ascilite
ascilite 2012
FUTURE CHALLENGES | SUSTAINABLE FUTURES
25 - 28 November 2012 | Wellington | New Zealand

Online training: Sustainability in the face of the unskilled, unsupervised, unmotivated "long tail"

Lisa Wise
Faculty of Higher Education, Lilydale
Swinburne University of Technology

Benedict Williams
Faculty of Life and Social Sciences
Swinburne University of Technology

Gregor McLean
Air Operations Division
Defence Science and Technology Organisation

Jason Skues
Faculty of Higher Education, Lilydale
Swinburne University of Technology

Communities of discourse and market places often follow power laws with "long tailed" distributions. The present student "voucher" system and uncapped demand in higher education leads to an analogous "long tail" in the tertiary sector. We argue that student attrition in the "long tail" is part of a natural democratic process of "churn" - legitimate peripheral participation which leads some students to the realisation that they are not sufficiently motivated or prepared for fuller participation in the academic community. We illustrate this idea with data from a task in an introductory psychology program to examine the relationship between persistence, performance, motivation and demographics. We found distinct subgroups whose performance is defined by levels of conscientiousness/motivation and initial skill. We suggest that it may be counterproductive to retain students who perform poorly or are disengaged. However, we also identified different patterns of performance and suggest that attrition can be attenuated by targeted interventions to improve initial performance of identified groups.

Keywords: online learning, long tail, motivation, mastery

Introduction

The long tail

Online social networks generate unequal distributions of audience in the face of broad availability and choice
Shirky (2008) in an influential blog entry entitled "Power laws, weblogs and inequality", argues that internet blog audiences follow a power law rather than a normal distribution, leaving a "long tail" of blogs with low readership, due to the nature of social networks rather than the intrinsic quality of these blogs. More popular blogs by definition have more incoming links as a direct function of their popularity, and thus will become even more popular. While the internet provides a vast potential readership for online material, blogs and websites that are already popular will have a greater influence in terms of determining what else their readers engage with. Although it is possible for new blogs to make an impact in the marketplace, it will be more difficult, and irrespective of quality, there will be a good deal of churn in new blog generation and new audience due to the nature of power laws that operate within social networks. This paper argues that the student "voucher" system and uncapped demand in higher education leads to an analogous "long tail" in burgeoning student enrolments and proliferation of courses in the tertiary education sector.

Non-traditional pathways to higher education, freedom of choice and uncapped student demand generates a "long tail"
The recent Bradley review of higher education (Bradley, Noonan, Nugent, & Scales, 2008) led to a student-centered funding model which has created an ongoing challenge for smaller universities to retain student numbers in the face of uncapped places available to students at "better" universities. Many universities are forced to dip lower into the entry score pool or actively pursue non-traditional student markets to maintain their student numbers and the funding attached to teaching load. This student "voucher" system and uncapped demand strives to uphold egalitarian principles, democratising choice of higher education in the same way the

internet democratises blog readership. This leads to a "long tail" both in terms of student enrolments and the popularity of tertiary institutions, courses, and modes of study.

Online cohorts

Online students play an increasingly important role in maintaining student numbers at some institutions, but there are important lessons to be learnt from teaching online students: many are under-prepared for tertiary studies and have difficulty completing tasks without real-time supervision, creating a "long tail" in the distribution of academic performance. Over several years of teaching online cohorts of psychology students, it has been noted that there is a very high dropout rate of first year students (approximately 40% of online students enrolled in the introductory unit of psychology withdraw from the unit or do not attempt all the assessments), whereas dropout rates in later stage online units are the roughly equivalent to rates for equivalent on-campus units. These online units have no entry requirements and many of the students who drop out appear to do so because tertiary study was not what they expected either in terms of the academic content, or in terms of the skill level and time required to engage appropriately with the learning material. Although we have used a range of teaching interventions to try to retain our students (Fleckhammer & Wise, 2010, 2011, 2012), it may be that high churn and attrition are perfectly acceptable as we move away from traditional student intakes to less well-prepared cohorts with long tails. If we understand student attrition from introductory units in the context of providing the opportunity to try something unfamiliar, the *availability of choice* is of primary importance, *not* the retention or attrition of the student. However, when the student progresses beyond first year units of study *then* retention becomes a critical focus as both student and institution have made a significant investment in the shared learning venture - the long tail has been trimmed. If on-campus student numbers in smaller universities are maintained by dipping lower into the tertiary entry pool and seeking students from non-traditional pathways, we may start seeing a much longer tail in these cohorts as well. While student attrition from on-campus study has been seen as a pressing problem in the past, is this way of thinking still appropriate in the context of the long tail?

Current study

This paper argues that student attrition in the "long tail" of enrolments is part of a natural democratic process of "churn" in which legitimate peripheral participation in a community of practice (tertiary studies) leads many students to the realisation that they are underprepared for, or uninterested in, fuller participation in this community. The question we are raising is *should* we aim to retain the "long tail" of poorly performing students? Is attrition *within the tail* a natural and appropriate form of selection for cohorts increasingly made up of students from beyond the traditional school-based entry pathways? In considering this question, we are aiming to understand specific patterns of performance that might call for specific interventions. For example, students performing poorly through lack of specific skills might benefit from learning interventions targeting those skills, whereas students performing poorly through lack of sufficient academic challenge would benefit from different interventions. Students performing poorly through lack of interest in learning (for whatever reason) might be allowed to drop out until they are sufficiently motivated to return. We use data from an experimental task in an undergraduate introductory psychology program as a micro-study of this idea, examining the relationship between persistence, performance, and demographics on a task which students should be motivated to complete.

The data set comes from an experimental study aimed at testing the efficacy of a brief (90 minutes) unsupervised web-based training package for learning scanning of aircraft instruments. The data was collected from the first year cohort who had to write the assignment, Preliminary data presented at last year's Ascilite conference suggested that unsupervised psychology students can readily learn the instrument scanning task, but we noted a surprisingly low participation rate and high dropout rate during the course of the experiment, despite the fact that the data students were providing the basis of their major assignment. This paper presents a more detailed analysis of a larger data set, and focuses specifically on performance characteristics and demographic details of students who did, and did not, complete the task. We examined whether (poor) performance predicted dropout from the study, and whether students who performed poorly in early stages of the experiment but persisted with the entire task improved their performance by the end. That is, what is the nature of the tail? Who are we teaching and how does this inform pedagogy?

Method

Participants

Participants were recruited from a pool of about 600 university students enrolled in a first year psychology program who were invited to participate in an experiment to collect data for their major assignment, a research report on perceptual learning.

Materials

The instrument scanning task was based on a task described by Kellman and Kaiser (1994) and was web-delivered using Inquisit software (by Milisecond.com). The main stimuli comprised a prototype of a standard 6 instrument panel (see Figure 1, left panel). Participants were also asked to complete a questionnaire providing brief demographic information and information on previous flying experience, gaming experience, and self-reported level of hand-eye coordination.

Figure 1: Experimental stimuli. Left panel shows the six main instruments (simulated) of Cessna cockpit used as stimuli in all experimental blocks except the "transfer" block. The instrument configuration indicates a "straight and level" situation. Right panel shows the instrument variation used in the transfer task. This configuration indicates a climbing left turn.

Procedure

After obtaining informed consent, participants were given a brief description of the study, and a sequence of training pages describing the function of each instrument of the instrument panel. They were then given a series of 30 instrument clusters depicting specific aircraft situations. Nine aircraft situations were depicted in the experiment including: *1- Straight and Level* (e.g., see Figure 1 left panel), *2- Level Climb*, *3 - Level Descent*, *4 - Level Left Turn*, *5 - Level Right Turn*, *6 - Climbing Left Turn*, *7 - Climbing Right Turn*, *8 - Descending Left Turn*, and *9 - Descending Right Turn*. An additional *10 – Incongruent condition* was also included, where one of the instruments was inconsistent with the other five instruments in the display. As soon as the participant identified the aircraft situation they were asked to press the spacebar. They were then presented with a list of the 10 options to describe the aircraft situation. The time to make each response was recorded.

After completing the 30 trial *Practice block*, participants were presented with a sequence of 90 instruments clusters (*Training block*) and were asked to identify the aircraft situation as rapidly as possible. To ensure that participants were not just guessing based on using the artificial horizon, we developed *Masked Instruments block* and *Integrated Instruments block* that exposed participants to the relationship between the artificial horizon and each other instrument and to relationships among instruments other than the artificial horizon respectively. These two tasks aimed to promote more rapid and effective information extraction using all the instruments via training tasks that were not directly related to actual tasks in the flying domain and are described in more detail in McLean, Wise, and Williams (2011). After each of the Masked and Integrated Instrument sequences participants completed a block of standard trials (*Test 1 block* and *Test 2 block*, respectively) to track performance at identifying the aircraft situation. To test the perceptual versus cognitive nature of the learning achieved, participants were finally tested on a set of instrument panel clusters in which the nature of the instruments and their interaction remained consistent with previous training, however the overall position and look of the instrument panel changed substantially (*Transfer block*) as shown in Figure 1 (right panel). Finally, the participants completed the demographic survey. The whole experimental sequence for each participant required approximately 90 minutes to complete.

Participants were asked to devote their full attention to the task, and work in a quiet, distraction-free environment. The use of web-based delivery in relatively uncontrolled circumstances mimics the conditions under which web-based and mobile training operates, such that each participant sets the parameters of their own working environment. The paper presents analysis of completion rates and patterns of performance for those who did not complete the task; data on accuracy and speed of performance, and patterns of performance that may be indicative of different levels of motivation with the aim of understanding attrition rates and potential learning interventions. The analysis of data to evaluate the potential of web-delivered instrument scanning tasks as portable learning modules for aviation training is addressed elsewhere (e.g., McLean et al, 2011; McLean, Wise, & Williams, 2012 this volume).

Results and discussion

Completion rates

Results for completion rates

There were 387 distinct sets of responses from a potential pool of around 600 respondents. Only 97 contain the full set of 510 trials. Of the incomplete data sets, 140 were deliberately aborted by the participant using a special quit command. Some of these were instructors demonstrating the system, but the majority were participants looking at the program but choosing not to complete the data collection (this was inferred from special commands by users to skip to specific stimuli without collecting data). The remaining 150 incomplete sets were terminated simply by closing the web browser. The relationship between response patterns and completion rates was examined by partitioning the cohort into three groups: i) those who dropped out during or at the completion the practice block, ii) those who completed the practice block, but did not complete the whole experiment, and iii) those who completed the whole experiment.

Table 1: Trials completed, valid keystrokes, and response accuracy for participants who dropped out at different phases of the experimental task

Group	Mean # trials completed	Median # trials completed	Mean % Invalid Keys	Median % Invalid Keys	Mean % correct	Median % correct
Dropped out in Practice block	8.68 (7.49)	7.00 (9)	18.28 (27.97)	5.56 (25.00)	60.65 (32.90)	62.50 (60.00)
Dropped out after Practice	169.04 (128.92)	140.00 (163)	5.65 (15.95)	0 (2.7)	59.90 (26.93)	65.12 (40.00)
Completed whole task	510 (-)	510 (0)	0.87 (6.63)	0 (0)	75.81 (20.27)	83.53 (26.00)

Note. Parentheses show SD for means and interquartile range for medians

Table 1 shows that those who dropped out in the first block completed very few trials and made many invalid keystrokes, that is pressing keys *irrelevant* to the task not just an incorrect response option. Those who dropped out after the first block progressed about a third of the way through the task, but had a significant number of invalid keypresses. The proportion of correct responses was markedly higher for the completers, but did not differ for those who dropped out in or after the first block. Six participants had more than 2.5% invalid responses, most of which turned out to be due to the numlock key being set to the incorrect state. We excluded from further analysis in case the incorrect (and thus uninformative) feedback they would have received adversely impacted their performance. This left 93 complete cases, 91 of which had zero invalid keypresses.

Discussion of completion rates

These data suggest that most of those who did not complete the task did not learn the task adequately. Moreover, many of those who did not show adequate performance on the task within the first block but persisted with the experimental task did not improve their performance in later blocks of trials. That is to say, failure to learn the task to a high level in the early stages decreased motivation to continue. In the light of these data, it is important to understand performance patterns of students who drop out of their studies in terms of capability and motivation. Reducing attrition from a long tail of poor performance may serve merely to increase the rate of poor performance. Preliminary data from an ongoing study comparing supervised and unsupervised training environments supports this conclusion. Anecdotal evidence from supervisors supports the idea that the presence (but non-intervention) of a supervisor for students undertaking web-based experimental tasks improves completion rates but decreases the overall quality of data by retaining more students who are performing poorly as a result of obvious lack of engagement or due to lack of task-related skill. The motivation for participation (e.g., slavish completion while being watched, reward for completion of the task, reward for performance on the task, compared with intrinsic interest in the task itself) may promote different patterns of performance, different patterns of engagement and different types of interventions to improve task performance.

Screening of data prior to analysis of response time and accuracy

Preliminary screening for outliers

Means, standard deviations (SDs), medians and ranges for response time and accuracy were computed for each aircraft situation within each block. Many such block/situation combinations contained substantial outliers – for example, while the mean response time was around 10 seconds for the first block and 4 seconds for the last, there were trials lasting over a minute. This is good evidence of data being corrupted by "coffee breaks" despite the fact that participants were measuring response times and accuracy for their own major assignment! The long tail and small numbers in some crossings made it difficult to determine whether a response time of, for example 15 seconds, was an error or simply a very slow response. Similarly, there were response times of only a few tens of milliseconds which cannot be meaningful in terms of the experimental task. Removing trials based on SDs was not practical since the SDs were heavily inflated by the outliers we were trying to identify. We tested a number of different screening methods before adopting the commonly used 5% trim, removing the top and bottom 2.5% of the entire distribution before calculating any other statistics.

Further screening of outliers

Even after the preliminary screening described above, there was quite considerable variation in performance. Responses ranged from 0% correct to 100% correct in nearly every block. The average correlation of the percent correct responses across all possible (21) block pairings was .72. The lowest correlation was .46 (between the first Practice and final Transfer block – those most temporally separated), the highest was .9 (between the Test1 and Test 2 blocks – the two most similar blocks). While these correlations indicate a high level of overall performance consistency, error rates varied markedly for some individuals. Deciding whether to exclude cases based on performance is an important issue in understanding learning and performance. By design, chance performance is ~10% for the standard blocks. Every standard block had a least one case that scored 0% correct, however, the lowest scoring individual was different in each block. There seemed to be "good performers" who scored more than 50% correct on most blocks but had a "bad" block, the location of which was not predictable. Six participants stood out as aberrant cases – performing acceptably in the practice block but declining to chance or near chance levels in the remainder of the experiment - and were removed from further analysis. All 87 remaining cases were retained for further analysis.

Discussion of deletion of cases

While the 6 extreme/atypical cases were removed from the analysis to prevent bias, it must be noted that these represent real performance of participants who completed the task. We do not know their situation or motivation, but we must acknowledge that they made several hundred deliberate keypresses. We must conclude that a small percentage of participants are simply going "through the motions" (e.g., a participant with some previous flying experience scored 77% correct on practice trials but responded at chance levels for the remaining trials) learning nothing from the task. This could be put down to motivation, distraction, or difficulty using the interface. This may also be remedied by providing different kinds of incentives, feedback, or supervision the training as will be discussed below.

Accuracy and speed

Results and discussion of accuracy

As can be seen from Table 2, mean accuracy was highest in the Practice block, dropped slightly during the training block then stabilized in the two test blocks at a level comparable to the Practice block. Accuracy also dropped from the Test 2 to the Transfer block – the performance level (and size of drop) was almost identical to the drop observed in the Training block. A repeated measures ANOVA indicated that accuracy varied significantly across blocks ($F(4, 83) = 12.30$, $p < .001$ – note that to stabilize variance and avoid overdispersion problems the arcsine transform was applied to accuracy scores before conducting the ANOVA). Four contrasts comparing successive blocks showed that, apart from the two test blocks, accuracy on every successive block was statistically different from that on the previous block – that is accuracy increased at each block except for the transfer block. This indicates that genuine learning is occurring, but an analysis of the learning rates is not the focus of this paper. The medians tell a slightly different story. The overall median accuracy is markedly higher than the mean accuracy for all blocks. Median accuracy peaks in the Test 2 block rather than the Practice block. It falls in both the Training and Transfer blocks, but the fall in absolute size is not as large as that observed for means. This mean/median discrepancy is consistent with a strong skew. This is confirmed by a large negative skew statistic for the distribution of each block in conjunction with positive kurtosis. This indicates a fairly tightly packed distribution of good performances and a long tail of poor performances.

Table 2: Aggregate accuracy (mean percent correct responses) across standard experimental blocks

	Practice Block	Training Block	Test 1 Block	Test 2 Block	Transfer Block
Mean (SD)	82 (17)	76 (18)	81 (23)	81 (20)	75 (23)
Median (Min/Max)	87 (10/100)	81 (11/100)	91 (8/100)	89 (8/100)	83 (100/100)
Skewness/Kurtosis	-2.36/6.44	-1.27/1.91	-1.85/2.67	-1.78/2.93	-1.15/.59

N=87

Results and discussion of response times

The overall pattern of response time data is shown in Table 3. Means and medians are comparable indicating that response time is less dependent on a person's place in the overall distribution. The distribution is still slightly negatively skewed, but less so than proportion correct. Mean response times begin at just under 7.5 seconds and steadily decrease to just under 4.5 seconds by Test 2 block indicating an overall 40% improvement in speed. A repeated measures ANOVA indicated that mean response time differed significantly across blocks ($F(4, 83) = 85.67$, $p < .001$). A set of four contrasts comparing successive blocks showed that responses got progressively quicker across each block, except for the transfer block where, as expected, responses were slightly, but significantly slower than the Test 2 block (all contrast p's < .001).

Table 3: Aggregate data for time to identify the aircraft situation in standard blocks and to identify congruent versus incongruent in Masked and Integrated Instrument blocks

	Practice	Training	Test 1	Test2	Transfer
Mean (SD)	7399 (2819)	5374 (1938)	4219 (1707)	3585 (1456)	4440 (2021)
Mean of Individual SDs	3123	2186	1718	1514	2311
Median (Range - Min/ Max)	7382 (799/13872)	5523 (493/9752)	4296 (498/8281)	3742 (332/6739)	4678 (428/8818)
Skewness/Kurtosis	-.08/-.49	-.56/.57	-.43/.27	-.48/.03	-.24/-.36

N=87

The SDs also steadily decrease, indicating more consistency *between* participants. As with error rate, performance is worse on the Transfer task – the mean response time for the Transfer block is halfway between the Training and Test 1 block. SDs also increased in this block, indicating a wider variation in the participants' ability to transfer training to the different display. Table 3 also shows the Mean SDs of individuals' time to identify the stimulus– that is how much an individual participant's performance varied within a block. Intra-individual variability on response time declined steadily over successive the blocks to around half its initial size. Response time variability increased in the transfer block similar to the increase in mean response time, such that the value was similar to that observed in the Training block.

Aggregate scores – aircraft situation and demographic data

Results and discussion of analysis by situation type as a function of demographic data

Responses were collapsed over blocks and aggregated by situation type. Across the aggregate data, accuracy, and response times tell similar story. In every block "straight and level" was the most quickly and accurately identified situation. Straight climb and descent were slightly less accurate and somewhat slower, while turns were slower and less accurate again. The instrument conflict situation had the second fastest response time but the lowest accuracy. For the purposes of this paper, we are less interested in the aviation-related aspects of identifying aircraft situations than patterns of responses that distinguish between different levels of learning or different levels of motivation to engage with the training task. Of specific interest for this paper is the idea that different patterns of results may correlate with different demographic factors, which in turn might lead to targeted interventions to improve learning.

Cluster analysis of – different patterns of responding evidence different learning strategies

Method

To empirically assess our intuitive hypothesis that there are literally different *types* of learners in this cohort we conducted a cluster analysis on the speed and accuracy data aggregated across situations within blocks. Since speed and accuracy measures had different scales, each variable was standardised before analysis. SPSS V20 hierarchical agglomerative clustering algorithm was used, with Euclidean distance as the distance measure and Ward's method (which minimises the within cluster variance thus favouring homogeneous clusters) used as the linking function. Other distance and linking functions were also examined, and yielded results consistent with those described below.

Results and discussion of cluster analysis

The dendrogram showed two widely separated clusters, the larger of which had about seven times as many members as the smaller. The dendrogram also indicated that the larger cluster could be split cleanly into two near equal sized groups, the larger of which also showed two subtly different subgroups. Two, three, and four cluster solutions were considered. The two, three, and four cluster solutions paint a similar picture, but the three and four cluster solutions provide some interesting insights into the microstructure of the larger group. Two and four cluster solutions are presented, with the four cluster solution discussed in detail. Figure 2 provides a visual depiction of the structure of this data tagged with a summary of the make-up of each cluster, which are described in detail below.

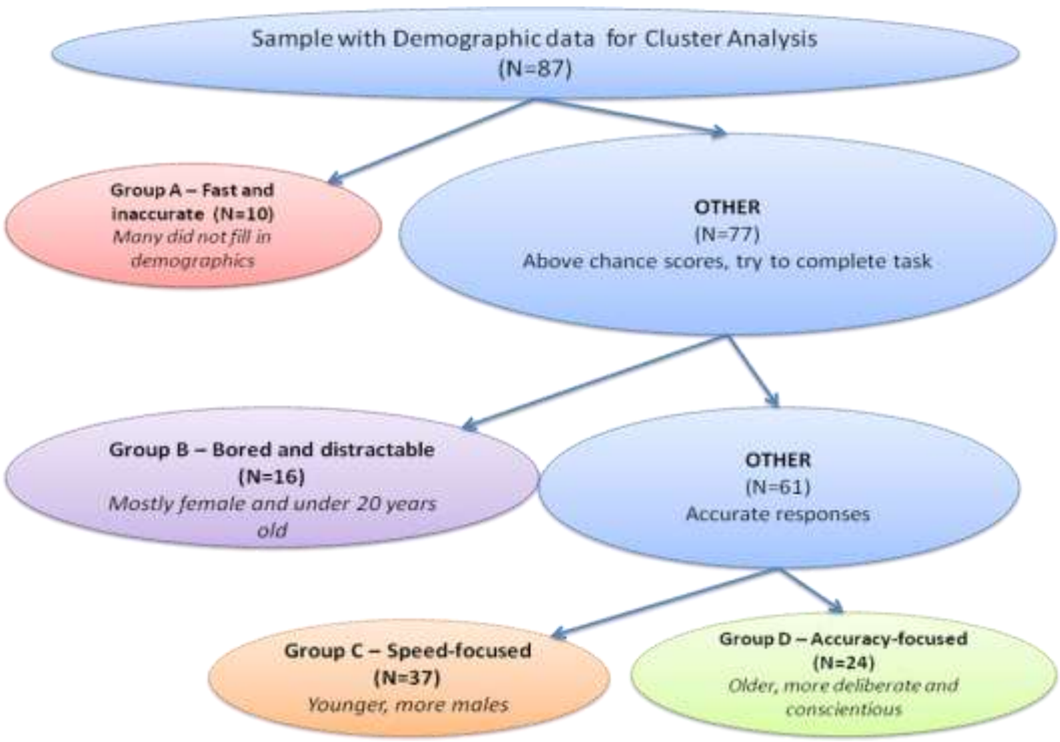

Figure 2: A visual depiction of the cluster analysis results with summaries of the demographic characteristics of each cluster.

Table 4: Two and four cluster solutions for speed and accuracy (see text for further description).

Group	Practice	Training	Test 1	Test 2	Transfer
RESPONSE TIMES (MILLISECONDS)					
A, N=10	4620	1680	783	806	1293
Other	7759	5853	4665	3946	4848
Which splits into					
B, N = 16	8563	6119	4850	3436	3577
C, N=37	6458	4908	3887	3451	4338
D, N = 24	9230	7134	5742	5048	6482
ACCURACY (PERCENT CORRECT)					
A, N=10	62	45	28	44	43
Other	85	80	88	86	79
B, N = 16	74	65	71	64	52
C, N=37	87	86	93	93	85
D, N = 24	87	81	91	91	88

As Table 4 shows, the two cluster solution can be described straightforwardly. Group A has low overall accuracy, averaging around 44%. They show the fastest initial response times and the largest speed increase, the mean response time on the Test 2 being only 17% of the practice response time. They also show the least accurate performance. These people could be described as "fast and loose" or "going through the motions", but with some sense that they should complete the task, despite low engagement. The "other" group display consistently high accuracy, and from a slower start, an overall doubling in response speed by the end of the test block, suggesting that these participants are actually engaged with the task and learning something across trials. "Other" breaks into three smaller groups. Two (labeled C and D) are superficially similar - both show high accuracy which improve marginally over the course of the task, group D, however, is about 3% better on most blocks and distinguishes itself with an outstanding performance on the transfer block. The response time of both groups have almost halved by the second test block, although group D's response time is about 40% longer than that of group C. On this basis we might call group C "speed-focused" and group D "accuracy-focused." Group B is an interesting group which lies somewhere between group A and groups C and D. Their initial speed and accuracy lie between those of A and C, but over the course of the task their accuracy falls and their responses speed up. Having begun the task reasonably well, they appear to lose interest and then try to rush through the remainder of the task and pay a steep price in terms of accuracy. On this basis we might call this group the "bored, distractable" group.

We examined the demographic data of each cluster (see Table 5). Note that the demographic data was not used to perform the cluster analysis, so some significance can be attached to these comparisons. The most significant feature of group A is that they did not provide responses for more than half their demographic questions. This profile would seem consistent with our characterisation of their style as disengaged or uninterested. Group B is almost exclusively female and mostly under 20 years of age. The majority reported playing no computer games, using a computer for non-study purposes (such as using Facebook, downloading music, doing assignments) at least five hours a week and doing little or no sport. This profile is consistent with the typical first year psychology student and their performance likely reflects little intrinsic interest in the task, but a high level of perceived conscientiousness, slavishly "finishing" all assigned tasks in the unit, despite not actually engaging in the task in any meaningful way.

Table 5: Clusters as a function of demographic variables (see text for further description). Numbers show % of cluster having that level of each attribute.

		A fast and loose	B bored, distractible	C speed focused	D accuracy focused
Age	**Under 20**	40	38	46	38
	21 - 24	10	25	24	4
	25 - 30	0	6	16	12
	31 - 35	0	13	5	17
	36 - 50	0	6	7	21
	Not given	50	12	6	8
Gender	**Female**	40	94	65	63
	Male	10	0	27	29
	Not Given	50	6	8	8
Computer Games	**1-2 hours**	0	0	3	13
	3-5 hours	10	0	8	0
	< 1 hour	20	94	84	75
	Not given	70	6	5	12
Using a Computer	**< 1 hour**	0	6	8	4
	1-2 hours	0	31	27	29
	3-5 hours	30	50	35	17
	> 5 hours	0	6	24	42
	Not given	70	6	6	8
Sport	**< 1 hour**	10	56	51	4
	1-5 hours	10	25	24	29
	11-20 hours	0	0	6	9
	6-10 hours	10	6	14	50
	Not given	70	13	5	8

The profile of groups C and D was almost identical. The main point of difference is that group D had many more older persons in it (> 31), while group C was much younger (mostly < 25). Both groups were approximately two thirds females, one third males, over-represents males compared with the gender profile of the psychology cohort. The vast majority reported playing less than 1 hour of computer games a week (although the younger group C had a small percentage playing several hours a week). Both groups used a computer for non-gaming purposes several hours a week (group D slightly more) and both groups reported doing some exercise, although group C had a large number that did only a small amount, while half of group D reported doing 6-10 hours per week. These profiles would seem stereotype-consistent with the younger group being considerably faster and slightly less accurate, and the older group being very conscientious and deliberate. It is also important to note that given the age profile of the undergraduate cohort, the older group D is heavily over-represented in our results, again speaking to maturity and conscientiousness in terms of completing work – rather than just completing the task (checking boxes), these participants engaged with the task and completed it "properly". Interestingly, when we re-ran this analysis with the six cases we had previously excluded for very poor performance included, we obtained the same cluster solutions with five of the six cases joining the "fast and loose" group, further supporting our interpretation of this data.

Concluding remarks – attrition of the long tail

Participants who dropped out of the experiment appear to be the ones who did not "get" the task in the first place, or who could do the task easily but got bored. These two groups form the long tail, but for substantively different reasons. An important finding is that if there was insufficient learning in the Practice trials, there was not much improvement irrespective of persistence at the task, suggesting that repetitive drills are not useful if the basic drill is not understood. Perhaps these participants would have benefited from more conceptual training before attempting the actual trials, or perhaps they did not have the appropriate skills for the task. Those that got bored with the task might just need clearer incentives to boost their performance, in the form of external supervision or extrinsic reward for high scores. The cluster analysis allowed a compelling narrative in terms of the long tail of performance on the instrument scanning task experiment, but also generated a compelling narrative via the demographic data in terms of stereotypical student characteristics within the cohort. Rather than attrition being a problem, it may be a natural outcome of a long tail that recruits tertiary students from a pool of people for whom tertiary education is a relatively unknown quantity – something that seems socially desirable, but about which little is known. Equal opportunity to participate in higher education is a cornerstone of social justice, but in reality, tertiary academic study is not the panacea for a better society. Should we rethink student attrition in terms of a natural process by which students without the appropriate skill level or intrinsic motivation for tertiary studies should be encouraged to follow other pathways (for example vocational pathways through TAFE)? It may not benefit students or the institution to retain them in program of academic study for which they are not well-matched. Conversely, changing programs of study to try to capture disengaged students may not serve students or the institution either. Legitimate peripheral participation in a community of practice can motivate a strong desire to become a full member of the particular community, but can equally legitimately motivate a strong aversion to further participation in the community of practice (Lave & Wenger, 1991).

References

Bradley, D., Noonan, P., Nugent, P., & Scales, W. (2008). Review of Higher Education final report. *Commonwealth of Australia*, http://www.deewr.gov.au/he_review_finalreport/

Fleckhammer, L. & Wise, L.Z. (2010). Providing timely assignment feedback to large online student cohorts. In C.H. Steel, M.J. Keppell, P. Gerbic & S. Housego (Eds.), *Curriculum, technology & transformation for an unknown future*. Proceedings ascilite Sydney 2010 (pp.343-352). http://ascilite.org.au/conferences/sydney10/procs/Fleckhammer-full.pdf

Fleckhammer, L. & Wise, L.Z. (2011). The role of tutors in facilitating online student engagement. In G.Williams, P Statham, N. Brown, B. Cleland (Eds.) *Changing Demands, Changing Directions*. Proceedings ascilite Hobart 2011. (pp.392-397). http://www.ascilite.org.au/conferences/hobart11/procs/Fleckhammer-concise.pdf

Fleckhammer, L. & Wise, L. (2012). Flexible learning outcomes offered by an online program in behavioural studies. In S. McCarthy, K.L.Dickson, J. Cranney, A. Trapp and V. Karandashev (Eds) *Teaching psychology around the world*, Vol 3, Chapter 9, (pp.116-132).

Kellman, P.J., & Kaiser, M.K. (1994). Perceptual learning modules in flight training. *Human Factors and Ergonomics Society Annual Meeting Proceedings, Training, 5*, 1183-1187.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge University Press, Cambridge, MA.

McLean, G., Wise, L., & Williams, B. (2011). A perceptual training module for pilot instrument scans. In G. Williams, P. Statham, N. Brown & B. Cleland (Eds.), *Changing Demands, Changing Directions.* Proceedings ascilite Hobart 2011. (pp.842-846) http://www.leishman-associates.com.au/ascilite2011/downloads/papers/Mclean-concise.pdf.

McLean, G., Wise, L., & Williams, B. (2012). Game-like digital training tools – do information-integrated skills transfer from static to dynamic interfaces. *Future Challenges, Sustainable Futures.* Proceedings of ascilite Wellington, New Zealand 2012.

Shirky, C. (2003). Power laws, weblogs, and inequality. *"Networks, Economics, and Culture" mailing list* http://www.shirky.com/writings/powerlaw_weblog.html.

Author contact details:

Lisa Wise, lwise@swin.edu.au
Ben Williams, bwilliams@swin.edu.au
Gregor McLean, greg.mclean@dsto.defence.gov.au
Jason Skues, jskues@swin.edu.au

Please cite as: Wise, L., Williams, B., McLean, G., & Skues, J. (2012). Online training: Sustainability in the face of the unskilled, unsupervised, unmotivated "long tail". *Future Challenges, Sustainable Futures.* Proceedings of ascilite Wellington, New Zealand 2012. (pp.1061- 1070).