Current Directions in Psychological Science
2017, Vol. 26(1) 59–67

DOI: 10.1177/0963721416672904
www.psychologicalscience.org/CDPS

# Plateaus and Asymptotes: Spurious and Real Limits in Human Performance

**Wayne D. Gray**
Department of Cognitive Science, Rensselaer Polytechnic Institute

**Abstract**
One hundred twenty years ago, the emergent field of experimental psychology debated whether plateaus of performance during training were real or not. Sixty years ago, the battle was over whether learning asymptoted or not. Thirty years ago, the research community was seized with concerns over stable plateaus at suboptimal performance levels among experts. Applied researchers viewed this as a systems problem and referred to it as the *paradox of the active user*. Basic researchers diagnosed this as a training problem and embraced *deliberate practice*. The concepts of plateaus and asymptotes and the distinction between the two are important as the questions asked and the means of overcoming one or the other differ. These questions have meaning as we inquire about the nature of performance limits in skilled behavior and the distinction between brain capacity and brain efficiency. This article brings phenomena that are hiding in the open to the attention of the research community in the hope that delineating the distinction between plateaus and asymptotes will help clarify the distinction between real versus "spurious limits" and advance theoretical debates regarding learning and performance.

**Keywords**
plateaus, asymptotes, expertise, performance, memory, spurious limits, cognitive skill acquisition

## Introduction

Since the beginning of experimental psychology, plateaus and asymptotes have confused or confounded our thinking about behavioral findings. One hundred twenty years ago, Bryan and Harter (1897, 1899) drew attention to plateaus of performance during training of telegraph operators. Sixty years ago, Underwood (1954) attacked the field of human learning for the long-standing generalization that "the fast learner retains more than the slow learner" (p. 276). Though not in these terms, Underwood argued that this spurious finding was based on mistaking a measurement asymptote as a theoretical asymptote on amount learned. Thirty years ago, Carroll and Rosson (1987) and Ericsson, Krampe, and Tesch-Römer (1993) each concluded that human experts were not performing at their optimal level and, without discussing plateaus and asymptotes, concluded that people could be more productive if we used our software more effectively (Carroll and Rosson) or if we practiced differently (Ericsson and many colleagues). Once again, without using the terms "plateau" and "asymptote," today's debates on *brain capacity* versus *brain efficiency* are really about escaping the plateau by acquiring new methods (brain efficiency) or redefining the asymptote through capacity gains from a general enhancement of various brain areas (brain capacity).

As ubiquitous as these dueling concepts have been throughout history, they are seldom contrasted. Indeed, although Bryan and Harter used "plateau" 14 times in their two articles, they used "asymptote" not at all. Likewise, Carroll and Rosslyn used "asymptote" once and "plateau" not at all. Ericsson et al. (1993) used "plateau" three times in a paragraph discussing Bryan and Harter's work but not anywhere else, nor "asymptote" at all. Underwood used "asymptote" once. And across five recent and prominent studies on brain training (Dunning & Holmes, 2014; Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Redick et al., 2012; Shipstead, Redick, & Engle, 2012; von Bastian & Oberauer, 2014), the word "plateau" has been used not at all and "asymptote" once.

**Corresponding Author:**
Wayne D. Gray, Department of Cognitive Science, Rensselaer Polytechnic Institute, 110 8th St., Troy, NY 12180
E-mail: wayne.gray.cogsci@gmail.com

### *Perspective and definitions*

> It seems to me therefore that mental training in schools, in industry, and in morals is characterized, over and over and over again, by *spurious limits*—by levels or plateaus of efficiency which could be surpassed. The person who remains on such a level may have more important things to do than to rise above it; the rise, in and of itself, may not be worth the time required; the person's nature may be such that he truly cannot improve further, because he cannot care enough about the improvement or cannot understand the methods necessary. But sheer absolute restraint—because the mechanism for the function itself is working as well as it possibly can work—is rare. (Thorndike, 1913, p. 181)

Following Thorndike's distinctions, I refer to spurious limits "which could be surpassed" as *plateaus* and to limits due to "the mechanism for the function itself" as *asymptotes*. To avoid any misunderstanding of the rigidity of these categories, let me acknowledge that any change to behavior or cognition must be accompanied by some change in the brain. If that change reflects a new method for completing an old task, it will most likely be accompanied by a sequential activation of brain areas different than for the old method—we will say that acquiring the new method enabled a spurious limit or plateau to be overcome. When that change reflects (a) the use of a new or modified tool or (b) a general enhancement in the functioning of a brain area, we will say that an asymptote has been raised.

### *Plan for this essay*

I proceed by providing a variety of examples that represent different causes of asymptotes and plateaus. There is no hidden agenda, no great conclusion, no great theoretical revolution. Rather, the goal is to bring phenomena that were hiding in the open to the attention of the research community in the hope that these distinctions will help to clarify and therefore advance theoretical debates regarding learning and performance.

## Asymptotes Due to Design

### *Cigar rolling in Cuba*

Crossman's (1959) study of cigar rollers, the results of which are illustrated in Figure 1, found a continued increase in performance over a 2-year period (estimated as 3 million cigars made) and then a flattening of the curve. Newell and Rosenbloom (1981, p. 7) attributed this flattening to a "known lower bound for the performance time"; namely, the "cycle time of the machine." Although it cannot be proved from these data, it would not be surprising to learn that the efficiency of eye, arm, hand, and finger movements continued to improve after performance had asymptoted; that is, any potential gains from

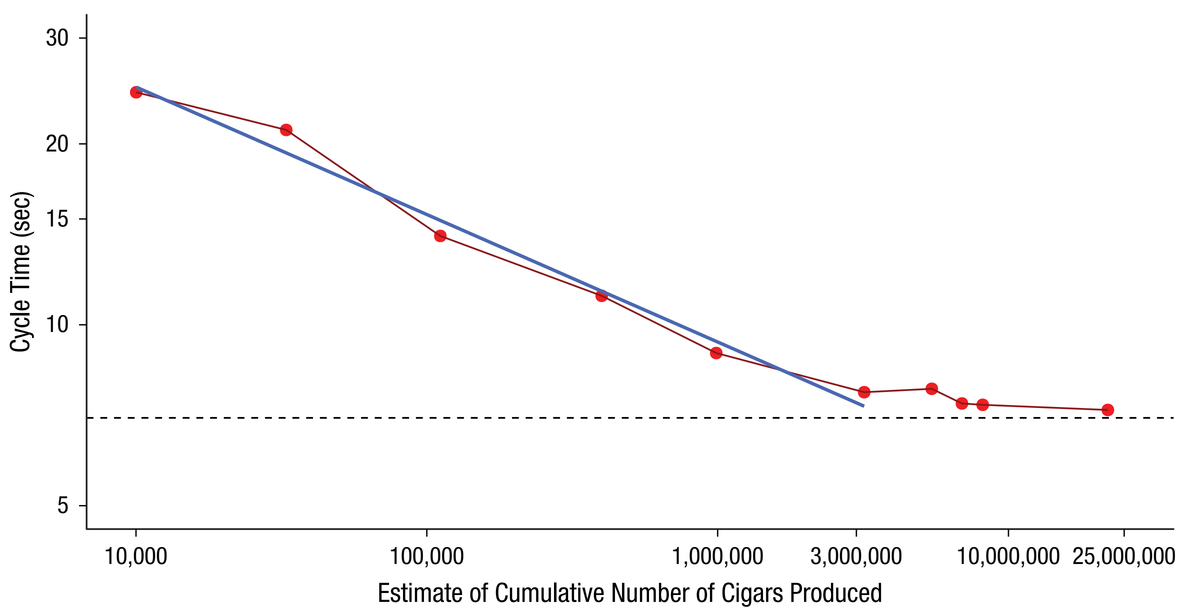


**Fig. 1.** A plot illustrating the productivity of Crossman's (1959) cigar roller, showing a flattening of performance after 2 years and 3 million cigars due to limits in the cycle time of the machine used for cigar rolling. Each of the 10 red dots represents the average cycle time and number of cigars produced over a 1-week period; the dashed line represents the machine cycle time. The 10 weeks plotted were sampled over a 7- to 8-year time period. The blue line shows the best-fitting regression line up to the point where the curve flattens. Crossman's original plot indicates that this would have been near the 2-year mark. The last dot plotted is beyond the 7-year mark. (Data were extracted from Fig. 2 of Newell & Rosenbloom, 1981, using Plot Digitizer software and should be considered as near approximations to the data in Crossman's Fig. 2.)

log-log speedups in perception, cognition, or motor processes would have been held hostage by the cycle time of the primitive machines used by the cigar rollers.

### *Ergonomically misengineered tools*

A higher-tech example of an asymptote due to design is provided by contrasting two workstations designed for telephone company Toll and Assistance Operators (TAOs; Gray, John, & Atwood, 1993). A 4-month field trial found that performance times on the *proposed* workstation stabilized within 2 months, with times per call slower than for the *current* workstation, and slow enough to increase annual operating costs by an estimated $5.8 million (in 2016 dollars). This performance asymptote was puzzling. Believing that call time was driven by the number of keys per call the TAO had to press, the designers eliminated keypresses across all call types for an expected decrease in mean call times of 4.1 seconds and a projected annual savings in operating costs of $24 million (in 2016 dollars).

CPM-GOMS[1] models (Gray et al., 1993) showed that the proposed workstation decreased the number of things the TAO could do in parallel. For example, the current workstation enabled TAOs to exchange information with the customer while keying or while waiting for verifications by external databases (see Fig. 2). (Hence, removing keystrokes from this part of the call did nothing to speed things up—compare Fig. 2 top vs. bottom.) Contrariwise, the proposed workstation moved required keystrokes to the end of the call (see Fig. 3), at which point the TAO could do nothing other than wait for external databases to return data. (Hence, moving one of the keystrokes from the beginning of the call, illustrated in Fig. 2, to the end, illustrated in Fig. 3, increased the call duration even though the total number of keystrokes decreased.) Although there were fewer keypresses per call, the critical path grew longer as (a) more keypresses and (b) more waiting time had been moved onto the critical path. Hence, within 2 months of practice, performance had asymptoted as the result of the interface design, not human cognitive, perceptual, or motor limits.

## Asymptotes Due to Measurement Method and Theory

### *Speed of learning and amount retained*

In her review of memory studies going back to Ebbinghaus, Gillette (1936) concluded that "the fast learner is the better retainer" (Gillette, 1936, p. 1). In reaching this conclusion, she used the "method of adjusted learning" to drop out paired-associate items from a list as soon as they had been correctly anticipated a certain number of times.

Underwood (1954) adopted the contrary hypothesis, maintaining that "the essential difference between fast and slow learners is that a reinforcement does result in more associative strength for a fast S than for a slow S" (p. 278).

The difference between these two views reflects a difference in theoretical perspective that in 2017 is hard to understand. Gillette seems to have believed that the binary outcome of "recalled" versus "not recalled" reflected an individual difference between fast and slow learners, such that people who learned fast remembered more but that learning was basically either all or none. Underwood maintained a "learning strength" perspective that meant that faster learners learned more on each correct trial than did slower learners. Hence, fast learners "overshot" the threshold between recall failure and recall success by more than did slow learners. For Underwood, forgetting was a gradual process of decline in associative strength, not an all-or-none phenomena.

Although Gillette believed that her method ensured that all learners learned equal amounts, as Underwood demonstrated, her equality across learners resulted from an asymptote due to measurement method. After adjusting for predicted "associative strength" after learning, Underwood found no difference between slow and faster learners after a 24-hour retention period.

In summary, Gillette's measurement method assumed an asymptote that Underwood's theory said was not there. That is, the ability to recall an item correctly once does not imply that all such items are of equal strength between Ss. Underwood's auxiliary assumption was that with no asymptote in item strength, a faster rate of learning would result in more learned during the last trial for fast than slow learners.

### *Space Fortress*

Space Fortress is a game created by experimental psychologists (Donchin, 1989, 1995) to be used as a common task in which diverse researchers can apply diverse "learning strategies" and compare their effectiveness in mastering "a rather complex task" (Donchin, 1989, p. 1). Most modern studies of Space Fortress (Donchin, 1995) have used the version crafted by Gopher, Weil, and Bareket (1994). This version includes the original measure (Points), three measures (Speed, Control [CNTRL], and Velocity) designed by Frederiksen and White (1989) as a means of focusing attention on different task components during training, and Total, which combines parts of the other four. Unfortunately, the three Frederiksen and White measures are sometimes used as independent measures of performance. There are two problems with that usage: First, they are not independent of each other, and second, two of these measures reach asymptote even as skilled performance continues to increase.

For example, the CNTRL measure was originally intended to encourage novices to fly close to the Fortress (Frederiksen & White, 1989). Although it may have served that purpose for novices, Figure 4 shows that the expert

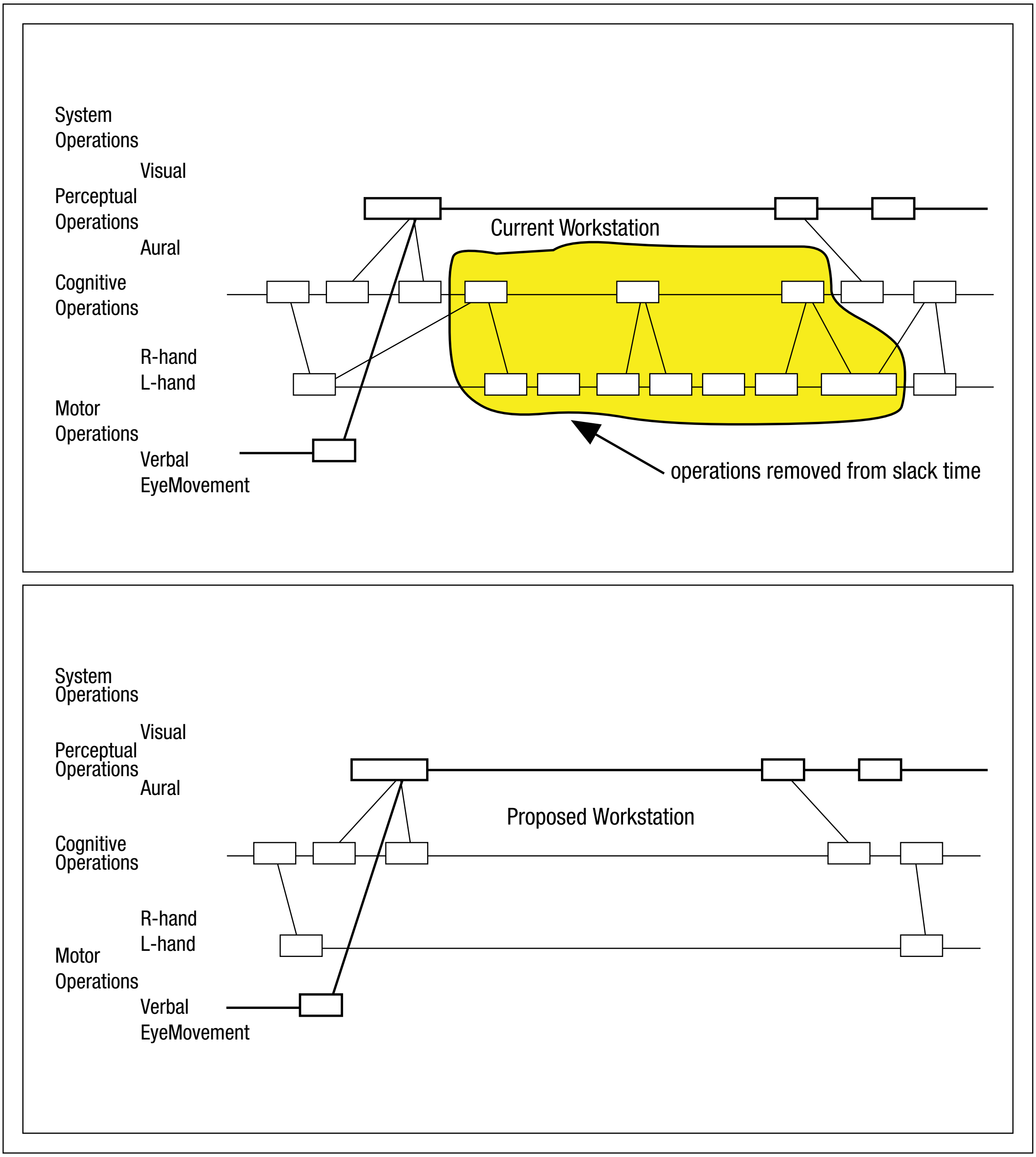


**Fig. 2.** Excerpts from a CPM-GOMS[1] model for the beginning of one category of telephone company operator-assisted phone calls. The *x*-axis represents time, and the *y*-axis represents the categorical variable of operator type. Each box represents an operation, which has a duration measured in seconds and milliseconds. Each line represents a different operator type. Within each line, the operator to the left must fire and complete before the operator to its right may be performed. The diagonal lines represent dependencies between different operator types. Bold lines show the "critical path" in both phone calls, which does not change from the "current" to the "proposed" design. However, what does change is the amount of work the TAO can do in parallel while waiting for the critical-path items to complete.

could fly closer and closer to the Fortress long after his CNTRL score asymptoted.

Unfortunately, many Space Fortress studies have attempted to relate improvements within the game to changes in external measures of cognitive abilities such as memory or attention (see, e.g., Boot et al., 2010); such efforts are doomed to false negatives, as the measure asymptotes long before the skill plateaus. Fortunately, researchers running studies of Space Fortress now recognize (Boot, Sumner, Towne, Rodriguez, & Ericsson, in press; Destefano, 2010) that flying closer to the Fortress than measured by the CNTRL score enables the expert

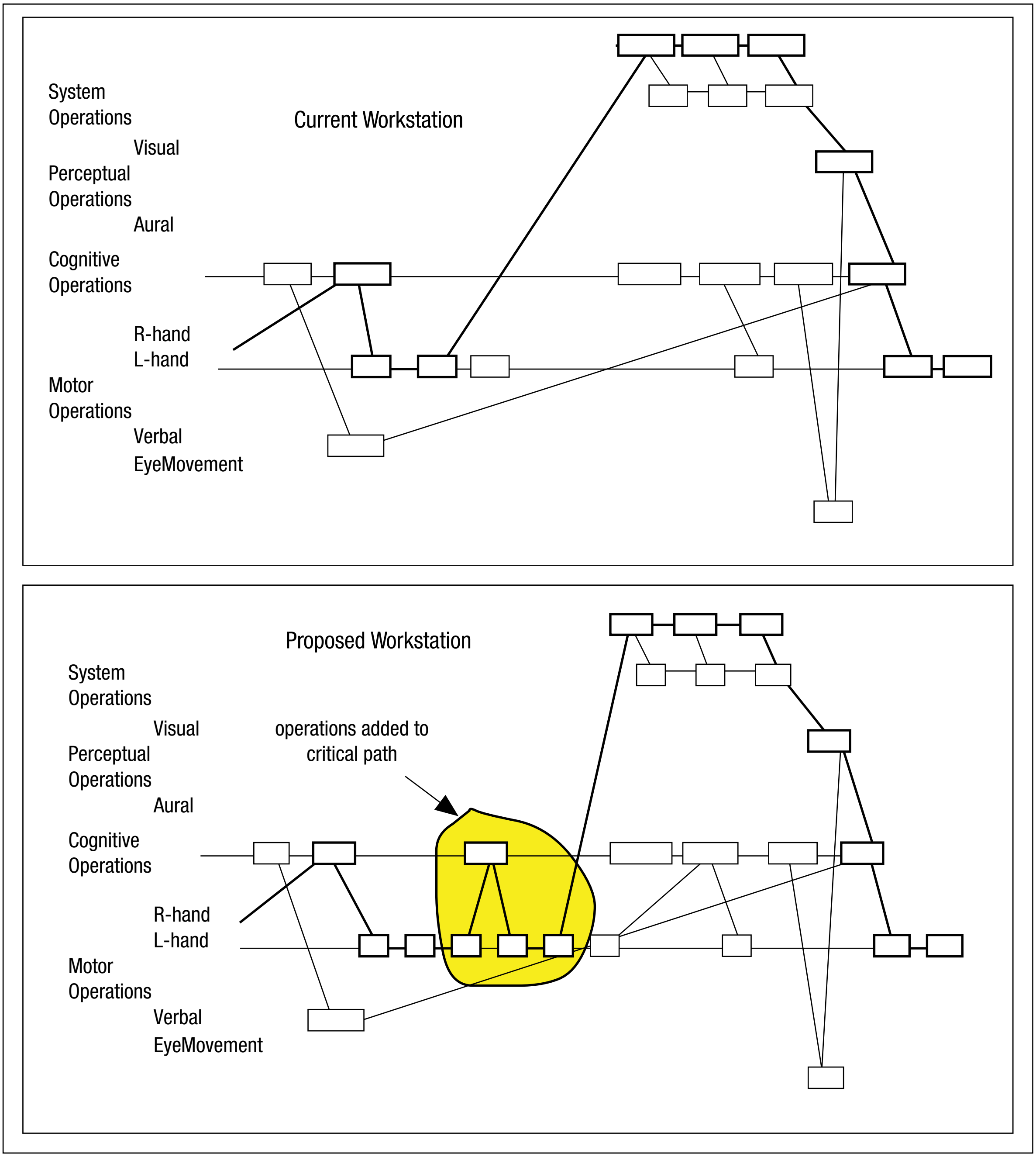


**Fig. 3.** Excerpts from a CPM-GOMS[1] model for one category of telephone company operator-assisted phone calls—same phone call as Figure 2; however, this time we focus on the end of that call. Note that for the proposed workstation, some of the work that the TAO would have done during slack time at the beginning of the call is now spliced into the critical path at the end of the call—thereby adding to the call's length.

player to gain points in other measures of skill by killing the Fortress faster while avoiding damage to the player's Space Ship.

## Plateaus

The distinction between a plateau and an asymptote may be hard to discern, especially in laboratory studies that collect only a few hundred performance trials. However, the distinction is clear in hindsight when it can be shown that groups of individuals performing at different skill levels are following different methods.

### *Plateaus due to self-inflicted limits*

***Typing.*** In a clear example of the use of different methods, *visually guided* typists[2] perform far below the level of touch typists. Yechiam, Erev, Yehene, and Gopher

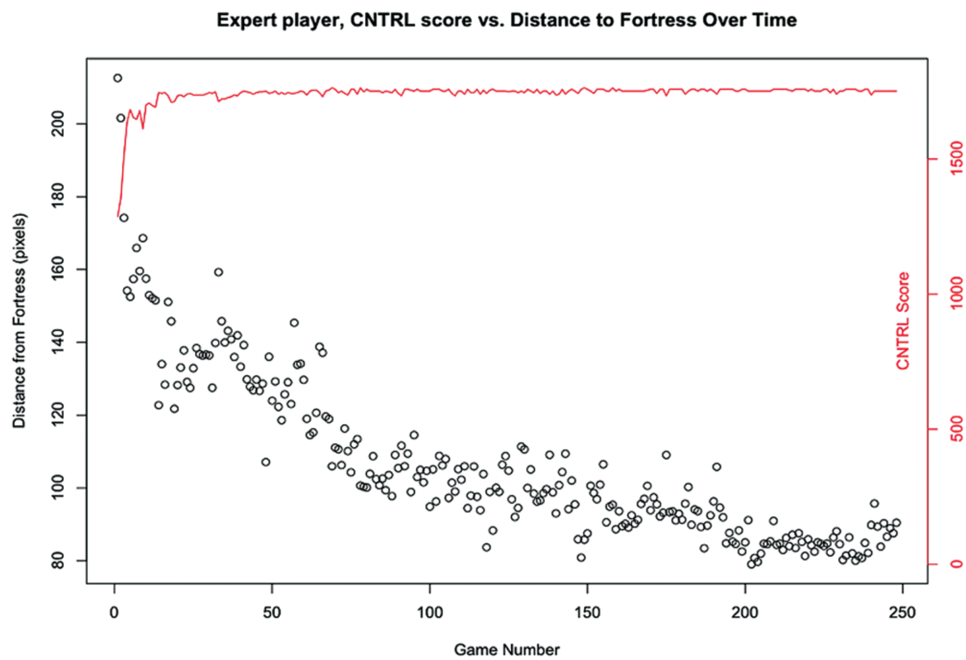


**Fig. 4.** Plot of one Space Fortress player's score on the CNTRL (Control) measure (shown in red at the top of the plot and on the right *y*-axis) versus the distance that player is actually flying from the Fortress (shown in black at the bottom of the plot and on the left *y*-axis). Although the player's distance follows the log-log law of improvement, the score rises rapidly across his first 10 games and then asymptotes from about Game 30 to Game 248. (Plot based on unpublished data collected by Destefano, 2010.)

(2003) tell us that "following a long period of touch-typing training, typists reach an average speed of 60 to 70 [words per minute (wpm)], whereas the average speed of very experienced visually guided typists is much lower (approximately 30–40 wpm)."

***Suboptimal strategies in architects.*** My colleague and I (Fu & Gray, 2004) reported the case of a professional architect who fit Carroll and Rosson's criteria of a professional daily user of productivity software and may have met Ericsson's criteria of an expert in his field. Many of the small, basic procedures he preferred to use in an architectural CAD/CAM system were an order of magnitude slower than the procedures recommended to him during training and in the manuals. In some cases, this would have resulted in a difference of 5 versus 50 seconds. Although 45 seconds wasted does not seem like a disaster, we estimated that if summed across the number of times daily in which the architect used the inefficient procedures, the total would add up to a daily waste of 30 minutes, or 6% of an 8-hour day.

### *Culturally inflicted limits: The case of digit-span experts*

What if we found that performance on an important IQ test, which purportedly measures individual differences within a population, reflected a population plateau due to the use of suboptimal methods, not individual asymptotes due to inherent mental capacity? Such is the case for digit span, and these facts have been well established for many years.

Following standard procedures, Chase and Ericsson read digits to their subject, SF, at the rate of one per second, followed immediately by ordered recall. When all digits were recalled correctly, the length of the next run of digits was increased by 1. Within 10 hours of training, SF improved from the population norm of 7 ± 2 to a span of 10. After 50 hours, he was at 20 and reached a span greater than 80 before leaving the study (Chase & Ericsson, 1982; Ericsson & Chase, 1982; Ericsson, Chase, & Faloon, 1980). This performance was later exceeded by that of another subject (Richman, Staszewski, & Simon, 1995). Performance on this task reflects plateaus due to culturally normative methods, not asymptotes due to limits in "the mechanism for the function itself" (Thorndike, 1913, p. 181).

## Asymptotes Versus Plateaus in Human Performance

Limits due to brain capacity reflect asymptotic performance. Increases in speed of processing, interconnections among cells in one part of the brain, and connections between different parts of the brain would provide the potential for a general increase in performance, much like the boosts when pole-vaulters switched from wood to bamboo to fiberglass poles.

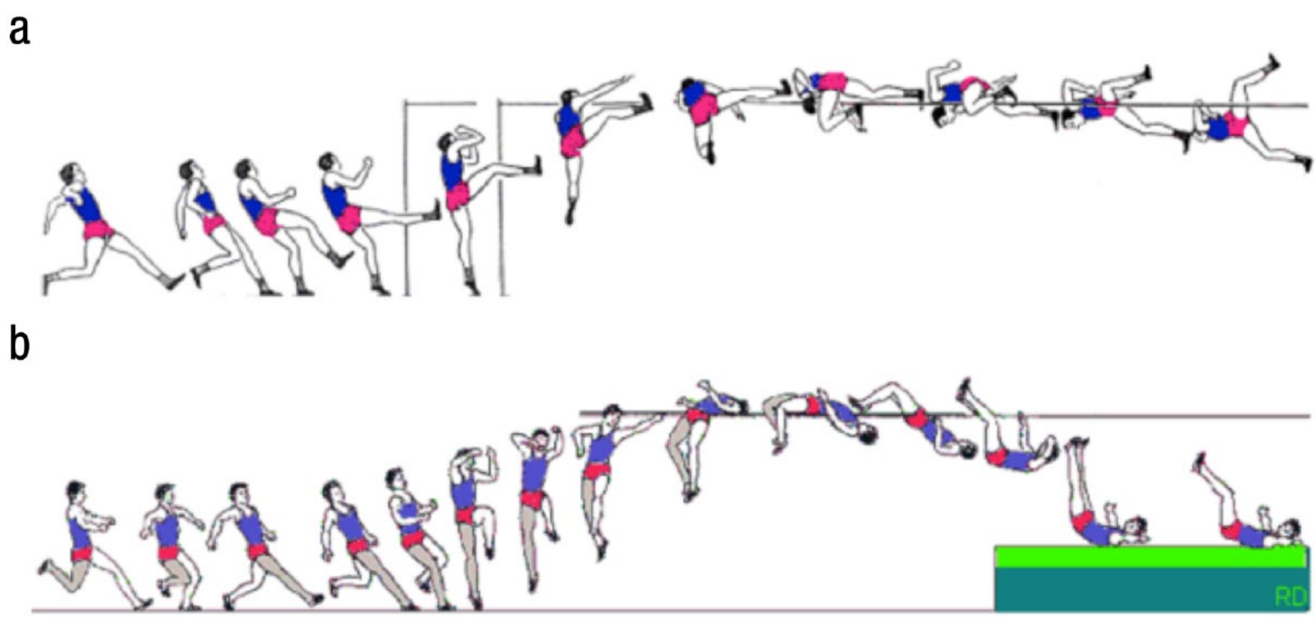


**Fig. 5.** Inventions in high jumping. See text for details. (Images reprinted from "Change Management and High Jump" by Carlos Lopez, December 28, 2011; retrieved from https://carloselopez.com/2011/12/28/change-management-and-high-jump. Reprinted with permission.)

Limits due to brain efficiency reflect plateaus in the performance of certain low-level cognitive tasks. If more efficient methods for performing those low-level tasks could be found and practiced, then higher-level skills that require these low-level components would be enhanced. In the sports world of high jumping, performance was limited by spurious limits (see Fig. 5a)—"levels or plateaus of efficiency which could be surpassed" (Thorndike, 1913)—and not by human physical abilities. The better method, the Fosbury Flop, first popularized by 1968 Summer Olympics gold medalist Dick Fosbury and now the dominant technique, represented an increase in efficient use of human abilities (see Fig. 5b) with no increase in human capacity.

The distinction between these two roads to increasing human performance can be very subtle but is theoretically and, maybe, practically important.

## Summary and Conclusions

The difference between a plateau and an asymptote is made clear by the existence of extreme experts with a known history of transcending the plateau of normal human performance. Hence, knowing that even one person was able to transcend "normal" strongly implies that the digit-span norm of 7 ± 2 is not an asymptote imposed by the hardware of the human brain. Rather, it is a spurious limit (in Thorndike's words) or plateau imposed by the methods used and the general unwillingness of people to put in the time and effort required to discover and practice new methods.

It is, of course, difficult to make any general recommendation for how to recognize a plateau in individual performance. One suggestion is to distinguish between asymptotes and plateaus by considering sudden drops, *dips*, and sudden increases, *leaps*, as well. My colleague and I discuss the plateau, dips, and leaps approach to the identification of periods of method discovery and change in Gray and Lindstedt (2016).

For researchers, the problem is not one of unwillingness to consider alternatives but unawareness that meaningful alternatives exist. For example, Keller (1958) excoriated Bryan and Harter (1899) for the training method they used and for the theory of hierarchical organization of learning they proposed. As a behaviorist, Keller would not have included hierarchies in his theoretical arsenal but would have included concepts that predict slow but steady increases in performance with practice. Today, almost 60 years after Keller's article, it seems a missed opportunity that rather than attacking the first theory of learning to emphasize hierarchical structures, Keller did not turn instead to investigating which conditions of practice led to plateaus and which led to asymptotes.

### Recommended Reading

Bryan, W. L., & Harter, N. (1899). (See References). A foundational work in experimental psychology, now much ignored, in which Bryan and Harter collected data from student telephone operators that was timestamped with millisecond accuracy; observing that "our evidence is that it requires ten years to make a thoroughly seasoned press

despatcher," they were the first to discover and announce the "10-year" rule of expertise.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). (See References). The foundational article for the concept of deliberate practice—a masterful, multifaceted data collection and analysis of the history and current habits of music students at the prestigious Hochschule der Künste in Berlin.

Foer, J. (2011, February 15). Secrets of a mind-gamer: How I trained my brain and became a world-class memory athlete. Retrieved from http://www.nytimes.com/interactive/2011/02/20/magazine/mind-secrets.html?_r=1&. Required reading for anyone who believes that memory experts are born, not made: a fascinating account by a journalist who took a year off to crash through his plateaus and win the U.S.A. Memory Championship.

Fu, W.-T., & Gray, W. D. (2004). (See References). Explores the mystery of stable but suboptimal performance from one expert, a group of graduate students, and subjects in a laboratory study.

Gray, W. D., & Lindstedt, J. K. (2016). (See References). Provides a briefer look at plateaus and asymptotes than the current article, but it quickly gets into the "dips" and "leaps" found when we stop averaging over groups of subjects and start focusing on trial-by-trial individual data.

## Declaration of Conflicting Interests

The author declared no conflicts of interest with respect to the authorship or the publication of this article.

## Funding

This work was supported in part by Grant N000141310252 from the Office of Naval Research (Dr. Ray Perez, Project Officer).

## Notes

1. "CPM" stands for both "critical-path method" and "cognitive, perceptual, motor," whereas "GOMS" stands for "goals, operators, methods, and selection rules."
2. Also known as *hunt-and-peck* or *Eagle Finger* typists.

## References

Boot, W. R., Basak, C., Erickson, K. I., Neider, M., Simons, D. J., Fabiani, M., . . . Kramer, A. F. (2010). Transfer of skill engendered by complex task training under conditions of variable priority. *Acta Psychologica*, *135*, 349–357.

Boot, W. R., Sumner, A., Towne, T., Rodriguez, P., & Ericsson, K. A. (in press). Applying aspects of the expert performance approach to better understand the structure of skill and mechanisms of skill acquisition in video games. *Topics in Cognitive Science*.

Bryan, W. L., & Harter, N. (1897). Studies in the physiology and psychology of the telegraphic language. *Psychological Review*, *4*, 27–53.

Bryan, W. L., & Harter, N. (1899). Studies on the telegraphic language: The acquisition of a hierarchy of habits. *Psychological Review*, *6*, 345–375.

Carroll, J. M., & Rosson, M. B. (1987). Paradox of the active user. In J. M. Carroll (Ed.), *Interfacing thought: Cognitive aspects of human-computer interaction* (pp. 80–111). Cambridge, MA: MIT Press.

Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. *The Psychology of Learning and Motivation*, *16*, 1–58.

Crossman, E. R. F. W. (1959). A theory of the acquisition of speed-skill. *Ergonomics*, *2*, 153–166.

Destefano, M. (2010). The mechanics of multitasking: The choreography of perception, action, and cognition over 7.05 orders of magnitude (Unpublished PhD thesis). Rensselaer Polytechnic Institute, Troy, NY.

Donchin, E. (1989). The Learning Strategies project: Introductory remarks. *Acta Psychologica*, *71*, 1–15.

Donchin, E. (1995). Video games as research tools: The Space Fortress game. *Behavior Research Methods, Instruments, & Computers*, *27*, 217–223.

Dunning, D., & Holmes, J. (2014). Does working memory training promote the use of strategies on untrained working memory tasks? *Memory & Cognition*, *42*, 854–862. doi:10.3758/s13421-014-0410-5

Ericsson, K. A., & Chase, W. G. (1982). Exceptional memory. *American Scientist*, *70*, 607–615.

Ericsson, K. A., Chase, W. G., & Faloon, S. (1980). Acquisition of a memory skill. *Science*, *208*, 1181–1182.

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

Frederiksen, J. R., & White, B. Y. (1989). An approach to training based upon principled task decomposition. *Acta Psychologica*, *71*, 89–146.

Fu, W.-T., & Gray, W. D. (2004). Resolving the paradox of the active user: Stable suboptimal performance in interactive tasks. *Cognitive Science*, *28*, 901–935.

Gillette, A. L. (1936). Learning and retention: A comparison of three experimental procedures. *Archives of Psychology*, *198*, 5–56.

Gopher, D., Weil, M., & Bareket, T. (1994). Transfer of skill from a computer game trainer to flight. *Human Factors*, *36*, 387–405.

Gray, W. D., John, B. E., & Atwood, M. E. (1993). Project Ernestine: Validating a GOMS analysis for predicting and explaining real-world performance. *Human–Computer Interaction*, *8*, 237–309.

Gray, W. D., & Lindstedt, J. K. (2016). Plateaus, dips, and leaps: Where to look for inventions and discoveries during skilled performance. *Cognitive Science*. Advance online publication. doi:10.1111/cogs.12412

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences, USA*, *105*, 6829–6833.

Keller, F. S. (1958). The phantom plateau. *Journal of the Experimental Analysis of Behavior*, *1*, 1–13.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1–55). Hillsdale, NJ: Erlbaum.

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., . . . Engle, R. W. (2012). No evidence of intelligence improvement after working memory training: A randomized, placebo-controlled study. *Journal of Experimental Psychology: General*, *142*, 359–379. doi:10.1037/a0029082

Richman, H. B., Staszewski, J. J., & Simon, H. A. (1995). Simulation of expert memory using EPAM-IV. *Psychological Review*, *102*, 305–330.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, *138*, 628–654. doi:10.1037/a0027473

Thorndike, E. L. (1913). *Educational psychology: Vol. II. The psychology of learning*. New York, NY: Teachers College, Columbia University.

Underwood, B. J. (1954). Speed of learning and amount retained: A consideration of methodology. *Psychological Bulletin*, *51*, 276–282.

von Bastian, C. C., & Oberauer, K. (2014). Effects and mechanisms of working memory training: A review. *Psychological Research-Psychologische Forschung, 78*(Suppl. 6), 803–820. doi:10.1007/s00426-013-0524-6

Yechiam, E., Erev, I., Yehene, V., & Gopher, D. (2003). Melioration and the transition from touch-typing training to everyday use. *Human Factors*, *45*, 671–684.