A visual way to represent content validity and item properties

Vanessa Sánchez-Mendoza[1], Encarnación Soriano-Ayala[2],
Pablo Vallejo-Medina[3]

Universidad de Almería[1,2] – España y Fundación universitaria Konrad Lorenz[3] –
Colombia

The purpose of this instrumental study was to show how to conduct an RStudio process to obtain a content validity visual representation through the alluvial function of the ggplot2 package to Condom Use Self-Efficacy Scale (This paper explains the stages to analyze content validity using a visual representation of an interrater agreement and item – factor belonging and focused mainly on qualitative aspects to evaluate the judge's behavior and items' weight. Furthermore, the paper discusses classical criteria to obtain content validity and its representation. The visual representation used shows high relevance to present the qualitative data associated with the content validity analysis since it allows to visualize in a differential way the weight of the dimensions in each item, the content validity coefficient in the set of dimensions, and the behavior of the judges in the overall evaluation of the scale.
Keywords: CUSES, content validity, RStudio, alluvial plot.

Una forma visual de representar la validez de contenido y las propiedades de los ítems

El propósito de este estudio instrumental fue mostrar cómo llevar a cabo un análisis a través de la plataforma RStudio para obtener una representación visual de validez de contenido usando la función aluvial del paquete ggplot2 para la escala Condom Use Self Efficacy. Este artículo explica las etapas para analizar la validez de contenido utilizando una representación visual del acuerdo entre evaluadores y la pertenencia de cada ítem enfocado principalmente en aspectos cualitativos que permiten a la vez evaluar el comportamiento del juez e identificar el peso otorgado a los ítems. Además, el documento discute acerca de los criterios clásicos para obtener la validez de contenido y su representación. La representación visual utilizada, muestra gran relevancia para presentar los datos cualitativos asociados al análisis de validez de contenido, ya que permite visualizar de manera diferencial el peso de las dimen-

[1]   Research Director at Fundación Universitaria Konrad Lorenz. Bogotá – Colombia. PhD Candidate at Universidad de Almería, Spain. Postal address: Carrera 10 No. 64-61. Bogotá, Colombia. Zip Code: 110231. Email: vsm646@inlumine.ual.es. https://orcid.org/0000-0001-7034-5889

[2]   PhD in Philosophy and Education Sciences. Professor of Humanities and Education. Universidad de Almería, Spain. Postal address: Carretera Sacramento, Calle San Urbano. La Cañada, Almería, España. Zip Code: 04120. Email: esoriano@ual.es. https://orcid.org/0000-0002-9506-0625

[3]   PhD in Psychology. Director of the human sexuality laboratory – Sexlab, Fundación Universitaria Konrad Lorenz. Bogotá – Colombia. Postal address: Carrera 10 No. 64-61. Bogotá, Colombia. Zip Code: 110231. Email: pablo.vallejom@konradlorenz.edu.co. https://orcid.org/0000-0002-8527-1584

siones en cada ítem, el coeficiente de validez de contenido en el conjunto de dimensiones y el comportamiento de los jueces en la evaluación general de la escala.
Palabras clave: CUSES, validez de contenido, RStudio, alluvial plot.

Uma maneira visual de representar a validade de conteúdo e as propriedades do item

O objetivo deste estudo instrumental foi mostrar como conduzir um processo RStudio para obter uma representação visual da validade do conteúdo por meio da função aluvial do pacote ggplot2 para a Escala de Autoeficácia do Uso do Preservativo. Este artigo explica as etapas para analisar a validade de conteúdo usando uma representação visual de um acordo inter-avaliador e item - pertencimento de fator e focado principalmente em aspectos qualitativos para avaliar o comportamento do juiz e o peso dos itens. Além disso, o artigo discute critérios clássicos para obter validade de conteúdo e sua representação. A representação visual utilizada apresenta grande relevância para apresentar os dados qualitativos associados à análise de validade de conteúdo, pois permite visualizar de forma diferencial o peso das dimensões em cada item, o coeficiente de validade de conteúdo no conjunto de dimensões e o comportamento dos juízes na avaliação geral da escala.
Palavras-chave: CUSES, validade de conteúdo, RStudio, alluvial plot.

Une façon visuelle de représenter la validité du contenu et les propriétés de l'item

L'objectif de cette étude instrumentale était de montrer comment mener un processus de RStudio pour obtenir une représentation visuelle de la validité du contenu par la fonction alluvionnaire de l'ensemble ggplot2 à l'échelle d'auto-efficacité de l'utilisation des préservatifs. Ce document explique les étapes de l'analyse de la validité du contenu à l'aide d'une représentation visuelle d'un accord entre les évaluateurs et de l'appartenance d'un élément à un facteur et se concentre principalement sur les aspects qualitatifs pour évaluer le comportement du juge et le poids des éléments. En outre, le document aborde les critères classiques pour obtenir la validité du contenu et sa représentation. La représentation visuelle utilisée montre une grande pertinence pour présenter les données qualitatives associées à l'analyse de la validité du contenu puisqu'elle permet de visualiser de manière différentielle le poids des dimensions dans chaque item, le coefficient de validité du contenu dans l'ensemble des dimensions et le comportement des juges dans l'évaluation globale de l'échelle.
Mots-clés : CUSES, validité du contenu, RStudio, parcelle alluviale.

Sexual intercourse is one of the main pathways of transmission for HIV and is the main cause of infection among young in Colombia (Instituto Nacional de Salud, 2019), several strategies are developed to recognize condom use as the only effective rate against both STIs, and pregnancy (UNFPA, 2015).

Studies showed that variables with greatest prediction about condom use include attitudes towards condom (Fergunson, 2011; Albarracin et al., 2001; Diez et al., 2000; Sheeran et al., 1999). Other psychosocial key variables are knowledge about HIV and other ITS´s, (Vallejo-Medina et al., 2018; Yi et al., 2018), subjective norms or beliefs (Vincent et al., 2016; Ellis et al., 2015) intention to use (Morales et al., 2019; Visser, 2017; Farrington et al., 2016; Widman et al., 2019), communication about condom (Cassar , 2019; Gause et al., 2018; Caouette et al., 2018; Horan & Cafferty, 2017) and condom use self-efficacy (Brien et al., 1994; Forsyth et al., 1997; Barkley & Burns, 2000; Asante & Doku, 2010; Shaweno & Tekletsadik, 2013; Sanchez-Mendoza et al., 2020). Intentions to use condom and subjective norms or beliefs are usually measured with single questions, while others are assessed trough standardized scales. In Colombia, some versions are adapted to measure such variables (Plaza-Vidal et al., 2020; Morales et al., 2019; Morales et al., 2018; Vallejo, 2017; Vallejo et al., 2016; Soler et al., 2016) but not to measure condom use self-efficacy yet.

Despite the importance of assessing the content validity, most psychometric studies skip this step (Vallejo-Medina et al., 2019; Mansouri et al., 2019; Gil-Llario et al., 2019) in fact, most psychometric works just give Cronbach's alpha and an Exploratory Factor Analysis - with main components - (Gil-Llario et al., 2019; Golub & Gamarel, 2017). Content validity is usually obtained through the inter-judge agreement, frequently three expert judges are needed (appointment) although this number of judges is not an agreement (Cronbach, 1971;

Carmines & Zeller, 1979; Haynes et al., 1995; Polit & Beck, 2006; Carretero-Dios & Pérez, 2007; LeBreton & Senter, 2008). In any case, it does not seem a difficult requirement to reach for a researcher. Some of the most popular estimators calculate consensus among judges are: a) the kappa statistic used to test interrater reliability, and represent the extent to which the data collected are correct representations of the variables measured, taking into account the possibility that two judges on categorical (or ordinal) data actually guess on at least some variables due to uncertainty (Cohen, 1960) is easily calculated through statistical programs like SAS, SPSS or STATA; b) iota coefficient, used when a variety of targets are rated by different set of judges, is an index of inter-rater agreement of quantitative or nominal multivariate observation (Janson & Olson, 2001); c) Kendall's coefficient of concordance, used as an index of interrater reliability of ordinal data (McGraw & Wong, 1996) and, d) Coefficient of content validity (Aiken's V), applied to summarize the ratings of expert judges through estimates of content validity by means of coefficients (Aiken, 1985).

At the mathematical or technical level, they do not seem difficult indicators to obtain either. So, why do so many researchers skip this step that minimizes the problems of construct validity so much? A possible explanation may be related to the number of words in the article, for example, an evaluation on a table of specifications that measures five attributes of 20 items by five judges, assumes at least 500 words for a table. It is possible to save words by highlighting trends, but much relevant information would be lost.

To achieve the aim, the authors made an interrater agreement using alluvial visual representation in R, a tool that shows a lot of relevant information about items marking and agreement between judges.

Method

Participants

On one hand, participants took place in this study as follows: two translators, five researchers with adequate expertise in sexuality, and three bilingual psychologists, whom translate and adipate linguistically the CUSES (see procedure). On the other hand, twelve professionals (three sexual health specialists, four psychometrics, one public health PhD and three sexologists) were asked to evaluate the translation and cultural adaptation of the CUSES in four attributes and they assessed the content validity, and some item properties through a specification table.

Instruments

Condom Use Self-Efficacy Scale (CUSES, Brien et al., 1994). Evaluates a person's feelings of confidence about the possibility of buying condoms, putting them on, taking them off, and negotiating their use with a new sexual partner. The CUSES, in its first version, had 28 items with five Likert-scale response alternatives ranging from "Strongly disagree" = 0, to "Strongly agree" = 4. It also had seven items coded in reverse. These 28 items grouped in the initial application into 15 factors. Cronbach's alpha was .91 and test re-test correlation after two weeks .81 (Brafford & Beck, 1991). Later studies propose changes in the factorial structure of the scale, reducing both, items, and factors (Brien et al., 1994; Forsyth et al., 1997; Barkley Jr. & Burns, 2000; Asante & Doku, 2010; Shaweno & Tekletsadik, 2013,) with Cronbach's alpha between .66 and .92 in all new versions.

Abbreviated version has 15 items with five response alternatives in Likert scale ranging from Strongly disagree = 0, to Strongly agree = 4, seven items coded in reverse. It has four factors, Mechanics (e.g., "*I feel confident in my ability to put a condom on myself or my partner*"; Cronbach´s alpha = .79), Partner's disapproval (e.g., "*I would not feel confident suggesting using condoms with a new partner because I would*

be afraid he or she would think I have a sexually transmitted disease”; Cronbach´s alpha = .71), Assertive (e.g., “*I feel confident in my ability to suggest using a condom with a new partner*”; Cronbach´s alpha = .73) and Intoxicants (e.g., “*I feel confident that I would remember to use a condom even if I were high*”; Cronbach´s alpha = .82).

Adapted version of the scale can be consulted in supplemental materials.

Procedure

In the first part of the study a team of psychologist and translators, translated and adapted linguistically the CUSES for use with Colombian Spanish speaking participants. Translators used intercultural adaptation guidelines for psychological measurement instruments based on Muñiz et al. (2013), as well as AERA, APA, and NCME (2014) standards.

Twelve professionals were asked to evaluate content validity and some item properties through the following criteria: a) representativeness, item’s representative degree of the self-efficacy variable in condom use; b) comprehensiveness, assessment of whether the item is properly understood; c) interpretability, number of item’s interpretations; d) clarity, item’s conciseness / direct; and e) belonging, construct’s component to which the item belongs. Four construct’s components were the options: Appropriation (confidence to acquire and use condom in a sexual intercourse), Partner’s disapproval (confidence to deal with the possible rejection of a sexual partner), Assertiveness (individual’s perceived ability to be assertive about the use of condoms) or Intoxication – “Self-control” in the Spanish version- (confidence to use a condom under the effect of substances that alter consciousness or sexual excitement). These evaluations enabled a final version of the scale to be reached.

Data analyses

This study was conducted with R (R Core Team, 2017) under RStudio interface (RStudio Team, 2016) content validity and other items properties were plotted using "ggplot2" package, "a system for declaratively creating graphics using the grammar of graphics" (Wickham et al., 2019. p. 1), an alluvial diagram, with "ggalluvial" function as a categorical data visualization technique generated its graphical representation, consist of axes, strata, alluvial and lodes (Brunson, 2019).

An axis is a dimension in which the data are vertically grouped at a fixed horizontal position, as is shown in the chart below (Figure 1) these correspond to representativeness, comprehensiveness, interpretability, and clarity. The strata are related to the categorical options for each dimension, for example, in representativeness, are four categories: none, low, moderate, and high. Finally, the alluvial is the horizontal-coloured spline and, each alluvium represents the fixed value of the dimension evaluated by the judge for each item.
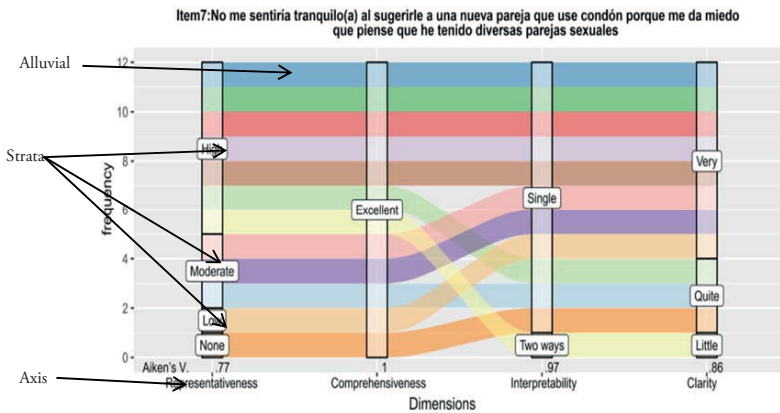

Figure 1.

The Alluvial diagram is a variant of a Parallel Coordinates Plot (PCP) but for categorical variables, these are assigned to vertical axes that are parallel. Values are represented with blocks on each axis and observations are represented with alluvia spanning across all the axis (Palma & Masera, 2018). An alluvial diagram allows us to visualize the interconnection of frequencies, for this study, those related to the categories stablished in the four dimensions of evaluation of the CUSES items. This function visualizes categorical data with accurately convey proportions to viewers from tabular (or array) form, popular for storing data with multiple categorical dimensions. Raw data and syntaxes can be consulted in: https://github.com/CUSES/CUSES-Validation

For a better understanding, we present the description of the alluvial data and the explanation of the syntaxes for this study. The image below (Figure 2) reflects the visual layout of an alluvial plot for item 1. Each row corresponds to a judge whose observations take a specific value in each category, and each category has its own column (R: representativeness; C: comprehensiveness; I: interpretability; Cl: clarity). This same layout of the information was done for each item in the same layout and later converted to the format in which the base function can be executed in R as.data.frame().

Judge	Item	R	C	I	Cl
1	1	4	4	3	3
2	1	4	4	4	4
3	1	4	4	4	3
4	1	4	3	3	4
5	1	3	4	4	4
6	1	4	4	4	4
7	1	4	3	3	3
8	1	4	4	4	4
9	1	3	3	3	3
10	1	4	4	4	4
11	1	4	4	4	3
12	1	4	4	4	4

Figure 2.

First, open the RStudio database, install and use the following packages from the library, *ggalluvial*, *ggplot2* and *ggpurb*. Now set factors in each item, as follow, in parenthesis are the categories chosen by judges for the item and organize the order in which the data must be presented, for this study begins with judge1 and ends with judge12. Add as many labels as need, ("High", "Moderate", "Low", "…"))

it1jueces$Rord <- factor(it1jueces$representativeness, c("High", "Moderate"))
it1jueces$Cord <- factor(it1jueces$Comprehensiveness, c ("Excellent", "Good"))
it1jueces$Iord <- factor (it1jueces$interpretability, c("Single", "Two ways"))
it1jueces$Clord <- factor (it1jueces$Clarity, c("Very", "Quite"))
it1jueces$Jord <- factor (it1jueces$Jugde, c("judge1", "judge2", "judge3", "judge4", "judge5", "judge6", "judge7", "judge8", "judge9", "judge10", "judge11", "judge12"))

Next, create the figure by applying the following steps, start naming the figure and using the database in the format as.data.frame; establish the axes names and characteristics; define the alluvium characteristics, as to fill, width, space between axis, labels, categories, title, and, the desired appearance like colour.

Figure1<-ggplot(as.data.frame(it1jueces),
aes(y = frec, axis1 = Rord, axis2 = Cord, axis3 = Iord, axis4 = Clord)) +
geom_alluvium(aes(fill = Jord), width = 1/8) +
geom_stratum(width = 1/8, alpha=0.5) +
geom_label(stat = "stratum", label.strata = TRUE, size = 3) +
scale_x_discrete(limits = c("representativeness", "Comprehensiveness", "interpretability", "Clarity"), expand = c(.1, .1)) +
scale_fill_brewer(type = "qual", palette = "Paired") +
ggtitle("Item2: Tengo confianza en mi capacidad de ponerme un condon o ponerselo a mi pareja")+ xlab("Dimensions")+

theme(plot.title = element_text(lineheight=.8, face="bold",hjust = 0.5, color = "black", size = 10)) + coord_cartesian(ylim=c(0, 12)) + scale_y_continuous(breaks=seq(0, 12, 2))

Include the Aiken's Value in each figure, it is obtained by means of a formulated Excel sheet, in which the category chosen was multiplied by the number of judges who awarded it, divided by the total number of judges minus 1 and, this value divided into 3. The resulting coefficient of this value can have rates between 0 and 1, the closer the value is to 1, the greater the validity of content the item will have. For example, for item 1, the formula was as follows:

*=((((category [4] * number of judges who awarded it [12])+(category [3]* number of judges who awarded it [1]))/12)-1)/3*

Then integrate all the independent figures into a single one using the following syntaxes, this makes it possible to choose the number of columns and rows to present the figures.

Sfig1<-ggarrange(FJ1, FJ2, FJ3, FJ4,FJ5,FJ6, FJ7, FJ8, FJ9, FJ10, FJ11, FJ12, FJ13, FJ14, FJ15, common.legend = TRUE, ncol = 3, nrow = 5)

Afterwards, analyse the data base to prepare the figure 2, include the factors information and the order in which the data will be displayed for example from least to greatest.

JA2$Itemreord <- factor(JA2$Item, c("item1", "item2", "item3", "item4", "item5", "item6", "item7", "item8", "item9", "item10", "item11", "item12", "item13", "item14", "item15"))
JA2$Factoreord <- factor(JA2$Factor, c("Appropriation", "Partner disapproval", "Assertiveness", "Self control" , "Missing"))
mutate(JA2$Item = stats::fct_relevel(JA2$Item, "item1", "item2", "item3", "item4", "item5", "item6", "item7", "item8", "item9", "item10", "item11", "item12", "item13", "item14", "item15")) %>%

Finally, make the figure 2 through the following syntaxes, define the information that goes on the axes and assign the desire labels, size and appearance characteristics.

```
F2<-ggplot(as.data.frame(JA2),
aes(y = frec, axis1 = Itemreord, axis2 = Factoreord)) +
geom_alluvium(aes(fill = Factoreord), width = 1/8) +
geom_stratum(width = 1/8, alpha=0.5) +
geom_text(stat = "stratum", label.strata = TRUE, size = 3) +
scale_x_discrete(limits = c("Factor"), expand = c(.1, 0)) +
scale_fill_brewer(type = "qual", palette = "Set3") +
ggtitle("Content Validity")+
xlab("Criterio")+
theme(plot.title = element_text(lineheight=.8, face="bold",hjust =
0.5, color = "black", size = 10))
```

Results

Figure 3 shows the observations of the expert judges and the coefficient of content validity for inter-observer reliability analysis. Visual representation illustrates the behaviour of each judge in the evaluation process as well as the evaluation of each judge against each criterion by item. For instance, we can observe how judges 1, 2, 4 and 6 had an acquiescent trend to value items as "perfect". While judges 8 and 11 had a more critical view of the items content. Furthermore, item 2 and 5 -despite Aiken V value is the same for comprehensiveness and interpretability (item 2) and, representativeness and clarity (item 5)- but labels of the characteristic representing those values with different weight.
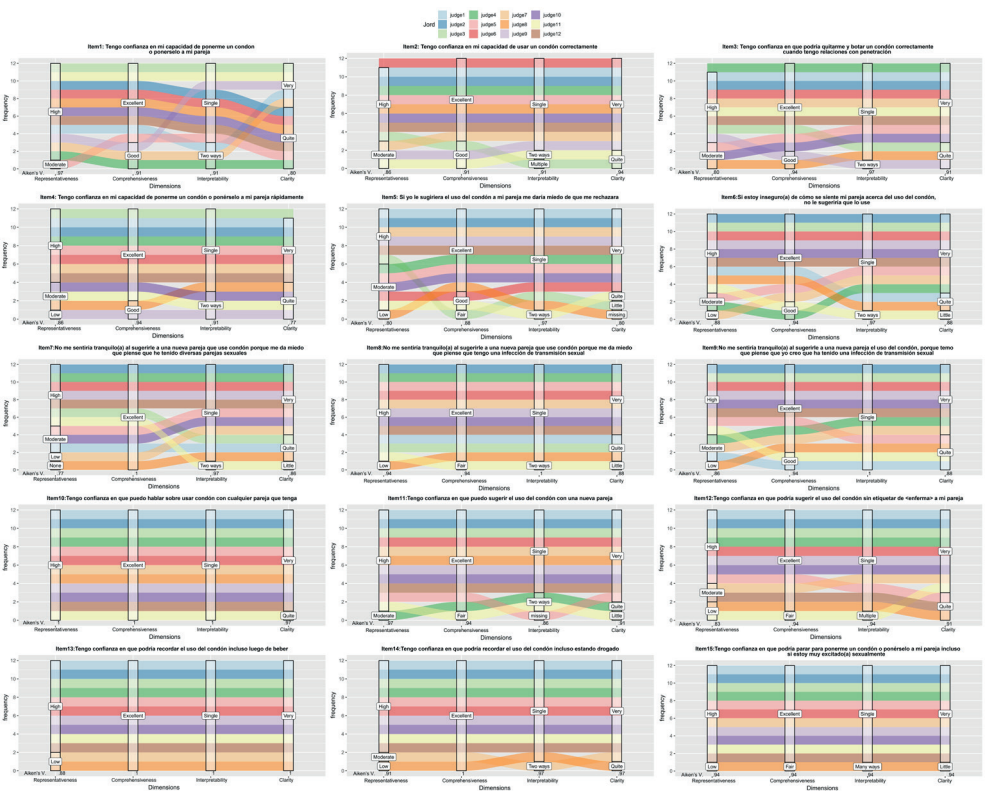
Figure 3. Judges agreement visual representation with ggalluvial

Figure 4 shows the percentage of agreement between judges aga-inst the factor to which each one belongs. An additional benefit of using this visualization is that is easier to observe items codominance in two dimensions, opposite to a classical numerical presentation. Thus, items 1, 2, 3 and 4 have mixed qualitative weights within appropriation and assertiveness, and item 15 with appropriation and self-control. Implications for FA will be discussed.
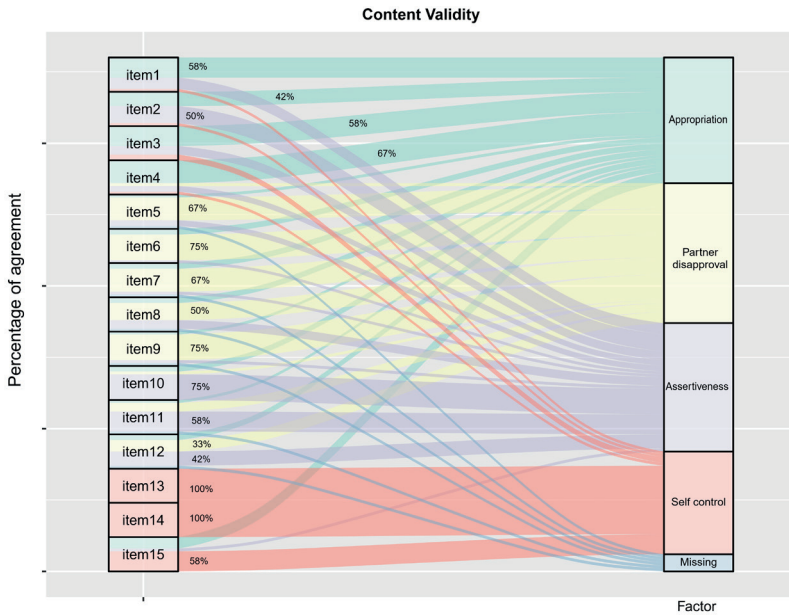
Figure 4*. Content validity visual representation with ggalluvial*

Discussion

The objective of the present study was to show a visual representation about some item's properties –content validity included- trough R studio. For those purpose, after translating and adapting the CUSES, twelve professionals were asked to evaluate its translation and cultural adaptation in four attributes and four factors. Semantic definition of the construct to be measured facilitates its syntactic definition and decreases the likelihood of ambiguities in the adaptation process; also state that its presentation through a specification table facilitates the evaluation that judges do, although it is common to ignore it (Carretero-Dios & Pérez, 2005).

Present study graphic representation provides helpful information about the distribution of the weights in each criterion evaluated by judges. It also displays trends of each judge evaluation and contrast the value of coefficient of content validity according to criteria. As could be seen here, three judges may not be a good number. Presence of missing data, acquiescence of some experts in the "best label", or an extreme criticism may determine the final agreement. Thus, we suggest using a higher number of experts. Additionally, the presentation of the test specifications is usually carried out through a table that summarizes the general data of the judge's evaluation. Perhaps limited by the manuscript words limitation, this prevents describing in depth the behaviour of the item in relation to the global scale and of each judge in relation to the set of judges. This feature is fixed with the use of the alluvial plot, since this allows identifying differentially the values attributed to each dimension and the relevance of each one in relation to the set of the items.

It is known that from a clear semantic definition of the construct it is possible to have better indicators of content validity, which in turn is related to the construct validity (Carretero-Dios & Pérez, 2007). Thus, with the results proposed here it is easy to see the possible overlap of some items in several factors. Despite, as happened with the versions by Brien, Thombs, Mahoney y Wallnau (1994), Asante and Doku (2010) and Shaweno and Tekletsadik (2013), four factors seems to set the scale (Appropriation, Partner´s disapproval, Assertiveness and Self-control), although this is different from results obtained in Thomas et al. (2000) with three factors. Visually is noted that some items may have relationships with two factors. In advance we could anticipate the using of Exploratory Structural Equation Modeling (ESEM) or by using rotations that force the positioning of an item on one of the factors. If item content is related to two dimensions, a rewording may help to clarify its factor. So that items with high scores are obtained even if judges do not agree about the attribute judged, or items with low scores can be obtained, although the judges are agreeing with item's attribute

This study suggests the use of a graphic representation to show the agreement between the judges, however the numerical indicators used

here were inserted afterwards. Thus, we still do not have an R package that matches the graphical display with the numerical descriptor, something that could be addressed in the future.

References

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement, 45*, 131-142. https://doi.org/ 10.1177/0013164485451012

Albarracin, D., Johnson, B., Fishbein, M., & Muellerleile, P. (2001). Theories of reasoned action and planned behavior as models of condom use: a meta-analysis. *Psychological Bulletin, 127*, 142-161. https://doi.org/10.1037/0033-2909.127.1.142

AERA, APA, & NCME. (2014). AERA, APA, and NCME. *Standards for educational and psychological testing.* American Educational Research Association.

Asante, K., & Doku, P. (2010). Cultural adaptation of the Condom Use Self-Efficacy Scale (CUSES) in Ghana. *BMC Public Health. 10:*227. https://doi.org/10.1186/1471-2458-10-227

Barkley, W. Jr,. & Burns, J. (2000). Factor analysis of the condom use self-efficacy scale among multicultural college students. *Health Education Research. 15*, 489-499. https://doi.org/10.1093/her/15.4.485

Brafford, L. J., & Beck, K. H. (1991). Development and validation of a condom self-efficacy scale for college students. *Journal of American College Health, 39*, 219-225. https://doi.org/10.1080/07448481.1991.9936238

Brien, T., Thombs, D., Mahoney, C., & Wallnau, L. (1994) Dimensions of Self-Efficacy among three distinct groups of condom users. *Journal of American College Health, 42*, 167-174. https://doi.org/10.1080/07448481.1994.9939665

Brunson, J. C. (2019). *Package ggalluvial.* Retrieved from https://cran.r-project.org/web/packages/ggalluvial/ggalluvial.pdf

Caouette, J. D., Hudson, K. A., Bryan, A. D., & Feldstein Ewing, S. W. (2018). Featured article: Adolescent condom use and connectivity in the social-planful brain. *Journal of Pediatric Psychology, 43*, 821-830. https://doi.org/10.1093/jpepsy/jsy025

Carmines, E. G., & Zeller, R. A. (1979). *Quantitative Applications in the Social Sciences: Reliability and validity assessment.* Thousand Oaks, CA: SAGE Publications, Inc. https://doi.org/10.4135/9781412985642

Carretero-Dios, H., & Pérez, C. (2005). Normas para el desarrollo y revisión de estudios instrumentales. *International Journal of Clinical and Health Psychology, 5*, 521-555.

Carretero-Dios, H., & Pérez, C. (2007). Standards for the development and review of instrumental studies: Considerations about test selection in psychological research. *International Journal of Clinical and Health Psychology, 7*, 863-882.

Cassar, J. (2019). There is condom-less sex but there are glow-in-the-dark condoms too: girls debating safe sex. *Sexuality & Culture, 23,* 1072-1088. https://doi.org/10.1007/s12119-019-09605-7

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37-46. https://doi.org/10.1177/001316446002000104

Cronbach, L. J. (1971). Test validity. In R. L. Thorndike, (Ed). *Educational measurement.* (pp.443-507). American Counsil on Education.

Diez, E., Juárez, O., Nebot, M., Cerda, N., & Villalbi, J. (2000). Effects on attitudes, knowledge, intentions and behaviour of an AIDS prevention programme targeting secondary school adolescents. *Promotion & Education, 7*, 17-22. https://doi.org/10.1177/102538230000700307

Ellis, E. M., Homish, G. G., Parks, K. A., Collins, R. L., & Kiviniemi, M. T. (2015). Increasing condom use by changing people's feelings about them: An experimental study. *Health Psychology, 34*, 941-950. https://doi.org/010.1037/hea0000205

Ellis, E. M., Rajagopal, R., & Kiviniemi, M. T. (2017). The interplay between feelings and beliefs about condoms as predictors of their use. *Psychology & Health, 33,* 176-192. https://doi.org/10.1080/08870446.2017.1320797

Farrington, E., Bell, D., & DiBacco, A. (2016). Reasons people give for using (or not using) condoms. *AIDS & Behavior, 20*, 2850-2862. https://doi.org/10.1007/s10461-016-1352-7

Forsyth, A. D., Carey, M. P., & Fuqua, R. W. (1997). Evaluation of the validity of the Condom Use Self-Efficacy Scale (CUSES) in young men using two behavioral simulations. *Health Psychology, 16*, 175-178. https://doi.org/10.1037/0278-6133.16.2.175

Gause, N. K., Brown, J. L., Welge, J., & Northern, N. (2018). Meta-analyses of HIV prevention interventions targeting improved partner communication: effects on partner communication and condom use frequency outcomes. *Journal of Behavioral Medicine, 41*, 423-440. https://doi.org/10.1007/s10865-018-9916-9

Gil-Llario, M. D., Ruiz-Palomino, E., Morell-Mengual, V., Giménez-García, C., & Ballester-Arnal, R. (2019). Validation of the AIDS prevention questionnaire: A brief self-report instrument to assess risk of HIV infection and guide behavioral change. *AIDS & Behavior, 23*, 272-282. https://doi.org/10.1007/s10461-018-2224-0

Gil-Llario, M. D., Morell-Mengual, V., Ruiz-Palomino, E., & Ballester-Arnal, R. (2019). Factorial structure and psychometric properties of a brief scale of the condom use self-efficacy for Spanish-speaking people. *Health Education & Behavior, 46*, 295-303. https://doi.org/10.1177/1090198118806964

Golub, S., & Gamarel, K. (2017). Psychometric evaluation of the condom carriers and motivations scale (CBMS). *Journal of Behavioral Medicine, 40*, 494-505. https://doi.org/10.1007/s10865-016-9815-x

Haynes, S. N., Richard, D., & Kubany, E.S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*, 238-247. https://doi.org/10.1037/1040-3590.7.3.238

Horan, S. M., & Cafferty, L. A. (2017). Condom communication: Reports of sexually active young adults' recent messages with new partners. *Journal of Health Communication, 22*, 763-771. https://doi.org/10.1080/10810730.2017.1355420

Instituto Nacional de Salud. (2019). *Informe de Evento. VIH/SIDA. Colombia 2018.* Retrieved from https://www.ins.gov.co/buscador-eventos/Informesdeevento/VIH-SIDA%20SEMESTRE%20I%202019.pdf

Janson, H., & Olsson, U. (2001). A measure of agreement for interval or nominal multivariate observations. *Educational and Psychological Measurement, 61*, 277-289. https://doi.org/10.1177/00131640121971239

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods, 11*, 815-852. https://doi.org/10.1177/1094428106296642

Mansouri, P., Nadrian, H., Allahverdipour, H., & Hejazi Bazargan, S. (2019). Condom use attitude and self-efficacy, and STIs risk perception among Iranian female sex workers: A psychometric study. *Health Care for Women International, 40*, 527-538. https://doi.org/10.1080/07399332.2018.1526288

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlations coefficients. *Psychological Methods, 1,* 30-46. https://doi.org/10.1037/1082-989X.1.1.30

Morales, A., Espada, J. P., & Orgilés, M. (2019). Mediation of an efficacious HIV risk reduction intervention for adolescents: A cluster-randomised controlled trial. *Journal of Health Psychology, 24*, 1884-1896. https://doi.org/10.1177/1359105317707256

Morales, A., Vallejo-Medina, P., Abello-Luque, D., Saavedra-Roa, A., García-Roncallo, P., Gomez-Lugo, M., García-Montaño, E., … & Espada, J.P. (2018). Sexual risk among Colombian adolescents: knowledge, attitudes, normative beliefs, perceived control, intention, and sexual behaviour. *BMC Public Health, 18*-1377. https://doi.org/10.1186/s12889-018-6311-y

Muñiz, J., Elosua, P., & Hambleton, R. (2013). Directrices para la traducción y adaptación de los test: segunda edición. *Psicothema, 25*, 151-157. https://doi.org/10.7334/psicothema2013.24

Palma, R. & Masera, G. (2018). Diagramas alluviales con R-CRAN. https://themys.sid.uncu.edu.ar/rpalma/Salta/Salta/Presentacion/R-CRAN/Aluvial/aluvial_Agricola.pdf

Plaza-Vidal, R., Ibagon-Parra, M., & Vallejo-Medina, P. (2020). Spanish Translation, Adaptation, and Validation of the Multi-

dimensional Condom Attitudes Scale with Young Colombian Men and Women. Archives of sexual behavior. Advance online publication. https://doi.org/10.1007/s10508-020-01759-y

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. https://doi.org/10.1002/nur.20147

Sanchez-Mendoza, V., Soriano-Ayala, E., Vallejo-Medina, P. Psychometric Properties of the Condom Use Self-Efficacy Scale among Young Colombians. *International Journal of Environmental Research and Public Health, 17*(11), 3762. https://doi.org/10.3390/ijerph 17113762

Shaweno, D., & Tekletsadik, E. (2013). Validation of the condom use self-efficacy scale in Ethiopia. *BMC International Health and Human Rights*, *13*, 22. https://doi.org/10.1186/1472-698X-13-22

Sheeran, P., Abraham, C., & Orbell, S. (1999). Psychosocial correlates of heterosexual condom use: A meta-analysis. *Psychological Bulletin, 125*, 90-132. https://doi.org/10.1037/0033-2909.125.1.90

Soler, F., Gómez-Lugo, M., Espada, J. P., Morales, A., Sierra, J. C., Marchal-Bertrand, L., & Vallejo-Medina, P. (2016). Adaptation and validation of the brief sexuality scale in Colombian and Spanish populations. *International Journal of Psychology and Psychological Therapy, 16*, 343-356.

Vallejo-Medina, P., Ramírez, C., Saavedra-Roa, D., Gómez-Lugo, M., & Pérez-Durán, C. (2019). Spanish validation of Female Condom Attitude Scale and female condom use in Colombian young women. *BMC Women's Health*, *19*, 128. https://doi.org/10.1186/s12905-019-0825-z

Vallejo-Medina, P., Saavedra-Roa, A., Gómez-Lugo, M., Morales, A., Abello-Luque, D., García-Montaño, E., Garavito-Zamudio, C. & Espada, J.P. (2018). Adaptación, fiabilidad y validez de una versión breve de la multicomponent AIDS phobia scale (MAPS) en adolescentes colombianos. *Revista Iberoamericana De Psicologia y Salud, 9*, 42-57. https://doi.org/10.23923/j.rips.2018.01.014

Vincent, W., Gordon, D. M., Campbell, C., Ward, N. L., Albritton, T., & Kershaw, T. (2016). Adherence to traditionally masculine norms and condom-related beliefs: Emphasis on African American and Hispanic men. *Psychology of Men & Masculinity, 17*, 42-53. https://doi.org/10.1037/a0039455

Visser, M. (2017). Rethinking HIV-prevention for school-going young people based on current behaviour patterns. *SAHARA-J: Journal of Social Aspects of HIV/AIDS*, 14, 64-76. https://doi.org/10.108 0/17290376.2017.1376704

Wickham, H., Chang, W., Henry, L., Lin Pedersen, T., Takahashi, K., Wilke, C., Woo, K., & Yutani, H (2019). Create Elegant Data Visualisations Using the Grammar of Graphics. Retrieved from https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf

Widman, L., Kamke, K., Evans, R., Stewart, J. L., Choukas-Bradley, S., & Golin, C. (2019) Feasibility, acceptability, and preliminary efficacy of a brief online sexual health program for adolescents, *The Journal of Sex Research*, *57*, 145-154. https://doi.org/10.108 0/00224499.2019.1630800

Yi, S., Te, V., Pengpid, S., & Peltzer, K. (2018). Social and behavioural factors associated with risky sexual behaviours among university students in nine ASEAN countries: a multi-country cross-sectional study. *SAHARA-J: Journal of Social Aspects of HIV/AIDS*, *15*, 71-79. https://doi.org/10.1080/17290376.2018.1503967