Performance Efficiency:
A Metric and Research Methodology for Task Analysis

David Lewis, Ph.D.
Nova Southeastern University, 1750 NE 167th St 409, Miami, FL 33162-3017

Abstract

This paper describes a metric called "performance efficiency," and its use in cognitive task analysis. This metric provides a means of determining the learning efficiency of instructional conditions. Performance efficiency will be described in the context of recording technologies that are often used in software usability studies. While usability is often considered in the programming of software environments, the "learnability" of these environments is more the concern of instructional designers. The advantages and disadvantages of these types of metrics and methodologies will be discussed in detail. Thus the purpose of this paper is to consider the applications of the "performance efficiency" metric to the design of instructional materials.

Introduction

Educational researchers have used a medical model to develop instructional materials; that is we hope to design instruction, which is both efficient and effective (Lewis & Barron, 2009). Gagné (1964) was one of the earliest educators to describe these two general categories of dependent variables. He proposed most educators are concerned with (1) "the rate of attainment of some criterion performance" (efficiency) and (2) "the degree of correctness of this performance" (effectiveness) (Gagné, 1964, p.295). It is an underlying theme of this paper that when these variables are applied to the design of instructional materials, we are considering the "learnability" of the instruction.

Nielsen (1993) defined usability by developing several subcomponents (learnability, efficiency, memorability, errors, and satisfaction). Soloway, Guzdial, and Hay (1994) called for Norman's "user-centered" design philosophy to be more "learner-centered." Nielsen's (1993) definition of learnability "How easy is it for users to accomplish basic tasks" is a subjective measure of "perceived usability," rather than a more objective measure. However, Nielsen (2001) proposed we should consider the user's opinions and suggestions, *but* only after actually watching them work with the software. That is we must start by observing learners before considering their perceptions. It's interesting that if we were to look to the international standards organization (ISO) for a definition of usability, we would find that they also chose to use Gagné's variables (ISO 9241-11, 1998).

While cognitive load may not seem related to usability, similarities reveal themselves if you consider the measures underlying this theoretical framework. Cognitive load theory is an instructional theory that is concerned with the learnability of instructional materials. This theoretical framework has become quite influential within the field of instructional design (Ozcinar, 2009; Paas, van Gog, & Sweller, 2010). Cognitive load theory is primarily concerned with procedural knowledge, task performances and problem-solving. Cognitive load measures are a combination of subjective data (mental-effort ratings) and performance scores (Tuovinen & Paas, 2004). These measures have been found to be reliable and correlated with error rates (Ayres, 2006) but not all cognitive load theorists agree with the use of subjective measures, and have proposed we consider more direct or objective measures (Brünken, Plass, & Leutner, 2003; Whelan, 2007). This concern has led to the impetus for this paper and the "performance efficiency" metric described in the next couple of sections. This type of research (task analysis) has a rich history and is certainly a metric to be used in cognitive task analysis.

Task Analysis

Task analysis researchers have used observation or photography/videography as a means of data collection for decades (Clark & Estes, 1996; Gilbreth & Gilbreth, 1917). Some of the earliest task analysis studies were made with stopwatches and the newly developed technology of chronocyclegraphy (Gilbreth & Gilbreth, 1917). This was the use of long exposure photography, which allowed for the detection of movements over time. While the Gilbreths were early pioneers of time motion studies, even they were aware of the underlying rationale for this type of research. They state it when they say "that the learner shall be taught the best way immediately, that is, from the beginning of his practice" (Gilbreth & Gilbreth, 1917, p.82). So they were amongst the first to promote efficient instruction, for it allows a learner to be more efficient with their time, and simply learn more.

In the 1970s and 80s, researchers began to realize cognitive processes controlled the behavior of those performing tasks (Clark & Estes, 1996). So it was then, when Psychology took a major step forward to develop cognitive task analysis. The Cognitivists analyzed task performance, but were aware of the decision making processes that occur during problem solving (van Merriënboer, 1997). Therefore cognitive task analysis takes in to account these cognitive processes, to subdivide complex tasks into their component parts in order to support a learner's performance (van Merriënboer, 1997).

Early researchers developed stepwise processes for analyzing learner behavior as they interacted with computers (Clark, Feldon, van Merriënboer, Yates, Early, 2008). Card, Moran, and Newell (1983) produced a seminal work in this arena as they developed a form of cognitive task analysis called GOMS. A GOMS analysis produces a step-by-step text-based description of the procedural knowledge required to accomplish a task (Card, Moran, & Newell, 1983; John & Kieras, 1996a). GOMS is an acronym which divides the components of computer-based problems and learner actions into goals, operators, methods, and selection rules. The GOMS model analyzes computer-based problems by reducing them to solution steps, known as problem solving operators (Dix, Finalay, Abowd, & Beale, 2003).

Even before GOMS existed, Merrill (1971) had proposed instructional designers use an "information processing" perspective toward task analysis as they design instructional materials. Later, researchers found that a GOMS analysis resulted in positive modifications to instructional materials (Elkerton & Palmiter, 1991; Steinberg & Gitomer, 1992; Sullivan, Ortega, Wasserberg, Kaufman, Nyquist, Clark, 2008). Therefore there is certainly precedence for the use of GOMS analyses in the design of instructional materials. One reason for this is because experts often have difficulty articulating exactly how they perform procedures (Villachica & Stone, 2010). This type analysis allows for the instructional designer to work with a series of subject matter experts and synthesize their input to develop a more comprehensive product (Sullivan, et al., 2008). Indeed several studies have found empirical evidence that support the efficacy of CTA-based instruction (Clark et al., 2008). It is with this thought in mind that researchers are beginning to consider cognitive task analysis within the context of cognitive load theory (e.g. Shachak, Hadas-Dayagi, Ziv, & Reis, 2009) for they can be used together within a GOMS-like analysis (Lewis, 2008). It is in this way that a cognitive task analysis may provide the means to test the predictions of cognitive load theory.

Efficiency and Effectiveness

So what is needed is a way of producing instructional design guidelines that are based on empirically driven studies. Researchers need an objective method of assessing and evaluating instructional materials which 1) is based upon Soloway's learner-centered design philosophy (Soloway et al., 1994), and 2) allows researchers to analyze instructional strategies, to find those which are the most efficient and effective. Cognitive load theory has provided us with several constructs that compare instructional conditions (Brünken, Seufert, & Paas, 2009; Tuovinen & Paas, 2004). Paas and van Merriënboer initially developed the primary cognitive load construct, "relative condition efficiency" (Paas & van Merriënboer, 1993; Sweller, van Merriënboer, & Paas, 1998). Relative condition efficiency combines learner performance scores and mental effort ratings to compare instructional conditions (See Equation 1).

$$\text{Relative condition efficiency} = \frac{Z_{Performance} - Z_{MentalEffort}}{\sqrt{2}} \qquad (\mathbf{1})$$

The resulting data is graphed on a biplot (See Figure 1) to allow researchers to compare the relative efficiency of instructional conditions. Since its development, relative condition efficiency has become an important basis for much of cognitive load research (Paas, Tuovinen, Tabbers, & Van Gerven, 2003). However, this measurement relies on indirect or subjective mental effort ratings (Brünken, R., Plass, J. L., & Leutner, D., 2003). Paas and van Merriënboer (1993) were aware of this limitation and even state in their original article that this construct should be qualified with performance data. While "relative condition efficiency" is a measure of the cognitive efficiency of an instructional condition, because it uses mental effort ratings, it does not include time in the equation. Therefore it is difficult to truly analyze the efficiency of the learner's performance given an instructional condition.

While learners are able to reliable report their perceptions of the instructional conditions (Ayers, 2008), it would be helpful to truly be able to gauge learner performance over time. Learners become more efficient each time they solve similar problem solving scenarios, as they acquire the underlying problem schema. Lewis (2008) synthesized each of these ideas to develop another construct, called "performance efficiency." Performance efficiency is similar to relative condition efficiency, but uses performance time rather than a mental effort rating (compare Equations 2 and 3).

$$\text{Relative condition efficiency} = \frac{Z_{Performance} - Z_{MentalEffort}}{\sqrt{2}} \qquad (\mathbf{2})$$

$$\text{Performance efficiency} = \frac{Z_{Performance} - Z_{PerformanceTime}}{\sqrt{2}} \qquad (\mathbf{3})$$

Performance efficiency was developed to complement rather than compete with relative condition efficiency. Therefore it is hoped that this separate efficiency metric may be used to strengthen cognitive load research and to analyze instructional conditions in a similar way. Performance efficiency contrasts instructional conditions in much the same manner as relative condition efficiency, to combine Gagné's dependent variables in a biplot (see Figure 2) with group performance times and performance scores.

A generalizable research methodology

Generalizablity is perhaps the most valued aspect of any task analysis methodology (Crandall, Klein, & Hoffman, 2006). The methodologies employed by the Gilbreths were generalizable and still useful today. Software task analysis studies are a modification of those early time motion studies, but now aimed at computer-based environments. Software packages like Techsmith Morae (Techsmith, 2011) are used to record user interactions with software. These software solutions can be used to compare the learnability of instructional conditions (Lewis, 2008). These computer-based solutions are a generalizable methodology for conducting a cognitive task analysis. Like the time motion studies of the early 20th century, today's computer based task analysis studies use time as a dependent variable. Certainly recordings have been used for years, but recordings of a learner's onscreen actions, during computer-based training is relatively new. Recordings free the researcher from the constraints of time, allow researchers to review, categorize and analyze the learner's performance. If necessary, they can watch a learner as they perform some behavior repeatedly, in order to document multiple aspects of that behavior (Lewis, 2008; Martin & Bateson, 2007). This is because a researcher can easily rewind the recordings to document multiple outcome variables, which may have occurred simultaneously. Thus researchers do not have to document learner behaviors as they occur, because this methodology allows these researchers to review learner actions weeks or months after the actual behavior.

Gagné's variables may be measured with recording software like TechSmith Morae (Techsmith, 2011). Learners are given credit for completing a problem solving operator (Lewis& Barron, 2009). To tally completion of a problem solving operator, researchers may use a rubric or checklist based upon the problem being solved (See Table1). In this table the problem solving operators are listed in the table to the left and then the operator is checked in the table cells (usually with a point value). This is made easier by using a spreadsheet application to tabulate scores automatically. Finally, performance time may be measured in seconds with the recording software. This software allows a researcher to mark the timeline of a learner's onscreen recording and to designate the beginning of a performance, in this case with the first initial mouse movement. Once data points are found per learner then a large spreadsheet of data points may be made. For more information on calculating the metric values consider other more detailed sources (Clark, Nuguygen, Sweller, 2006; Lewis, 2008).

Conclusions

The most important implication of this research is that it produced a new generalizable metric for contrasting instructional conditions. This metric considers performance outcomes as they relate to the design of instructional materials. This is as opposed to the cognitive efficiency of the materials which is the role of relative condition efficiency. It is hoped that these metrics can be used together to promote cognitive load research. This performance efficiency metric was an outcome of a cognitive task analysis, and a development of a dissertation study (Lewis, 2008; Lewis & Barron, 2009). An important implication of this study was that it demonstrated a new methodology for analyzing both learner tasks, and on-screen behaviors of the learner. This analysis provided instructional design researchers a generalizable methodology of evaluating learner responses (documenting learner errors & problem solving operators) within a computer based learning environment.

Recording learner onscreen actions is generalizable to any e-learning environment. It uses “off-the-shelf” software (Techsmith Morae) to document and analyze learner behavior. Along with the performance efficiency metric this methodology allowed for a comparison of instructional conditions. Each of these can be used together with cognitive load measures as a research methodology to compare and contrast instructional materials. Task analysis researchers have used observation (via photography/videography) for decades, this is yet another methodology. It is hoped that others make uses of these methods and find them useful as a means of improving instructional materials.

References

Ayers, P. (2006). Using subjective measures to detect variations of intrinsic cognitive load within problems. *Learning and Instruction 16* (5) 389-400.

Brünken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning, *Educational Psychologist, 38*(1), 53–61.

Brünken, R., Seufert, T., & Paas, F. G. W. C. (2009). Measuring Cognitive Load. *In Cognitive Load Theory.* Plass, Jan L., Moreno, R., Brunken, R. (Eds). Cambridge: Cambridge University Press

Card, S.K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Clark, R. E., Feldon, D., Van Merriënboer, J. J. G., Yates, K., & Early, S. (2008) Cognitive Task Analysis. In Spector, J. M., Merrill, M. D., van Merriënboer, J. J. G., & Driscoll, M. P. (Eds.). *Handbook of research on educational communications and technology (3rd ed.).* Mahwah, NJ: Lawrence Erlbaum Associates.

Clark, R.C., Nguyen, F., and Sweller, J. (2006). *Efficiency in learning: evidence-based guidelines to manage cognitive load*. San Francisco: Pfeiffer.

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

Gilbreth, F.B. & Gilbreth, L. M. (1917). *Applied motion study: a collection of papers on the efficient method to industrial preparedness*. New York: Sturgis & Walton Company

Elkerton, J. & Palmiter, S. (1991). Designing Help Using a Retrieval GOMS Model: An Information Evaluation. *Human Factors, 33*(2) 185-204.

ISO 9241-11 (1998). *Ergonomic requirements for office work with visual display terminals (VDTs) — Part 11: Guidance on Usability*. Retrieved from http://www.iso.org/iso/catalogue_detail.htm?csnumber=16883

Kirschner, P, A., Ayres, P., & Chandler, P. (2011). Contemporary cognitive load theory research: The good, the bad and the ugly. *Computers in Human Behavior, 27*(1) 99–105

Lewis, D. (2008). *The acquisition of procedural skills: An analysis of the worked-example effect using animated demonstrations*. (Doctoral dissertation) (UMI No. AAT 3376198)

Lewis, D., & Barron, A. (2009, July). Animated demonstrations: Evidence of improved performance efficiency and the worked example effect. *Paper presented at the HCI International Conference 2009, San Diego, CA, USA*. doi: 10.1007/978-3-642-02806-9_29

Merrill, P.F. (1971). *Task analysis: an information processing approach. CAI Center Tech Memo Number 27.* (ERIC Document Reproduction Service No. ED050554)

Martin, P., & Bateson, P. (2007). *Measuring behaviour: an introductory guide*. (3nd ed.) Cambridge: Cambridge University Press

Nielsen, J. (1993). *Usability engineering*. Boston: Academic Press.

Nielsen, J. (2011) *First rule of usability? Don't listen to users.* Retrieved from http://www.useit.com/alertbox/20010805.html

Ozcinar, Z. (2009). The topic of instructional design in research journals: A citation analysis for the years 1980–2008. *Australasian Journal of Educational Technology, 25*, 559–580.

Paas, F., and van Merriënboer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combine mental-effort and performance measures. *Human Factors, 35*(4), 737-743.

Paas, F., Tuovinen, J.E., Tabbers, H.K., & Van Gerven, P.W.M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38* (1) 63–71. doi:10.1207/S15326985EP3801_8

Paas, F., van Gog, T., & Sweller, J. (2010). Cognitive load theory: New conceptualizations, specifications, and integrated research perspectives. *Educational Psychology Review 22,* 115–121 doi: 10.1007/s10648-010-9133-8

Shachak, A. Hadas-Dayagi, M. Ziv, A., & Reis, S. (2009). Primary Care Physicians' Use of an Electronic Medical Record System: A Cognitive Task Analysis. *Journal of General Internal Medicine, 24* (3) 341-348. doi: 10.1007/s11606-008-0892-6

Soloway, E., Guzdial, M., & Hay, K. H. (1994). Learner-centered design: The challenge for HCI in the 21st century. *Interactions, 1*(2), 36-48.

Steinberg, L.S., and Gitomer, D.H. (1992). *Cognitive Task Analysis, Interface Design, and Technical Troubleshooting*. (ERIC Document Reproduction Service No. ED384677)

Sullivan, M.E., Ortega, A., Wasserberg, N., Kaufman, H., Nyquist, J., Clark, R. (2008). Assessing the teaching of procedural skills: can cognitive task analysis add to our traditional teaching methods? *The American Journal of Surgery* 195, 20–23

Sweller, J., Van Merriënboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.

Tuovinen, J. E. & Paas, F. G. W. C. (2004). Exploring multidimensional approaches to the efficiency of instructional conditions. *Instructional Science 32*, (1-2) 133–152.

Techsmith (2011). TechSmith Morae 3.2 [Computer program] Okemos, MI

van Merriënboer, J. (1997). *Training complex cognitive skills: A four-component instructional design model for technical training*. Englewood Cliff, New Jersey. Educational Technology Publications.

Villachica, S. W., & Stone, D. L. (2010). Cognitive task analysis. *In K. H. Silber & W. R. Foshay (Eds.), Handbook of improving workplace performance* (Vol. 1: Instructional design and training delivery, pp. 227-258). San Francisco, CA: Jossey-Bass/Pfeiffer.

Whelan, R.R. (2007). Neuroimaging of cognitive load in instructional multimedia. *Educational Research Review 2* (1) 1-12

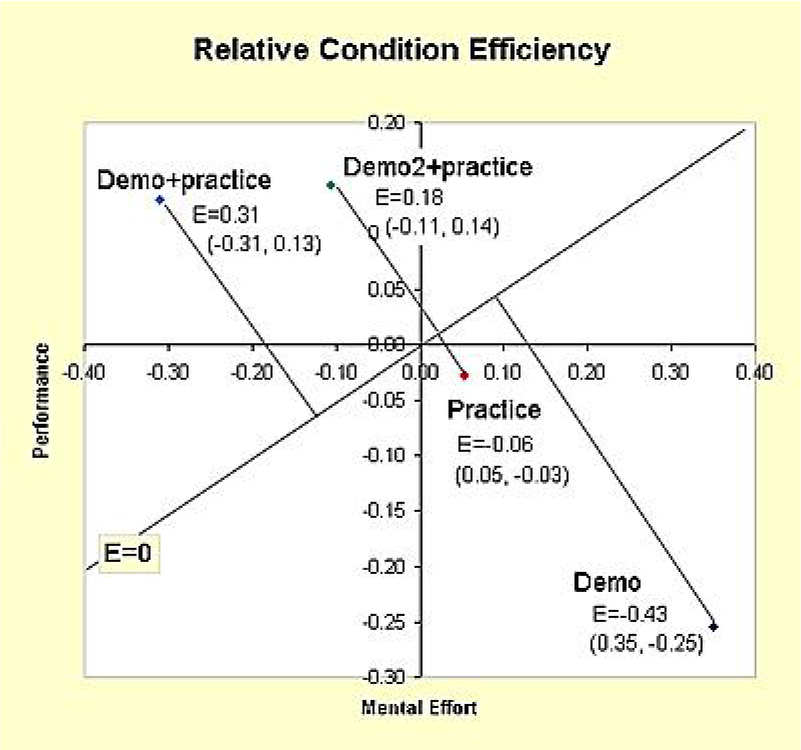

Figure 1. A biplot of Relative Condition Efficiency

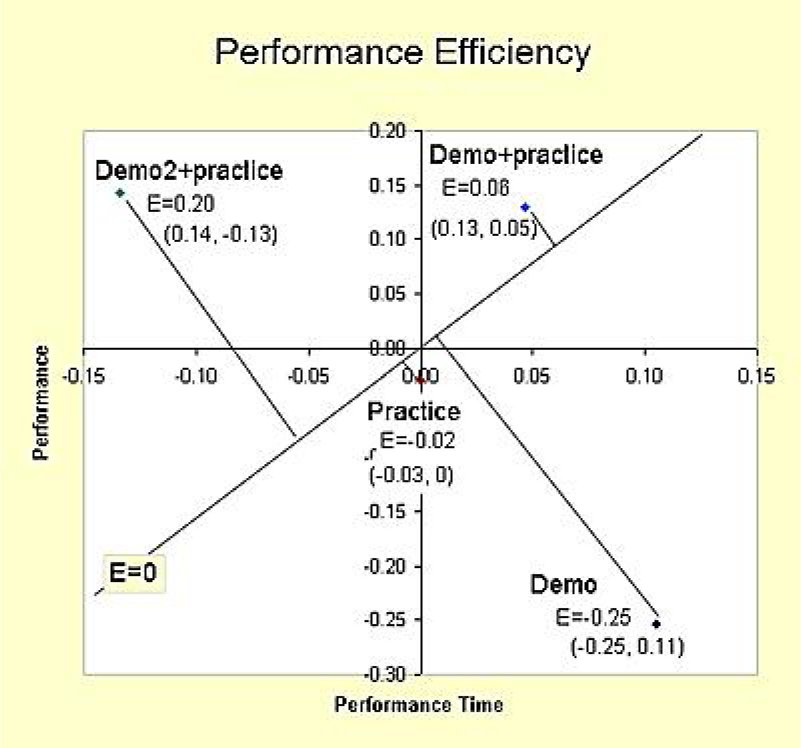

Figure 2. A biplot of Performance Efficiency

Table 1

Picnic problem accuracy rubric

flip	**layer**	**rotate**	**move**	**item**
	***			umbrella
	***			tshirt
***	***			head
***	***	***		right leg
***	***			head 2
***	***			purple shirt
***				hat
***	***	***		s left leg
***	***	***		bent right leg
***	***	***		left leg
***				green shorts
***	***	***		arm 2
***				pink shorts
***	***	***		left arm
***	***			body
***	***			picnic basket
***	***	***		arm
***	***			right arm
***	***			torso
***	***			table
***	***			bird3
***	***	***		bird2
***	***			bird1
0	0	0	0	0