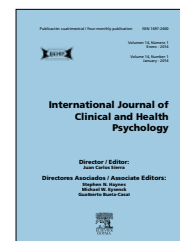


International Journal of Clinical and Health Psychology

www.elsevier.es/ijchp



BRIEF REPORT

Some thoughts on Likert-type scales

James Hartley

Keele University, United Kingdom

Received August 8, 2013; accepted October 9, 2013

KEYWORDS

Attitude
measurement;
Likert scales;
Reverse-scoring;
Scale formats;
Descriptive study

Abstract Many articles published in the International Journal of Clinical and Health Psychology describe the results obtained in studies that have used Likert-type scales. Such scales have advantages and disadvantages and in this article I comment on some of these. In particular I comment on the effects of using scales that start with the positive end and finish with the negative one, and on those that present high ratings on the left - as opposed to the right. I also consider that scores on negatively-worded items differ from those obtained on positively-worded ones and that it is not, therefore, appropriate to simply reverse the scores obtained on such items. Finally, I note that some items on these scales present more than one issue and it is hard to know whether when responding to these items, respondents are responding to all of the item or just one part.

© 2013 Asociación Española de Psicología Conductual. Published by Elsevier España, S.L.
All rights reserved.

PALABRAS CLAVE

Medida de actitudes;
Escala tipo Likert;
Puntuaciones
invertidas;
Formatos de escala;
Estudio descriptivo

Resumen Muchos de los artículos publicados en *International Journal of Clinical and Health Psychology* describen los resultados obtenidos en estudios que han utilizado escalas tipo Likert. Estas escalas tienen una serie de ventajas y de desventajas. En este artículo se comentan algunas de ellas. En particular, se discute sobre los efectos del uso de escalas que comienzan en un extremo positivo y terminan en el negativo, así como las que presentan altos índices a la izquierda a diferencia de la derecha. También se señala que las puntuaciones en los ítems redactados de forma negativa difieren de las obtenidas en ítems positivos y que, por tanto, no es apropiada una simple inversión de las puntuaciones obtenidas en dichos ítems. Por último, cabe señalar que algunos ítems de estas escalas plantean más de una cuestión, siendo difícil saber si a la hora de contestarlos, los encuestados responden a todo el ítem o sólo a una parte.

© 2013 Asociación Española de Psicología Conductual. Publicado por Elsevier España, S.L.
Todos los derechos reservados.

*Corresponding author at: School of Psychology, Keele University, Staffordshire, ST5 5BG, United Kingdom.
E-mail address: j.hartley@keele.ac.uk (J. Hartley).

Much research in clinical and health psychology involves measuring peoples' opinions and attitudes. Different people employ different ways of doing this, and there are controversies over which are the most appropriate measures to use. One the one hand there so-called 'objective measures' where individuals complete measurement scales, such as Likert-type scales and the Semantic Differential, and on the other there are 'subjective measures', where individuals are interviewed, or complete tests such as the Rorschach and the Thematic Apperception Test. Likert-type scales are ones that more or less follow the original format devised by Rensis Likert in the 1930s. Two typical examples (from Likert, 1932) are as follows:

Our country should never declare war under any circumstances.

Strongly approve (5) Approve (4) Undecided (3) Disapprove (2) Strongly disapprove (1)

We should be willing to fight for our country whether it is right or wrong.

Strongly approve (1) Approve (2) Undecided (3) Disapprove (4) Strongly disapprove (5)

Although such scales are popular, I was somewhat surprised to note the popularity of Likert-type scales in a recent issue of International Journal of Clinical and Health Psychology. Volume 13, number 3, September 2013, contains ten articles, and nine of these describe the use and/or development of Likert-type scales. I thought it might be useful, therefore, to comment on these different articles and then to make some suggestions about the difficulties of using such Likert-type scales, given their popularity.

Table 1 shows a summary of the key features of the scales used in the nine studies. It can be seen that:

1. Generally speaking, the sample sizes are good.
2. The number of different scales used in any one study varies from one to seven.
3. Almost all of the scales are Likert-type scales.
4. Most of them use 4 or 5 scale points.
5. Most of them start with 0 (or 1) - the negative end of the scale - and progress to 5 (or 7) - the positive end.
6. Few state in these papers whether or not any of the items are negatively worded and are 'reverse scored'.

Here are some brief remarks about these features:

1. Likert-type scales are used regularly, and much is known about their properties (for useful reviews, see Hartley & Betts, 2010; Krosnick & Fabrigar, 1997; Oppenheim, 2000).
2. It is good to see large sample sizes: this increases the validity of the findings, and sub-groups can be meaningfully compared. It also allows for parametric analyses of the results.
3. Some investigators create their own scales. This involves them in a lot of preparatory work (assessing the reliability and validity of the individual items). Others use or adapt scales developed by previous researchers – but these

may not always be entirely appropriate. In this case, most of the authors of the nine studies adapted previous scales for use in their own studies.

4. Using several different scales in one study provides a bigger picture of the issues in question. A problem arises, however, especially with children, when respondents are asked to complete several different scales in a single session, and when the layouts of these scales vary (e.g., see Betts & Hartley, 2012). Respondents do not always notice.

'After I read your debrief...I went back and corrected my responses. Originally I had put 9 for every answer but, when I went back, I realised that I had assumed 'clear' was on the left (at 10) as opposed to 'unclear'. I therefore corrected my responses.'

This example also shows that some respondents simply tick the same box for every item, perhaps without considering carefully enough the meaning of each one.

5. Scales using 5 or 7 scale points are common. Some use only two points: some use 100 (percentages). The research does not suggest any real reason for favouring any one number of scale points over any other. There is some research that examines the use of say, four points as opposed to five, thus eliminating neutral mid-points and forcing the respondents to make a choice (Garland, 1991). Table 1 shows the wide variation in the number of scale points used in these studies.
6. Scale items are usually rated from low to high, with the negative pole on the left and the positive one on the right. Research has shown, however, that slightly higher scores are obtained on the topic in question with English speaking respondents, when this is reversed – that is when the scales start with the high positive values on the left (Hartley & Betts, 2013).
7. It is common to have some items in a scale that are 'negatively worded' and respondents have to reverse their thinking when saying that they agree or disagree with them. The responses to such items are usually 'reversed scored' so that they can be included in the total scale scores. There are three difficulties with this procedure. First, it is not easy (in English) to write exactly equivalent items in a positive and negative form (Bozin, Berman, & Boyzman, 2010); second, respondents have difficulty in reverse thinking; and third, different ratings are obtained on positive and negative versions of the same item (Hartley & Betts, 2013; Yorke, 2009). It is best either (a) to remove negatively worded items from a scale, or (b) to present the results for such items separately.
8. One additional item (not considered above because I have not seen the actual scales used in the nine studies) is that some of the items in some scales ask about more than one thing, and this causes difficulties for the respondents. Suppose, for example, you were asked to rate how far you agree with the statement, 'All research impacts society one way or the other: the real challenge is to limit possible negative impacts'. You might agree 'yes' to the first part, but 'not necessarily so' to the second. Items with more than one idea should be clarified so that each one has only one issue. It is best to check

Table 1 Key features of the scales used in nine out of ten studies published in Volume 13 (No. 3) of the International Journal of Clinical and Health Psychology, 2013.

Study Number	Authors	Participants	Number of Likert scales	Description of scales	Direction of scores	Negatively worded items?
1	Cerezo, Méndez, & Ato (2013)	1,239 adolescents (50.6%female)	1	90 items 3 or 5 points	Varies	Not stated
2	Ramiro, Teva, Bermúdez, & Buela-Casal (2013)	1,005 adolescents (52.9%female)	3	12 items 5 points 10 items 4 points 8 items 4 points	Low-high	Not stated
3	Moyano & Sierra (2013)	1,500 adults 18-72 years (57.2%female)	2	28 items 7 points 36 items 5 points	Low-high	14/ 36 in scale 2
4	Ariza, Quevedo-Blasco, Ramiro, & Bermúdez (2013)	1,361 University teachers (50.1%female)	1	65 questions in 7 blocks	-	Not stated
5	Fonseco-Pedrero, Paino, Lemos-Giráldez, & Muñiz (2013)	1,443 adolescents (52%female)	3	97 items 5 points 25 items 5 points 12 items 5 points	Low-high	Not stated
6	De la Fuente et al. (2013)	2,403 police (84.2%male)	4	26 items 5 points 22 items 7 points 60 items 5 points 50 items 5 points	Low-high Not stated Not stated Not stated	Yes, in one scale
7	Fernández-Jiménez et al. (2013)	221 multiple sclerosis patients (mean age = 40.61 years; 63.03%females)	1	20 items 5 points	Low, borderline and high	5
8	Ginting, Näring, van der Veld, Srisayekti, & Becker (2013)	720 healthy individuals 215 coronary heart disease patients 102 depressed patients	5	21 items 4 points 14 items 5 points 21 items 4 points 8 items 5 points 12 items 7 points	Low-severe	Not stated
9	Zenger et al. (2013)	1,500 adults (51.70%female)	7	10 items 5 points 14 items 4 points 20 items 5 points 30 items 8 items 12 items 4 points 10 items 4 points	Low-severe — — — High-low Low-severe High-low	Yes, on some scales

first for this before using questionnaires to see if this problem has been eliminated.

These are all small points, but I think that they are important.

Acknowledgements

I am grateful for comments offered on an earlier version of this paper by the editor and the authors of the nine studies.

References

- *Indicates that this is one of the nine studies referred to in this article.
- *Ariza, T., Quevedo-Blasco, R., Ramiro, M. T., & Bermúdez, M. P. (2013). Satisfaction of Health Science teachers with the convergent process of the European Higher Education area. *International Journal of Clinical and Health Psychology*, 13, 197-206.
- Betts, L., & Hartley, J. (2012). The effects of changes in the order of verbal labels and numerical values on children's scores on attitudes and rating scales. *British Educational Research Journal*, 38, 319-331.

- *Cerezo, F., Méndez, I., & Ato, M. (2013). Moderating role of family and friends' factors between dissocial behaviour and consumption in adolescents. *International Journal of Clinical and Health Psychology*, 13, 171-180.
- *De la Fuente, E. I., Lozano, L. M., García-Cueto, E., San Luis, C., Vargas, C., Cañadas, G. R., Cañadas-De la Fuente, G. A., & Hambleton, R. K. (2013). Development and validation of the Granada Burnout Questionnaire in Spanish police. *International Journal of Clinical and Health Psychology*, 13, 216-225.
- *Fernández-Jiménez, E., Pérez-San-Gregorio, M. A., Taylor, G. J., Bagby, R. M., Ayeart, L. E., & Izquierdo, G. (2013). Psychometric properties of a revised Spanish 20-item Toronto Alexithymia Scale adaptation in multiple sclerosis patients. *International Journal of Clinical and Health Psychology*, 13, 226-234.
- *Fonseco-Pedrero, E., Paino, M., Lemos-Giráldez, S., & Muñiz, J. (2013). Maladaptive personality traits in adolescence: Psychometric properties of the Personality Diagnostic Questionnaire-4+. *International Journal of Clinical and Health Psychology*, 13, 207-215.
- Garland, R. (1991). The mid-point on a rating scale: Is it desirable? *Marketing Bulletin*, 2, 66-70.
- *Ginting, H., Näring, G., van der Veld, W. M., Srisayekti, W., & Becker, E. S. (2013). Validating the Beck Depression Inventory-II in Indonesia's general population and coronary heart disease patients. *International Journal of Clinical and Health Psychology*, 13, 235-242.
- Hartley, J., & Betts, L. (2010). Four layouts and a finding: The effects of changes in the order of the verbal labels and numerical values on Likert-type scales. *International Journal of Social Research Methodology*, 13, 17-27.
- Hartley, J., & Betts, L. (2013). Let's be positive: The effects of the position of positive and negative values and labels on responses to Likert-type scales. *Chinese Journal of Psychology*, 55, 291-299.
- Krosnick, J. A., & Fabrigar, L. R. (1997). Designing rating scales for effective measurement in surveys. In L. E. Lyberg, P. Biemer, M. Collins, E. D. De Leeuw, C. Dippo, N. Schwarz, & D. Trewin (Eds.), *Survey Measurement and Process Quality* (pp. 141-164). New York: Wiley.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22, 1-54.
- *Moyano, N., & Sierra, J. C. (2013). Relationships between personality traits and positive/negative sexual cognitions. *International Journal of Clinical and Health Psychology*, 13, 189-196.
- Oppenheim, A. N. (2000). *Questionnaire design, interviewing and attitude measurement* (2nd edition). London: Bloomsbury Academic.
- *Ramiro, M. T., Teva, I., Bermúdez, M. P., & Buela-Casal, G. (2013). Social support, self-esteem and depression: Relationship with risk for sexually transmitted infections/ HIV transmission. *International Journal of Clinical and Health Psychology*, 13, 181-188.
- Rozin, P., Berman, L., & Ryzman, E. (2010). Biases in the use of positive and negative words across twenty natural languages. *Cognition and Emotion*, 24, 536-548.
- Yorke, M. (2009). 'Student experience' surveys: Some methodological considerations and an empirical investigation. *Assessment and Evaluation in Higher Education*, 34, 721-739.
- *Zenger, M., Finck, C., Zanon, C., Jimenez, W., Snger, S., & Hinz, A. (2013). Examination of the Latin American version of the Life-Orientation Test-Revised. *International Journal of Clinical and Health Psychology*, 13, 243-252.