The Impact of Task Difficulty and Performance Scores on Student Engagement and Progression

Dr. Raymond Lynch¹, James Vincent Patten², Dr. Jennifer Hennessy¹.

¹Department of Education and Professional Studies, University of Limerick, Ireland.

²Department of Computer Science and Information Systems, University of Limerick, Ireland.

Corresponding author:
Dr. Raymond Lynch,
Department of Education and Professional Studies,
University of Limerick,
Limerick.
E-mail: Raymond.Lynch@ul.ie

Abstract

Background  This article considers the impact of differential task difficulty on student engagement and progression within an Irish primary school context. Gaining and maintaining student engagement during learning tasks such as homework is a significant and understandable on-going challenge for teachers. The findings of this study hold the potential to support teachers’ decision making processes regarding the development of student tasks.

Purpose  The research study aimed to explore the impact of task difficulty on student engagement and subsequent progression in the computerised navigation task Pac-Man. The central research questions addressed in this article were: do subtle variances in task difficulty impact on student volition and consequently, will this result in a significant variance in students’ levels of improvement?

Sample  Sixty students from a large urban, coeducational primary school in the south of Ireland were identified as a suitable sample cohort. All students were in their final year of primary school within the Irish education system and were between 11 and 12 years of age.

Design and Methods  The study employed the use of the popular arcade game Pac-Man. In a test-retest approach, sixty primary school students completed the standard computerised navigation task with a seven day interval. Between testing, participants were randomly subdivided into three cohorts. Each cohort of twenty participants received a different version of the Pac-Man game to practice with for one week. Cohort A received a version of the Pac-Man game of lesser difficulty, Cohort B received the standard Pac-Man game, and Cohort C received a version of greater difficulty. A paired-samples t-test (repeated measures) was employed to compare the scores achieved by each of the three cohorts both pre and post practice. As an indication of the resulting effect size for each cohort the eta squared statistic was subsequently calculated. In order to support any future meta-analysis, Cohen’s d statistic is also provided in this paper. Analysis of variance (ANOVA) was employed to explore
differences between groups with regard to progression scores and number of games played when practicing.

**Results**  The results of this small scale study found the cohort who received the easier version of the task presented the greatest overall improvement in performance between the pre and post tests. No statistically significant difference was found in the change in scores of the three cohorts – potentially due to the small sample size. However, paying attention to the size of the effect indicated that, over seven days, there was an 80% improvement in performance for Cohort A, 63% improvement for Cohort B, and 26% improvement for Cohort C. The results highlight the negative impact of increased task difficulty on students’ volition and consequently, on overall progress in the task.

**Conclusions**  Further research with larger student populations would be needed to assess the generalisability of the results. However, the findings suggest that when designing tasks to promote student learning, particularly self-directed tasks such as homework, it is important that teachers afford ample opportunity for student success.

**Keywords:** Task Difficulty, Student Performance, Engagement, Self-efficacy, Homework

**Introduction**

Students’ perception and categorisation of tasks based on the potential for success operates as a valuable heuristic for filtering information and focusing attention on material germane to respective learning goals. However, in recent years as a result of advances in technology, communication and globalisation, students are increasingly exposed to greater levels of complex sensory stimuli (Robinson 2011). The recent information and communication revolution places greater demands on teachers. Teachers today must compete with computer games, music videos, films, television shows, widespread internet access, and smart phones, for students’ attention and engagement (Barkley 2009, p.xii). Consequently, gaining and maintaining student engagement during learning tasks is a significant and understandable on-going challenge for teachers. This is not to suggest that being entertained is the same as being engaged. Engaging students requires the promotion of critical thinking related to the learning task being completed or having been completed. It is not surprising
then that “student engagement is generally considered to be the better predictor of learning and personal development” (Carini, Kuh and Klein 2006, p.2). This premise is predicated on the simple but reasonable supposition that the more students study or practice a task, the more they tend to learn. This premise is also the principal proponent for the allocation of homework to students and research in relation to this area would appear to promote it.

Significant research exists to support the allocation of homework. Homework in general appears to have positive effects on student performance at all levels of achievement, where an increase in the amount of time spent on homework results in improved performance. In one of the largest ever collections of evidence-based research into the influences on academic achievement in school aged students, Hattie (2009) found that on average homework resulted in a positive effect-size of 0.43. However, the influence of homework on students’ academic achievement, as measured by standardised tests, varies across disciplines. For example both Aksoy and Link (2000) and Betts (1997) found significant positive effects of homework on Mathematics test scores. Betts (1997) estimated that an extra half an hour of homework per night for American students in grades 7-11 would advance a student almost two grade equivalents. Equally, in a study by Natriello and McDill (1986) of 12,146 students across 20 public high schools in America, homework was shown to have a much smaller effect on performance in standardised English examinations. The findings of Natriello and McDill (1986) show that an additional hour of homework each night results in an increase of just 0.130 in students’ cumulative grade point average (GPA) in English. Not only does the influence of homework vary greatly across disciplines but the effects are also not uniform across different student cohorts. Additional homework is shown to be most effective for lower achieving students with extra time spent on homework often merely having a compensatory effect (Keith 1982; Purdie and Hattie 1999; Eren and Henderson 2008). Stanley (1980, p. 164) found that the “main variable differentiating the
successful students, who move ahead at astonishing speeds, from the unsuccessful ones is homework”. However, there exists a dearth in research into the relationship between the learning task or homework and student engagement. Why do some students engage better with the task and what happens if the level of task difficulty is varied?

**Task difficulty and student performance**

Research suggests an inverted-U relationship between task difficulty and performance. Achievement goal theory posits that students’ performance goals are expected to increase and then decrease as individuals experience increased difficulty (Dowson and McInerney 2001). This is supported by the findings of Kumar and Jagacinski (2011, p. 677) who found that “participants exposed to increasing levels of difficulty reported declining levels of perceived ability and performance approach goals and increasing levels of work-avoidance goals”. Duda and Nicholls (1992) suggest that students’ goals change once the individual decides that they are unable to demonstrate high ability and as a result switch to the goal of avoiding a demonstration of low ability. In short, as the task becomes more difficult, students’ perceived ability in the respective task decreases, which in turn leads to decreasing commitment to the goal of demonstrating high ability (Nicholls 1984). The proposed invert-U relationship is also suggested to be a corollary of an arousal effect. Research shows that as task difficulty increases it initially results in a corresponding increase in ‘mental effort’ and task arousal (Huber 1985; Barch et al. 1997, p. 1373). However, as highlighted by Hubar (1985, p. 493), “while increased arousal facilitates performance on simple tasks, at some level of task complexity, arousal begins to impair performance”. This occurs because elevated levels of task arousal helps support dominant responses (frequently
an instinctive, learned or conditioned response), which are often appropriate and correct on simple tasks but incorrect on complex tasks (Organ 1977).

The switch from positive to negative task valence (intrinsic attractiveness or aversiveness to a task) as difficulty levels increase, can occur at any time depending on the individual (Silvestrini and Gendolla 2009). As the task difficulty increases, the probability of failure also increases. Once failure is viewed as a strong possible outcome students rapidly become understandably disengaged (Bandura 1977; Bandura 1986). For some this switch from positive to negative task valence may never occur, while for others it may occur at the very beginning of the task (low task difficulty), depending on the individual’s ability in the respective task. Such disparity and varied patterns in students’ responses to differential task difficulty makes it difficult for a teacher to discern appropriate pedagogical strategies to cater for the mixed ability of most contemporary classrooms (Ireson and Hallam 2005; Ireson, Hallam and Hurley 2005). The task facing teachers becomes even more daunting given the panoply of alternative stimuli available to students at present. The level at which one should pitch learning tasks within mixed ability groups is a matter of considerable concern for educators at all levels (Kerry 1982; Sands and Kerry 1982; Boaler 1997a; Boaler 1997b). Teachers must assign tasks at a level that ensures pupils are challenged but still capable of succeeding, that engages but still remains pertinent. This considerable challenge is somewhat more attainable within the classroom where the teacher can monitor students and continually differentiate for ability and levels of engagement. However, it is appreciably more difficult for the teacher to ensure student engagement in learning tasks assigned as homework. Under the autonomous, self-directed learning conditions afforded through homework, the teacher has very little influence on student engagement once the learning task has been assigned.

The importance of continually monitoring performance and levels of engagement is well recognised by the computer and video gaming industry. Modern intelligent ‘adaptive
computer games’ provide players with immediate feedback and continually monitor and differentiate for varying ability to ensure maximum engagement with the game (Squire 2003; Charles et al. 2011). As highlighted by Gee (2003) video game designers face and for the most part solve an “intriguing educational dilemma, one also faced by schools and workplaces: how to get people, often young people, to learn and master something that is long and challenging - and enjoy it, to boot”. It is not surprising then that educators at all levels are turning to ‘smart’ computer games as an educational tool for promoting student engagement and learning (Gee 2003; Charles, Bustard and Black 2011). However, within the current educational milieu, the assignment of traditional learning tasks in preparation for summative examinations or standardised tests is broadly advantageous (Eren and Henderson 2008) and arguably requisite practice for students. Consequently, the study outlined in this article set out to examine what impact, if any, the level of task difficulty that students practice with has on progression on a standardised computer navigation task (Pac-Man). The place and impact of standardised tests and summative examinations on student learning is too broad and lengthy a subject to address in this paper (Suggested reading: Black and Wiliam 2003; Ryan and Weinstein 2009). Therefore, for the purpose of this paper the debate surrounding standardised tests and their impact on student learning is momentarily deferred. Moreover, this study was predicated on the position that such tests and examinations persist as a fixed educational axiom throughout contemporary schooling and as a prerequisite under many university matriculation systems (Black and Wiliam 1998; Trigwell, Prosser and Waterhouse 1999). Within this context, this study seeks to explore the impact of task difficulty on student engagement and subsequent progression. The authors do not suggest the Pac-Man task employed in this study to be generalisable to alternate learning tasks but assert that the findings from this study provide information on differentiated pupil engagement in an ‘optimised learning context’. As noted by Brophy and Kher (1990, p.264) “student
motivation to learn is optimised when it has the qualities associated with positive task-endogenous”, as elicited through the use of the computer navigation task in this study. This information holds the potential to better support teachers’ decision making processes concerning the effectual development of student tasks, such as homework.

Method

In an attempt to examine the proposed relationship between task difficulty, student engagement and progression, this study employed the use of the popular arcade game Pac-Man\(^1\). A version of this well-known computerised navigation task with adjustable game settings was specifically developed for this study. The game was developed to allow its settings to be specified in an XML file which was loaded when the game started. Extensible Markup Language (XML) is a widely used markup language that provides a set of rules for encoding documents in a format that is both human-readable and machine-readable. XML was designed to store and transport data, making it an ideal choice for storing both the settings that the Pac-Man game uses and the results generated when running the experiments. The software automatically stored the participants’ score and performance information (silently in the background) every 30 seconds, as well as at the end of every game. By modifying the standard Pac-Man settings (increasing or decreasing the speed of the characters), two alternate versions of the game were subsequently created. The degree of deviation from the standard computerised navigation task was equal, but in opposing directions, for both the easier and the harder version of the task.

\(^1\)A non-deterministic version of Pac-Man, often referred to as Ms. Pac-Man, was employed and adapted for this study as it ensured that participants could not learn or predict the movement of the game’s characters. The ability to predict the characters’ movements would have reduced the skill level required to complete the task and distorted the findings. It was developed by extending the C# implementation of the game provided by Flensbank & Yannakakis (Available at: http://mspacmanai.codeplex.com/).
Pilot Studies

In order to ensure that any disparity in the level of task difficulty was not transparent to participants in the study when practicing the task, pilot studies were conducted with twelve postgraduate students. Postgraduate students were employed at this stage of the research due to ease of access for pilot testing of multiple iterations of the computerised navigation task. Also, as postgraduate students are much further into Piaget’s formal operational stage of cognitive development (Piaget 1967), the authors were confident that if the differences in the levels of task difficulty were not discernible to them then they would not be detectable by primary school students. In a series of randomized blind pilot studies, the twelve postgraduate students were subdivided into three groups of four students. All students first performed the standard computerised navigation task and registered a score at the end of the task. Subsequently, they were randomly provided with one of the three versions of the task on a portable USB ‘memory stick’ to practice with for one week (unaware that there was any variance in the tasks assigned). After seven days the twelve students were then invited back to repeat the standard computerised navigation task. Directly after repeating this task, students were asked if they ‘observed any difference in the level of difficulty’ between the task they practiced with and the one they just completed, and to rate the task they practiced with on a Likert scale. The levels on this Likert scale were delineated as follows: 1) a lot easier, 2) easier, 3) equal, 4) harder, and 5) a lot harder. The results of this survey were used to amend and adjust the task settings, reducing the degree of deviation from the standard task. After adjusting the settings, the blind pilot study was repeated twice more following identical steps to those previously outlined until no statistically significant correlation existed between the task level assigned and the resulting Likert scale data. Once the task had been calibrated so that differences in levels of task difficulty were not discernible to postgraduate students it was deemed ready for implementation with primary school students.
Participants

For the purpose of this study a group of sixty students from a large urban primary school were identified as a suitable sample cohort. All sixty students were in their final year of primary school within the Irish education system (between 11 and 12 years of age). Students in this year of study were identified for two principal reasons. Firstly, as suggested by Piaget students of this age are entering the formal operational stage of cognitive development (Piaget 1967; Piaget 1972). Consequently they will have just begun to develop the logical thought, deductive reasoning, systematic planning and problem-solving skills required to complete this computerised navigation task (Snowman and Biehler 2000). Secondly, students of this age will have developed the requisite computer literacy skills to complete the Pac-Man task. As highlighted by Pane and Myers (2006, p.46), participants from age 9 years and upwards possess the necessary skill sets to partake in studies that utilise computerised navigation tasks comparable to the Pac-Man task used as part of this research.

Procedure

A test-retest approach was employed with a one week interval between tests. Students were first invited to complete the standard Pac-Man task on the 24th November 2011 during their first three class periods within the setting of their primary school. In groups of two, students were in turn excused from class and proceeded to simultaneously complete the standard computerised navigation task on two identical laptops in a spare classroom. Echoing the pilot studies, upon completing the standard task once and registering a score, participants were provided with a version of the task on a portable USB ‘memory stick’ and asked to practice with it for one week. This was to be part of students’ homework for the week and cooperating teachers in the study supported this by requesting students to write it into their homework notebooks and outlining that no one else should play the game provided
on their respective USB ‘memory sticks’. There was no minimum practice time requirement set for students each evening, however, as suggested by Sharif and Sargent (2006) students were asked to restrict their time practicing Pac-Man to less than 1 hour a day. Unknown to the participants, 20 of them had received the easier version of the task (cohort A), 20 had received the standard task (cohort B), and 20 had received the harder version of the task (cohort C). Each time a participant practiced the task, as part of their homework, the software would output their performance information in a separate file stored on the memory stick, which was automatically both date and time stamped. This enabled the authors to track participants’ performance and time spent practicing over the seven days between testing. The principal investigator returned to the school on the 1st December 2011 where once more participants completed the standard Pac-Man task and registered a repeat score.

Full ethical approval for this study was granted by the University of Limerick, Faculty of Science and Engineering Research Ethics Committee on the 21st September 2011.

**Analysis of data**

A paired-samples t-test (repeated measures) was employed to compare the scores achieved by each cohort both prior and post practice, using the version of the Pac-Man task they received on a portable USB ‘memory stick’ as part of their homework for one week. As an indication of the resulting effect size for each cohort the eta squared statistic was subsequently calculated. For the purposes of further analysis and in an attempt to support any future meta-analysis of these results, Cohen’s d statistic is also provided in this paper. As suggested by Rosnow and Rosenthal (1996), Cohen’s d was calculated based on the pooled standard deviation of both the pre-test and post-test scores registered by participants. A one-way between-group analysis of variance (repeated measures) was conducted to examine the levels of variance in the mean scores of each cohort and once again the eta squared statistic
was calculated to further explore trends in the data. A correlation analysis was employed to investigate the relationship between the number of games played during practice and the level of improvement in participants overall performance.

Results

A one-way between-group analysis of variance was conducted to compare the mean scores for all three cohorts in the initial task (pre-test scores). There was no statistically significant difference ($p = 0.845$) in the initial scores for the three cohorts: $F (2, 57) = 0.169$. This is supported by the extremely small effect size, calculated using eta squared, of 0.0059 and the relatively small differences between the pre-test mean scores of each cohort (see Table 1). This is important as it suggests homogeneity across groups and therefore supports group comparability.

A second one-way between-group analysis of variance was conducted to compare the mean scores for all three cohorts in the final task (post-test scores). No statistically significant difference ($p = 0.212$) was found in the post test scores for the three cohorts: $F (2, 57) = 1.593$. However, the medium effect size (Cohen 1988, p. 284-7), calculated using eta squared, of 0.0661 reflects the fact that the actual difference in mean scores between the three cohorts was quite large in the post-test, especially when compared to the effect size from the pre-test scores. A final one-way between-group analysis of variance was conducted to compare the change in scores (the difference between the pre and post test scores) for all three cohorts. Again, although close, no statistically significant difference ($p = 0.059$) was found in the change in scores for the three cohorts: $F (2, 57) = 2.967$. Consequently, as no significant difference was found, the hypothesis was not met. This is potentially due to sample size discussed later under the limitations of the study. However, Carver (1993, p.288)
insists “that attention be paid to the size of the effect, whether it is statistically significant or not”, a position also supported by Fritz et al. (2012, p.15). Therefore, as with previous tests, the effect size was once more calculated (eta squared = 0.1441). In this case, the actual difference between the change in scores for each cohort was substantial, as reflected in the large effect size (Cohen 1988, p. 284-7) of 0.1441.

The results of this study found that Cohort A presented the greatest overall improvement in performance between the pre and post tests. There was a statistically significant increase in students’ Pac-Man scores from the initial test to the repeat test a week later, \( t (19) = 3.766, p = 0.001 \). The mean difference in their scores was 1533, an overall improvement of 80%. This resulted in a large effect size (Cohen 1988, p.22), calculated using eta squared, of 0.427. Cohort B also produced a statistically significant increase in their performance between tests, \( t (19) = 2.203, p = 0.04 \). Cohort B was shown to have made the second largest overall improvement between tests with a mean difference of 1136, corresponding to an average improvement of 63.1%. Again, the eta square statistic (0.203) indicated a large effect size. While not statistically significant \((p = 0.257)\), Cohort C also demonstrated an improvement in performance between tests, \( t (19) = 1.168 \). However, of the three groups, Cohort C was shown to have made the smallest improvement over the week with a mean difference in scores between tests of 463, an overall improvement of 25.8%. The eta squared statistic (0.067) indicated a medium effect size. These results, including the Cohen’s d statistic for each cohort are shown in Table 1.

Insert Table 1 Here
In an attempt to uncover the cause of the significant discrepancy in performance across each cohort, an evaluation of the number of games played over the week by each participant from the respective cohorts was conducted. The total daily number of games played by each cohort is graphed in Figure 1. Participants practiced most frequently with their version of the Pac-Man task on the evening of the initial test (Thursday the 24th November 2011), with Cohort A playing 170 games, Cohort B playing 163 games, and Cohort C playing 153 games in total. While the number of games played by participants in the subsequent days fell rapidly for each cohort, the drop was demonstrably sharper for Cohort C, with the numbers of games played by members of this cohort continuing to fall throughout the week. By comparison, the number of games played by members of Cohort A began to increase after the weekend, starting on the Monday. The number of games played by Cohort B continued to fall until the Sunday, after which time they began to stabilise. A Pearson product-moment correlation coefficient highlighted the importance of the relationship between games played and improvement in performance. There was a strong positive correlation \( r = 0.853, n = 60, p< 0.0005 \) between these two variables, suggesting that the number of games students played when practicing had a significant impact on their performance in the post-test.

**Insert Figure 1 Here**

A one-way between-group analysis of variance found that there was no statistically significant difference \( p = 0.200 \) between the average number of games played for the three cohorts: \( F (2, 57) = 1.657 \). Therefore, once more the hypothesis was not met. However, again the medium effect size (Cohen 1988, p. 284-7), calculated using eta squared, of 0.0639
reflects the fact that the actual difference in the mean number of games played between the groups was noteworthy. On average, each participant from Cohort A played 38.9 games over the seven day period, participants from Cohort B played 30.8 games, and participants from Cohort C played the game 21.4 times. Multiple linear regression was used to assess the ability of the independent variable (numbers of games played) to predict participants’ level of improvement in performance, after controlling for any direct or indirect effects resulting from the group they were assigned to i.e. Cohort A, B or C. Dummy variables identifying the cohort that students were assigned to were created and entered into the regression model in order to control for any such direct or indirect effects. For the purpose of this model Cohort C was used as a reference variable and dummy variables were created for cohorts A and B. The total variance explained by the model as a whole was 71.3%; Adjusted R squared = .713, $F (3, 56) = 49.962, p < .001$. Upon evaluating the independent variables, the number of games played was the only variable shown to be statistically significant (Beta = .847, $p < .001$). The dummy variables created for Cohort A and B were not statistically significant in predicting participants’ level of improvement in performance (beta = .025, $p < .759$ and beta = .037, $p < .650$ respectively). This suggests that there were no other direct or indirect effects of the varying levels of task difficulty (as determined by the cohort students were randomly assigned to) other than the indirect impact this had on the number of games students played when practicing over the week. This is further supported by a second multiple linear regression analysis conducted, this time without the dummy variables, which resulted in a very similar adjusted R square value of 0.722; $F (1, 58) = 154.423, p < .001$. The Beta value also remained largely unchanged at .853 ($p < .001$).

**Limitations of the study**
This study and the analysis of the resulting data, was concerned with producing type II errors (beta errors) and failing to reject the null hypothesis when the null is false. Consequently, these initial findings and the outlined effects need to be replicated in larger and systematic studies across diverse student cohorts. Therein, one of the clear limitations of this study was its size which makes it very difficult to generalise the resulting findings. On average the participants in this study scored greater than the 80th percentile for students across Ireland on standardised tests for numeracy and literacy. Consequently, in general, participants in this study are arguably academically ambitious students and similar studies with alternate student populations are required to assess the generalisability of these findings. Finally, the computerised navigation task employed in this study was chosen because it is intrinsically motivational in nature and provided an ‘optimised learning context’ (Brophy and Kher 1990). Consequently, and as highlighted previously, the Pac-Man activity itself cannot be generalised across other activity modes without further research. However the findings of this study may serve to better inform teachers’ decision making processes around the design of tasks that can enhance student motivation. Regardless of the task or assignment provided to students, the findings of this study highlight the importance of the potential to succeed in the promotion of student engagement and volition.

Discussion

At a minimum, the results of this study highlight the impact of differential task difficulty on student performance. Within this study, task difficulty was noted to be inversely proportional to student volition. As task difficulty increased, student volition decreased. Equally as task difficulty decreased student volition increased. Therein the potential for success was noted to be fundamental, not alone to students' engagement with the task, but
also to their subsequent performance on the standardised test. It follows that homework, as a tool for the consolidation and advancement of learning, should include provision for student success if it is to be effectual.

The results of this research also generate many questions. It is clear that Cohort A, who had received the easier version of the task, practiced more and played more games between tests than both other cohorts. Consequently, they also demonstrated the greatest increase in task performance. Conversely, Cohort C received the harder version of the task to practice with and was shown to have improved the least amount. It seems reasonable to imagine an alternate outcome if on average each cohort were to have practiced an equal number of times. Under such conditions where participants would be moving from practicing with a harder version of the task down to the standard task, one would be excused for expecting them to out-perform those going from an easier version up to the standard task. However, as highlighted by Martin, Hodges, and Kulinna (2002, p. 18), “student improvement in many activities is contingent on effort and practice, which rely on student volition”. The findings of this study suggest that while the provision of an alternate learning task was successful in the initial motivation of pupils, as task difficulty increased, student volition decreased. Again, it is important to note that this was a gradual, and for cohort C, a continual decrease over the seven days of practicing with the task. This would suggest that feedback in the form of task scores and results had a significant impact on student volition and task valence. By propensity, as task difficulty increases it becomes equally harder to gain a high score in that task. This in turn can negatively impact on students’ self-efficacy and as suggested by flow theory (Csikszentmihalyi 1975; Csikszentmihalyi 1988) the task can then lose its intrinsic enjoyability. As highlighted by Shernoff et al. (2003, p. 160) “the flow experience is believed to occur when one’s skills are neither overmatched nor underutilized to meet a given challenge.” However, this balance between the challenge of a
task and the skill required to complete it is fragile. Even minor variances in perceived task difficulty have been shown to potentially disrupt ‘flow’ (Csikszentmihalyi, 1988), resulting in student anxiety if the challenge is perceived to be too difficult or even apathy if not perceived as challenging.

This disruption to flow as a result of differential task difficulty could potentially explain a certain percentage of the variance in student engagement and, consequently, performance. However, this position fails to acknowledge the complex social milieu in which this study was conducted. After practicing with the task each evening as part of their homework, students came to school the following morning and discussed their performance and results with their peers. Informal discussions with participants while administering the post-test exercise confirmed that a competitive environment had evolved over the course of the week, with many of the participants endeavouring to out-perform each other. While under controlled settings competition has been shown to promote student learning (Attle and Baker 2007; Burguillo 2010), if not structured around and supported by appropriate pedagogy it has also been shown to have a consistently negative impact on task enjoyment and achievement attribution (Lam et al. 2004). The competition described by participants in this study was an organic and unstructured development which resulted in many students (especially those who had received the harder version of the task) believing they were not competent at the Pac-Man task. Social cognitive theory (Bandura 1986) predicts that these beliefs matter and that the resulting lower self-efficacy would impact on student motivation and performance. This would explain the consistent decline in the number of games played by Cohort C when practicing over the seven days, especially when compared to Cohort A, who saw an increase in the number of games played after the weekend suggesting higher self-efficacy. Vancouver et al. (2002, p. 509) found that “beliefs in capacity to organize and execute behaviours (i.e., self-efficacy beliefs)” can have a significant impact on student
learning and performance. These beliefs are often socially constructed (Bandura 1993), with students readily comparing their performance in a task to that of their peers. It is for comparable reasons that formative feedback on examination performance, for example, has been shown to promote student learning over a formal grade or score (Butler 1988; Black and Wiliam 1998).

The findings of this study suggest the importance of the development of positive, supportive learning environments which promote student self-efficacy regardless of the learning task assigned. As highlighted by Rosenthal (1994), the sheer belief that a student can succeed and improve is enough to alter a teacher’s approach to that student, resulting in genuine improvement. The results of this study indicate that self-efficacy beliefs produce similar results, with students altering their engagement with a task depending on perceived ability to succeed in that task.

Conclusion

The results of this study suggest that when designing tasks to promote student learning, particularly self-directed tasks such as homework, it is important that teachers afford ample opportunity for student success, especially with respect to their peers. The provision of opportunity here is founded in the teacher’s understanding of pupils’ individual levels of ability. In order to ensure all students experience success, task differentiation should be provided for at all levels of academic ability. This is also pertinent for the provision of homework. As proposed in the opening sentence of this paper, students’ perception and categorisation of tasks is an essential heuristic. However, as highlighted in this study, if students perceive a task to be beyond their ability and once failure (with respect to their peers) becomes a strong possible outcome, the task can soon be relegated in favour of
alternate assignments where their perceived strengths lie. This is not to suggest that by simply promoting success, all students will succeed. Nevertheless, it does indicate that success is an important advocate for enhanced student volition, self-efficacy and consequently, performance.

One of the strengths of this study was in the task assigned to students. By propensity, this task was designed to be intrinsically motivational. In spite of this, if students did not believe they could compete with their peers and succeed in the task, they slowly became disengaged and ceased to practice with the computer game Pac-Man. Given the array of alternate sensory stimuli available to students at present, maintaining engagement and motivation through essential learning tasks is becoming more challenging. It is difficult to compete with contemporary computer and video games which provide participants with instant feedback and are able to continually differentiate according to ability. However, if educators are to be successful they need to create a positive learning environment which promotes success irrespective of initial ability. As highlighted by Rodgers (1951, p. 64) “The educational situation which most effectively promotes significant learning is one in which (a) threat to the self of the learner is reduced to a minimum, and (b) differentiated perception of the field is facilitated.”

References


Boaler, J. 1997b. When even the winners are losers: Evaluating the experiences of top set's students. *Journal of Curriculum Studies* 29, no. 2: 165-182.


### Table 1 Comparison of Means - Repeated Measures

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>σ</th>
<th>Mean Difference</th>
<th>t</th>
<th>Sig. (2-tailed)</th>
<th>η²</th>
<th>Cohen's d</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cohort A</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-Test</td>
<td>1917</td>
<td>961.655</td>
<td>1533</td>
<td>3.766</td>
<td>0.001</td>
<td>0.427</td>
<td>0.873</td>
</tr>
<tr>
<td>Post-Test</td>
<td>3450</td>
<td>2289.337</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cohort B</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-Test</td>
<td>1800.5</td>
<td>733.122</td>
<td>1136</td>
<td>2.203</td>
<td>0.04</td>
<td>0.203</td>
<td>0.709</td>
</tr>
<tr>
<td>Post-Test</td>
<td>2936.5</td>
<td>2144.806</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cohort C</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-Test</td>
<td>1792</td>
<td>519.854</td>
<td>463</td>
<td>1.168</td>
<td>0.257</td>
<td>0.067</td>
<td>0.329</td>
</tr>
<tr>
<td>Post-Test</td>
<td>2255</td>
<td>1921.719</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1 Numbers of games played