

# Ecological Validity and “Ecological Validity”

**John F. Kihlstrom**

Department of Psychology, University of California, Berkeley

## Abstract

Egon Brunswik coined the term *ecological validity* to refer to the correlation between perceptual cues and the states and traits of a stimulus. Martin Orne adapted the term to refer to the generalization of experimental findings to the real world outside the laboratory. Both are legitimate uses of the term because the ecological validity of the cues in an experiment determines the ecological validity of the experiment itself.

## Keywords

demand characteristics, ecological validity, behavioral methodology, neuroscience methodology, scientific methodology

Recently in this journal, Shamay-Tsoory and Mendelsohn (2019) called attention to the “limited ecological validity that characterizes the bulk of the paradigms and settings” (p. 841) in cognitive neuroscience. Holleman et al. (2021), while supporting their call for more lifelike stimuli and tasks in brain-imaging studies, criticized them for misusing the concept of *ecological validity*, a term coined by Egon Brunswik (e.g., 1955, 1956). They pointed out, correctly, that Brunswik referred to the ecological validity of *cues* (i.e., the extent to which information available in the environment provides valid information about a distal stimulus) rather than the ecological validity of *experiments* (i.e., the extent to which experimental findings can generalize to the “real world” situation that a researcher wishes to understand; see also Holleman et al., 2020). Hammond (1998), himself a former student of Brunswik, likewise referred to the “thoroughly corrupted . . . flagrant misuse” (para. 2) of *ecological validity* in an unpublished article circulated widely in *samizdat* (see also Dunlosky et al., 2009; Hammond & Stewart, 2001; Schmuckler, 2001).

The alleged misuse of the term *ecological validity* has its origins in an article by Martin T. Orne (1962), originally delivered at a symposium (“On the Social Psychology of the Psychological Experiment”) at the 1961 convention of the American Psychological Association. In that article, Orne proposed that the essence of the experimental method is to reproduce elements

of complex natural situations in the laboratory to determine cause-and-effect relations. But he also argued that artificial laboratory situations have features of their own that may differ from those found in the natural environment and that threaten generalization from the laboratory back to the real world. Orne also argued that psychologists err in treating human subjects as passive responders to experimental stimuli. Instead, human subjects are sentient beings who are trying to understand the situation they are in, whether in the laboratory or the real world. To this end, they pick up on the *demand characteristics* of the experimental situation, which he defined as “the totality of cues which convey an experimental hypothesis to the subject” (Orne, 1962, p. 779). The term itself came from *Aufforderungscharaktere* (Koffka, 1935, p. 345; Lewin, 1935, p. 51)—the same word from which Gibson (1966) derived his concept of *affordances*. The subject’s understanding of the experiment might differ from that intended by the experimenter, in which case, a laboratory experiment would lack ecological validity, which Orne (1962) defined as “appropriate generalization from the laboratory to nonexperimental situations” (p. 776; see also Orne, 1969, 1970a, 1970b, 1973, 1981).

---

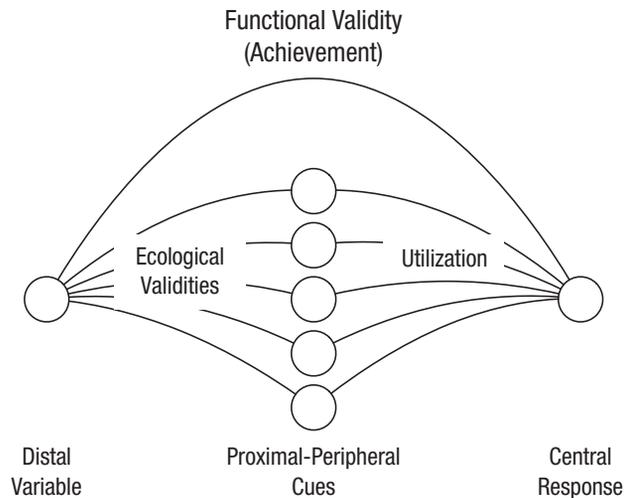
## Corresponding Author:

John F. Kihlstrom, Department of Psychology, University of California, Berkeley  
E-mail: jfkihlstrom@berkeley.edu

Trained as a social psychologist as well as a psychiatrist, Orne was especially interested in the ecological validity of experiments in social psychology—particularly those involving deception (Orne & Holland, 1968; for a response, see Milgram, 1972; see also Mohl, 2017). But ecological validity is not just for social psychologists anymore. Almost 100 years after Ebbinghaus (1885/1964), for example, Neisser (1978) notoriously suggested that studies employing nonsense syllables and other verbal-learning paradigms lacked ecological validity and had taught us virtually nothing about memory in real life (for responses, see Banaji & Crowder, 1989; Dunlosky et al., 2009; Kihlstrom, 1996). Clinical neuropsychologists have long worried that elderly subjects' performance on laboratory tests does not correlate well with subjective complaints of memory function (e.g., Perlmutter, 1978; Sbordone, 2008). The real-life significance of performance on laboratory tasks has also been of concern to developmental psychologists (e.g., Bronfenbrenner, 1977)—especially those concerned with developmental disabilities (e.g., Brooks & Baumeister, 1977; Ford & Gaylord-Ross, 1991). Forensic psychologists, not to mention judges and juries, are often concerned with the generalization of laboratory findings to the courtroom (e.g., Goodman & Hahn, 1987; Rose, 2017). As the popularity of the concept grew, the “ecological validity” of an experiment came to be equated with its resemblance to real life. This was never Orne's intention.

Orne's article, with its emphasis on the individual subject's perception of the situation, foreshadowed the cognitive revolution in social psychology. But it also precipitated social psychology's first great crisis: a debate over whether social psychologists should do laboratory experiments at all or confine themselves instead to descriptive and correlational studies situated in the real world outside the laboratory (e.g., Gergen, 1973; Harre & Secord, 1972). Many experimental social psychologists cried foul, fearing that such criticisms would discredit the entire discipline (e.g., Berkowitz & Donnerstein, 1982; Sharpe & Whelton, 2016). This turn of events distressed Orne greatly: He never intended his ideas about ecological validity and demand characteristics to be used in a reflexive, nihilistic fashion. As he noted in a later commentary,

Unfortunately, this work has also been used as the basis for criticizing all experimental research in psychology and as an argument to abandon such efforts. I cannot share this view since I would not have been concerned about analyzing the nature of the psychological experiment if it were not an essential tool to elucidate psychological processes. (Orne, 1979, p. 56)



**Fig. 1.** Schematic diagram of the lens model, after Brunswik (1955, Fig. 8). Reproduced from Holleman et al. (2021, Fig. 1).

Far from being a misuse of the term, Orne's notion of ecological validity is thoroughly grounded in Brunswik's (1955, 1956) “lens model” for perception. Brunswik argued that the individual perceives the world through a “lens” of imperfect cues, as depicted in Figure 1. The observer's goal is to form an accurate perception of a distal stimulus, in which accuracy is defined in terms of the match between the actual nature of the stimulus (size, distance, form, etc.) and the subject's perception of it. Brunswik's lens model was subsequently generalized to the realm of judgment and decision-making, which is where it has been most frequently applied (e.g., Hastie & Dawes, 2010). There is an object about which a judgment is to be made: The validity of the judgment depends on the match between the features displayed by the target of the judgment and those considered by the judge. Accuracy, whether in perception or judgment, reflects the degree of correspondence between the object and the inference.

On the stimulus side of the lens model, the distal stimulus presents various proximal cues that reflect the state of the target. These cues vary in terms of their ecological validity, or the strength of their association with the stimulus. The cues themselves have objective values, which can be thought of as the correlation between each cue and the target's state. On the subject side of the model is the subset of proximal cues actually available to the perceiver. These cues have a subjective value in terms of the subject's beliefs about their relevance to the judgment he or she is trying to make. The cues vary in their cue utilization, or the extent to which they play a role in the subject's judgment. Cue utilization can be thought of as the correlation between the cue and the judgment.

Brunswik's lens, then, serves as a kind of filter between the cues presented by the target and those used by the subject. If the cues used by the subject match the ecologically valid cues presented by the target, then the percept will be accurate, or the judgment will be valid. But if there is a discrepancy between the cues presented by the target and the way they are processed by the subject, then the percept will be inaccurate and the judgment invalid. For example, ecologically valid cues may be unavailable for some reason. Other opportunities for error arise from the subject's judgment policy. For example, subjects may fail to use ecologically valid cues that are available to them, use cues that are not ecologically valid, or overvalue or undervalue cues relative to their actual ecological validity.

Traditionally, Brunswik's lens model applies to one target and one perceiver (or judge). If subjects attend to ecologically valid cues, then their percepts and judgments will be accurate. When there is a single judge and two or more targets, the judgments of the various targets will be similar provided that the cues provided by the targets, and used by the subject, are also similar. When there are two or more judges of a single target, they may make different judgments depending on the cues available to them or their judgment policies concerning cue utilization—which cues they pay attention to, how they weigh them, and the like. Orne understood that experimenters and subjects may have different perceptions of the experimental situation; to understand the subject's behavior in an experiment, the experimenter must understand the subject's point of view.

This last point brings me to the matter of the ecological validity of experiments—that is, whether one can generalize from the results of an experiment to the real-world situation that inspired it. In this case, there are two different targets, the real-world situation and the laboratory analogue. To the extent that the experiment and the real-world situation present the same ecologically valid cues to the subject, one can say that the experiment itself is ecologically valid. If the cue information presented by the experimental situation differs significantly from that presented by the real-world situation, however, the cues available in the experimental situation will lack ecological validity with respect to the real-world situation. Whatever behavior occurs in the laboratory will not generalize to the real-world situation that the laboratory situation is intended to model. Even an experiment containing very “real-life complex, dynamic, naturalistic stimuli” (Shamay-Tsoory & Mendelsohn, 2019, p. 851) can lack ecological validity if the laboratory environment contains cues that are not representative of the real-world situation that inspired the experiment.

This is the case in some visual illusions. In the Ames room (Ittelson, 1952), for example, the experimenter has altered the geometry of the room—the lengths of walls, the shape of the windows, the slant of the ceiling—to make it appear to be rectangular when in fact it is trapezoidal and objects in it appear as different in size when in fact they are identical (Epstein, 1962). The visual cues available to and used by observers are ecologically valid for a normal rectangular room but are completely misleading when it comes to perceiving the Ames room as it really is. This is also the case in some social-psychological experiments involving deception. In Milgram's (1963) experiments on obedience to authority, there are cues that suggest that the experimental situation is not what it appears to be. For example, there is nothing the “teacher” does in that experiment that the experimenter could not do, and the fact that the experimenter's attention is focused on the teacher rather than the “learner” clearly communicates that the teacher, not the learner, is the actual subject of the experiment.

Such highly informative demand characteristics, cues that are critical to accurate perception of the laboratory situation, are absent in the real world. Put another way, the ecology of the experiment does not match the ecology of the real world. Because the cues unique to the laboratory situation lack ecological validity with respect to the real-world situation, the experiment itself lacks ecological validity. This is not just a matter of ecological generality—of whether the conditions in an experiment constitute an adequate sample of conditions in the real world. To the extent that the experimental situation contains cues that simply would not be found in the real world, the experiment itself lacks ecological validity.

Ecological validity, in Orne's sense, is sometimes conflated with other aspects of validity. Ecological validity is an aspect of external validity, defined as the ability to generalize experimental results to different populations, situations, and variables (Campbell, 1957). Will scores on a personality inventory predict how individuals will actually behave in some specific situation? Will the performance of American college students in an experiment generalize to other populations? But ecological validity in Orne's sense is a narrower concept, having to do with whether an experiment contains cues that preclude generalization from the laboratory to real life.

This kind of generalizability is not guaranteed by mundane realism, defined as the extent to which events in an experiment are likely to occur in the real world (Aronson & Carlsmith, 1968). An experiment can employ extremely lifelike stimulus materials in an extremely lifelike setting, collecting actual behavioral data—for example, four individuals engaged in a conversation in

a park while wearing wireless EEG caps, as in Shamay-Tsoory and Mendelsohn's (2019) Figure 1h. But the experiment would still lack ecological validity if it also contained demand characteristics that are not present in the nonexperimental situation that it is intended to represent—cues that indicate that the experimental setup is not what it seems to be, what the experiment is really all about, what the experimenter's hypothesis is, and how the subject is expected to behave. These cues, part and parcel of the experimental setting, are not present in the real world outside the laboratory: They might, for example, lead one of the subjects to suspect that the other three participants are confederates of the experimenter. An experiment lacks ecological validity, in Orne's sense, when the cues presented by the experimental situation lack ecological validity with respect to the real-world situation that the experiment is intended to model.

In introducing his ideas about ecological validity, Orne (1962) cited an early article by Brunswik (1947/1949) on experimental design. In it, Brunswik criticized the systematic designs favored by most experimental psychologists of his day (and today) involving only a very limited and somewhat arbitrary range of independent variables—for example, an experiment on size constancy with a “Noah's ark” design consisting of two levels of size and two levels of distance. Instead, Brunswik argued for a representative design entailing a much broader, even random sampling of stimulus conditions, which he hoped would enhance the ecological generality of experimental results. Illustrating his approach, Brunswik cited an earlier experiment on size constancy in which he obtained size and distance estimates in 180 different, arbitrarily selected, indoor and outdoor nonlaboratory situations (Brunswik, 1944).

For Brunswik, ecological generality was a quantitative matter: An experiment should include enough different stimulus conditions to ensure that it adequately samples the stimulus conditions found in the world outside the laboratory. In the experiment on size constancy just described, Brunswik sought to sample a wide range of stimulus conditions to ensure that both the cues he sampled and their intercorrelations were representative of those found in the real world. In his view, an experiment whose cues possess low ecological validity could still have high ecological generality provided that low ecological validity was also characteristic of the cues available in the natural environment. Brunswik thought this might be especially true in certain cases of social perception, such as physiognomy or judgments of personality from photographs.

Ecological validity, in Orne's revisionist sense, is a qualitative matter: Experiments lack ecological validity when they contain cues (demand characteristics)

unique to the experimental setting that may threaten the ecological validity of the experiment itself. This is because those demand characteristics that provide so much information about the nature of the experiment have no counterpart in the real world that the experiment is intended to model. With respect to the laboratory, they have a high degree of ecological validity, but they lack ecological validity with respect to the real world.

Orne did not limit his concern to experiments involving deception. He invented the “real-simulator” paradigm in an attempt to disentangle those phenomena that were intrinsic to hypnosis from those that were artifacts of the demand characteristics of the hypnotic situation (Orne, 1959). Using this paradigm, he was able to show that the ostensible power of hypnosis to compel antisocial and self-injurious behavior was a product of demand characteristics unique to the experimental situation such that the behavior of subjects in the laboratory would not generalize to the real world (Orne & Evans, 1965). Beyond hypnosis, he showed that many of the ostensible effects of sensory deprivation were produced by the demand characteristics of the testing situation rather than deprivation per se (Orne & Scheibe, 1964). Far from using ecological validity as a cudgel to undermine experimental research, he showed how researchers could use information about demand characteristics to improve their experiments, rendering them more ecologically valid (Evans & Orne, 1971; Orne & Evans, 1966). In these and other studies, Orne evinced no concern with whether his stimulus materials or tasks were lifelike. He was interested only in whether his experiments contained cues unique to the experimental situation that would compromise his ability to generalize from the laboratory to real life.

All experiments have demand characteristics: They are the price one pays for doing behavioral research with human subjects (real-world situations have demand characteristics, too; see Kihlstrom, 2002). The only question is whether, and to what extent, those cues that are unique to the experimental situation—characteristic of the experiment but not of the real world—render it ecologically invalid as a representation of the real-world situation one is trying to understand.

Language evolves, and the technical vocabulary of psychology is no exception. The term *ecological validity* has now taken on three different meanings: Brunswik's original concept, referring to cues; Orne's revisionist construal, referring to experiments; and the more commonplace equation with mundane realism. This last sense has no warrant in the work of either Brunswik or Orne: Even experiments employing extremely lifelike situations and tasks can lack ecological validity, depending on their demand characteristics. Orne admittedly employed

*ecological validity* in a somewhat different sense than Brunswik, but his use of *ecological validity* was faithful to Brunswik's because the ecological validity of the cues in the experimental setting determines the ecological validity of experiment itself.

### Transparency

Action Editor: Laura A. King

Editor: Laura A. King

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

### Acknowledgments

I thank Reid Hastie for introducing me to Brunswik's lens model; he and two anonymous reviewers provided very helpful comments on an earlier version of this article.

### References

- Aronson, E., & Carlsmith, J. M. (1968). Experimentation in social psychology. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (2nd ed., Vol. 2, pp. 1–79). Addison-Wesley.
- Banaji, M., & Crowder, R. G. (1989). The bankruptcy of everyday memory. *American Psychologist*, *44*, 1185–1193. <https://doi.org/10.1037/0003-066X.44.9.1185>
- Berkowitz, L., & Donnerstein, E. (1982). External validity is more than skin deep: Some answers to criticism of laboratory experiments. *American Psychologist*, *37*, 245–257. <https://doi.org/10.1037/0003-066X.37.3.245>
- Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, *32*, 513–531. <https://doi.org/10.1037/0003-066X.32.7.513>
- Brooks, P. H., & Baumeister, A. A. (1977). A plea for consideration of ecological validity in the experimental psychology of mental retardation. *American Journal of Mental Deficiency*, *81*, 406–416.
- Brunswik, E. (1944). Distal focussing of perception: Size-constancy in a representative sample of situations. *Psychological Monographs*, *56*, i–49. <https://doi.org/10.1037/h0093505>
- Brunswik, E. (1949). Systematic and representative design of psychological experiments. With results in physical and social perception [University of California Syllabus Series #304]. In J. Neyman (Ed.), *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability* (pp. 143–202). University of California Press. (Original work published 1947)
- Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, *62*, 193–217. <https://doi.org/10.1037/h0047470>
- Brunswik, E. (1956). *Perception and the representative design of psychological experiments* (2nd ed.). University of California Press.
- Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin*, *54*, 297–312. <https://doi.org/10.1037/h0040950>
- Dunlosky, J., Bottiroli, S., & Hartwig, M. (2009). Sins committed in the name of ecological validity: A call for representative design in education research. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of meta-cognition in education* (pp. 430–440). Routledge.
- Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology*. Dover. (Original work published 1885)
- Epstein, W. (1962). A test of two interpretations of the apparent size effects in a distorted room. *Journal of Experimental Psychology*, *63*, 124–128. <https://doi.org/10.1037/h0040345>
- Evans, F. J., & Orne, M. T. (1971). The disappearing hypnotist: The use of simulating subjects to evaluate how subjects perceive experimental procedures. *IJCEH*, *19*, 277–296. <https://doi.org/10.1080/00207147108407173>
- Ford, J., & Gaylord-Ross, R. (1991). Ecological validity revisited: A 10-year comparison of two journals. *American Journal on Mental Retardation*, *96*, 95–98. <https://doi.org/10.1177/154079699101600204>
- Gergen, K. J. (1973). Social psychology as history. *Journal of Personality & Social Psychology*, *26*, 309–320. <https://doi.org/10.1037/h0034436>
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Houghton Mifflin.
- Goodman, G. S., & Hahn, A. (1987). Evaluating eyewitness testimony. In I. B. Weiner & A. K. Hess (Eds.), *Handbook of forensic psychology* (3rd ed., pp. 258–292). Wiley.
- Hammond, K. R. (1998). *Ecological validity: Then and now*. <https://www.brunswik.org/notes/essay2.html>
- Hammond, K. R., & Stewart, T. R. (Eds.). (2001). *The essential Brunswik: Beginnings, explications, applications*. Oxford University Press.
- Harre, R., & Secord, P. F. (1972). *The explanation of social behaviour*. Blackwell.
- Hastie, R., & Dawes, R. M. (2010). *Rational choice in an uncertain world: The psychology of judgment and decision making* (2nd ed.). SAGE.
- Holleman, G. A., Hooge, I. T. C., Kemner, C., & Hessels, R. S. (2020). The 'Real-World Approach' and its problems: A critique of the term ecological validity. *Frontiers in Psychology*, *11*, Article 721. <https://doi.org/10.3389/fpsyg.2020.00721>
- Holleman, G. A., Hooge, I. T. C., Kemner, C., & Hessels, R. S. (2021). The reality of "real-life" neuroscience: A commentary on Shamay-Tsoory and Mendelsohn (2019). *Perspectives on Psychological Science*, *16*(2), 461–465. <https://doi.org/10.1177/1745691620917354>
- Ittelson, W. H. (1952). *The Ames demonstrations in perception*. Princeton University Press.
- Kihlstrom, J. F. (1996). Memory research: The convergence of theory and practice. In D. Hermann, M. Johnson, C. McEvoy, C. Hertzog, & P. Hertel (Eds.), *Basic and applied memory: Theory in context* (Vol. 1, pp. 5–25). Erlbaum.
- Kihlstrom, J. F. (2002). Demand characteristics in the laboratory and the clinic: Conversations and collaborations with

- subjects and patients. *Prevention & Treatment [Special Issue Honoring Martin T. Orne]*, 5, Article 36c. <https://doi.org/10.1037/1522-3736.5.1.536c>
- Koffka, K. (1935). *Principles of gestalt psychology*. Lund Humphries.
- Lewin, K. (1935). *A dynamic theory of personality*. McGraw-Hill.
- Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal & Social Psychology*, 67, 371–378. <https://doi.org/10.1037/h0040525>
- Milgram, S. (1972). Interpreting obedience: Error and evidence. A reply to Orne and Holland. In A. G. Miller (Ed.), *The social psychology of psychological research*. Free Press.
- Mohl, J. (Producer). (2017, June 30). *Debate between Stanley Milgram and Martin T. Orne (University of Pennsylvania, February 16, 1969)* [Video]. YouTube. <https://youtu.be/Sr5dkBauVgg>
- Neisser, U. (1978). Memory: What are the important questions? In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 3–24). Academic Press.
- Orne, M. T. (1959). The nature of hypnosis: Artifact and essence. *Journal of Abnormal and Social Psychology*, 58, 277–299. <https://doi.org/10.1037/h0046128>
- Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 17, 776–783. <https://doi.org/10.1037/h0043424>
- Orne, M. T. (1969). Demand characteristics and the concept of quasi-controls. In R. Rosenthal & R. Rosnow (Eds.), *Artifact in behavioral research* (pp. 143–179). Academic Press.
- Orne, M. T. (1970a). From the subject's point of view, when is behavior private and when is it public: Problems of inference. *Journal of Consulting and Clinical Psychology*, 35, 143–147. <https://doi.org/10.1037/h0030088>
- Orne, M. T. (1970b). Hypnosis, motivation, and the ecological validity of the psychological experiment. In W. J. Arnold & M. M. Page (Eds.), *Nebraska symposium on motivation* (pp. 187–265). University of Nebraska Press.
- Orne, M. T. (1973). Communication by the total experimental situation: Why it is important, how it is evaluated, and its significance for the ecological validity of findings. In P. Pliner, L. Krames, & T. Alloway (Eds.), *Communication and affect* (pp. 157–191). Academic Press.
- Orne, M. T. (1979). Commentary on "Citation Classic - On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications." *Current Contents/Social & Behavioral Sciences*, 13, 56. <https://garfield.library.upenn.edu/classics1979/classics1979.html>
- Orne, M. T. (1981). The significance of unwitting cues for experimental outcomes: Toward a pragmatic approach. *Annals of the New York Academy of Sciences*, 364, 152–159.
- Orne, M. T., & Evans, F. J. (1965). Social control in the psychological experiment: Antisocial behavior and hypnosis. *Journal of Personality and Social Psychology*, 1, 189–200. <https://doi.org/10.1037/h0021933>
- Orne, M. T., & Evans, F. J. (1966). Inadvertent termination of hypnosis with hypnotized and simulating subjects. *International Journal of Clinical and Experimental Hypnosis*, 14, Article 71078. <https://doi.org/10.1080/00207146608415895>
- Orne, M. T., & Holland, C. H. (1968). On the ecological validity of laboratory deceptions. *International Journal of Psychiatry*, 6, 282–293.
- Orne, M. T., & Scheibe, K. E. (1964). The contribution of nondeprivation factors in the production of sensory deprivation effects: The psychology of the "panic button." *Journal of Abnormal and Social Psychology*, 68, 3–12. <https://doi.org/10.1037/h0048803>
- Perlmutter, M. (1978). What is memory aging the aging of? *Developmental Psychology*, 14, 330–345. <https://doi.org/10.1037/0012-1649.14.4.330>
- Rose, M. R. (2017). How typical is Lockhart v. McCree? Ecological validity concerns in court opinions. In M. B. Kovera (Ed.), *The psychology of juries* (pp. 227–253). American Psychological Association.
- Sbordone, R. J. (2008). Ecological validity of neuropsychological testing: Critical issues. In A. M. Horton & D. Wedding (Eds.), *The neuropsychology handbook* (3rd ed., pp. 367–394). Springer.
- Schmuckler, M. A. (2001). What is ecological validity? A dimensional analysis. *Infancy*, 2, 419–436. [https://doi.org/10.1207/S15327078IN0204\\_0](https://doi.org/10.1207/S15327078IN0204_0)
- Shamay-Tsoory, S. G., & Mendelsohn, A. (2019). Real-life neuroscience: An ecological approach to brain and behavior research. *Perspectives on Psychological Science*, 14, 841–859. <https://doi.org/10.1177/1745691619856350>
- Sharpe, D., & Whelton, W. J. (2016). Frightened by an old scarecrow: The remarkable resilience of demand characteristics. *Review of General Psychology*, 20, 349–368. <https://doi.org/10.1037/gpr0000087>