Sequential Analyses of Error Rate: A Theoretical View

Ronald Mellado Miller and Richard J. Sauque

1Brigham Young University – Hawaii,
55-220 Kulanui Street #1896, Laie,
HI 96762 USA
millerr@byuh.edu

Abstract. Though error rate is a ubiquitous measure of human performance, as typically measured in terms of overall error rate or percentage, there are a number of predictive variables lost by summing or averaging the errors made. In this paper, we present a sequential analysis of error rate, where the pattern of errors is analyzed. By examining such concepts as the number of transitions from incorrect responses (I) to correct responses (C) or IC transitions as well a concept called I-length, which refers to the number of incorrect responses followed by a correct response, valid ordinal predictions of persistence in the face of continuous failure can be made. This paper develops this theoretical construct in the hopes that utilizing such data will facilitate the analysis and predictive quality of error rate data.

Keywords: error rate, sequential, performance, percentage, extinction, persistence.

1 Introduction

If a person had bought a high quality new car and that car had worked flawlessly for a year, but, one morning, simply stopped functioning, how long would that person persist in trying to start that car? It had always worked with no difficulties over an extended period of time, its reliability had been shown to be high, etc. Under such circumstances, most people would try a few times and then realize they needed to take the car to the mechanic. On the other hand, think of the car that person might have used as a student during their undergraduate days. It was probably unreliable and may not have started each time—it also may have had other quirks that made it difficult to use. With persistence and the right amount of luck, the car would start. If that car stopped working that morning and, unbeknownst to the driver, it would never start again, how long would that driver continue to try to start their car? It had been unreliable, so it was clear that more effort would be needed to be sure the car really was inoperable. In this situation, the driver might attempt many times to start the car, and certainly many more times than did the driver of the high quality car.

The difference in persistence between the high quality and low quality cars exemplify at least two important points. The first is that that we may spend more time working with a historically unreliable system than with one which has given us better

D. Harris (Ed.): Engin. Psychol. and Cog. Ergonomics, HCII 2007, LNAI 4562, pp. 375–378, 2007.

results in the past. This argues against most rational explanations of behavior because we are expending the least amount of effort on the system that has given us the most success in the past. However, when we look at it as a subset of pattern learning, in that the respective drivers had learned what patterns of behavior for a particular system lead to success, then the actions make sense. The other point is that error rate alone does not tell us the nature of the problem. The unreliable car had a higher error rate, in that it started less often than the reliable car, yet it produced the most persistence in the face of failure. For many human-computer interaction studies, the analysis of error rates plays a key role in evaluating the utility and functionality of processes, components, and users. While the concept of utilizing error rates to determine such functionality is clear, as, generally, lower error rates are desirable and processes are designed to monitor and minimize errors, for the end-user, error rate as typically measured becomes problematic. This paper seeks to establish a theoretical model of a sequential analysis of error rates with its major purpose being to increase the utility and informativeness of error rates in general.

2 Sequential Theory

Error rates are often given either as frequency data (how many errors occurred) or as a percent of errors (what proportion of responses were, in some way, erroneous), with those items producing lower error rates being preferred over those with higher error rates. However, even examining simple sequences made up of 50% error, CCII, IICC, ICIC, where "C" refers to correct performance trials and "I" to incorrect performance trials, though having equal percentages and frequencies of error, can lead to distinct behaviors in human end-users. For example, recent studies have indicated that not only do these patterns influence instrumental conditioning (Capaldi & Miller, 2004), where a participant must perform an action to cause the trial to be labeled as correct or incorrect, but can also influence variables dealing with Pavlovian or classical conditioning as well (Miller & Capaldi, 2006).

In the example that follows, extended training, such that the subject learns the pattern of the responses (correct/incorrect) is assumed. In this sense, we are describing expert performance. For example, given the series CCII, subjects learn that their correct responses are followed by incorrect, or that success is followed by failure. Given this pattern, it would be assumed that once an incorrect trial was encountered, the subject would then anticipate further failure, allowing their effort to lag. In other words, the memory stimulus of having a correct response would come to predict future incorrect responses, $S^{Correct} \longrightarrow$ Incorrect. In the case of an IICC series, the opposite would be assumed to occur, as, here, incorrect trials are followed by correct ones, persistence in the face of failure is followed by success, $S^{Incorrect} \longrightarrow$ Correct. Thus, in the IICC series, persistent and continuous effort in the face of failure would be trained, which would not be the case in the CCII series. The ICIC series would be assumed to create an intermediary condition, where persistence would be trained but not to the extent of the IICC series, where more failure was encountered before success. It is important to note that the IICC, CCII, and ICIC series each have the same percent correct, the same error rate percentage, and have an equal number of successful and unsuccessful trials. It is the pattern of those responses which influence consequent persistence. Here, a concept which can be called I-length, or the number of incorrect trials followed by a correct trial may be useful. With sufficient training, a would be expected in expert performance, longer I-lengths should lead to greater persistence in the face of constant incorrect responses (failure). As an additional, albeit extreme condition, a series such as IIIIIIC (I-length 6), $S^{6\ \text{Incorrect}} \longrightarrow$ Correct, with only a 16.7% success rate would be expected to produce more persistence than an I-length 0 series (CIIIIII) with the same percentage of success, an I-length 3 series (IIICIIIC) with a 25% success rate, or a CCCCCCCC (100% success) series, $S^{\text{Correct}} \longrightarrow$ Correct, where no experience with incorrect answers occurs and where success is only followed by additional successes. In the 100% success series, if switched to a situation where incorrect responses occurred, as the incorrect responses would be novel stimuli with no training, the participant would be expected to not persistent and cease responding rapidly. Thus, it is possible, and has been shown, that under the proper conditions, even series which produce the lowest success rates can also produce the highest sustained persistence. Interestingly, this indicates that participants sometimes persist least in conditions where they have enjoyed the greatest success in the past, while they may persist, and thus put forth the most effort, in cases where, historically, they have had the most failure. This is in contrast to some predictions of various optimization theories which suggest that participants would be the most persistent in tasks which, historically, have given them the greatest percentages of successes over time.

Additionally, sequential theory (Capaldi, 1994) also suggests that if the subject receives very little training, such that the pattern of responses (incorrect/correct) is not yet learned, as when a novice undertakes a task, then the patterns described reverse in their ordinal position. Thus, in this condition, for example, the IIIC pattern would be expected to produce less persistence than an ICIC one. The basic concept here is one of IC transitions or, how many times has the subject been exposed to incorrect responses followed by correct responses. Early in training, therefore, the group with the most IC responses is expected to persist the longest. Thus, being trained ICICIC would lead to more persistence than an IIICCC series, because in the former, three IC transition occurred and In the latter, only one was trained. Only after sufficient training do the I-lengths supplant the IC transitions in determining ordinal persistence in the face of extinction (all incorrect responses).

Data that support such a sequential view range from animal studies (Capaldi & Miller, 2004) to human studies (Miller, Terry, & Johnson, 2005) from simple alleyway running in rats to software usage in humans, and, anecdotally, in aviation maintenance training regimens. In general, it may be stated that such an analysis can inform training regimens across a wide field of areas. The major effect of taking into account such sequential variables would be to better understand why participants persist as some tasks more than others, even when error rates/frequencies are the same or similar, as well as ensuring that training takes into account the sequential effects of error. For example, in tasks where persisting in a task might damage equipment or cause other untoward consequences, the training itself can be built around sequential contingencies designed to minimize persistence. In the case where persistence in the face of repeated failures would be valuable, such sequential contingencies can also be programmed as part of the training regimen. In each case, it would be important to determine if the subjects of the training were considered to be novices or experts so that the training would emphasize IC transitions for novices, moving to a focus on I-length series for experts.

As one example of a training contingency, a simple machine system might be used. If the machine was unreliable or required a great deal of effort or steps to accomplish then training should include the worker laboring until the task was finished without interruption. In this way, an undetermined I-length would be conditioned. A common mistake in training novices in this situation would be having the expert interrupt them and finish the work, thus making it so that the novice is not conditioned to experience success (IIIII… only and not III…C) and comes to rely on the expert to finalize their labors. Similarly, an expert who has been trained in long I-lengths and so persists in accomplishing a task might make mistakes with a new process where continual unsuccessful attempts may damage the machinery, etc., and so would need to be retrained in the new CCCIIII… contingency where incorrect responses or failure is only followed by more failure, etc. While overall percentages of error rate might be useful to distinguish worker's quality, only by examining the sequence of their errors would other useful information predictive of their behaviors in the face of machine/system failures be gathered.

3 Conclusion

Thus, it is hoped that such an introduction to sequential theory, IC transitions, and I-length will be of benefit to those whose needs include anticipating human responding in the face of failure or continued incorrect responses. While error rate as whole is a valuable measure, by examining the particular sequences that make up the overall error patterns, error rate becomes at once more predictive and utilitarian.

References

1. Capaldi, E.J.: The sequential view: From rapidly fading stimulus traces to the organization of memory and the abstract concept of number. Psychonomic Bulletin & Review 1, 156–181 (1994)
2. Capaldi, E.J., Miller, R.M.: Serial learning in rats: A test of three hypotheses. Learning and Motivation 35, 71–81 (2004)
3. Miller, R.M., Capaldi, E.J.: An Analysis of Sequential Variables in Pavlovian Conditioning Employing Extended and Limited Acquisition Training. Learning and Motivation 37, 289–303 (2006)
4. Miller, R.M., Terry, R., Johnson, E.: An analysis of error rate as a predictor of user persistence in machine systems. In: Proceedings of the 11th International Conference on Human-Computer Interactions, vol. 4, Lawrence Erlbaum Press, Mahwah (2005)
5. Miller, R.M., Olds, K., Benson, J., Strong, N.: Question Order Effects on Task Persistence in Computerized Testing. In: Proceedings of the 10th Annual Global Chinese Conference on Computers in Education (2006)