EAM2025 XI Conference

23RD - 25TH JULY 2025

European Association of Methodology

Spain Tenerife Canary Islands

SCIENTIFIC REASONING SCALE (SRS): A MULTI-METHOD APPROACH PROVIDING EMPIRICAL EVIDENCE OF VALIDITY

ROSSELLA CALICIURI, MARGHERITA LANZ
UNIVERSITÀ CATTOLICA DEL SACRO CUORE, MILANO

INTRODUCTION

THE SCIENTIFIC REASONING SCALE (DRUMMOND AND FISCHHOFF, 2017) IS A SCALE TO TO MEASURE THE ABILITY TO THINK LIKE SCIENTISTS, THAT IS ABILITY TO EVALUATE SCIENTIFIC EVIDENCE

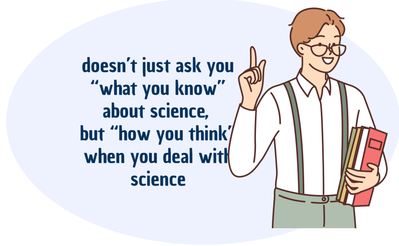
THEORETICAL BACKGROUND:

- PHILOSOPHY AND METHODOLOGY OF SCIENCE: *HOW SCIENCE SHOULD WORK*
- PUBLIC UNDERSTANDING OF SCIENCE: *WHAT PEOPLE ALREADY KNOW ABOUT SCIENCE*
- COGNITIVE DEVELOPMENTAL PSYCHOLOGY: *HOW PEOPLE LEARN SCIENTIFIC THINKING*

11 TRUE/FALSE ITEM, CONCEPT PER ITEM:

1. Blind / Double Blind
2. Causality
3. Confounding Variables
4. Construct Validity
5. Control Group
6. Ecological Validity
7. History
8. Maturation
9. Random Assignment to Condition
10. Reliability
11. Response Bias

In an experiment to evaluate coffee taste preferences, a researcher puts brand A coffee and brand B coffee in identical cups, placing identifying labels on the bottom to distinguish the coffee types. A lab assistant then distributes one of the cups to the tasters, while the researcher observes their facial expressions.
A. The lab assistant should not look at the cups while they are being filled.
B. The lab assistant should look at the cups while they are being filled.



- NO STUDIES HAVE EMPLOYED SEM TO VALIDATE THE SRS, NOR HAS MEASUREMENT INVARIANCE EVER BEEN TESTED, MEANING THAT NO EMPIRICAL EVIDENCE EXISTS TO SUPPORT THE ASSUMPTION OF MEASUREMENT EQUIVALENCE ACROSS GROUPS.
- NO VERSION OF THE SRS IS AVAILABLE FOR THE ITALIAN CONTEXT. ESTABLISHING VALIDITY EVIDENCE ACROSS DIVERSE CULTURAL SETTINGS IS ESSENTIAL BOTH TO BROADEN THE CROSS-NATIONAL APPLICABILITY OF THE CONSTRUCT AND TO ALIGN WITH CONTEMPORARY VIEW ON VALIDITY (HUBLEY AND ZUMBO, 2011).

TO ADDRESS THIS GAPS, THE PRESENT AIMS TO PROVIDE VALIDITY EVIDENCES OF THE SRS EMPLOYING A MULTI-METHOD APPROACH:

1. PRELIMINARY STUDY:
   - backtranslation
   - pilot test
   - CFA (n = 337)
2. QUALITATIVE PHASE
   - Linguistic review
   - Expert interviews (n =6)
   - Cognitive interview (n = 5)
3. QUANTITATIVE PHASE (n = 897)
   - Collection of validity evidences in line with *contemporary view of validity*:
     - FACTOR STRUCTURE: CFA
     - RELIABILITY: omega coefficient
     - GENERALIZABILITY: measurement invariance
     - KNOWN-GROUP: comparing factor means
     - CRITERION: correlation
     - CONVERGENT: regression

ALL THE MATERIALS ARE AVAIABLE ON OSF

METHODS

SAMPLE (n = 897):
representative of the Italian population aged 18 to 60 years (Istat, Istituto Nazionale di Statistica, 2024).

- GENDER: Female 50.5%
- AGE-GROUP: 45-60 (genY) 35.12%
- REGION OF RESIDENCE: Northwest 25.7%
- EDUCATIONAL LEVEL: High School Diploma 52.86%
- EMPLOYMENT STATUS: Employed 67.20%
- RELIGION AFFILIATION: *Yes* 52.8%
- POLITICAL ORIENTATION: Left 37.1%

MEASURES:

- *SRS* - italian version, 11 items
- COVERGENT MEASURES:
  - *Cognitive Reflection Test-Long* (CRT-Long; Frederick, 2005; Primi et al., 2016) - 6 items
  - *Probabilistic Reasoning Scale* (PRS; Primi, Morsanyi, Galli, and Chiesi, 2017) - 14 items
- CRITERION MEASURES:
  - *Paranormal Health Beliefs Scale* (PHBS, Donizzetti and Petrillo, 2017) - 31 items
  - *Climate Change Beliefs* - 2 items
  - *Climate Change Awareness* (Robba et al., 2024) - 6 items
- *Socio-demo*: gender, age, region of residence, education level, and employment status
- *Political orientation* - 1 item
- *Religious affiliation* - 1 item

RESULTS

SRS - % correct responses	
1. Double blind	69,40%
2. Causality	58,10%
3. Confounding variables	74,80%
4. Construct validity	83,00%
5. Control Group	69,50%
6. Ecological validity	72,50%
7. History	67,50%
8. Maturation	75,10%
9. Random assignment to condition	82,10%
10. Reliability	68,40%
11. Response bias	69,40%

DIFFERENT KINDS OF VALIDITY EVIDENCES through SEM (contemporary view of validity):

- FACTOR STRUCTURE: fit indices were good: [χ^2 (35) = 61.690, p = .004; RMSEA = .029 (.017 .041), p = .999; CFI = .966; WRMR = .953]. All factor loadings were high (greater than .3) and significant (p < .001), indicating a robust mono-factor structure.
- RELIABILITY: ω = .612.
- GENERALIZABILITY AND KNOWN-GROUP: measurement invariance was fully supported across gender, age, employment status, religious affiliation, and political orientation — showing invariant structure, factor loadings, thresholds, residuals, factor variance, and factor means. In contrast, for education level, only measurement invariance was established, with no evidence of structural invariance, indicating differences across educational groups.
- CONVERGENT: [χ^2 (53) = 88.14, p = .002; RMSEA = .027 (.017 .037), p = 1; CFI = .973; WRMR = .909]; higher SR scores are associated with higher scores in both cognitive reflection (.523) and probabilistic reasoning (.578), and vice-versa.
- CRITERION: [χ^2 (98) = 123.85, p = .040); RMSEA = .017 (.004 .026), p = 1; CFI = .989; WRMR = .750]; negative correlations with all the paranormal health belief factors (-.316; -.394; -.398; -.406; -.421); positive correlations with climate change awareness (.118) and beliefs (.131).

DISCUSSION

- We collected evidences to support the validity of inferences drawn from SR scores, employing a rigorous methodology (CTT and SEM) consistent with contemporary views of validity. This is crucial, as identifying and defining constructs is foundational for theory development in psychology: reliable measurement must precede empirical investigation. Thus, construct validation—gathering evidence that instruments truly capture the intended constructs—is both a challenging and essential step in the research process (Cronbach and Meehl, 1955; Flake and Fried, 2020).
- Demonstration of full measurement invariance strengthens the scale's generalizability and applicability to the Italian adult population,
- Education-related differences suggest a need for formal training in scientific reasoning during compulsory schooling;
- Strong correlations with related measures (CRT, PRS) and predictive power for paranormal health beliefs and climate change awareness confirm construct validity and practical relevance.

PLEASE, FOR ANY QUESTION:
rossella.caliciuri@unicatt.it

scan me to the poster