Perceptual (Re)learning: A Leverage Point for Human-Centered Computing

Robert Hoffman, Institute for Human and Machine Cognition
Stephen M. Fiore, University of Central Florida

Vol. 22, No. 3
May/June 2007

This material is presented to ensure timely dissemination of scholarly and technical work. Copyright and all rights therein are retained by authors or by other copyright holders. All persons copying this information are expected to adhere to the terms and constraints invoked by each author's copyright. In most cases, these works may not be reposted without the explicit permission of the copyright holder.
At least since Adrianus Dingeman de Groot conducted his pioneering work on the reasoning of chess masters, perceptual skill has been regarded as key to the advantage of experts. Here we explore the conjunction of two facts:

1. Experts can perceive things that are invisible to the novice.
2. It takes a decade or more for someone to become an expert in most significant domains.

This conjunction represents a leverage point for intelligent systems—not from the Turing Test perspective of building machines that emulate humans but from the human-centered computing (HCC) perspective of amplifying and extending human capabilities.

Perceptual learning
James and Eleanor Gibson, champions of “ecological psychology,” emphasized the importance of perceptual learning. It’s an interesting notion because it cuts across a Cartesian distinction between two processes that are believed to be distinct and fundamental to how the mind works. Today, there is a growing literature on psychophysical changes, including demonstrations of the neural basis for cue imprinting. With regard to cognition, however, the phenomenon of perceptual learning has been difficult to capture in the laboratory. It’s not easily reproducible in a controlled experiment that can last only as long as a college class period and that uses undergraduates as “subjects.” The only recourse has been to use simple, artificial materials (for example, geometric forms or schematic faces that fall into categories based on artificial rules) and relatively simple discrimination tasks or similarity judgment tasks.

Researchers have conducted studies on somewhat more realistic tasks, such as learning to discriminate flavors of beer or learning how to determine the sex of newborn chicks. In one study, college students received six training sessions on recognizing species of birds (using photographs as stimuli). Next, the students were tested in a same/different discrimination task. Those who had received training on discriminating species according to functional considerations (such as wading birds versus types of owls) rather than visual considerations showed better discrimination performance on a transfer task involving novel stimuli. A clear conclusion is that perceptual learning doesn’t happen by mere exposure to exemplars but is acquired through deliberate acts of discrimination and differentiation combined with corrective feedback.

Overall, the drive to capture phenomena in replicable, controlled circumstances has led psychological research to eliminate much of the richness of perceptual learning. Research hasn’t looked much at the perception of real moving faces but has looked at the perception of pictures or static cartoons of faces. It hasn’t looked much at the perception of real handled chicks or flying birds but photographs of chicks and birds. Precious little research has attempted to capture perceptual learning across time spans longer than those of typical laboratory experiments, and only a fraction of studies have trained the participants for more than one day.

Perceptual knowledge, perceptual skill, and expertise
Although capturing perceptual learning in the lab is challenging, studies of expert-novice differences have illustrated perceptual learning in a variety of domains ranging from neonatal critical-care nursing to commercial fishing. Studies have looked at domains where perceptual skill is paramount, such as radiology and baggage screening. We know that experts’ knowledge organization involves finer gradations of functional categories—their “basic object-level” categories fall at a finer level than for nonexperts. For example, for most people, limestone is simply a kind of rock, but to the expert, many variants (for example, tilted thinly interbedded limestone shale with limestone predominating) inform of geological dynamics. Experts’ categories fall also at a functional rather than literal surface feature level. So, experts can rapidly evalu-
ate a situation and determine an appropriate plan of action, a phenomenon called recognition-primed decision making.\footnote{14} Within the first few seconds of exposure to a novel chess position, chess experts can perceive important information about the relations of the chess pieces’ positions and begin identifying promising moves.\footnote{15}

This begins to get us closer to the heart of the matter: Perceptual learning isn’t just about the perception of cues or reckoning of variables—it’s about their meaningful integration. For example, a story is told in which an expert intelligence analyst inspected photos that had been used to show that Iraq was developing biological weapons. The expert disagreed with the original assessment that the photo showed a decontamination vehicle parked outside a chemical bunker: “I don’t think it is—and I don’t see any other decontamination vehicles down there that I would recognize.”\footnote{16} He explained that the standard decontamination vehicle was a Soviet-made box-body van; this truck was too long. Another expert agreed, saying “If you are an expert, you can tell one hell of a lot from pictures like this.”\footnote{16} The experts didn’t just detect cues but understood the meaning and significance of cues that were present, and ones that were absent.

Here’s another illustrative case. Terrain analysts interpret aerial photographs to assess soils, bedrock, vegetation patterns, drainage patterns, and so forth. In one experiment, an expert had two minutes to inspect a stereo photograph.\footnote{17} Ordinarily, the full systematic terrain-analysis process can take hours. After the viewing period, the expert had five minutes to report everything he could remember about the photo. In one particular trial, the expert began his retrospection by asserting that any personnel sent to the depicted area should be prepared for certain types of bacterial infection. The experimenter asked, “You can see bacteria in a photo of a pond taken from 40,000 feet?” The expert recounted the following reasoning sequence: The photo covered an area of tropical climate. The forest was mature and uniform, which meant that the contour of the tree canopy fairly mirrored the contours of the terrain surface (that is, the ground) and the underlying bedrock. Specifically, the expert could tell that the terrain was based on tilted interbedded limestone. The bedrock also determined the pattern of the streams and ponds, and it seemed one pond didn’t have a tributary running away from it. Given the climate, vegetation (tropical legumes), and stagnant water, the presence of bacteria was a virtual certainty.

This appears to be a long chain of inferences, dependent on a great deal of declarative knowledge. Yet, in the actual interpretation exercise, the expert’s judgment was the sort that you might be inclined to call direct or immediate, a perceptual phenomenon rather than a linear, deliberative process.

Psychological research on cue utilization illustrates the knowledge-based integration of cues. Expert decision makers don’t always rely on all relevant cues; sometimes they seem to make decisions based on a limited number of available cues.\footnote{18, 19} This finding feeds into the “biases and limitations” view of human cognition.\footnote{20} But what makes experts unique is “their ability to evaluate what is relevant in specific contexts. It’s the study of that skill, not the number of cues used, that should guide future research on experts.”\footnote{21} Furthermore, it’s also likely that studies showing cue underutilization might be epiphenomenal to the subtle ways in which experts integrate information.\footnote{22} In other words, the informative cues might be separable in the sense that they can beoperationally defined and measured independently of one another, but they link together for principled reasons and relate to one another in meaningful ways. In functional terms, they’re integral.

Consider a study involving livestock judges.\footnote{23} After viewing photos of young female pigs (gilts), the judges had to rate each gilt on its breeding quality. This is a typical task—agricultural extension experts can’t possibly travel to all the regional farms to evaluate animals. By tradition and (extensive) training, judges rely on a set of 11 gilt features (including weight, length, ham thickness, heaviness of bone structure, and freeness of gait). However, the study’s results suggested that the judges weren’t relying on all 11 features. Sometimes, they seemed to ignore information. In a second condition, judges were given verbal descriptions of the gilts (reflecting a telephone conversation) that listed the values of the 11 attributes for each animal. This revealed the basis of the judges’ reasoning: The cues interact. In the photo-based evaluations, the judges perceptually collapsed the 11 dimensions into three main judgments of size, meat quality, and breeding potential. Then they combined these to form an overall judgment. They could integrate the relevant stimulus attributes because they were, by nature, partly correlated for meaningful reasons (for example, tall gilts tend to be heavier and wider). Thus, even though a judge might have seen all the cues, he or she perceived meaningful cue integrations. The verbal descriptions, however, separated the cues, forcing a deliberative analysis and revealing the cue relations and the judges’ sequential collapsing strategy. Robert Goldstone has referred to this as a process of unification, in which a “single constructed unit represents a complex configuration.”\footnote{16}

Now we begin to see the crux of the matter. Often, the patterns that bear meaning can’t be defined in terms of the simple presence or absence (or values) of separable cues. Meaningful patterns are sometimes defined by the relations among functionally integral cues. Ludwig Wittgenstein was getting at this with his notion of featureless family resemblances. For instance, an expert weather forecaster can see a “gate-to-gate signature” in radar that is a clue to tornado formation. This signature is a function of a difference (which is a relation) in relative velocity (a second relation) of proximal (a third relation) winds, with strong (fourth relation) winds moving toward the radar (fifth relation) in very close proximity (sixth relation) to strong (seventh relation) winds that are moving away from (eighth relation) the radar.\footnote{24} Clearly, a considerable nexus of relations exists.

This is well known in other domains too. In medicine, it’s captured by the phrase, “Diseases do not read the textbooks.” Here, for example, is a passage from Malcolm Gladwell’s interview with an expert radiologist about the process of interpreting mammograms:
Some calcium deposits are oval and lucent. “They’re called eggshell calcifications ... and they’re basically benign.” Another kind of calcium runs like a railway track on either side of the breast’s many blood vessels—that’s benign too. ... “There are certain kinds that are always associated with cancer. But those are the ends of the spectrum, and the vast amount of calcium is somewhere in the middle. And making that differentiation ... is not clear-cut.”16

Mammogram interpretation is made complex and difficult for a number of reasons. This isn’t just a matter of making better cameras or taking better pictures:

You can build a high-tech camera ... [But] even then the pictures are not self-explanatory. They need to be interpreted, and the human task of interpretation is often a bigger obstacle than the technical task of picture-taking.16

As one expert said in discussing a tough case:

That cancer shows up only because it is in the fatty part of the breast. If you take that cancer and put it in the dense part of the breast, you’d never see it. ... If the tumor was over there [by a few centimeters], it could be four times as big and we still wouldn’t see it.16

You might expect such interpretive uncertainty and ambiguity to be frequently associated with things that are abnormal. However, “the overwhelming number of ambiguous things really are normal.”16 Perceptual learning is what makes expertise possible: “There is good evidence that with more rigorous training and experience radiologists can become much better at reading X-rays.”16 For example, Alan Lesgold and his colleagues found that novice radiologists rather quickly acquired the ability to identify problem areas in an X-ray but were much slower in their ability to accurately interpret and diagnose the problem.11

From this brief review we can summarize that the patterns that are meaningful to experts sometimes involve:

• individual cues,
• cues that are absent,
• sets of separable cues with some cues being necessary and some being sufficient,
• patterns that can be defined in terms of combinations of cues,
• patterns defined in terms of relations among cues,
• patterns defined in terms of the relations of cues that are present to cues that are absent,
• featureless family resemblances where cues are neither necessary nor sufficient when considered individually, and
• meaning that resides in the relations among cues that are integral, what Goldstone calls unitized cue configurations.

Now, let’s up the stakes.

Dynamic cue configurations

When civil engineers who are experts in interpreting aerial photographs look at such photos to determine, for example, the best site for a dam, engineering requirements for building a new road, or a housing development’s environmental impact, they see terrain features. However, they perceive geological dynamics: the complex, long-term processes that led the terrain to take its current form.25 When expert radiologists look at mammograms, they see patterns of shades of white, gray, and black. But they perceive processes such as calcification.

The patterns that experts perceive, even in static images, are dynamic. Experts perceive processes. When expert weather forecasters look for tornadoes in a radar image, they see a pattern of colors and shapes, but that isn’t what they perceive. The gate-to-gate signature we mentioned earlier looks like an owl’s head (in its prototypical manifestation), but it rarely shows up in individual radar scans. More often, it appears only across a series of radar scans over time. The expert firefighter can determine a fire’s location and cause by the movement of its flame and smoke; the expert bird watcher can identify a species even when all there is to see is a fleeting shadow of movement in flight. Meaningful patterns sometimes exist only over time.

Let’s up the stakes again.

Dynamic cue configurations that exist across multiple data types

So far, we’ve talked about how experts perceive patterns in such things as x-ray films and radar images. But the patterns that experts perceive sometimes don’t exist in individual data types. Indeed, the really critical information often is “transmodal”—that is, it exists only across data types. For instance, in weather forecasting, the radar images aren’t the only thing that is guiding sensemaking activity and shaping the forecaster’s formation of a mental model. A great many other data types are involved, such as satellite images, computer model outputs, wind fields, and pressure data. For example, indications of severe weather might lie in a combination of:

• Satellite image loops. These show, on a space-time scale on the order of continent/week, the convergence of air masses of differing pressure, temperature, and moisture.
• Wind fields as a function of height in the atmosphere. These show, on a space-time scale on the order of states/days, the localized regions of potential instability.
• Observational data. These show, on a space-time scale on the order of regions/hours, where layers of the atmosphere are about to undergo inversion, releasing potential energy and triggering storm formation.

The “Aha!” moment might come when viewing radar, but the mind had been prepared in the sense that a mental model was based on an integration of meaningful patterns that only exist across data types. Likewise, the expert terrain analyst makes determinations when viewing aerial photos but also engages in the systematic analysis of other data, such as maps.

We should note that “multiple data types” can mean multiple perceptual modalities. In the case of a breast cancer diagnosis, “a skilled pair of fingertips can find out an extraordinary amount about the health of a breast, and we should not automatically value what we see in a picture over what we learn from our other senses.”16 Now let’s up the stakes even more.
Perceptual relearning

This final step involves bringing into the mix the Moving Target Rule: 26 The socio-technical workplace is constantly changing, and constant change in environmental constraints (such as technologies in the workplace) might entail constant change in cognitive constraints (the work to be accomplished), even if the domain constraints remain constant. Change in cognitive work comes because of changes in goals (such as new tasks or challenges) but especially because of changes in technology, including changes in data types and display types. For instance, the NEXRAD radar has revolutionized radar meteorology and forecasting, and new radar algorithms are being introduced all the time, resulting in new data products (hence, new displays) and new combinations of data types (such as those that combine satellite and radar images, called “Sat-Rad” displays). In the modern sociotechnical work context, the expert must engage in frequent, if not nearly continuous, perceptual relearning. Patterns previously learned and perceived in one way come to be perceived in a new way. New patterns pique the muse.

The crux of the matter

We can now return to the conjunction of the two facts with which we began this essay.

Fact one: Experts can perceive things that are invisible to the novice. We can be a bit more specific now: Experts engage in perceptual relearning of dynamic information defined over sets of integral cues that are transmodal (they exist over different data types). This is a holy grail for expertise studies because perceptual skill is critical for many, if not most, domains of expertise. A critical scientific gap is the paucity of research attempting to capture the perceptual relearning process as it occurs in the practice of domain experts. We need to know more about how it happens psychologically. Available research gives only the barest of clues.

Perceptual relearning of dynamic integral transmodal cue configurations (yes, we could give it an acronym, but we won’t) is also important for intelligent systems. One way of thinking about it is that it takes the notion of pattern recognition to entirely new levels. Another way of thinking about it is that it’s a cautionary tale about the rather nebulous yet popular notion of “information fusion.” 27 We’re reminded of the Sacagawea Law of HCC, 28 which asserts that effective complex cognitive systems support the integration, search, and active exploration of meaning. The law was invoked in the context of decision-aiding, but here we apply it to perception.

Fact two: It takes years to achieve expertise—thousands of hours of deliberate practice to reach world-class caliber in chess, 15 sports, 29 or weather forecasting. 22

Now here is how these two facts come together: Any method for accelerating the achievement of expertise will hinge on the ability to support the processes of learning and perceptual relearning of dynamic cue configurations, including those that exist across multiple data types. Although the value of “on-the-job” training is clear, 30 it generally doesn’t explicitly focus on perceptual learning and often lacks adequate organizational procedures and adequate management or organizational support. 31

Thus, we find a role for intelligent, human-centered systems. A key challenge is determining whether any shortcuts to mastery exist—that is, determining whether we can accelerate the perceptual learning process and facilitate the perceptual relearning process.

In domains of expertise where perceptual skill is paramount, such as radiology and baggage screening, 32 it seems reasonable to speculate that providing critical exemplars of targets makes the perceptual learning process possible. But it might not accelerate it. It might take upwards of 10 years to achieve expertise because experts (by definition) can deal with tough discriminations and challenging cases that are (by definition) rare. 2 If we could somehow compress the time needed to experience such cases, we might be able to accelerate perceptual learning. You might refer to this as “tough-case time compression.” We predict that engaging in perceptual relearning will require effort and be disruptive even for experienced domain practitioners. But experienced practitioners should be able to reacquire expertise in less time than it takes less-experienced practitioners (students, journeymen) to acquire it.

Intelligent technologies will need to support the conduct of such research. Time-compressed case-based practice will depend on having at hand a large and navigable corpus of cases, in all their rich detail. Furthermore, the full corpus, and therefore both the training and testing sets of cases, will need to include cases that are routine and frequent, nonroutine and rare, easy and simple, and complex and tough. Generating sets of scenarios is a job for cognitive task analysis, which we know how to do. 33 What we need now are technologies through which the learner can experience the full range of meaningful event patterns.

The dynamics associated with this experience must be manipulable. Here is another role for computer science: Packaging case information and meta-information so that cases could be relived, explored as they are relived, and compared as they are explored... but in “compressed” time. We might compress time through splicing or judicious speeding up to remove chunks of time or by shortening delays (for example, letting thunderstorms develop over a span of minutes rather than hours), but such an approach won’t likely be the only—or even the consistently best—way to compress time. Cases find their meaning in how sequences or parallelisms of events hang together and unfold across time, thus making time one of the cues within a configuration. So “dynamic” isn’t just a qualifier but also a variable that the learner might need to manipulate. In other words, we’ll likely need multiple methods for compressing time, some of which might be domain or event dependent.

We hypothesize that a benefit from the research and development activity we’re suggesting would be intelligent systems that have been referred to in this HCC department as Janus Machines—human-centered machines that can facilitate learning (as training aids) and improve performance (as performance...
support systems). Acquiring the ability to make perceptual judgments will correlate with increasing knowledge about the underlying meanings, dynamics, and causal relations that are formative of the perceptible patterns. Conversely, increasing perceptual skill relates to a more sophisticated understanding and integration of the relevant perceptual dimensions. Strategies embodied in intelligent systems might help learners at all levels (initiates, apprentices, journeymen) ramp up their knowledge and skill more quickly when they move into an unfamiliar zone or when the nature of the work changes abruptly. Janus Machines for perceptual relearning would be of considerable benefit in many work domains, perhaps especially ones that are linked to significant workforce issues of our time, such as the loss of expertise.

References


Robert R. Hoffman is a senior research scientist at the Institute for Human and Machine Cognition. Contact him at rhoffman@ihmc.us.

Stephen M. Fiore is an assistant professor of cognitive science in the Department of Philosophy and a research scientist at the Institute for Simulation and Training, both at the University of Central Florida. Contact him at sfiore@ist.ucf.edu.