Journal of Expertise
2023. Vol. 6(3)
ISSN 2573-2773

Identifying Experts for the Design of Human-Centered Systems: The Pentapod Principle

Robert R. Hoffman
Institute for Human and Machine Cognition, USA

Correspondence: Robert Hoffman, rhoffman@ihmc.us

Abstract

The identification of experts is crucial in many research projects and domain application areas. However, research reports often assert that the research participants were experts, when in fact the participants were graduate students, or individuals having only a few years of professional experience. This essay briefly discusses the conceptual definition of expertise, and questions the tendency of researchers to bifurcate humanity into novices versus experts. The essay then addresses the matter of how to identify experts, offering a method that is more robust and scientifically grounded than the common, and questionable reliance on the so-called “10-year” or "10,000 hours" rules for deciding who is, and who is not an expert. The approach presented here is based on five distinct classes of methods. The Pentapod Principle asserts that rigorous proficiency scaling should rely on methods from at least three of the classes. This approach should be useful in any investigation that intends to study experts and present conclusions about expertise.

Keywords

defining expertise, measurement, proficiency scaling, methodology, experts vs. novices

Introduction

The identification of experts is crucial in many research projects and domain application areas (e.g., Brooks, 2019). In support of such activities as work design or the development of intelligent systems, how many experts are sufficient for knowledge elicitation or cognitive task analysis? This question presupposes that, in the first place, one has a robust method for determining who qualifies as a genuine expert (Crispen & Hoffman, 2016). In the various literatures in which the concept of expertise is referenced, including human factors psychology and cognitive systems engineering, we still see quite often that the method often used to determine whether an individual is an expert is the so-called 10-year rule of thumb.

The rule originated in an examination of the careers of famous musicians by John R. Hayes (1985), who found that masterful works were created only after about 10 years of intensive effort. The rule also originated in the work of Herbert Simon and Kevin Gilmartin (1973) on the development of mastery at chess. Since those studies, a number of technical writings in the cognitive and computational sciences, and articles in the popular press, have not only mythologized the rule but have laid claim to it (e.g., Gladwell, 2008). Thus, it has become possible for researchers—even researchers who should know better—to assert that their research participants were experts simply because they had been doing their job for at least 10 years (sometimes fewer) or because they knew more about a domain than other people (e.g., Fischer & Keil, 2016). This crops up in studies in which “experts” conduct usability analysis, studies in which “experts” are the participants in human factors experiments on human-machine performance, and studies in which “experts” are the participants in cognitive task analysis or knowledge elicitation.

The 10-year rule is questionable on a number of grounds (see Baer, 2014; Miller, 2023; Stillman, 2016). For example, the 10-year figure represents an average (e.g., over a number of studied musicians). There are measurable individual differences in terms of how much, and what kinds of practice are required for individuals to achieve expertise (Macnamara et al., 2014; Hambrick et al., 2014). If the 10,000 hours rule is to have value, the question is not whether an individual has conducted the task for 10,000 hours, but at what point in their career did they reach an expertise benchmark?

It has been claimed that 10 years of practice is an absolute minimum for the achievement of expertise (Ericsson et al., 1993). But the 10-year criterion is insufficient. The notion of "deliberate practice," which has been seen as foundational to the 10-year rule (Ericsson & Harwell, 2019) misses the fact that many individuals do achieve high levels proficiency in fewer than 10 years. More fundamentally, individuals in many professions simply do not have time to "practice.". Instead, it is "deliberate performance" that makes possible an increase in proficiency (Fadde & Klein, 2010). And practice alone is not sufficient since a person can progress to the journeyman level of proficiency and stay there (LaDue et al., 2019).

This article next elaborates a grounded method for identifying experts.

Method for Identifying Experts

At a theoretical or conceptual level, experts have been defined by reference to the concepts of the craft guilds of the Middle Ages, which distinguished a number of levels of proficiency (novice, beginner, apprentice, journeyman, expert, master) (Hoffman, 1998). The craft guild scheme reminds us that humanity does not neatly bifurcate into people who are novices versus people who are experts—this being another myth that appears in both the popular press and even (unfortunately) in the technical literatures. Additionally, the craft guild scheme provides a good conceptual definition of expertise:

> The expert is a distinguished or brilliant journeyman, highly regarded by peers, whose judgments are uncommonly accurate and reliable, whose performance shows consummate skill and economy of effort, and who can deal effectively with certain types of rare or "tough" cases. Also, an expert is one who has special skills or knowledge derived from extensive experience with subdomains (Hoffman, 1998, p. 85).

This definition of expertise alludes to measures of peer review, professional judgement, experience, performance, skill, and knowledge. It also piggybacks on the notion of a journeyman, which is itself defined as a person who can perform competently and reliably, and without supervision.

Opportunities, Challenges, and Constraints

The domain of chess presents an opportunity. Studies of chess have relied on the formal ranking process (Gobet & Charness, 2006). Estimates have also been made of the numbers of years and hours spent playing chess. Thus, two types of measurement methods have been utilized: (1) performance and (2) hours of experience. Relying on rankings is convenient for the researchers, and along with hours of experience certainly adduces convincing evidence that the individuals identified as expert do qualify as such. However, in many other domains, performance data are not available, and it is not feasible for researchers to attempt to develop them. Furthermore, for many jobs hours of experience can only be roughly approximated since time-at-work will rarely equate with time-on-tasks (Gilbreth, 1911).

A project conducted for electric utilities typifies the "real world" constraints that can be involved (Moon et al., 2009). Within the scope and time frame of this project, creating a full proficiency scale was not feasible. Pausing to design and conduct measures and longitudinal experiments to demonstrate reliably superior performance on a representative task would have precluded progress on the main problem: the pressing need to capture undocumented and soon-to-be-lost expert knowledge and skill.

The primary motivation for the project was the "grey tsunami" (Hoffman and Hanes, 2003). This was the demographic trend circa 1990 in which a boomer generation was entering their retirement years. There was considerable concern among private sector businesses and other organizations about the loss of critical expertise (DeLong, 204). In 2002 the Electric Power Research Institute conducted a survey of managers representing 21 electric utilities. They found that 92 percent of managers believed that loss of expertise would pose a problem within the next five years, but only 30 percent indicated that a planning effort was in place to retain knowledge. Indeed, one professional asserted that within the electric utilities broadly, 67 percent of the senior professionals would be retiring. Since then, major initiatives were taken to mitigate the consequences (Gross, 2001; Gross et al., 2002).

Many companies established knowledge management programs and "chief learning officers" (O'Dell & Grayson, 1998; Pringle, 2003). Utilities in the southeast United States were interested in developing a cadre of individuals who were trained to conduct knowledge elicitation with retiring professionals so that the organization might capture and hence share their knowledge and make major changes to the approach taken in training.

This presented a challenge for identifying experts (Ziebell, 2005).

> Expertise in many industries is the result of unstructured, ad-hoc sequences of experiences by relatively rare individuals. This rarity probably is the result of being in the right place to experience infrequent and challenging events several times. As a result, typically 20 years or more goes into the making of a recognized expert. Of course, many people in the workforce have 20 years of experience, and most of these become very good at what they do. But only a few are recognized as having expertise that exceeds that of others. Mere time in grade does not enable just anyone to adequately fill a unique and mission-critical function, so we cannot simply use this criterion to identify true experts (Moon et al., 2009, p. 23).

Additionally, there were many particular jobs, for both fossil and nuclear power generation utilities, where knowledge and skill were undocumented: equipment troubleshooting, instrumentation and control, control center operations, security, telecommunications, and others. Each of those jobs, if deconstructed, would map onto many dozens of particular tasks. Although experimental study of performance on some selected, specific task might have adduced some evidence about expert knowledge and skill, the pressing need would have remained largely unresolved.

In the utilities work, the challenge of identifying experts was actually solved in a simple and direct way. Management leadership in a number of utilities companies selected them. The managers were instructed to identify the following:

- Individuals who were recognized by their managers or by other managers and peers as being the only expert on something of high importance.
- Individuals who were one of only a few local site experts.
- Individuals with expertise in handling rare or infrequent events (e.g., repair of a unit that fails on average once every 10 years, or handling extensive repairs necessitated by a hurricane in areas not normally experiencing hurricanes).

The managers also down-selected for expertise that did not need to be maintained: Individuals with expertise for systems, etc., that were going to be replaced with different technology involving different skills (e.g., an "old" computer system being replaced about the same time the expert on that system retires).

The focus was on processes of power generation, power coordination, and environmental monitoring. The individuals selected were the "go-to" people for particular jobs. While they had worked in a variety of jobs within their company, learning from the ground up as it were, they were uniquely capable of conducting particular job tasks (such as the elaborate process of deconstructing and reconstructing a turbine in a power plant). On such tasks, they were indispensable (see Table

1, below). While one might argue that this by itself does not guarantee that the appellation of "expert" is definitive, there can be no doubt that the selected individuals were accomplished, highly experienced, and highly proficient.

Using methods of concept mapping and the Critical Decision Method (see Crandall et al., 2006), knowledge elicitation was utilized to capture expert knowledge, but also enabled a hindsight analysis, verifying that the participants manifested the qualities seen in experts in other domains. These are listed in Table 1.

Table 1. Qualities of experts in professional domains

Cognitive Capacities	Experts "see the invisible." They recognize the uncommon or irregular.
	Experts rarely say, "I don't know." In the occasional case in which they find themselves at a loss, experts engage in problem-solving techniques to make sense of a situation.
	Experts create comprehensive and thorough mental models. They anticipate not only consequences throughout a system, but also the collateral consequences to other systems.
Action Capacities	Experts are willing to improvise. They know how and when to improvise, particularly when situations go beyond the typical.
	Experts rarely say, "This is what I believe." They are constantly on the hunt for formal, empirical evidence.
	Experts create and rely upon "treasure maps." They develop and use memory artifacts that are unique organizing schemes and reinforce the structure of their knowledge.
Affective Capacities	Experts "live for the edge." They recognize that in order to achieve the mission, work needs to be done at the edge of the familiar.
	Experts revel in tough cases.
	Experts have unique incentives. Experts expect to be compensated, but typical compensation packages are not the only, or even most important, "carrots" they seek.
Social Capacities	They are an "ad-hoc solution provider." By virtue of their continuously demonstrated success, they become the "go-to" pro.
	The absence of the experts causes trauma. For colleagues who rely on experts, their absence can be a disruptive event.
	Experts have the ability to consider the perspectives of others involved in the situation. Unlike the expert who may become engrossed in the problem at hand, experts have the ability to consider the perspectives of others involved in the situation.
	Experts lead, but often only by example. They have the admiration of their peers and subordinates and develop knacks for employing their special position in furtherance of the mission. While they may find themselves in management roles, they are not always comfortable there.

The social qualities in Table 1 motivated a notion of the "franchise expert" (Hoffman, et al., 2011). These individuals are not only expert at their primary jobs but are also expert with regard to their knowledge of their organization and skill at working within the organization.

These case study domains—chess and electric utilities— are contrastive in many ways. But they both involved the discovery of convincing evidence that the designation of "expert" could be applied to certain individuals. They also point to the challenges of identifying experts, and proficiency scaling in general. They also illustrate the sorts of practical constraints that can be inescapable in the identification of experts.

Best practice in experimental psychology advocates for a reliance on more than one independent measure for any given theoretical concept. (This is another problem with the sole reliance on the 10-year rule.) A proficiency scale for a given domain should be based on more than one of a number of general classes of measures, each having associated measurement methods. Five classes of methods are described in Table 2. The classes are illustrated primarily by another case study, of expertise in weather forecasting (Hoffman et al., 2017; Hoffman et al., 2006; LaDue et al., 2019), but the sociometric method is best illustrated by the work in the electric utilities.

Table 2. Pentapod classes of methods that can contribute data for the creation of a proficiency scale

Method	Yield	Example
In-depth career interviews about education, training, hours of experience.	Ideas about breadth and depth of experience. Estimate of hours of experience and the actual primary domain tasks (versus hours on the job).	Weather forecasting in the armed services, for instance, involves duty assignments having regular hours and regular job or task assignments that can be tracked across entire careers. The amount of time spent at actual forecasting or forecasting-related tasks can be estimated with some confidence.
Professional achievements, standards, or licensing	Criteria about what it takes for individuals to be licensed, to be qualified to mentor apprentices, or to have reached the top of their field.	The study of weather forecasters involved individuals who had qualified to issue forecasts, including senior meteorologists from the U. S. National Oceanic and Atmospheric Administration (NOAA) and the National Weather Service. One participant was one of the forecasters for space shuttle launches; another was one of the designers of the first weather satellites.
Measures of performance at the familiar tasks	Can be used for convergence on scales determined by other methods. One should never assume that the ostensive primary task is the task at which the individual is expert. Furthermore, one should never assume that performance-based proficiency scaling should be based on performance on a single task.	Weather forecasting is again a case in point since records can show for each forecaster the relation between their forecasts and the actual weather. In fact, this is routinely tracked in forecasting offices by the measurement of “forecast skill scores.”
Social Interaction Analysis (Sociometry)	Who talks to whom? Who goes to whom for particular problems? Proficiency levels in some group of practitioners or within some community of practice (Mieg, 2000; Stein, 1997)	In the project on knowledge preservation for electric power utilities, experts at particular jobs were identified by plant managers, trainers, and engineers. The individuals identified as experts had been performing their jobs for years and were known among company personnel as *the* person in their specialization; e.g., “If there was that kind of problem I'd go to Ted. He's the turbine guy.”
Cognitive Task Analysis and Knowledge Elicitation	Models of knowledge, strategies	Examples would include all the applications of the critical decision method, (Hoffman et al., 1998), knowledge modeling using concept maps (Clark & Estes, 1996). Models can be compared for concordance across experts (Crispen & Hoffman, 2016).

Based on these classes of methods is the Pentapod Principle: Always use methods that are from at least two and ideally three of the five distinct methods classes to converge upon and validate a proficiency scale that is appropriate to the given domain.

Example: Weather Forecasting

The studies of proficiency in the domain of weather forecasting, cited above, illustrate the multi-method approach. In-depth interviews relied on the career records of civilian and military weather forecasters. It was possible to describe the depth and diversity of forecaster training and experience, and also estimate the amount of time that had been spent at actual forecasting tasks on work shifts. This included determination of the amount of time it took to qualify as a forecaster (that is, being allowed to issue official forecasts). Another method was performance analysis. Forecasts are routinely evaluated post hoc in terms of what is (somewhat misleadingly) called a “skill score.” This is the value added by a forecast over and above the accuracy that would derive from a forecast based solely on climatological data. Finally, knowledge was evaluated by having the forecasters engage in concept mapping of their domain’s concepts, principles, and atmospheric dynamics. The propositions in the knowledge models were cross-validated by having an experienced forecaster review the concept maps proposition by proposition.

As the data from these measures showed, and, as one would hope and expect, the individuals who were identified as experts demonstrated these characteristics:

- Had more diverse experiences (e.g., forecasting at diverse locations having differing climates and weather tendencies)
- Knew more about domains concepts and principles, with about 90% of the knowledge propositions cross-validating (disagreements mostly involve wordsmithing)
- Were identified in social network analysis as “go-to” persons for special skill at particular forecasting problems, (e.g., hurricane tracks)
- Had spent more time at actual forecasting tasks (in some cases, well over 10,000 hours)
- Showed reliably superior performance (e.g., accuracy of 85% on the particularly difficult task of forecasting summertime thunderstorms)
- Had developed forecasting procedures that were more refined and seasonally-dependent than those of apprentices and journeymen (who tend to over-rely on the outputs of the computer models)

Additionally, the data led to the conclusion that it is valuable, and not merely possible, to distinguish grades within levels of proficiency (e.g., junior journeyman, journeyman, senior journeyman, junior expert, expert, senior expert).

Conclusion

The assumption that 10,000 hours (or 10 years) experience is enough to qualify a person as expert, and the equally flawed assumption that humanity neatly bifurcates into novices versus experts, are assumptions that feed the “war on expertise” (Klein et al., 2019). Criticism of experts in particular domains, and the very concept of expertise itself, seems to be a part of the *zeitgeist*, especially in the politicized popular press but also in the technical literatures. Certain claims need to be countered and disavowed, claims—such as “people are surprised by the limitations in their understanding” (Fischer & Keil, 2016, p. 1251)—that are asserted in studies that are ostensibly about experts, but actually are about college freshmen who are subjects in laboratory experiments, and whose only claim to expertise is that they were “familiar” with the problem domain.

In the field of knowledge elicitation, there are instances of studies that used a multi-method approach and were arguably successful (see Table 2, above; see also Hoffman, 1987). There are numerous cases where a single “hours or years” rule was used but the consequent claim to have clearly bifurcated experts versus novices remained dubious, or at least arguable. What is lacking, and would be interesting, are cases of failure using a Pentapod approach. While those may be difficult to find at the present time, it will be important in establishing the multi-method approach for there to be criteria for evaluating its success and robustness.

Author's Declarations

The author declares that there are no personal or financial conflicts of interest regarding the research in this article.

The author declares that the research reported in this article was conducted in accordance with the Ethical Principles of the *Journal of Expertise*.

The author declares that the dataset is not publicly available but can be provided upon request.

ORCID iD

Robert R. Hoffman
https://orcid.org/0000-0002-1387-7659

References

Baer, D. (2014, June 2). Malcolm Gladwell explains what everyone gets wrong about his famous "10,000 Hour Rule." *Business Insider*. https://www.businessinsider.in/Malcolm-Gladwell-Explains-What-Everyone-Gets-Wrong-About-His-Famous-10000-Hour-Rule/articleshow/35964144.cms

Brooks, R. (2019, March 19). A better lesson. *MIT Robust AI*. https://rodneybrooks.com/a-better-lesson/

Clark, R. E., & Estes, F. (1996). Cognitive task analysis. *International Journal of Educational Research*, *25*(5), 403–417.

Crandall, B., Klein, G., & Hoffman, R. R. (2006). *Working minds: A practitioner's guide to cognitive task analysis.* Cambridge, MA: MIT Press.

Crispen, P., & Hoffman, R. R. (2016). How many experts? *IEEE Intelligent Systems*, *31*(4), 57–62.

DeLong, D. (2004). *Lost knowledge.* New York: Routledge.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406.

Ericsson, K. A., & Harwell, K. W. (2019). Deliberate practice and proposed limits on the effects of practice on the acquisition of expert performance: Why the original definition matters and recommendations for future research. *Frontiers in Psychology*, *10*, 2396. https://doi.org/10.3389/fpsyg.2019.02396

Fadde, P., & Klein, G. (2010). Deliberate performance: Accelerating expertise in natural settings. *Performance Improvement*, *49*(9), 5–14.

Fischer, M., &. Keil, F. C. (2016). The curse of expertise: When more knowledge leads to mis-calibrated explanatory insight. *Cognitive Science*, *40*(5), 1251–1269.

Gilbreth, F. (1911). *Motion study*. NY: Van Nostrand.

Gladwell, M. *Outliers*. (2008.) New York: Little, Brown and Company.

Gobet, F., & Charness, N. (2006). Expertise in chess. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 597–615). Cambridge: Cambridge University Press.

Gross, M. M. (2001). Capturing undocumented worker-job-knowledge: Overview and 2000 status report. Technical Publication, Electric Power Research Institute, Palo Alto, CA.

Gross, M., Hanes, L., & Ayres, T. (2002, September 15-19). *Capturing undocumented worker-job-knowledge at electric utilities: The EPRI strategic project* [Conference presentation]. IEEE 7th Conference on Human Factors and Power Plants, Scottsdale, AZ.

Hambrick, D. Z., Ferreira, F., & Henderson, J. M. (2014, September 28). Practice does not make perfect. *Slate*. https://slate.com/technology/2014/09/malcolm-gladwells-10000-hour-rule-for-deliberate-practice-is-wrong-genes-for-music-iq-drawing-ability-and-other-skills.html

Hayes, J. R. (1985). Three problems in teaching general skills. In S. F. Chipman, J. W. Segal, & R. Glaser (Eds.), *Thinking and learning skills, vol. 2: Research and open questions* (pp. 391–405). Mahwah, NJ: Erlbaum.

Hoffman, R. R. (1987). The problem of extracting the knowledge of experts from the perspective of experimental psychology. *AI Magazine*, *8*(2), 53–67.

Hoffman, R. R. (1998). How can expertise be defined?: Implications of research from cognitive psychology. In R. Williams, W. Faulkner, & J. Fleck (Eds.), *Exploring expertise* (pp. 81–100). New York: Macmillan.

Hoffman, R. R., Crandall, B., & Shadbolt, N. (1998). A case study in cognitive task analysis methodology: The Critical Decision Method for the elicitation of expert knowledge. *Human Factors, 40,* 254–276.

Hoffman, R. R., & Hanes, L. F. (2003). The boiled frog problem. *IEEE Intelligent Systems*, *18*(4), 68–71.

Hoffman, R. R., LaDue, D., Mogil, H. M., Roebber, P., & Trafton, J. G. (2017). *Minding the weather: How expert forecasters think*. Cambridge, MA: MIT Press.

Hoffman, R. R., Coffey, J. W., Ford, K. M., & Novak, J. D. (2006). A method for eliciting, preserving, and sharing the knowledge of forecasters. *Weather and Forecasting*, *21* (3), 416–428.

Klein, G., Shneiderman, B., Hoffman, R. R., & Wears, R. L. (2020). The “war” on expertise: Five communities that seek to discredit experts. In P. Ward, J. M. Schraagen, J. Gore, & E. Roth (Eds.), *The Oxford handbook of expertise* (pp. 1157–1192). United Kingdom: Oxford University Press.

Hoffman, R. R., Ziebell, D., Feltovich, P. J., Moon, B. M., & Fiore, S. F. (2011). Franchise experts. *IEEE Intelligent Systems*, *26* (5), 72–77.

LaDue, D. S., Daipha, P., Pliske, R. M., & Hoffman, R. R. (2019). Expertise in weather forecasting. P. Ward, J. M. Schraagen, J. Gore, and E. Roth (Eds.), *The Oxford handbook of expertise* (pp. 872–902). United Kingdom: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198795872.013.38

Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis. *Psychological Science*, *25* (8), 1608–1618.

Miller, M. (2023) The great practice myth: Debunking the 10,000-hour rule. Six Seconds. https://www.6seconds.org/2022/06/20/10000-hour-rule/

Moon, B., Hoffman, R. R., & Ziebell, D. (2009). How did you do that? Utilities develop strategies for preserving and sharing expertise. *Electric Perspectives, 34* (1), 20–29.

O’Dell, C., & Grayson, C. (1998) If only we knew what we know: Identification and transfer of internal best practices. *California Management Review*, *40*, 154–174. https://doi.org/10.2307/41165948

Pringle, D. (2003, January 7). Learning gurus adapt to escape corporate axe. *The Wall Street Journal*, B1.

Simon, H. A., & Gilmartin, K. (1973). A simulation of memory for chess positions, *Cognitive Psychology*, *5*(1), 29–46.

Stillman, J. (2016). What Malcolm Gladwell got wrong about the 10,000-hour rule. Inc. https://www.inc.com/jessica-stillman/10-000-hours-researcher-actually-tons-of-practice-isn-t-enough.html

Ziebell, D. (2005). Capturing undocumented expert knowledge. Report 1012127, Electric Power Research Institute, Palo Alto, CA.

Received: 6 March 2023
Revision received: 9 May 2023
Accepted: 10 May 2023