DOI: 10.61782/fa.2023.0457

forum acusticum 2023

EVALUATING THE ECOLOGICAL VALIDITY OF SPATIAL HEARING

Michael A Akeroyd[1,2*] **Louise Hickson**[3] **Inga Holube**[4]
Gitte Keidser[5,6] **Graham Naylor**[1,2] **Karolina Smeds**[7]

[1] Hearing Sciences, School of Medicine, University of Nottingham, Nottingham, U.K.
[2] NIHR Nottingham Biomedical Research Centre, Nottingham, U.K.
[3] School of Health and Rehabilitation Sciences, The University of Queensland, Australia.
[4] Inst. Hearing Technology & Audiol, Jade University of Applied Sciences, Oldenburg, Germany.
[5] Eriksholm Research Centre, Oticon A/S, Snekkersten, Denmark.
[6] Department of Behavioral Sciences and Learning, Linkoping University, Linkoping, Sweden.
[7] ORCA Europe, WS Audiology, Stockholm, Sweden.

ABSTRACT

A working group in the International Collegium of Rehabilitative Audiology is considering the problem of developing a checklist for evaluating the ecological validity (EV) of experiments intended to reflect real-life hearing-related function, activity, or participation with or without hearing devices. We envisage that the checklist would help inform the design of experiments in meeting their goals or in interpreting the meaning or limitations of the results. This paper focuses on an experiment that aimed at investigating spatial hearing and we will show how our putative checklist can be used for that experiment. Keidser et al. (Ear & Hearing, 2020, 41:5S-19S) [1] described five dimensions that are central to consider when designing an experiment. Two of these dimensions, "sources of stimuli" and "environment" can be reproduced in the laboratory to an accuracy imperceptibly different from the real world, given technology trends in HRTF recordings, ambisonics, motion tracking, and computer models. But a perceptually accurate spatial acoustical environment is neither necessary nor sufficient for EV; the other three dimensions ("context", "task", and "individual") also should be considered when assessing the EV of any particular experiment.

[*]***Corresponding author***: mszmaa@nottingham.ac.uk.

Keywords: *ecological validity, spatial hearing, hearing impairment.*

1. INTRODUCTION

Keidser et al [1] defined "ecological validity" as referring to *"the degree to which research findings reflect real-life hearing-related function, activity, or participation"*. They also noted that it is not a "*binary phenomenon that is either present or absent .. but each study represents a certain level of ecological validity"*. This latter statement implies that ecological validity could be quantified or scored, and that experiments can have a greater or lesser extent of it. It is with that purpose, using spatial hearing as an example, that we are presently concerned.

Importantly, this definition does not specify that an accurate reproduction of real-world environments is sufficient or even necessary for a high degree of ecological validity. Today's (and tomorrow's) technology would likely enable an accurate recreation of even the most complex, dynamic spatial situation. For example, with a head tracker, individualized HRTFs, smooth interpolation of moving sound sources, a treadmill to walk on, and plenty of real-time signal processing power, even the situation of walking with a friend along a busy street while listening to them tell a story could be built in the laboratory. But, though this may be a very realistic spatial-acoustic field, the inevitable design choices made when implementing it as an experiment may reduce its likeness to the everyday situation

it is targeting, and potentially invalidate the findings with respect to a research purpose. A second example is the detection of a brief complex tone in an interaurally uncorrelated noise. This is a spatial-hearing experiment that might typically be regarded as unrealistic. But, if one's purpose is to develop a better mobile-phone alerting sound that someone inside a steadily-driven car could easily detect, then arguably it has sufficient ecological validity for its stated purpose.

These two thought experiments highlight that perfect realism in spatial fidelity is not by itself necessary nor sufficient for ecological validity. Instead, what is more interesting is how realistic the experiment needs to be, how realistic it is, and, if there's a gap here, why. Generalizing this, the critical step becomes comparing what everyday situation an experiment is designed to emulate versus what constraints or limitations were needed to make it a viable experiment, all in light of what research purpose it aims to address. We propose that a simple reporting table, giving a clear itemisation of the purpose, everyday situation, implementation, and limitations, could therefore prove useful for gauging ecological validity. Here we report on this work-in-progress..

2. THE EV REPORTING TABLE

We illustrate the idea using an experiment designed by the late Stuart Gatehouse [2]. The purpose was "*to measure the benefit that helping attention (by visual cueing) would have on word identification*" in *"a complex listening situation such as a multiperson conversation"*. Characteristics of this situation assumed to be salient to the purpose included: *"Whoever is talking at any one moment will vary among the people present, and there will often be interjections from others or extraneous sounds of momentary importance. The participants will be receiving sounds whose source, direction, and content can randomly change, often unpredictably. Occasional other sounds will also occur, again often unpredictably, but other times with a visual cue to the interjection [..] There may well be random, unwanted background sounds masking the targets."*. The experiment that was built used a pseudo-continuous stream of sentences presented to a participant (sentences from two standard speech-test databases, spoken by two different people, from any one of 24 loudspeakers in a circular ring around the participant), masked by a continuous diffuse noise, with occasional single words presented (from another standard database), whose timing and direction was sometimes cued by lights. The listener was tasked with responding whenever one of the main sentences was about food or drink and also to identify the occasional extra word and from where it was presented.

Somewhat obviously, from the moment the very first listener started the experiment it became a real scenario. But of course it does not follow that therefore this experiment has perfect ecological validity. We suggest that the definition we use, *"the degree to which research findings reflect real-life hearing-related function, activity, or participation"* [1], implies that it is the translatability of the findings to the real-world function, activity, or participation under investigation that matters. Keidser et al. [1] described five dimensions of independent variables that should be considered when designing or evaluating an experiment in relation to EV, namely "Sources of stimuli", "Environment", "Context of participation", "Task", and "Individual". Within the realm of spatial hearing (weighting primarily on the "Sources of stimuli" and "Environment" dimensions), the experiment gains EV by being designed to include speech sources from any direction that could overlap, with random, unpredictable changes in direction, mostly unpredictable timings, and with some visual cues to where something new was about to happen. But EV related to spatial hearing is limited by the directions being quantized to 15° azimuths on the horizontal plane, the timing of changing of directions being partly predictable, there being a stationary, diffuse noise field, and there not being any reverberation, talker movements, or spatial overlap of sentences. These pros and cons are summarized in Table 1 (which includes a partial itemisation of the Design and Limitations for the other three dimensions of Ecological Validity that need to be considered, namely "Context of Participation", "Task", and "Individual" [1]).

3. DISCUSSION

For ecological validity to be high, the outcomes/findings of the experiment must translate (and be meaningful) to the particular everyday scenario that the experiment was designed to create knowledge for (and relate to real-life hearing-related function, activity, or participation).

We suggest that the author of a paper claiming high ecological validity should discuss to what extent the identified design limitations may affect the ecological validity of the study outcome. This is a purely subjective exercise – at least until such time that we have evidence about the importance (or not) of such limitations.

There is considerable debate about "ecological validity", both as a concept and a term. We use the definition from [1] (but for an alternative definition, see [3]). Some may argue that ecological validity is unreachable; as close as one

Table 1. Suggested EV reporting table. A study of Gatehouse and Akeroyd [2] is used as an example. Two design domains, "Sources of stimuli" and "Environment" have been prioritized in the table for the potential risks.

Purpose	To measure the benefit that helping attention (by visual cueing) would have in word identification		
Everyday situation	A conversation among a group of people		
Design domains	***Implementation***	***Limitations*** *(re. target everyday scenario)*	***Potential risk to EV of outcome***
Sources of stimuli	A. Full sentences, single words B. Audio only	1. Standard speech-test sentences (and words) with uniform syntactic structure. 2. Only 3 voices throughout	1. High 2. Low
Environment	A. Free-field presentation B. Speech from any direction. C. Spatial separation of sound sources. D. Random, unpredictable changes in direction. E. Temporal overlap of targets. F. Visual cue on/off.	1. Directions quantized to 15° azimuths. 2. Stationary, diffuse noise field. 3. No reverberation. 4. No talker movements. 5. No spatial overlap of sentences. 6. Timing of changing of directions partly predictable.	1. Low 2. Medium 3. High 4. Low 5. Low 6. Medium
Context of participation	A. Recruited from institutional pool.	1. No screening of the scenario's relevance for individual participants.	
Task	A. Sustaining, monitoring and switching of attention. B. Identification of content of some sentences, of some words, and of some directions. C. Head movement allowed.	1 Mostly recognition, little understanding of meaning needed. 2. No conversational interaction of the listener with the stimuli.	
Individual	A. Experienced adult hearing-aid users.	1. No normal-hearing reference group. 2. No measures of other relevant variables (e.g. cognition, fluency in test language).	

thinks one has got to perfect ecological validity, then it is as far away as ever. Nevertheless, if we allow that experiments can have greater or lesser ecological validity, then the proposed table could be a potential way of gauging it. The table makes clear what everyday situation an experiment was designed to emulate versus the constraints that were necessary to make it happen, all within view of its stated purpose.

In many ways our proposed table simply collates information that many papers already report, though it is typically spread across the whole paper with the Purpose and Everyday situation in the Introduction, the Implementation in the Method, and Limitations towards the end of the Discussion. Our table collates this into an easy-to-access form. There are other advantages too. Encapsulating ecological validity in a simple reporting tool will encourage experimenters to properly consider any strengths or weaknesses of their experiment design. We hope that these will then simplify the categorization of experiments in future meta-analyses and perhaps inspire experimenters to improve on particular limitations and so increase an experiment's ability to fulfill its stated purpose.

Finally, we note an indirect feature. Some experimenters (and some future readers of their papers) may place more weight on one design domain or instead may be more concerned by ten limitations. Ecological validity remains

subjective. The table therefore represents, to quote [1], that "*... a study's level of ecological validity is ultimately based on subjective judgment, so the concept of ecological validity cannot be used to provide comprehensive objective criteria for experimental designs.*"

4. REFERENCES

[1] G. Keidser, G. Naylor, D.S. Brungart, A. Caduff, J. Campos, S. Carlile, M.G. Carpenter, G. Grimm, V. Hohmann, I. Holube, and S. Launer. "The quest for ecological validity in hearing science: What it is, why it matters, and how to advance it," *Ear and hearing*, 41(Suppl 1):5S-19S, 2020.

[2] S. Gatehouse and M.A. Akeroyd. "The effects of cueing temporal and spatial attention on word recognition in a complex listening task in hearing-impaired listeners," *Trends in amplification,* 12(2):145-161, 2008

[3] T. Beechey. "Ecological validity, external validity, and mundane realism in hearing science," *Ear and Hearing,* 43(5):1395-1401, 2022.