

Record details

IDESR ID

IDESR000086

Title

Protocol for a scoping review of L2 learners' cognitive processes research in language testing

Main Contact/Corresponding Author

Sathena Chan, University of Bedfordshire. sathena.chan@beds.ac.uk
Rm 119 Putteridge Bury Luton LU2 8LE UK

Additional Authors

Sin Wang Chong, University of St Andrews. swc5@st-andrews.ac.uk

Review Question

- 1.What language skills do studies on test takers' cognitive processes focus on?
- 2.What are the assessment contexts of these studies? What are the levels of test takers in these studies?
- 3.What research designs do these studies adopt? What research tools do these studies use to reveal test takers' processes?
- 4.What information about test takers' cognitive processes do these studies report?

Rationale

An understanding of the processes that underpin skilled writing can provide test developers with a systemic view of what a learner is likely to be capable of at different levels of proficiency (Field, 2004). Weir's (2005) seminal socio-cognitive framework (SCF) brings social, cognitive and evaluative dimensions of language together to inform language test development, research and validation. SCF identifies "cognitive validity" as one of the major components of language test validity. It concerns "how closely a task represents the cognitive processing involved in contexts beyond the test itself (Shaw & Weir, 2007, p.34)". Since the introduction of SCF, more researchers have started to examine test takers' cognitive processes with most of the early studies using think-aloud protocols. However, the fact that concurrent think-aloud protocols inevitably intervene with test taker's processes brings certain challenges for this line of research (Stratman & Hamp-Lyons, 1994). Until recently, the advancement in process-tracking technology, such as screen capture software, keystroke logging and eye-tracking, opens up opportunities to collect cognitive validity evidence for language tests. For example, keystroke logging provides an unobtrusive record of the moment-by-moment composition of a text. There has been a marked increase in language testing research involving process-tracing tools. Some researchers have also called for mixed-method design which combines some of these process-tracing tools with concurrent think-aloud or retrospective stimulated recall. While there has been methodological development in test-taker processes research, one major challenge is to relate measures generated by these tools and verbal reports to models of language use. Due to the range of foci and methods used, there is no clear consensus about how best to analyse process data to reveal test takers' level of proficiency. In addition, the range of terminologies such as process, strategy, subskill and function, to name a few, used in these studies has also added to the difficulty of conceptualisation of the second language

constructs involved and generalisation of the findings (Aryadoust, & Luo, 2022). To synthesize the development in test takers' cognitive processes research in the past two decades, a scoping review is proposed to identify current research trends in terms of language skills involved, contexts of assessment, data collection methods and data analysis frameworks. The review also aims to pinpoint limitations of the current research and to provide an agenda for future studies of test takers' cognitive processes.

References

- Aryadoust, V., & Luo, L. (2022). The typology of second language listening constructs: A systematic review. *Language Testing*. <https://doi.org/10.1177/02655322221126604>
- Field, J (2004) *Psycholinguistics: The Key Concepts*, London: Routledge.
- Stratman, J., & Hamp-Lyons, L. (1994). Reactivity in concurrent think-aloud protocols: Issues for Research. In P. Smagorinsky (Ed.) (pp. 89-112). Sage Publications, Inc.
- Weir, C. (2005). *Language testing and validation: An evidence-based approach*. Palgrave MacMillan.

Inclusion Criteria

Definition

Include: Studies which investigate L2 students' cognitive processes when completing a language test task, e.g. reading, writing, speaking, listening or integrated tasks, through data collection methods such as think-aloud protocols, questionnaire and process-tracking software. Examples of writing processes include planning, formulating and revising. We include "cogniti*" in the search string, for details see below. We do not include "SCF" in the search string as we don't want to exclude studies which use other theoretical frameworks of test validation.

Exclude: L2 students' cognitive processes of performance a language task in non-assessment context; L1 students' cognitive processes

Rationale: The focus of this scoping review is on L2 students' cognitive processes in language testing.

Language

Include: English

Exclude: Languages other than English

Rationale: The primary focus of this review is on the relevant international literature which is published in English.

Type of publication

Include: Journal articles, technical reports books, book chapters, and doctoral theses

Exclude: Master's dissertations

Rationale: Given the scoping and exploratory nature of the review, we intend to survey various types of publications. Master's dissertations are excluded because, unlike doctoral theses, not all can be found in university repositories.

Publication date

Include: Publications between 2005 and 2023

Exclude: Publications published before 2005

Rationale: Weir's (2005) seminal SCF for the first time incorporates the component of cognitive validity into conceptualisation and validation of L2 language assessment. The scoping review aims to establish an understanding of the research on L2 test takers' processes conducted since then.

Information Sources

Given the scoping nature of this review, we aim to focus on different types of research publications including journal articles, technical reports books, book chapters, and doctoral theses. For journal publications, we will search on Directory of Open Access Journals (DOAJ), ERIC, PsycINFO, Scopus,

and Social Sciences Citation Index. As for the other types of publications, they will be retrieved from EBSCO, Electronic Theses Online Service (EThOS) (for UK-based doctoral theses), and ProQuest.

Search Strategy

This scoping review uses the following search string:

("test*" OR "assess*" OR "exam*") AND ("cogniti*" OR "process*" OR "strateg*" OR "skill*")

Data Management

We will review the titles and abstracts on the database websites and those that are deemed relevant, their full-texts will be downloaded and imported to Mendeley for second-level screening.

For identification and first-level screening, potential papers will be identified using the above search string by reviewing the papers' titles and abstracts against the inclusion and exclusion criteria. Papers that are borderline in terms of relevance will still be retained in this stage for a more thorough screening of their full texts. Papers deemed to be of interest to the present review will be downloaded and their PDFs will be imported to Mendeley. Second-level screening will take place using Mendeley by reviewing the full texts of the downloaded PDFs using the stipulated inclusion and exclusion criteria. Following the second-level screening, the excluded papers will be moved to a separate folder on Mendeley. As for data extraction, the completed data extraction forms will be imported to NVivo, a qualitative data analysis software for data synthesis.

Selection Process

Following the latest PRISMA's statement (Moher et al., 2009), the selection process will be conducted comprising three components: identification, screening, and inclusion of extracted publications.

For identification, major electronic databases listed in "Information Sources" will be searched. Studies published after 2005 till July 2023 will be searched using the aforementioned search string (see "Search Strategy") to identify potential papers. For first-level screening, titles and abstracts will be screened against the inclusion and exclusion criteria. After that, second-level screening will take place where full texts of potential papers will be screened using the stipulated inclusion and exclusion criteria. Second-level screening will be conducted by two reviewers, who will screen 20% of the downloaded papers separately using the inclusion and exclusion criteria. The two reviewers will then meet and discuss the outcomes and resolve any discrepancies. To adhere to the reflexivity principle of conducting qualitative synthesis of research (Liu & Chong, 2023), a researcher logbook will be kept to document the process. Afterwards, the remainder of the second-level screening will be performed by one of the reviewers. Only studies that meet the criteria will be included in the present scoping review.

Liu, Q., & Chong, S. W. (2023). Bilingual education in China: A qualitative synthesis of research on models and perceptions. *Applied Linguistics Review*, advanced online publication. <https://doi.org/10.1515/applirev-2022-0194>

Page, M.J., McKenzie, J.E., Bossuyt, P.M. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev* 10, 89 (2021). <https://doi.org/10.1186/s13643-021-01626-4>

Data Collection Process

The present review is a scoping review. A scoping review refers to a comprehensive, systematic

literature review on emerging substantive topics (Tricco et al., 2016). Comparing with other types of research synthesis, for example, systematic literature reviews, scoping reviews usually adopt a more inclusive approach to study selection, and they are more exploratory in nature (Chong & Plonsky, 2023; Chong & Reinders, 2022).

The more encompassing nature of scoping review is reflected in the design of the data extraction form items. To ensure that the data extraction form is validated, an existing data extraction form for conducting a scoping review on a topic in the field of language education (Chong & Reinders, 2022) will be modified.

The first author who has expertise in the subject area and the second author who has expertise in the research method will have shared responsibility in modifying the data extraction form. The data extraction form will be piloted on approximately 20% of the included publications, covering journal articles, books/book chapters, and doctoral theses, by both authors. Any further changes will be made after the piloting phase. They meet regularly during the data collection and selection process to report progress and discuss any issues. This final version of data extraction form will be used to extract information from the remaining publications.

References

Chong, S. W., & Plonsky, L. (2023). A typology of secondary research in applied linguistics. *Applied Linguistics Review*, advanced online publication. <https://doi.org/10.1515/applirev-2022-0189>

Chong, S. W., & Reinders, H. (2022). Autonomy of English language learners: A scoping review of research and practice. *Language Teaching Research*. Advanced online publication. <https://doi.org/10.1177%2F13621688221075812>

Data Items

The following items are based on the validated data extraction form of Chong and Reinders (2022) and are tentative. The items may be revised following the validation process mentioned in "Data Collection Process".

Characteristics of included studies:

- Type of publication
- Year of publication
- Location of authors
- Location of study (only applicable to primary studies)
- Research questions
- Research design (qualitative, quantitative, mixed)

1. What language skills do studies on test takers' cognitive processes focus on?

- language skills (reading, writing, speaking, listening or integrated)

2. What are the assessment contexts of these studies? What are the levels of test takers in these studies?

- type of assessment (formative, summative, standardised, classroom-based)
- proficiency level of test takers

3. What research designs do these studies adopt? What research tools do these studies use to reveal

test takers' processes?

- Research design (qualitative, quantitative, mixed)
- Instruments (keystroke logging, eye-tracking, stimulated recall, verbal report, questionnaire)
- Conceptual framework of processes

4. What information about test takers' cognitive processes do these studies report?

- Measures of process (fluency, pauses)
- Measures of product (linguistic features)
- Test takers' perception

Reference

Chong, S. W., & Reinders, H. (2022). *Autonomy of English language learners: A scoping review of research and practice*. *Language Teaching Research*. Advanced online publication. <https://doi.org/10.1177%2F13621688221075812>

Risk of bias/trustworthiness of individual studies

The risk of bias and trustworthiness of the included studies, when appropriate, will be evaluated using the Mixed Methods Appraisal Tool (Hong et al., 2018).

Reference

Hong, Q. N., Gonzalez-Reyes, A., & Pluye, P. (2018). Improving the usefulness of a tool for appraising the quality of qualitative, quantitative and mixed methods studies, the Mixed Methods Appraisal Tool (MMAT). *Journal of evaluation in clinical practice*, 24(3), 459-467.

Data Synthesis

To synthesise the charted data in the extraction forms, we will employ thematic analysis, which include five stages: "familiarizing yourself with the data", "generating initial codes", "searching for themes", "reviewing themes", "defining and naming themes", and "producing the report" (Nowell et al., 2017, pp. 4-8). This systematic yet flexible methodological framework is deemed suitable for this review, which is qualitative in nature, to maintain trustworthiness in the data analysis process (Lincoln & Guba, 1985).

Operationally, thematic analysis will first be conducted by the all authors on 20% of the included publications independently. The resultant coding scheme will be compared and discussed in a meeting, with an aim to reach a consensus on interpretations. Adhering to the reflexivity principle of conducting research synthesis (Liu & Chong, 2023), meeting notes, which detail discussion process and outcomes as well as the researchers' reflections, will be kept. After the meeting, the remainder of the data extraction forms will be synthesised in like manner by the third author, with regular input from and discussions with the first and second authors.

A scoping review, unlike a meta-analysis and a qualitative evidence synthesis, can focus on both qualitative and quantitative data. To ensure that various research designs and data sources can be adequately represented and synthesised, a narrative approach to reporting the synthesised results

will be adopted (Schwarz et al., 2019). A narrative synthesis enables a fit-for-purpose and eclectic combination of textual and visual representations of data, anchored in the main narrative.

References

- Liu, Q., & Chong, S. W. (2023). Bilingual education in China: A qualitative synthesis of research on models and perceptions. *Applied Linguistics Review*, advanced online publication. <https://doi.org/10.1515/applirev-2022-0194>
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis. *International Journal of Qualitative Methods*, 16(1), 160940691773384. <https://doi.org/10.1177/1609406917733847>
- Lincoln, Y. S. & Guba, E. G. (1985). *Naturalistic Inquiry*. SAGE Publications.
- Schwarz, C. M., Hoffmann, M., Schwarz, P., Kamolz, L. P., Brunner, G., & Sendlhofer, G. (2019). A systematic literature review and narrative synthesis on the risks of medical discharge letters for patients' safety. *BMC Health Services Research*, 19(1). <https://doi.org/10.1186/s12913-019-3989-1>

Meta-biases

Our research process and reporting standard are informed by the PRISMA 2020 Checklist (Page et al., 2021). For specifics, refer to the "Selection Process" section of this protocol.

Reference

- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, n71. <https://doi.org/10.1136/bmj.n71>

Confidence in cumulative evidence

To gauge the confidence in cumulative evidence, the Weight of Evidence Framework will be applied to the appraisal of the completed review (Gough, 2007) to inform the conclusions drawn from the synthesised evidence base. The focus of this appraisal is threefold: on the quality of how the included studies were conducted, how appropriate the research methods of the included studies are, and the extent to which the included studies respond to the review questions.

References

- Gough, D. (2007). Weight of Evidence: a framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education*, 22(2), 213-228. <https://doi.org/10.1080/02671520701296189>

Sources of Funding

N/A

Role of Funders

N/A

Anticipated or actual start date:

2023-10-01

**Anticipated completion date:**

2024-07-31

Other language resources

N/A

Current Status

Ongoing

Details of Published Review**IDESR URL**

<https://idesr.org/article/IDESR000086>