Integrating Fantasy Role-Play into the Programming Lab: Exploring the 'Projective Identity' Hypothesis

Michael James Scott
Information Systems, Computing & Mathematics
Brunel University
Uxbridge, Middlesex, UB8 3PH
United Kingdom
michael.scott@brunel.ac.uk

Gheorghita Ghinea
Information Systems, Computing & Mathematics
Brunel University
Uxbridge, Middlesex, UB8 3PH
United Kingdom
george.ghinea@brunel.ac.uk

ABSTRACT
It has been claimed that learning can be facilitated by a positive academic self-concept. Therefore, reinforcing this construct may benefit students and the application of 'projective identity' in educational multimedia could be a means of achieving this. To test this hypothesis, two versions of a debugging exercise were developed, with one incorporating elements of fantasy role-play. They were compared through a double-blind parallel-group randomised trial using a sample of 36 undergraduate computing students. Factor scores for academic self-concept in programming were imputed from responses to a 5-point Likert scale, validated through a confirmatory factor analysis of 91 responses. An ANCOVA revealed that students using the fantasy role-play learning activity developed a stronger self-concept than the control group, with respective gains of 2.4% and 1.1%. However, further work is required to determine if such modest gains are practically significant, can be further enhanced and maintained.

Categories and Subject Descriptors
K.3.1 [**Computers and Education**]: Computer Uses in Education – *computer-assisted instruction (CAI).*

General Terms: Design, Experimentation, Human Factors

Keywords
Academic Self-Concept, Programming, Fantasy, Roleplay, Projective Identity, Learning, Education.

1. INTRODUCTION
Positive psychology claims that a reciprocal relationship exists between achievement and academic self-concept (ASC), defined here as "a person's self-perceptions that are formed through experience with and interpretations of one's environment" in relation to a specific academic context [10]. Therefore, reinforcing this construct during the delivery of teaching material could lead to the emergence of more effective learners. However, it is not clear what practices could be applied in learning environments for introductory programming courses to achieve this. Due to the increasing adoption of educational multimedia in such environments, it would seem sensible to identify how such tools can be leveraged to enhance academic self-concept effectively.

SIGCSE'13, March 6–9, 2013, Denver, Colorado, USA.

The use of fantasy role-play [7,8] within a dramatic narrative [5] could be one such practice. It has demonstrated some qualitative success in *Space Mission: Ice Moon*, enabling students to "think and act like scientists" [4]. Among creative learning environments [14], a similar approach can also be seen in *GameStar Mechanic*. There, novice designers are immersed in a series of story-driven scenarios while they learn basic concepts and skills before proceeding onto more practical design activities. While the latter tool addresses games design, rather than programming, a similar approach could effectively prepare students for creative programming activities, as in [13]—but conveying preparatory material in an engaging manner, rather than using just-in-time delivery where the pace can be uncomfortable for some students, which is undesirable [1]. Thus, to what extent could the application of fantasy role-play contribute to the development of academic self-concept?

During an exploration of learning in games, Gee [7] considered the role of identity. A three-way relationship was proposed between: real-world identity; virtual identity; and a mediating projective identity. It was argued that the virtual and projective identities formed when immersed in a fantasy role could reinforce the real-world identity relating to the activity of that role. During a separate investigation of self-representation in virtual reality, Yee and Bailenson [16] empirically demonstrated a somewhat similar phenomenon relating to identity and self-beliefs: the Proteus Effect. Using specialist equipment under experimental conditions, it was shown that manipulating the avatars of participants immersed in a virtual environment had some impact on their attitudes and behavior within the virtual world. Subsequently, some of this difference was maintained when measured a short time after the experiment. However, the extensibility of this effect to other domains, such as educational multimedia, is unclear.

If the phenomenon does extend to educational multimedia, it could have significant implications for the design of learning environments and the delivery of e-learning material. Thus, this paper describes an initial experimental study, in which fantasy role-play is integrated into a prototype virtual lab exercise that aims to enhance debugging skills. It then explores the following research question: assuming an equal baseline, does the incorporation of fantasy role-play in an e-learning activity increase its impact on academic self-concept development for undergraduates enrolled on a programming course?

2. DEBUGGING EXERCISE
The e-learning activity used in this study was inspired by Logo Geometry [11] and Karel the Robot [12]. It aims to increase student confidence by teaching how to identify and correct simple syntax and logic errors in snippets of Java code. Students are first

shown an instructional video within the tool, explaining how to trace code to identify errors. A fragment of code containing faulty instructions to navigate a maze is then displayed. Once a student has identified and corrected the mistakes, an animation shows an object moving through the maze, revealing whether all of the errors have been removed. Advice is offered at various intervals, helping students to improve their ability in analyzing faulty code and enabling them to identify common mistakes in their own programs more readily. The interface is shown below in Figure 1:

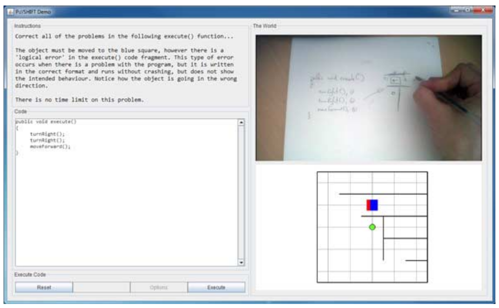

Figure 1. Learning to Trace Code

In order to embed virtual and projective identities within the tool, elements of fantasy role-play were incorporated. In this new version, students would select an avatar, assuming the role of a computer systems specialist on an advanced interstellar spaceship. In this role, students would program repair robots to navigate the ship's maintenance areas in order to fix all of the problems. A video is shown as the tool is launched, setting the scene, and several graphical improvements were added to the user interface, as shown below in Figure 2:

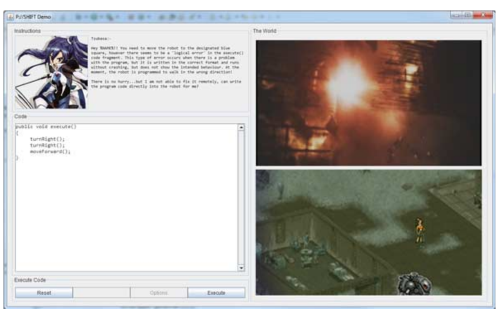

Figure 2. Integrating Fantasy Role-Play

This version of the tool is almost functionally identical to the original. The exercises are the same, requiring about 20 minutes to complete, however each problem is recast into a narrative context. This is conveyed by characters who relate the learning content, instructions and the story through dialogue. The virtual environment is also updated with new graphics, similar to those in a 2D computer game. These changes were necessary to make the exercise more immersive.

3. METHOD

A between-subjects experimental design was adopted due to the potential for preference to bias the self-belief measure. The design consisted of a parallel-group double-blind randomised trial, incorporating balanced allocation between two groups (1:1). Two versions of the virtual lab activity were compared, one incorporating fantasy role-play (experimental condition) and an ablated version, omitting elements of fantasy role-play (control condition). Scores were imputed from a self-completed measurement of academic self-beliefs towards programming, which were measured at pre-test and post-test. The difference between the two groups were compared using an ANCOVA on the post-test scores, using pre-test scores as a covariate.

3.1 Sample

An *a priori* power analysis in G*Power 3.1 showed that for a study of typical significance (α = .05) attempting to detect a "large" effect (f = .5) [2,3], a sample size of 34 would be sufficient to achieve acceptable power (1-β = .80). A sample size of 36 was obtained for this study.

The participants were recruited from a pool of students that had registered interest online (http://www.p-shift.org) in April 2012. To be eligible, participants had to be first or second year undergraduate students enrolled on programming modules at the authors' institution. Furthermore, they had to be aged 18 or over. Although some details were omitted in order to practice blinding, each participant provided informed consent.

3.2 Measurement

A psychometric scale was adapted from an exploration of mathematic self-beliefs by Ferla, Valcke and Cai [6]. This included items relating to academic self-concept (ASC), subject interest (SI), academic self-efficacy (ASE) and learning anxiety (ANX), which were repurposed for use within the programming context. It was not expected that any meaningful change in terms of overall programming could occur within the short timescale of the experiment, so the measure was calibrated to the specific learning domain of the proposed activity: errors and debugging. The items were arranged as a 5-point Likert scale, from strongly disagree to strongly agree, consisting of 22 items.

A pilot study was conducted to validate the scale as a research tool prior to the trial, in which the proposed items were administered to 91 undergraduate students that satisfied the eligibility criteria. A confirmatory factor analysis was performed in AMOS 18.0.0 for Windows. Maximum likelihood estimation was applied to the covariance matrix, but as some multivariate non-normality was discovered (kurtosis = 12.52, c.r = 2.35) a 2000-sample Bollen-Stein bootstrap was applied. During the analysis, several indicators were removed: one due to significant divergence from normality (c.r > 4.0); three due to poor path estimates (r < .5); and one due to excess standard residual covariance (*cov* > 4.0). Covariance was allowed between pairs of error terms on the same factor, as recommended by the modification indices. The model demonstrated adequate fit to the data (χ^2 = 153.56, df = 108, p = .003), as determined by the fit indices presented below in Table 1:

Table 1. Model Fit Indicies and Fit Criteria

Fit Index	Value	Fit Criteria [9]
SRMR	.057	< .08
CFI	.957	> .90
RMSEA	.068	< .08
Bollen-Stein *p*	.149	> .05

A *post-hoc* analysis of sample size and power using Westland's tool [15] indicated the data was sufficient for model testing. Common method bias, measured by common latent factor, was small. Furthermore, adequate reliability, convergent validity and discriminant validity were found in terms of composite reliability (CR), average variance explained (AVE), maximum shared squared variance (MSV), and average shared squared variance (ASV) [9], as shown on the following page in Table 2:

Table 2. Support for Construct Validity in the Factor Model

	Variance and Reliability				Factor Correlations[(a)]			
	CR	*AVE*	*MSV*	*ASV*	*ASC*	*ASE*	*SI*	*ANX*
ASC	.881	.654	.358	.266	.809			
ASE	.876	.588	.540	.404	.395	.767		
SI	.890	.672	.540	.397	.533	.735	.820	
ANX	.865	.617	.517	.414	-.598	-.719	-.605	.786

(a) values on the diagonal indicate √AVE

Consequently, factor score weights generated by AMOS can be utilized to impute composite scores, by a weighted sum procedure using the Likert-type responses, as shown below in Table 3:

Table 3. Factor Score Weights for Each Subscale[(a)]

Indicator	*ASC*	*ASE*	*SI*	*ANX*
I am just not good at programming	**-.035**	.001	-.001	.004
I learn programming quickly	**.583**	-.010	.023	-.074
I have always believed that programming is one of my best subjects	**.063**	-.001	.003	-.008
In my programming labs, I can solve even the most challenging problems	**.017**	.000	.001	-.002
I am confident with organising code into a readable format	-.001	**.051**	.004	-.010
I am confident I can understand Java exceptions (e.g. NullPointerException)	-.001	**.040**	.004	-.008
I am confident I can solve simple problems with my programs	-.005	**.335**	.029	-.066
I am confident I can implement a method from a description of a problem or algorithm	-.007	**.416**	.036	-.083
I am confident I can debug a program that calculates prime numbers	-.001	**.042**	.004	-.008
I enjoy reading about programming	.003	.007	**.116**	-.001
I do programming because I enjoy it	.016	.036	**.638**	-.003
I am interested in the things I learn in programming classes	.003	.006	**.104**	-.001
I think programming is interesting	.003	.007	**.123**	-.001
I often worry that it will be difficult for me to complete debugging exercises	-.005	-.008	.000	**.137**
I often get tense when I have to debug a program on my own	-.007	-.011	.000	**.179**
I get nervous when trying to solve programming bugs	-.011	-.019	-.001	**.302**
I feel helpless when trying to solve some bugs in my programs	-.003	-.004	.000	**.070**

(a) significant loadings have been given in bold face, all items used to impute factor scores

3.3 Procedure

As participants registered interest in the experiment, they immediately completed the online pre-test questionnaire. Once an appropriate number had signed up, they were assigned to one of two groups, of equal size, and then emailed a link to download the relevant version of the tool.

The allocation procedure was performed in Microsoft Excel 2007 (12.0 SP3). First, the identity of each participant was obfuscated using an identification number and arbitrarily allocated a row. The rows were then sorted according to a RAND() value in a separate column. Then, the worksheet was divided into two halves, representing each group.

Participants were free to complete the exercise and the online post-test questionnaire within 10 days. Although this resulted in less control over experimental conditions, it is more representative of how a virtual lab exercise may be used by students in practice, thus enhancing ecological validity. It should be noted, however, that no formal teaching (such as lectures, seminars, or labs) occurred during the time period of the trial. Furthermore, while the pre-test and post-test questionnaires contained the same items, the order in which they were presented was randomised.

4. RESULTS

The data was analyzed using PASW 18.0.3 for Windows. All cases were included in the analysis, with any missing ASC values in the raw pre-test data being replaced by the sample mean and any missing ASC values in the raw post-test data being replaced by the allocation mean.

4.1 Data Analysis

After the allocation, one-way ANOVAs demonstrated that no significant differences existed in terms of pre-test ASC score ($F[1,35] = .214$, $p = .646$, $\eta_p^2 = .006$) or academic grade profile ($F[1,35] = .332$, $p = .568$, $\eta_p^2 = .010$) between the two groups. Descriptive statistics are shown below in Table 4:

Table 4. Descriptive Statistics for Each Group

	Variable	Mean	Std Dev	Confidence Intervals	
				Lower	*Upper*
Experimental	Pre-Test ASC	1.7056	.8411	1.2873	2.1239
	Post-Test ASC	1.7443	.8303	1.3313	2.1572
	ASC Gain	.3866	.0352	.0211	.0561
	Grade	3.8900	1.1320	3.3300	4.4500
Control	Pre-Test ASC	1.5804	.7807	1.1922	1.9686
	Post-Test ASC	1.5991	.7739	1.2142	1.9839
	ASC Gain	.0186	.0291	.0041	.0331
	Grade	4.1100	1.1830	3.5200	4.7000

An ANCOVA examined the impact of fantasy role-play on developing academic self-concept. Assumptions of normality, homogeneity of variance and homogeneity of regression were verified. The results are summarized below in Table 5:

Table 5. ANCOVA Results (*dv* = Post-Test ASC)

Source	SS	*df*	MS	F	*p*	η_p^2
Pre-Test ASC	21.872	1	21.872	5.679	.000	.999
Allocation	.004	1	.004	4.181	.049	.112
Error	.032	33	44.326			

Corrected Model R Squared =.999 (Adjusted R Squared =.998)

Hence, after controlling for pre-test scores, there was a significant difference ($p < .05$) in the post-test ASC between the two groups, which constitutes an effect size between "medium" and "large" [2,3]. This difference can be demonstrated by the estimated marginal means, as shown below in Table 6:

Table 6. Estimated Marginal Means (*dv* = Post-Test ASC)

GROUP	Mean[(a)]	Std Error	Confidence Interval	
			Lower	*Upper*
Experimental	1.682	.007	1.667	1.698
Control	1.661	.007	1.646	1.676

(a) Pre-Test ASC Score = 1.643 (Covariate)

Based on the means and baseline score (1.643) presented above, the gain in academic self-concept for those allocated to the fantasy role-play condition (2.4%) was greater than those in the control condition (1.1%).

5. LESSONS LEARNED

Although a significant difference was detected ($p < .05$), the effect was lower than predicted, resulting in low power ($1-\beta = .54$). This effect, from the η_p^2 value, being $f = 0.355$, and from the observed difference in gain scores, being $d = 0.619$. These results support the notion that fantasy role-play can enhance the development of self-concept, as shown by the gain scores in Figure 3 below:

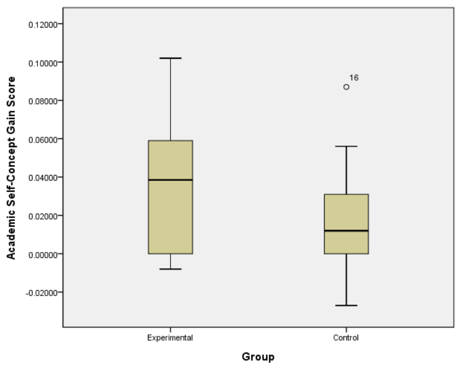

Figure 3. Distribution of ASC Gain Scores

However, there was no significant difference between the two groups at post-test ($F[1,35] = .227$, $p = .591$, $\eta_p^2 = .082$). Thus, the isolated exercise was not sufficient to produce a meaningful difference in academic self-concept. Nevertheless, paired t-tests showed that small gains were present ($t[17] = 4.660$, $p < .001$, $d = .327$; and $t[17] = 2.720$, $p = .015$, $d = .240$). If these small effects can be sustained, then it might result in a more practically significant difference over multiple tasks. Thus, a longitudinal trial is appropriate. Assuming that the observed difference remains consistent ($d = .087$), then perhaps eight activities could produce a more meaningful change ($d \approx .69$).

Of note, the item that emerged as the greatest contributor to academic self-concept was "I learn programming quickly" (.583). This may imply that experiences providing a sense of learning something quickly contribute the most to ASC development. Thus, unnecessary narrative may impede this aim. On the other hand, fantasy role-play can increase engagement [7]. Potentially, maintaining interest and supporting persistent attitudes, enabling more learning to occur within the time available. A goal for further study could be to investigate whether this assertion holds and how to convey such an experience.

The item with the largest mean increase at post-test in the experimental group was "I can solve the most challenging problems". Perhaps the fantasy context can encourage senses of pride and fiero in some students. Thus, helping them to feel they can overcome greater challenges. However, further enquiry is required to corroborate this interpretation.

6. CONCLUSION

An early prototype of a debugging exercise that incorporates elements of fantasy role-play has been shown to strengthen academic self-concept to a greater extent than a conventional approach. This supports the hypothesis that projective identities enhance the self-concept development of novice programmers on a computing course. However, the difference was only of medium size and neither condition had a practically significant impact. Due to the limited scope of this initial study, it remains unclear whether such a modest gain could be further enhanced over time and then maintained. Thus, additional longitudinal study is appropriate. Moreover, qualitative enquiry is necessary to reveal how the fantasy role-play could be improved for greater impact. Future research should anticipate "medium" differences and "small" effects for individual tasks [2,3].

7. REFERENCES

[1] Barker, L., McDowell, C. and Kalahar, K. 2009. Exploring Factors that Influence Computer Science Introductory Course Students to Persist in the Major. ACM SIGCSE Bulletin. 41, 1 (Mar. '09), 153-157. DOI = http://dx.doi.org/10.1145/1539024.1508923.

[2] Cohen, J. 1992. A Power Primer. Psychological Bulletin. 112, 1 (Jul. '92), 155-159. DOI = http://dx.doi.org/10.1037/0033-2909.112.1.155.

[3] Cohen, J. 1977. Statistical Power Analysis for the Behavioral Sciences (Revised Edition). New York: Academic Press.

[4] Daanen, H. and Grant, L. 2007. Space Mission: Ice Moon. In Proceedings of the ACM SIGGRAPH Educators Conference (San Diego, USA, 5-9 Aug., '07), GRAPH'07, ACM, NY, 19. DOI = http://dx.doi.org/10.1145/1282040.1282060.

[5] Erion, P. 1996. Drama in the Classroom: Creative Activities for Teachers, Parents & Friends. New York: Lost Cost.

[6] Ferla, J., Valcke, M., and Cai, Y. 2009. Academic Self-Efficacy and Academic Self-Concept: Reconsidering Structural Relationships. Learning and Individual Differences. 19, 4 (Jun. '09), 499-505. DOI = http://dx.doi.org/10.1016/j.lindif.2009.05.004.

[7] Gee, J.P. 2003. What Video Games Have to Teach Us About Learning and Literacy. New York: Palgrave/Macmillan.

[8] Habgood, M.P., Ainsworth, S.E., and Benford, S. 2005. Endogenous Fantasy and Learning in Digital Games. Simulation and Gaming. 36, 4 (Dec. '05), 483-498. DOI = http://dx.doi.org/10.1177/1046878105282276.

[9] Hair, J., Black, W., Babin, B. and Anderson, R. 2010. Multivariate Data Analysis, Seventh Edition. Upper Saddle River, New Jersey: Prentice-Hall.

[10] Marsh, H. and Martin, A. 2011. Academic Self-Concept and Academic Achievement. British Journal of Educational Psychology. 81, 1 (Mar. '11), 59-77. DOI = http://dx.doi.org/10.1348/000709910X503501.

[11] Papert, S. 1980. Mindstorms: Children, Computers, and Powerful Ideas. New York: Basic Books.

[12] Pattis, R., Roberts, J., and Stehlik, M. 1995. Karel the Robot: A Gentle Introduction to the Art of Object-Orientated Programming, Second Edition. New York: John Wiley & Sons.

[13] Repenning, A. 2012. Programming Goes Back to School. Communications of the ACM.55, 5, (May '12), 38-40. DOI = http://dx.doi.org/10.1145/2160718.2160729.

[14] Resnick, M. et al. 2009. Growing Up Programming: Democratizing the Creation of Dynamic, Interactive Media. In Extended Abstracts from the SIGCHI Conference on Human Factors in Computing Systems (Boston, USA, 4-9 Apr. '09), CHI'09, ACM, NY, 3293-3296. DOI = http://dx.doi.org/10.1145/1520340.1520472.

[15] Westland, J.C. 2010. Lower Bounds on Sample Size in Structural Equation Modeling. Electronic Commerce Research and Applications. 9, 6 (Jul. '10), 476-487. DOI = http://dx.doi.org/10.1016/j.elerap.2010.07.003.

[16] Yee, N. and Bailenson, J. 2007. The Proteus Effect: The Effect of Transformed Self-Representation on Behavior. Human Communication Research, 33, 3 (Jul. '07), 270-290. DOI = http://dx.doi.org/10.1111/j.1468-2958.2007.00299.x.