwankung der Passungen im Einzelfall aufweist (Gemessen anhand der Rangkorrelation zwischen Item und Ja-Prozenten).

Um Ergebnisse von Befragungen, bei denen 6er-Skalen eingesetzt wurden, mit den Ergebnissen der gängigeren 5er-Skalen zu vergleichen, müssen aber anstatt der Ja-Prozente die Mittelwerte der Items herangezogen werden. Zumindest rechnerisch gewährleisten diese allein einen direkten Vergleichsmaßstab.