

EVALUATING ECOLOGICAL VALIDITY OF OUTCOMES (EVO)

Karolina Smeds^{1,2}, Michael Akeroyd², Louise Hickson³, Inga Holube⁴, Gitte Keidser⁵, Graham Naylor²

1 WS Audiology, ORCA Europe; 2 University of Nottingham, Hearing Sciences; 3 University of Queensland, Health and Behavioural Sciences; 4 Jade University of Applied Sciences, Hearing Technology and Audiology; 5 Oticon, Eriksholm Research Centre

Background

Keidser et al. (2020) defined ecological validity as referring to “the degree to which research findings reflect real-life hearing-related function, activity, or participation”. To emphasize that we are focused on research findings, we use the nomenclature **Ecological Validity of Outcomes (EVO)**.

In the 2020 paper, it was noted that EVO is not a “binary phenomenon that is either present or absent ... but each study represents a certain level of ecological validity”, which implies that EVO can be quantified or scored. A working group within the International Collegium of Rehabilitative Audiology (ICRA) is now considering how to evaluate EVO.

The main problem is that we usually don’t have valid real-life benchmarks to evaluate EVO against. Instead, our approach builds on a comparison of the everyday situations an experiment is designed to emulate and the compromises that were made to make it a viable experiment.

References

Keidser, G., Naylor, G., Brungart, D. S., Caduff, A., Campos, J., Carlile, S., Carpenter, M. G., Grimm, G., Hohmann, V., Holube, I., Launer, S., Lunner, T., Mehra, R., Rapport, F., Slaney, M., & Smeds, K. (2020). The quest for ecological validity in hearing science: What it is, why it matters, and how to advance it. *Ear Hear, 41 Suppl 1, 5S-19S*.

Smeds, K., Larsson, J., Dahlquist, M., Wolters, F., & Herrlin, P. (2021). Live Evaluation of Auditory Preference, a Laboratory Test for Evaluating Auditory Preference. *J Am Acad Audiol, 32(8), 487-500*.

Methodological dimensions

Keidser et al. (2020) listed commonly used independent variables. These are here used to create an EVO checklist.

Methodological dimensions	Independent variables (abbreviated version)
Sources of stimuli	Characteristics of stimulus sources, e.g., speech/other, diversity, familiarity, ...
	For multimodal stimuli, which modalities are subjected to controlled manipulations...
	...
Environment	Acoustic field, e.g., levels, SNRs, spatial fidelity... Interaction of environment and hearing devices... Incorporation of dynamic aspects ...
	...
Context of participation	Participant preparation, e.g., instructions ... Motivation to take part, e.g., reimbursement...
	...
Task	Nature of task, e.g., speech communication vs. environmental monitoring/detection Nature of task if speech, e.g., repeat, recall, comprehend Complexity, e.g., single vs multiple tasks Predictability e.g., limited response options...
	...
Individual	Personality, e.g., open, agreeable, extrovert... Hearing health, e.g., type, degree and configuration... Sensory, cognitive, motor abilities ... Competency in task language, e.g., native... Disease burden, e.g., frailty, multimorbidity
	...

Conclusions

Potential use of the EVO checklist:

- Help with the interpretation of reported research findings
- Inspire research design improvements to increase experiments’ ability to fulfill stated purposes
- Simplify categorization of experiments in future meta-analyses

Suggestion:

- If claiming high EVO, use the checklist to discuss how design limitations may affect the EVO.

Suggested reporting tool for EVO

STATED PURPOSE	
OUTCOME MEASURE(S)	
EVERYDAY SITUATION(S)	
STIMULI	
Implementation	
Limitations relevant to EVO	Potential risk to EVO
ENVIRONMENT	
Implementation	
Limitations relevant to EVO	Potential risk to EVO
CONTEXT OF PARTICIPATION	
Implementation	
Limitations relevant to EVO	Potential risk to EVO
TASK	
Implementation	
Limitations relevant to EVO	Potential risk to EVO
INDIVIDUAL	
Implementation	
Limitations relevant to EVO	Potential risk to EVO

Example (Smeds et al. 2021)

STATED PURPOSE	
OUTCOME MEASURE(S)	
EVERYDAY SITUATION(S)	
STIMULI	
Implementation	
Limitations relevant to EVO	Potential risk to EVO
To evaluate a new laboratory test, developed to investigate hearing-instrument performance in test scenarios that represent everyday listening situations. (Only lab part of paper included in this example.)	
Preference (and strength of preference) for one of two hearing-aid programs, assessed by in-the-moment paired comparisons.	
Six mandatory scenarios : 1. Conversation two people in “quiet”, 2. Conversation two people in car noise. 3. Conversation three people in restaurant noise, 4. Focused listening to TV, 5. Focused listening to music, 6. Passive listening. 7. Plus individual scenarios.	
Implementation	
A. Talking to one test leader (♀) in “quiet” [noise @ ~40 dB(A)*]	
B. Talking to one test leader (♀) in car noise [~65 dB(A)*]	
C. Talking to two test leaders (♀ or ♂) in canteen noise [~67 dB(A)*]	
D. TV (prerecorded nature documentary), self-selected SPL*	
E. Music (prerecorded vocal jazz), self-selected SPL*	
F. Paper rustling in “quiet”, ~40 dB(A)*	
Limitations relevant to EVO	Risk
1. Limited variety of talkers and only one sample of TV program and Music	1. Med
2. Unfamiliar talker(s)	2. Low
3. Unknown familiarity of TV program and Music	3. Med

* Information not provided in the paper

Example continued

ENVIRONMENT	
Implementation	
A. Furnished and ventilated rather small office room	
B. Realistic presentation levels (see STIMULI) and SNRs	
C. Basic loudspeaker setup with two loudspeakers for delivering noise at realistic sound pressure levels	
Limitations relevant to EVO	
1. Only 2 loudspeakers for the background noise	1. Med
2. Just one room used, with its specific room acoustics, used for all mandatory scenarios.	2. Low
3. Car background noise played back in a room.	3. High
CONTEXT OF PARTICIPATION	
Implementation	
A. Recruitment from participant database*. Participants paid.	
B. Not familiar with the LEAP test.	
C. Mandatory scenarios selected based on commonness but were used for everyone.	
Limitations relevant to EVO	
1. Volunteers motivated by research interest or willingness to contribute to solutions for people with hearing loss. Potentially also motivated by payment for participation.	1. Low
2. No check if the mandatory scenarios were relevant to participants	2. Med
3. Data collection after short period of HA use	3. Low
TASK	
Implementation	
A. Engage in different activities (conversation, focused listening or sorting exercise).	
B. Compare HA programs, select preferred program, and rate degree of preference.	
Limitations relevant to EVO	
1. Systematic comparisons of HA programs to find “the best” and reporting preference ratings are not everyday tasks.	1. Low
2. Conversations with unknown persons.	2. Low
INDIVIDUAL	
Implementation	
A. 19 participants	
B. 42-90 years old	
C. Symmetric sensorineural hearing loss	
D. Experienced hearing aid users	
E. No previous experience with LEAP	
Limitations relevant to EVO	
1. Unknown demographics, health, cognition etc. of participants	1. Low
2. Only symmetric sensorineural hearing loss	2. Low
3. Only experienced hearing aid users	3. Low