How variability shapes learning and generalization

May 13 2022

An example of visual data augmentation techniques used in machine learning, which captures the main principle of variability effects: exposure to variation along non-discriminative dimensions (i.e., by rotating, changing the color, or partially masking the target image) improves the neural networks' ability to generalize (in this case - to identify a fox), but at the cost of slowing down initial learning. Humans show a similar effect: more variable input is harder to learn, but eventually boosts our ability to generalize the knowledge we learned to new contexts. This is because variability helps to highlight which features of the category are actually relevant, and which are not. Credit: Limor Raviv

Variability is crucially important for learning new skills. Consider learning how to serve in tennis. Should you always practice serving from the exact same location on the court, aiming at exactly the same spot?

Although practicing in more variable conditions will be slower at first, it will likely make you a better tennis player in the end. This is because variability leads to better generalization of what is learned.

Chihuahuas and great danes

This principle is found in many domains, including speech perception, grammar, and learning words and categories. For instance, infants will struggle to learn the category "dog" if they are only exposed to chihuahuas, instead of many different kinds of dogs (chihuahuas, poodles and great danes).

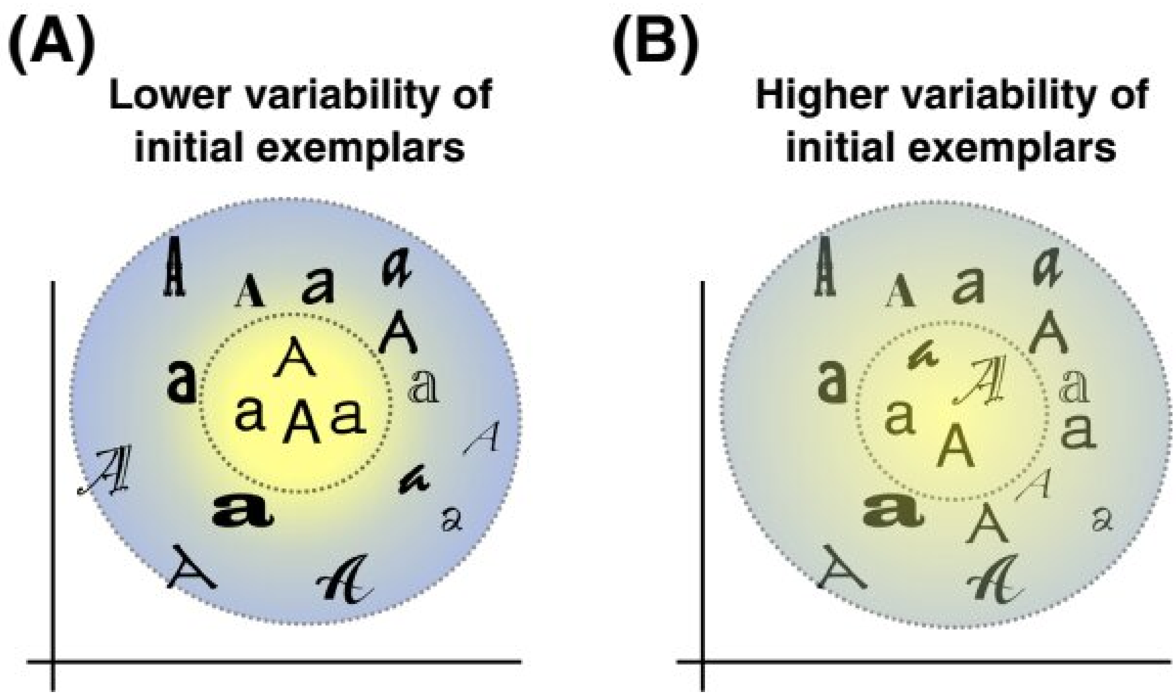
"There are over 10 different names for this basic principle," says MPI's Limor Raviv, the senior investigator of the study published in *Trends in Cognitive Sciences*. "Learning from less variable input is often fast, but may fail to generalize to new stimuli. But these important insights have not been unified into a single theoretical framework, which has obscured the bigger picture."

To identify key patterns and understand the underlying principles of variability effects, Raviv and her colleagues reviewed over 150 studies on variability and generalization across fields, including computer science, linguistics, categorization, motor learning, visual perception and formal education.

Mr. Miyagi

The researchers discovered that, across studies, there are at least four different kinds of variability, such as set size (e.g., the number of different examples or locations on the tennis court) and scheduling (e.g. practice schedules with different orders or time lags). "These four kinds of variability have never been directly compared, which means that we currently don't know which is most effective for learning," says Raviv.

The impact of variability depends on whether it is relevant to the task or not (arguably, the color of the tennis court is not relevant to serving practice). But according to the "Mr. Miyagi principle" (inspired by the 1984 classic movie "The Karate Kid"), practicing seemingly unrelated skills (such as waxing cars) may actually benefit learning of other skills (such as martial arts).



An example of the effect of exposure to more or less variability when learning to identify what the letter ‘A’ looks like. Initial training items are shown in the center circle of each panel, and the color gradient symbolizes generalization performance: greater accuracy and/or certainty in our generalization is represented by shades of yellow, while lower accuracy and/or certainty in our generalization is represented by shades of blue. Less variability during initial training (Panel A) can cause learners to form more conservative hypotheses about what the letter ‘A’ can look like, resulting in narrower generalization to less frequent instances of the letter 'A'. More variable examples during initial training

(Panel B) will result in broader hypotheses/categorization, and will enable learners to more accurately and/or more certainly classify different instances of the letter 'A' encountered later. Credit: Limor Raviv

Competing theories

But why does variability impact learning and generalization? One theory is that more variable input can highlight which aspects of a task are relevant and which are not (color is useful for distinguishing between lemons and limes, but not for distinguishing between cars and trucks).

Another theory is that greater variability leads to broader generalizations. This is because variability will represent the real world better, including atypical examples (such as Chihuahuas).

A third reason has to do with the way memory works: when training is variable, learners are forced to actively reconstruct their memories.

Face recognition

"Understanding the impact of variability is important for literally every aspect of our daily life. Beyond affecting the way we learn language, motor skills, and categories, it even has an impact on our social lives," explains Raviv. "For example, face recognition is affected by whether people grew up in a small community (fewer than 1,000 people) or in larger community (more than 30,000 people). Exposure to fewer faces during childhood is associated with diminished face memory."

"We hope this work will spark people's curiosity and generate more work on the topic," concludes Raviv. "Our paper raises a lot of open questions. For example: Is the relationship between variability and learning broadly similar across species, or are there species-specific adaptations? Can we find similar effects of variability beyond the brain, for instance in the immune system?"

More information: Limor Raviv et al, How variability shapes learning and generalization, *Trends in Cognitive Sciences* (2022). DOI: 10.1016/j.tics.2022.03.007

Provided by Max Planck Society