EDUR 8331 Applied Measurement
Activity 2: Validity - Answers

(a) Content Validity

Read the following publication.

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.

See especially Content Validation Guidelines starting page 244, and the summary in the Appendix.

The direct link appears below.

http://www.bwgriffin.com/gsu/courses/edur9131/2018spr-content/07-validity/07-1995-Haynes-Content-Validity-Assessment.pdf

Note that the content validity guidelines they outline can be summarized in the following steps.

1. Define/describe construct to be measured
2. Identify domains, or dimensions, of construct
3. Describe item selection and/or generation – creating pool of items, using various approaches (e.g., logic, experience, theory, literature, other scales, experts, targeted population participants)
4. Describe item fit to construct/dimensions, and assess number of items to ensure adequate measurement of construct and dimensions
5. Address assessment of item suitability - wording clarity, appropriateness for targeted population (e.g., low reading level for young folks)
6. Employ expert and participant review of items (and scale itself, with instructions, etc.)
7. Review field test results and make revisions, if needed. This step should be repeated, along with expert and participant reviews, until a final scale is judge satisfactory.

After reviewing the material above, read the following study with particular attention to their development of the Cyber Incivility scale.

Lim, V. K., & Teo, T. S. (2009). Mind your E-manners: Impact of cyber incivility on employees' work attitude and behavior. Information & Management, 46(8), 419-425.

http://www.bwgriffin.com/gsu/courses/edur8331/edur8331-activities/EDUR-8331-01-2009-Lim-workplace-uncivil.pdf

Of the 7 content validation steps outlined above, how were each addressed by Lim and Teo when developing the Cyber Incivility scale? List each step and explain how it was, or was not, addressed.

1. Define/describe construct to be measured

Defined construct (p 419): Cyber incivility is communicative behavior exhibited in computer-mediated interactions that violate workplace norms of mutual respect.

Described construct (p 419): Cyber incivility could be intentional or unintentional and causes harm, detrimental to victims work behavior and productivity if incivility is perceived poorly

2. Identify domains, or dimensions, of construct

Two dimensions identified, active and passive (p 420).

The authors did not define these in terms of incivility in the form of emails.

Neither did the authors formally identify which of the 14 items were examples of active or passive incivility.

However, combining information from Table 2 and Appendix A, it is possible to determine that items 1 to 7 are active examples (i.e., said something hurtful, wrote negative emails, made demeaning remarks, uses sarcastic or mean comments, condescending, used rude tone, used CAPS), and items 8 to 14 are the passive examples (i.e., no reply, ignored request, replied but gave no answer, used email for time sensitive message, showed little interest, not acknowledge email receipt, used email instead of face-to-face when latter preferred).

3. Describe item selection and/or generation – creating pool of items, using various approaches (e.g., logic, experience, theory, literature, other scales, experts, targeted population participants)

Item development followed input from a focus group of 10 employed individuals in an MBA program (p 420) – goal was item formation by identifying uncivil behaviors committed via email, 20 items resulted.

Focus was on supervisor conduct, supervisor uncivil behavior via email.

Additionally, the draft of 20 items initially developed from the focus group was "pre-tested" by two professors and 20 MBA students (p 421). Six items were eliminated due to their feedback.

Final scale contained 14 items.

4. Describe item fit to construct/dimensions, and assess number of items to ensure adequate measurement of construct and dimensions

Not formally addressed; informally, the reader can deduce that both the passive and active dimensions have 7 items each given information in Table 2, Appendix A, and brief descriptions of active and passive definitions.

Also, item fit here means that items appear to be relevant to the construct (e.g., if the construct is cyber incivility, the item should focus on that), and whether the items fit on of the dimensions of the construct. For example, cyber incivility has two dimensions, active and passive. If one of the items focused on the use of threatening words, would that fit with the active or passive dimension – that is the assessment of fit referenced in content validity.

5. Address assessment of item suitability - wording clarity, appropriateness for targeted population (e.g., low reading level for young folks)

 The "pre-testing" with two professors and 20 MBA students (p 421) was used to identify issues and eliminate problematic items. Additionally, 20 employees pretested the questionnaire (p 420) to assess "clarity of the instructions and overall presentation…"

6. Employ expert and participant review of items (and scale itself, with instructions, etc.)

Before employing questionnaire, and incivility scale, authors obtained feedback from 20 employees with focus on "clarity of instructions and overall presentation of the survey prior to its administration" (p 420).

> **Additionally, the draft of 20 items initially developed from the focus group was "pre-tested" by two professors and 20 MBA students (p 421). Six items were eliminated due to their feedback.**

7. Review field test results and make revisions, if needed. This step should be repeated, along with expert and participant reviews, until a final scale is judge satisfactory.

> **This implies administration of the questionnaire followed by revision if weaknesses were found. To some extent the "pre-test" with two professors and 20 students served as an initial field test, although no discussion of validity or reliability is presented. The administration to collect data for the study, therefore, is the initial field test, and the reliability and validity information provided can be viewed as the field-test analysis to use for future revisions, if any.**

(b) Empirical Validity
In my notes on validity, linked below, I explain that evidence for validity can take several forms. The most common are internal structure (i.e., assessing structure of scale items via factor analysis or similar procedures) and relational assessment (i.e., do scores from a scale correlate with other variables as hypothesized, do scores show mean differences as hypotheses among groups, do scale scores predict some future behavior as predicted, etc.).

http://www.bwgriffin.com/gsu/courses/edur7130/2018spr-chats/EDUR-7130-Chat-6b-Validity.pdf

Again, using the Lim and Teo study above, what evidence, if any, did they provide for empirical validity of the Cyber Incivility scale? Address this question by answer the following.

8. Reliability – for scores to demonstrate validity, they must demonstrate reliability. What evidence did Lim and Teo provide for the reliability of the Cyber Incivility scale?

> **They report reliability to be .95 (Table 2 and p 421). They do not report the type of reliability calculated, although it is likely Cronbach's alpha.**

9. Internal structure – my notes are incomplete on this form of validity assessment, so no need to explain whether this was addressed (p.s., they did address, to some extent, internal structure, see table 1 factor analysis).

> **Internal consistency was addressed via factor analysis, and results are shown in Table 1 (p 421). Two factors for incivility were found, one aligning with the active items, thus the active dimension, and the second aligning with the passive items, which signifies the passive dimensions. Factor loadings (not clearly identified as factor loadings in table 1 or text) ranged from .71 to .84 for both dimensions/factors. The cross-loadings – active item loadings on the passive factor, and passive item loadings on the active factor – were weaker than the primary loadings. One would expect a similar patterns of factor loadings if there is good internal structure, so this is sufficient evidence of the internal structure of the cyber incivility items.**

> **The authors also report factor loadings on items designed to measure other constructs. For these, one would hope to see clear distinctions among factor loadings with items designed to measure one construct loading heavily on that construct while loading weakly on other constructs. Loadings of the incivility items were weak for the other factors investigated (Organizational Commitment, Job Satisfaction, etc.). This too is solid evidence for the internal structure of the incivility item responses. Additionally, as the authors note, these findings can be interpreted as a form of discriminate validity since the disparate items loaded as they should**

(on their own factors, but not on others) thus showing that various items discriminated among the various constructs.

10. Relational validity evidence – what evidence, if any, did Lim and Teo provide of relational validity evidence for the Cyber Incivility scale? Remember, for validity evidence, there must be hypotheses about how scores from a scale will relate to other variables.

(a) Primary Evidence

The primary evidence stems from the hypotheses presented (p 420) and the results obtained (Table 2 and p 422-423).

Hypotheses
1a. Cyber Incivility negatively related to job satisfaction
1b. Cyber Incivility negatively related to organizational commitment
2a. Cyber Incivility positively related to workplace deviance
2b. Cyber Incivility positively related to quit intention
3a. Male supervisors more likely to use active cyber incivility
3b Female supervisors more likely to use passive cyber incivility

Correlations in Table 2 address hypotheses 1a, 1b, 2a, and 2b. Each of the hypotheses were supported. Correlations are reported below for each hypothesis. All correlations were in the direction hypothesized, and significant, which provides good evidence for the validity of these cyber incivility scores.

r = -.40 1a. Cyber Incivility negatively related to job satisfaction
r = -.44 1b. Cyber Incivility negatively related to organizational commitment
r = .38 2a. Cyber Incivility positively related to workplace deviance
r = .19 2b. Cyber Incivility positively related to quit intention

Means and t-test results are reported in Table 4 and address hypotheses 3a and 3b. As predicted, male supervisors were more likely to engage in active incivility with both female and male employees (means of 1.94 vs 1.37 and 1.70 vs 1.26), and female supervisors were more likely to use passive incivil email with female and male employees (means of 2.56 vs 1.90 and 2.30 vs 1.80). As above, evidence for validity of these scores.

(b) Secondary Evidence

The authors report results of regression with cyber incivility as a predictor in addition to two covariates (predictors), age and social desirability. After controlling for these, each of the four outcomes remained related to cyber incivility as predicted in the hypotheses above. Further support for construct validity of the cyber incivility items.

Additionally, the passive and active dimensions of cyber incivility were used as predictors in the same regression equation as above, and the authors found that active cyber incivility was a stronger predictor than passive incivility, as one might expect.

Lastly, the authors examined specific items to learn whether preferences for certain types of cyber incivility differed by supervisor sex. Results are shown in Figure 3. As hypothesized, males were more likely to use active, and females more likely to use passive acts of incivility in emails. As above, this finding supports validity of scores.