Evaluating the Content Validity of a Psychological Capital Scale (PCQ-12): A Prospective Study of Higher Education Faculty in Pakistan

Afia Saleem[*], Atta Ullah Shah[†], Marium Durrani[‡]

Abstract

The aim of our study was to obtain the evidence of content validity in the test construction/adaptation process. Our paper assesses the content validity of the Psychological capital Scale (Luthans, Avolio, Avey & Norman, 2007). Based on the results obtained for PCQ-12, the study concludes the significance of content validity during the scale development/adaptation. This study mainly focused on the expert judgment of 6 judges for the items constructed in order to determine the content validity. The assessments of the judges in order to analyze the PCQ-12 that collectively comprised of 12 items, resulted in elimination of 2 items with insufficient content validity (CVI < 0.70 and Kappa < 0.40). Our study confirmed that PCQ-12 demonstrated a good content validity.

Keywords: content validity, Kappa Index, psychological capital questionnaire, construct validity

Introduction

The content validity ensures the omission of any logical errors when the conclusions are drawn from the data (Gregory, 1992, p.117). The evaluation of a scale's content validity is a foremost requisite to ameliorate the construct validity of an instrument (Haynes, Richard, & Kubany, 1995). Content validity determines the successful definition of the elements and dimensions of the concept under study (Sekran, 2011). Additionally, it validates all the items relevant to their constructs (Miller et al, 2013). Moreover, Construct validity is related with the logic of items of an instrument/measure and is subject to how accurately the research concept, idea or behavior (called as construct) has been operational zed or transformed in to reality (Trochim, 2006).

The content validity is a matter of concern for the researchers who crave for high quality measurements because it increases the confidence of the researchers in the validity of the results obtained from the instrument being developed (Johnson, 2012). Adhering to this, our research aims to measure the content validity of teachers' psychological capital with a

[*] Lecturer, Institute of Management Sciences, Peshawar, Pakistan
[†] Professor, Institute of Management Sciences, Peshawar, Pakistan
[‡] Lecturer, Institute of Management Sciences, Peshawar, Pakistan

shorter version of PCQ-12 (Luthans, Avolio, Avey & Norman, 2007). The theory underpinning the PsyCap constructs was based on positive organizational behavior (POB) that has sufficient reliability and is commonly used by the scholars (Lu, Wang, Xu, Teng, Li & Guo, 2023). According to (Luthans, Youssef et al., 2007), people differ in their psychological abilities and their access to resources in different cultures. For example, in a western culture that is more individualistic, a person is viewed more independent and autonomous than a collectivist culture such as Pakistan that takes in to account group interests and values. Therefore, the concept of PsyCap that emerges from the United States that is an individualistic culture, may be completely different in Pakistan. (Ali, Khurshid, Shahzad, Hussain & Abu Bakar, 2018). Hence, in order to tap this differentiation, various researchers have examined the effects of PsyCap on individual's job attitude, workplace behaviors and performance in groups belonging to different cultural and professional grounds. For example, Rego et al (2010) investigate PsyCap among Portuguese civil servants. Similarly, Zhang et al. (2010) surveyed college student's PsyCap Moreover, Cui et al. (2021) developed a PsyCap inventory for cancer patients. Likewise, Lou et al. (2022), constructed the PsyCap scale for male nurses in Taiwan. Summing up, the core dimensions of PsyCap have been widely verified by many studies. However, little attention has been paid on assessment of the interpretations of the individuals based on the cultural differences regarding the structure of psycap. Therefore, our study fills this gap by taking in to account the psychological capital beliefs of the teachers of a collectivist culture (e.g. Pakistan) subject to the higher education institutions of Pakistan.

Our research took substantial steps to confirm the validity of the research instrument under examination. The subsequent paragraphs will ad hoc the key stages including 1) operational definitions of the construct to be assessed 2) Inventory for the assessment of the construct and 3) Expert judgments of the items of the inventory, underpinning the construct under investigation. The stages for the content validation of the PCQ-12 of are in accordance with the view point of the researchers who ratify that the process of content validation is all based on judgment and is mainly composed of two stages. The initial stages are ancillary to scale development in light of previous researches and the stages henceforth are pervasive to relevance estimation of its contents by the experts through their assessments (Mastaglia et al., 2003).

Moreover, the expert judgment determines the content validity of the instrument in a way that it ensures that the items of the inventory are relevant to the domain of the construct to which they belong (Mastaglia, Toye, and Kristjanson, 2003). As a reference our research adhere to the

criteria proposed by Angleitner, John, and Löhr (1986) for expert judgment. Their criteria include the assessment of the instrument items around three characteristics, i.e. *comprehension* that evaluates if the item is properly understood, *clarity* which is the degree to which the item concise, correct and direct and *ambiguity* which is an evaluation of the chances that the item can be interpreted in many other ways. Further details about the expert judgments of regarding the inventory under examination are detailed in the sections below.

Psychological Capital Questionnaire

Luthans, Youssef and Avolio (2007) define Psycap as a positive state of mind of a person in form of their high levels of confidence to succeed while performing challenging tasks. Moreover, their positive attribution and determination towards goals and when needed redirecting themselves towards the paths towards goals. Furthermore, bouncing back during hardships to achieve success. Thus, the four constructs included in this second order construct of Psycap, is self-efficacy, optimism, hope and resilience.

The PCQ-12 is made up of positive constructs including 4 items of hope (adapted from Snyder et al., 1991), 3 items for self-efficacy (originally adapted from Parker, 1998), 3 items of resilience (adapted originally from Wagnild & Young, 1993) and 2 items of optimism (originally adapted from Scheier & Carver, 1985). Previously, the reliability and validity of both shorter as well as long versions of the scales have already been established (Avey, Avolio, & Luthans, 2011), however, our research attempts to investigate it further keeping in view the cultural context of Pakistan.

Participants

In an endeavor to seek the expert judgment the key stone is to decide the number as well as the attributes of the judges/raters. Hence for the purpose of obtaining the adequate inter judge/rater agreement the recommendation given by Crocker, Llabre, and Miller (1988) was considered according to which at least three judges should be selected for every item (Lynn, 1986). Besides, another matter of concern that is equally important and inescapable is the characteristics of judges/raters. Studies have delineated the significance of engaging the experts in test construction or adaptation. The involvement of those judges has also been emphasized who are though not the experts in the measure yet who carry ample knowledge of the discipline it forms part of or are if the specialized people in the domain of the assessed construct (Davis, 1992).

Hence in accordance with the criterions aforementioned, we selected 6 judges to assess each dimensions of the constructs underlying the items of the inventory. The basic credentials of the judges included their expert knowledge and teaching experience moreover their published recent scientific works complemented their eligibility to assess and refine the inventories employed for the present study. A booklet was prepared including the randomly arranged items and aspects for the assessment by the judges. The judges were asked to rate 12 items on relevance by using 4- point Likert scale (Davis, 1992). An open-ended question was asked from the judges about their assessment regarding clarity and completeness of the questions pertinent to each construct with the intentions to identify the need for eliminating or revising them. Additionally, each item included the possibility of being assigned alternative wordings by the judges. Furthermore, judges were requested to provide their overall opinion about items and any other specific observations that can guide the researcher to ensure the excellence of the tools for collection of data.

Methods

The debate concerning the method of calculation to find out the degree of consistency or congruence among the ratings of the raters is debatable (Landsheer and Boeije, 2010). However, Stemler (2004) proposed consensus, consistency and measurement estimates as the most valid methods. The Content Validity Index (CVI; Polit and Beck, 2006) according to the categories defined by Stemler (2004) falls in to consensus estimates.

The key feature of consensus estimates is that they identify the levels to which the judges present a common interpretation of the construct (Stemler, 2004). A low level of consensus among judges means that they differ in their interpretations, thus leading to decisions like refining or discarding the items.

CVI is the widely used method for multi items scales, which has been used traditionally to estimate comprehension, representativeness, clarity and ambiguity (Polit and Beck, 2006). Although, it can be determined in many other ways however, we calculated it with a method according to the recommendations of Rubio, Berg- Weger, Tebb, Lee, and Rauch's (2003). In lines with their proposition the CVI for the respective item, commonly referred as Item-level CVI (I-CVI), should be computed by dividing the number of judgments issued as 3 or 4 on the analogous Likert scale by the total number of judges. CVI can also be calculated for over all scale and is referred as CVI for the scale (S-CVI), which is discussed in the following paragraphs. The content validity of the research

instruments of this study was determined at both the scale as well as the item level. Hence in order to determine the relevance of items of this study, the I-CVI for each item corresponding to their respective construct was computed by dividing the number of judgments that favored the relevance of the item to the intended facet by the total number of raters. Lynn's (1986) widely- cited guidelines for acceptable I-CVI were followed conferring that the I-CVI should be 1.00 with 5 or fewer judges, which implies that all experts should agree that the item is content valid. However modest amount of disagreement is acceptable when there are 6 or more judges and as a general rule the acceptable I-CVI in that case is ≥ 0.70 (Tilden, Nelson, and May, 1990).

Apart from the I-CVI, the S-CVI for the overall scales was computed in accordance with the method described by Polit & Beck (2006). According to them S-CVI/average can be computed in different ways and application of each methods yield same results. The S-CVI/average was calculated for the scales used in this study as the Mean I-CVI value that is obtained by summing the I-CVIs and dividing it by total number of items. The guideline offered by Polit & Beck (2006) was followed with values ≥ 0.90 as standard criterion for acceptability of S-CVI.

Synchronously, one of the reasons of selecting CVI as a criterion for inter-rater agreement in present study was its ease of computation, understandability, and provision of item as well as scale level diagnosis. Moreover, the focus of our research was consensus rather than the consistency estimates. Furthermore, the reason for preference of consensus estimates over consistency estimates is described in the paragraph ahead.

Unlike consensus estimates, consistency estimates focus on the extent to which judges are consistent or reliable in their ratings for item relevance. However, one of the limitations of the consistency approach like alpha is that it provides limited information for evaluation of individual item or individual judges since alpha is computed across all items and judges. Additionally, a high alpha value can even be obtained in situations where the consensus about content validity is low (Waltz, et al., 2005). Hence, for consistency indices the central point is the internal consistency, i.e. inter-rater reliability instead of agreement across judges. Comparably, CVI (a consensus estimate) holds up well against most of the aforementioned limitations of consistency estimates for inter-rater agreement.

Despite its manifold advantages, however, CVI is equally subject to criticism. One of the most recurrent criticisms is its inability for adjustment of chance agreement that results in to inflated values (Wynd *et*

al., 2003). The Kappa statistic turns out as a support against such arguments. Kappa is a consensus index that adjusts for chance agreement and has been applied along with CVI in content validation of scales (e.g. Rico, Dios and Ruch, 2012; Wynd *et al*., 2003). Subsequently in the present research I-CVI for each item of the scales were computed along with the Kappa index that is applied to categorical judgments issued by multiple experts (Fleiss, 1971).

Results

The assessments of the 6 judges resulted in elimination of 2 items from the PCQ-12 with insufficient content validity (CVI < 0.70 and Kappa < 0.40). The process of the calculations of I-CVI and S-CVI of PCQ-12 is mentioned in the coming paragraphs. I-CVIs of the 12 items of the Psychological Capital ranged from 0.5 to 0.833, however after excluding the two items below the standard CVI (i.e. <0.70), the I-CV of the remaining items improved and ranged between 0.66 to 0.833. Table 1, samples I-CVIs with Kappa coefficients of the PCQ-12 items that were retained and some of those items which were considered to be eliminated due to poor ratings by the judges on relevance.

Table 1
Results of Content Validity of PCQ-12 on Relevance

Facet	Examples of items as per the original PCQ-12	I-CVI	Kappa	Action taken
Hope	1	1	1	Retained
	2	1	1	Retained
	3	1	1	Retained
	4	0.833	0.66	Retained
Self-Efficacy	1	0.833	0.66	Retained
	2	0.333	0.19	Eliminated
	3	0.833	0.66	Retained
Resilience	1	0.5	0.3	Eliminated
	2	1	1	Retained
	3	0.7	1	Retained
Optimism	1	1	1	Retained
	2	1	1	Retained

Note: I-CVI= ≥ 0.70; Kappa= ≥ 0.40

The exclusion of the two items less than the expected I-CVI further improved the S-CVI value, which was raised from 0.833 to 0.91

(Marry, Billeen, Kruszewski, Sims, Fettes and Wilson, 2013). Table 2 provides information regarding the S-CVI values of the PCQ-12. Moreover, the comments from the experts with regard to clarity, ambiguity and wordings of the items lead to minor revision of many of the scales items that were either re-titled or re-worded in order to clarify the respondents' ambiguities.

Table 2

S-CVI/Ave of PCQ-12

Scale	Source	S-CVI/Ave
Psychological Capital 12-Item Questionnaire	(Luthans, Avolio, Norman and Avey, 2007)	0.91

Note: S-CVI/Ave= scale-level content validity index, ≥ 0.90

Conclusion and Future Research Directions

Our study established the content validity of Psychological capital scale. Almost all the items of PCQ-12 obtained the adequate value of I-CVI ranged between 0.66 to 0.91hence, it was proved that the items were well operationalized and well conceptualized. Based on the findings of our research, PCQ-12 is proposed as a good measure teachers' psychological capital. The findings of our research also go consistent with some other research studies where PCQ-12 successfully taped the teachers' perception regarding their psychological capital beliefs (e.g. Zhou & Zheng, 2022; Zewude & Hercz, 2021; Rehman & Ali, 2017).

However, further investigation can be held regarding the validity of this scale such as convergent, discriminant and construct validity. In addition to it, in future, more sophisticated statistical analysis such as structural equation modelling can also be taken in to account to establish measurement models for gaining deeper insights to the internal consistency of the items of the scale.

References

Ali, R., Khurshid, K., Shahzad, A., Hussain, I & Abubakar, Z. (2018). Nature of Conceptions of   Learning in a Collectivistic Society: A Qualitative Case Study of Pakistan. *Journal of Mathematics, Science and Technology Education,* 14(4), 1175-1187.

Angleitner, A., John, O.P., & Löhr, F.J. (1986). It's what you ask and how you ask it: An   item metric analysis of personality questionnaires. In A. Angleitner and J.   S. Wiggins   (Eds.),

Personality assessment via questionnaires. Current issues in theory and measurement (pp. 61-108). Berlin, Germany: Springer.

Avey, J. B., Avolio, B. J., & Luthans, F. (2011). Experimentally analyzing the impact of leader positivity on follower positivity and performance. *The Leadership Quarterly, 22*(2), 282-294.

Crocker, L., Llabre, M., & Miller, M.D. (1988). The Generalizability of Content Validity Ratings. *Journal of Educational Measurement, 25*, 287-299.

Cui CY, Wang Y, Zhang Y, et al. (2021). The development and validation of the psychological capital questionnaire for patients with Cancer the psychological capital questionnaire. *BMC Cancer.* 21(1), 1194.

Davis, L.L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research, 5,* 194-197.

Fleiss, J. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin, 76*, 378–382.

Gregory, R.J. (1992). *Psychological Testing: History, Principles and Applications,* Allyn and Bacon: Boston.

Haynes, S.N., Richard, D.C.S., & Kubany, E.S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*,238-247.
Johnson, B. & Christensen, L. (2012). *Educational research quantitative, qualitative, and mixed approaches*. Alabama: SAGE Publications Inc.

Landsheer, J.A., & Boeije, H.R. (2010). In search of content validity: Facet analysis as a qualitative method to improve questionnaire design. *Quality and Quantity, 44,*59-69.
Lou J, Li R, Chen S. (2022). Development of the psychological capital scale for male nursing students in Taiwan and testing its measurement invariance between genders. *Int J Environ Res Public Health.*19(6).

Lu, X, Wang, L, Xu, G, Teng, H, Li, J & Guo, Y. (2023). Development and initial validation of the psychological capital scale for nurses in Chinese local context. *BMC Nursing,* 22:28.

Luthans, F., Avolio, B.J., Avey, J.B., & Norman, S.M. (2007). Positive psychological capital: Measurement and relationship with performance and satisfaction. *Personnel Psychology, 60*(3), 541-572.

Luthans, F., Youssef, C. M., &Avolio, B. J. (2007). *Psychological capital: developing the competitive edge*. New York: Oxford University Press

Lynn, M.R. (1986). Determination and quantification of content validity. *Nursing Research, 35,* 382-385.

Mary, B., Billeen, Kruszewski, A., Sims, K., Fettes, S. & Wilson, S. (2013). Content validity of the Wilson-Sims falls risk assessment tool to measure fall risk of psychiatric inpatients.

Mastaglia, B., Toye, C., & Kristjanson, L.J. (2003). Ensuring content validity in instrument development: Challenges and innovative approaches. *Contemporary Nurse, 14,*281-291.

Miller, L.A., et.al. (2013). *Foundations of psychological testing: a practical approach.*   California: SAGE Publications Inc.

Norman, S. M., Avolio, B. J., & Luthans, F. (2010). The impact of positivity and transparency on trust in leaders and their perceived effectiveness. *The Leadership Quarterly, 21*(3), 350-364.

Parker, S. (1998). Enhancing role-breadth self-efficacy: The roles of job enrichment and other organizational interventions. *Journal of Applied Psychology,83,* 835–852.

Polit, D.F., & Beck, C.T. (2006). The content validity index: are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health, 29,* 489-497.

Rico, E.D., Dios, H.C., & Ruch, W. (2012). Content validity evidences in test development: An applied perspective. *International Journal of Clinical and Health Psychology*, *12*(3), 449-460.

Rego A, Marques C, Leal S, et al. (2010). Psychological capital and performance of Portuguese civil   servants: exploring neutralizers in the context of an appraisal system. *Int J Hum Resour Manag.* 21(9), 1531–52.

Rehman & Ali, A. (2017). Relationship between psychological capital and teachers' job   commitment. *Journal of Applied Environmental and Biological Sciences,* 8(3), 148-154.

Rubio, D.M., Berg-Weger, M., Tebb, S.S., Lee, E.S., and Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research, 27,* 94-104.

Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: Assessment   and   implications   of   generalized   outcome expectancies. *Health Psychology,4,* 219–247.

Sekaran, U. & Bougie, R. (2011). *Research methods for business: A skill building approach.*   New Delhi: John Wiley & Sons.

Snyder, C. R., Irving, L. M., & Anderson, J. R. (1991). Hope and health. In C. R. Snyder (Ed.), *Handbook of social and clinical psychology* (pp. 295–305). Oxford: Oxford University Press.

Stemler, S. (2004). A comparison of consensus, consistency, and measurement approaches to estimating inter-rater reliability. *Practical Assessment, Research, and Evaluation, 9*(4).

Tilden V.P., Nelson C.A., & May B.A. (1990). Use of qualitative methods to enhance content validity. *Nursing Research, 39,* 172-175.

Trochim, W. M. K. (2006). Introduction to Validity. *Social Research Methods*, retrieved from www.socialresearchmethods.net/kb/introval.php.

Wagnild, G.M., & Young, H. M. (1993). Development and psychometric evaluation of the resiliency scale. *Journal of Nursing Management,1,* 165–178.

Waltz, C. F., Strickland, O. L., & Lenz, E. R. (2005). *Measurement in nursing and health research* (3rded.). New York: Springer.

Wynd, C.A., Schmidt, B., and Schaefer, M.A. (2003). Two quantitative approaches for estimating instrument content validity. *Western Journal of Nursing Research, 25,* 508-518.

Zewude.G.T. T & Hercz. (2021). Psychological capital and teacher well-being: The mediating role of coping with stress. *European journal of educational research,* 10(3), 1227-1245.

Zhang K, Zhang S, Dong Y. (2010). Positive psychological capital: measurement and relationship with mental health. *Stud Psychol Behav*, 8(1), 58–64.

Zhou, X & Zheng, S. (2022). Psychological capital relates with teachers' enjoyment: The mediating role of reappraisal. *Frontiers in Psychology,* 13, 879312.