



Referencia: Rey-Sáez, R. (2022). Réplica a: validación cruzada sobre una misma muestra: una práctica sin fundamento. *Revista Electrónica de Metodología Aplicada*, 24(1), 41 – 44. <https://doi.org/10.17811/rema.24.1.2022.41-44>

***Correspondencia:** ricarrey@ucm.es

Recibido: 21/11/2021
Aceptado: 16/12/2021
Publicado: 15/01/2022

Réplica a: validación cruzada sobre una misma muestra: una práctica sin fundamento

Ricardo Rey-Sáez¹

¹Universidad Complutense de Madrid (Madrid, España)

Resumen: La práctica de validación cruzada sobre una misma muestra es un procedimiento habitual en la investigación aplicada, y se sugiere que no debe ser aplicada porque el contraste que se está realizando finalmente no es el adecuado. Es cierto que su aplicación conlleva ciertas consecuencias, pero ahondar en ellas puede ser más beneficioso que intentar establecer reglas de oro. El conocimiento en profundidad de qué se está utilizando para conseguir un determinado objetivo de investigación (funcionalidad y limitaciones) unido al contexto habitual en investigación aplicada (escasez de recursos) puede ayudar a los investigadores a ver hasta dónde pueden llevar sus conclusiones. Por tanto, en lugar de "validación cruzada: sí o no", se plantean algunos aspectos que deben tenerse en consideración, siendo finalmente obligación del investigador aplicado el justificar adecuadamente la ejecución de una validación cruzada en la misma muestra.

Palabras clave: Validación cruzada; Análisis Factorial Exploratorio; Análisis Factorial Confirmatorio

Reply to: cross validation on the same sample: an unfounded practice

Abstract: The practice of cross-validation on the same sample is a common procedure in applied research, and it is suggested that it should not be applied because the contrast that is ultimately being made is not adequate. It is true that there are certain consequences to its application, but further exploration of these may be more beneficial than trying to establish golden rules. In-depth knowledge of what is being used to achieve a given research objective (functionality and limitations) coupled with the usual context in applied research (scarcity of resources) can help researchers see how far they can take their findings. Therefore, instead of "cross-validation: yes or no", there are some aspects to be considered, and it is ultimately the applied researcher's obligation to adequately justify the execution of cross-validation on the same sample.

Keywords: Cross-validation; Exploratory Factor Analysis; Confirmatory Factor Analysis



Los autores argumentan que las únicas diferencias entre ambas soluciones factoriales serán debidas a la aleatorización previa de los participantes. Considero que el mensaje es correcto, y que grandes autores en el campo llevan mucho tiempo recomendando esta aplicación en distintas publicaciones (Byrne et al., 1989; Fabrigar et al., 1999), y manuales de referencia (Brown, 2015). También se plantean distintas estrategias para validación cruzada, como bootstrap (Thompson, 2004), pero no hay una reflexión propiamente dicha sobre la muestra como tal. No obstante, la afirmación sobre diferencias debidas al azar entre ambas soluciones factoriales es susceptible de llevar a confusión, y puede ser beneficioso ahondar más en el problema de la validación cruzada en la misma muestra.

Supongamos que contamos con dos muestras distintas, representativas de la misma población, y aplicamos un Análisis Factorial Exploratorio (AFE) en una y un Análisis Factorial Confirmatorio (AFC) en la otra. Si logro replicar la estructura factorial al aplicar la validación cruzada, las únicas diferencias que habrá entre ambas estructuras factoriales también serán debidas a fluctuaciones del azar vinculadas con el proceso de muestreo. Que en ambos casos (dos muestras o dos submuestras) las diferencias entre las estructuras puedan deberse a lo mismo (aleatorización) siendo en una de ellas un resultado deseable y en otra un posible inconveniente puede dificultar el identificar el problema subyacente. ¿Es lo mismo coger dos muestras representativas de la población que una única muestra representativa y dividirla aleatoriamente en dos muestras, también representativas de la población? Depende. Si el objetivo es agrupar a los participantes en distintas condiciones en una tarea experimental, entonces sí. Si la división se realiza después de recoger las medidas con el objetivo de realizar una validación cruzada, sí que nos encontramos con una diferencia clara: las condiciones de medida.

Ahora, supongamos que administramos un cuestionario a una muestra representativa grande de nuestra población de interés. Tras recoger los datos, subdividimos aleatoriamente la muestra original en dos submuestras distintas, y aplicamos una validación cruzada. Si hubiese algún tipo de contaminación durante el proceso de muestreo que pudiese afectar a las respuestas de los participantes de forma sistemática, este se encontraría en las dos submuestras. Esta situación derivaría en dos estructuras factoriales que podrían no diferir entre sí y que se hubiesen visto afectadas por las mismas posibles fuentes de error sistemático (Alvarado, 2019). Debido a que no podemos controlar todas las fuentes de variabilidad error que pueden afectar al proceso de medida, siempre es recomendable que las muestras sean independientes, pero la independencia de las observaciones no es lo único que debe tener en cuenta el investigador aplicado. Los constructos de interés en distintos campos de investigación suelen tener características distintas. Por ejemplo, la estabilidad del constructo puede variar entre distintos fenómenos teóricos en la población general: no es lo mismo medir personalidad o inteligencia que ansiedad o depresión. En lo referente a este problema, la investigación se ha centrado en estudiar la estabilidad de la recuperación de ciertos parámetros en modelos factoriales (Ondé & Alvarado, 2020). Por otro lado, aunque frecuentemente pensemos en pruebas de rendimiento típico, no podemos obviar los matices que las diferencian de las pruebas de rendimiento óptimo. Las personas se encuentran bajo circunstancias diferentes a la hora de enfrentarse a uno u otro tipo de prueba, y también deben tenerse en consideración cuando busquemos replicar una estructura interna con una validación cruzada.

Todas estas consideraciones deben ir acompañadas de las condiciones habituales en investigación aplicada. Es posible que la validación cruzada en la misma muestra fuese un problema inexistente si contásemos con suficientes recursos como para reclutar muestras grandes en distintos momentos temporales. Por tanto, los investigadores aplicados tienen que adaptarse y buscar alternativas que permitan llegar a algún tipo de conclusión. Teniendo la situación en cuenta, es posible que aplicar una validación cruzada en la misma muestra sí nos esté dotando de cierto grado de evidencia. Al aplicar un AFC sobre la estructura obtenida en un AFE podemos comprobar si las restricciones que aplicamos sobre ciertos parámetros derivan en una solución estable o no, y este resultado es informativo de por sí.

En última instancia, es el investigador aplicado quien debe tomar la decisión de aplicar o no una validación cruzada, teniendo en cuenta distintos aspectos: el objetivo de aplicar dicho procedimiento, la

funcionalidad y limitaciones de una validación cruzada en dos submuestras, su fenómeno de estudio, su disposición de recursos y el grado de evidencia al que puede llegar.

Conflicto de Intereses: El autor no declara ningún conflicto de intereses.

Aprobación Ética: No procede.

Consentimiento Informado: No procede.

Referencias

- Alvarado, J. (2019). *Resumen de malas prácticas en Análisis Factorial*.
<https://doi.org/10.13140/RG.2.2.34491.36646>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications.
- Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: the issue of partial measurement invariance. *Psychological Bulletin*, 105(3), 456-466. <https://doi.org/10.1037/0033-2909.105.3.456>
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299. <https://doi.org/10.1037/1082-989X.4.3.272>
- Ondé, D y Alvarado, J. (2020). Reconsidering the Conditions for Conducting Confirmatory Factor Analysis. *The Spanish Journal of Psychology*. 23, e55, 1-15.
<https://doi.org/10.1017/SJP.2020.56>
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. American Psychological Association. <https://doi.org/10.1037/10694-000>