Investigating the Impact of Personality and Cognitive Efficiency on the Selection of Exercises for Learners

Juliet Okpo
University of Aberdeen
King's College
Aberdeen, UK AB24 3UE
r02jao15@abdn.ac.uk

Judith Masthoff
University of Aberdeen
King's College
Aberdeen, UK AB24 3UE
j.masthoff@abdn.ac.uk

Matt Dennis
University of Portsmouth
Lion Terrace
Portsmouth, UK PO1 3HE
matt.dennis@port.ac.uk

Nigel Beacham
University of Aberdeen
King's College
Aberdeen, United Kingdom AB24 3UE
n.beacham@abdn.ac.uk

Ana Ciocarlan
University of Aberdeen
King's College
Aberdeen, United Kingdom AB24 3UE
ana.ciocarlan@abdn.ac.uk

ABSTRACT

Adapting to learner characteristics is essential when selecting exercises for learners. This paper investigates how humans adapt next exercise selection to learner personality and invested mental effort to enable a future Intelligent Tutoring System to use these adaptations. Participants were presented with validated stories of a learner's personality at polarised levels, a validated story conveying the mental effort invested in carrying out a given task and an indication of a previous performance (just passing) at a simple arithmetic exercise. Participants were also shown a selection of validated exercises of varying difficulty levels and asked to select the exercise which they thought the learner should do next. We found that overall more difficult exercises were selected for learners who used little effort than for learners who used more effort. We found that although an exercise of slightly harder difficulty remains the most popular choice in the high and low self-esteem conditions, for low self-esteem, participants picked an exercise of lower or the same difficulty more often than in the high condition.

CCS CONCEPTS

•**Human-centered computing** → User models; •**Applied computing** → Computer-assisted instruction;

KEYWORDS

Exercise Selection, Adaptation, Personality, Mental effort, Performance, Cognitive Efficiency

1 INTRODUCTION

Capturing the effective behaviours of human tutors to create optimal learning tools has been a major motivation for work in Intelligent Tutoring Systems. Human tutors can effectively sense when a learner is challenged and provide individualized tutoring through careful considerations of the factors that affect the learning process of the learner. Tutoring is tailored to individual characteristics of the learner in order to improve learner performance. It is believed that an intelligent tutoring system can combine specific characteristics of a learner just like a human tutor and adapt teaching and learning to these characteristics for better learning outcomes. One major learner characteristic which is closely related to a learner is the learner's personality which is often overlooked in intelligent tutoring. Research in adaptive learning has considered aspects of the learners such as learner performance, learner mental effort, learning support and several other learner characteristics [10, 24]. In order for an Intelligent Tutoring System (ITS) to select the next exercise for learners, the selection process should be effective such that the relevant characteristics of the learner should be taken into consideration [2, 9, 10]. Very few works have used personality as a learner characteristic for adaptive task selection.

We believe that learner personality is worthy of investigation for adaptive exercise selection, as previous research has shown it to be important [7] in other areas of learning. In addition we will investigate cognitive efficiency (which is a balance between learner effort and performance, see below). In this paper, we will use a particular level of learner performance (just passing) and vary learner effort.

In this paper, we describe the construction and validation of a set of exercises with varying difficulties (Section 3), the validation of stories that express a learner's self-esteem (Section 4) and the construction of statements which describe the mental effort invested by learners (Section 5). We then use these in a user-as-wizard study which investigates the effect of learner personality and mental effort on task selection by participants playing the role of a teacher. The results will be used to inform the creation of an algorithm to allow a future intelligent tutoring system to consider learner personality and cognitive efficiency when selecting exercises for learners to complete.

2 RELATED WORK

There is considerable research on adapting learning content to different learner characteristics [9, 33]. In the area of task selection, the focus has been on the design of intelligent tutors that select learning tasks for the learner based on the learner's past performance, available learning support and recently, cognitive load (e.g. [1, 3, 4, 18, 30, 31]). [32] explores how the activities and principles in expert performance research can be used to design instructional formats based on cognitive load theory for skills mastery. In this work, they showed that learning tasks can be selected on the basis of an assessment of a learner's level of expertise.

Table 1: Exercises Table. All exercises used 2 baskets. Six exercises were used of each type.

Type	No. of Balls	Ball Distribution	Basket total	Example exercise	Validated Difficulty
1	3	1 − 2	<10	(4)(9)(5)	1
2	4	1 − 3	<10	(3)(1)(5)(9)	2
3	4	2 − 2	<10	(6)(5)(3)(4)	2
4	4	2 − 2	≥10	(7)(5)(4)(6)	3
5	5	1 − 4	<10	(1)(1)(9)(2)(5)	3
6	5	2 − 3	<10	(6)(2)(3)(4)(3)	4
7	5	2 − 3	≥10	(2)(8)(6)(9)(3)	5

Personality refers to an individual's whole psychological structure and this includes his or her temperament, character, intelligence, sentiments, attitudes, interests, beliefs, ambitions and ideals. A person's personality is shown by his/her disposition [16]. There have been several studies on adaptation to personality [7]. Self-esteem is defined as how favourably a person regards oneself [27]. For a learner to achieve better learning outcomes in a specific domain, they must believe in their abilities and this belief in the fact that they can produce a favourable outcome will in turn serve as motivation to learn. Self-esteem is seen as an important component of personality [20]. Self-esteem is one of the most widely studied personality concepts in psychology such that in 2001, 20,203 articles had self-esteem as a keyword which made self-esteem the most researched personality concept in comparison with concepts like neuroticism with 20,026 articles and locus of control with 13,428 articles [17]. Significant associations can be found between self-esteem and all personality traits such as openness, conscientiousness, extraversion, agreeableness and neuroticism [12]. For our study, it was decided that self-esteem was a good personality characteristic to investigate first.

Cognitive efficiency, the combination of invested mental effort and performance [5, 26], is also an important aspect of a learner's learning and learning outcomes in the context of assessment, so may need to be taken into account for exercise selection. Additionally, effective exercise selection may need to take learners' mental well-being into account, to keep them motivated. As research has shown that personality can predict and justify 32 to 56 percentages scores' variant of mental well-being [29, 34], this may provide another reason for taking personality into account.

Good estimation of difficulty level of exercises will allow good exercise sequencing [23] as well as better adaptation to the cognitive level of the learners. Previous research has investigated the estimation of difficulty level of exercises using different approaches. Quiz difficulty levels have been evaluated by [19] using similarity measures. While the study investigates the estimation of difficulty of exercises in the learning domain, several other works on the estimation of the difficulty of exercises have been done in various domains using different estimation methods. A graph based strategy for difficulty level estimation for chemistry was used by [35]. Foteini et al presented a neuro-fuzzy approach for the estimation of exercises on search algorithms [13, 14]. In this method, specific characteristics of the exercises are taken as inputs to provide the difficulty of the exercises as output. Also in a report, the level of the difficulty of the exam was based on the item analysis approach where the difficulty of an item was understood as the proportion of persons who correctly answered a test item. This has to do with an inverse relationship such that the higher the proportion of those that answered the test item correctly, the lower the difficulty of the test item [11]. In this paper, we will not investigate how to automatically detect exercise difficulty, but validate exercise difficulty through human studies to allow us to use these validated difficulties in our main study.

3 EXERCISE DIFFICULTY VALIDATION

To investigate the effect of mental effort and personality on the difficulty of exercises, we first needed a set of exercises with validated difficulty levels for participants to select from. In this paper, we have used a simple addition exercise which asks learners to place differently weighted balls into a set number of baskets so that they weighed the same. For example, a learner might be given three balls: (2), (3) and (5) and asked to place them into two baskets. This is a very easy exercise as the heaviest ball is also the sum of the remaining two. Exercises can be made more challenging in many ways. For example, requiring more than one ball in both baskets: (2)(3) in one basket and (4)(1) in the other; increasing the basket total so that more balls are needed: (5)(4)(7)(9) in one and (3)(8)(2)(4)(6) in the other; and by increasing the number of baskets. With this in mind, we wanted to create a systematic way of increasing difficulty using these methods. To achieve this, we designed a validation study described in the next section.

3.1 Study Design

Participants were recruited via Amazon's Mechanical Turk [22], a crowdsourcing tool. Participants required an acceptance rate of 90% to ensure good quality of responses and had to pass a Cloze test for English fluency to ensure they had enough literacy skills to understand the language used for the study. 155 participants took part in the study (74 females and 81 males).

Six pairs of two exercises with different estimated difficulty levels were shown to the participants. These exercises were generated
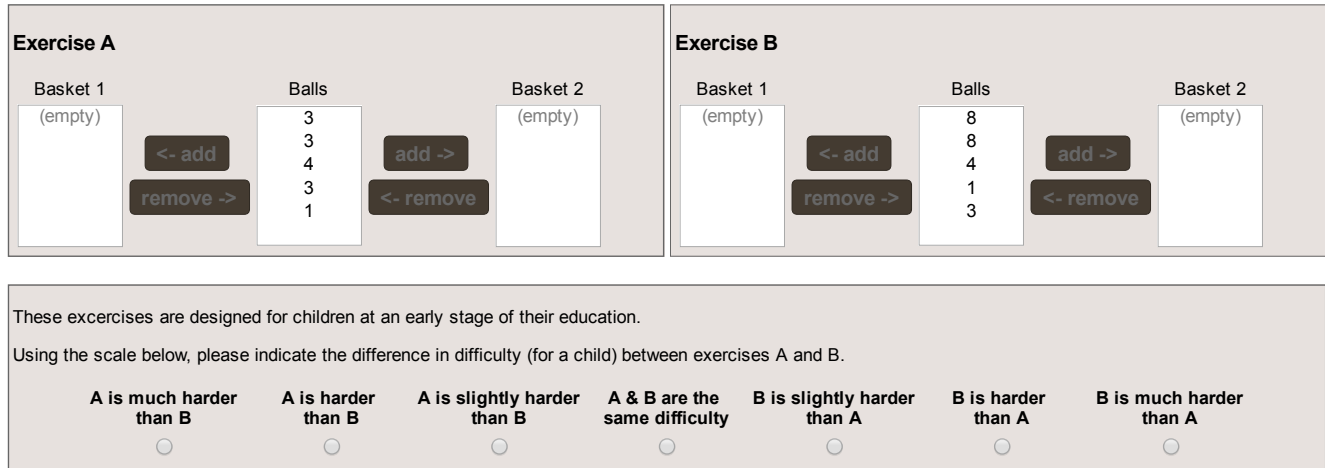
Exercise A

Basket 1		Balls		Basket 2
(empty)	<- add	3	add ->	(empty)
	remove ->	3	<- remove	
		4		
		3		
		1		

Exercise B

Basket 1		Balls		Basket 2
(empty)	<- add	8	add ->	(empty)
	remove ->	8	<- remove	
		4		
		1		
		3		

These excercises are designed for children at an early stage of their education.

Using the scale below, please indicate the difference in difficulty (for a child) between exercises A and B.

A is much harder than B	A is harder than B	A is slightly harder than B	A & B are the same difficulty	B is slightly harder than A	B is harder than A	B is much harder than A
○	○	○	○	○	○	○

Figure 1: Interface for Difficulty Validation Study

Table 2: Results of Difficulty Validation

Study	Exercise Types	Num. Participants	Mean	Std. Dev	ztest
1	1 & 2	25	0.74	0.944	0.000
2	2 & 3	25	-0.13	0.808	0.978
3	3 & 4	25	0.20	0.993	0.006
4	4 & 5	25	-0.29	1.006	0.999
5	4 & 6	25	0.43	1.226	0.000
6	6 & 7	25	0.50	1.184	0.000

using the rules described in Table 1. As previously discussed, we varied the complexity of the exercises by adjusting the number of balls required to be placed in each basket to solve the exercise, and by increasing the basket total.

We had six variants of the study in a between-subjects design, each investigating the difference in difficulty between two types of exercises. Each variant had twelve exercises (six of each type) for participants to solve, shown in pairs (one exercise of each type), where one exercise was expected to be easier than the other. Participants then rated which exercise they thought was easier on a scale as shown in Figure 1. Within each pair, the order was randomized so that the exercise we expected to be easier could appear on the left or right.

3.2 Results

To calculate a relative difficulty score between two exercises of different types, we transformed the scale shown in Figure 1 into numbers. If the exercise with the harder expected difficulty was deemed slightly harder this was scored as 1, harder as 2 and much harder as 3 (and corresponding negative scores for the reverse). If the exercises were deemed of the same difficulty a score of 0 was used. The score was averaged over the six comparisons of the two exercise types each participant did.

From our results in Table 2, the Z-test of the grand mean scores compared to 0 shows that exercise types 1 and 2 had a significant difference and therefore type 2 is more difficult than type 1. Exercise types 2 and 3 showed no significant difference. Exercise type 3 was then tested with exercise type 4 which gave a significant result showing that type 4 was more difficult than type 3 (or 2). Testing exercise type 4 against type 5 yielded no significant difference indicating that types 4 and 5 were of the same difficulty. We then tested exercise type 4 with type 6 and this returned a significant difference in difficulty with type 6 being more difficult than type 4 (or 5). Finally, exercise type 6 was tested with type 7 and the result showed a high significant difference proving that type 7 was more difficult than type 6. Consequently, we have been able to derive 5 levels of difficulty of exercises (see Table 2).

4 SELF-ESTEEM STORY VALIDATION

This section describes the creation and validation of stories which express learner self-esteem at polarized levels (high and low), following a similar approach to [6]

4.1 Story Development

To construct the stories, we used the well-established State Self-Esteem scale [15]. The SSES consists of 20 items that measure momentary fluctuations in self-esteem. For each story, we changed a selection of the questionnaire items into third person, inverting them where necessary. In trying to make the story real, we linked it with a character, a student called Nancy. The resulting stories are shown in Table 4.

4.2 Story Validation

40 participants saw one of the two stories in a between-subjects design (following a similar approach to [6]). Participants were crowd-sourced on Amazon's Mechanical Turk [22]. They were adults based in the US, had to pass an English fluency test and have

Table 3: Mental Effort Statements Mapped to Effort Rating

Description	Effort rating by participants (%)										Average Rating
	1 (none)	2	3	4	5	6	7	8	9	10 (max)	
no effort	95%		5%								1.10
minimal effort	24%	**52%**	14%	10%							2.10
little effort		24%	**71%**			5%					2.95
token effort		35%	10%	10%	15%	20%		5%		5%	4.25
some effort		5%	14%	52%		24%	5%				4.38
average effort				10%	71%	10%	10%				5.19
moderate effort				14%	**48%**	24%	10%			5%	5.52
quite a lot of effort		5%			10%	14%	33%	33%		5%	6.90
considerable effort					10%	24%	38%	14%	14%		7.00
substantial effort					10%	25%	30%	25%	10%		7.00
much effort			5%			14%	**52%**	24%	5%	0%	7.00
very much effort						10%	10%	48%	29%	5%	8.10
ultimate effort						10%	5%	0%	10%	76%	9.38
maximum effort							5%	5%		10%	81%	9.57
all possible effort							5%	5%	10%	**80%**	9.65

an approval rating of 90%. To validate that the stories conveyed self-esteem at the desired level, participants were asked to rate the self-esteem of Nancy using the Rosenberg self-esteem questionnaire [28], which uses different terms to measure the same concept.

A between subjects T-test was performed on the self-esteem score between the high and low self-esteem stories, which was significant at $T(38) = 13.93, p < 0.001$. The mean self-esteem score for high was 24.6 ± 3.68 SD and the mean for low was 8.0 ± 3.87 SD. The Rosenberg scale ranges from 0-30. Scores between 15 and 25 are within the normal range, while scores below 15 suggests low self-esteem [28]. The level of self-esteem for the low story is lower than the normal range. The level of self-esteem for the high story is at the top end of the normal range.

Given the large difference in the self-esteem levels between the two stories, we believe the stories are suitable for future studies to express high and low learner self-esteem.

Table 4: Stories depicting High and Low Self Esteem

SE Level	Story
High	Nancy is a learner who is confident about her abilities. She is satisfied about the way she looks and feels good about herself. She thinks she is as smart as others and believes that others admire and respect her. She feels that she has a good understanding of things.
Low	Nancy is a learner who worries about the impression she makes and whether she is regarded as a success or a failure. She feels like she is not doing well and she believes she cannot understand the things she reads. Nancy thinks she is unattractive and is displeased with herself. She feels inferior to others.

5 MENTAL EFFORT STATEMENT VALIDATION

5.1 Statement Generation

Many statements can be used to describe the amount of mental effort that has been invested by a learner in carrying out a task. Unfortunately, we have not found a list that clearly defines varying levels of mental effort. To have a clear definition of invested mental effort, we wanted to map mental effort statements to numbers indicating the effort used, so that we could use a selection of statements in our studies. In a brain-storming session, three lecturers and a research student came up with 15 invested mental effort statements (see statements in Table 3).

5.2 Statement Validation

26 participants (staff and students of the University) completed the survey using an on-line survey tool (Survey Monkey) in about 5 minutes. The data from 21 participants were used (16 males, 4 females, 1 undisclosed). The other 5 participants were excluded due to the low quality of their responses (providing the same response for almost all statements, or not mapping 'no effort' to 1 as was indicated on the scale).

Participants read the mental effort statements and mapped them to numbers from 1 to 10 with 1 representing no effort and 10 representing maximum effort. The order of the mental effort statements was randomized for each participant.

Table 3 shows the percentage of participants who mapped a statement to a particular number. Some statements (e.g. 'token effort') showed little agreement between participants, whilst others showed better agreement. We decided to use five statements (shown in bold) for the main study.

Now that you have completed some exercises, read the story about Nancy below and recommend an exercise for her to do next.

Meet Nancy

Nancy is a learner who worries about the impression she makes and whether she is regarded as a success or a failure. She feels like she is not doing well and believes she cannot understand the things she reads. Nancy thinks she is unattractive and is displeased with herself. She feels inferior to others.

Nancy completed a set of mathematics exercises of difficulty level 2. Each exercise required the addition of 4 balls representing 1 digit numbers. The totals in the baskets are 1 digit numbers.(For example: ④, ①, ②, and ⑦).

Ten exercises were given to Nancy and she **just passed.** Nancy put **all possible effort** into solving the exercises.

Assuming you were Nancy's Mathematics teacher, which next exercise would you give to Nancy from the list of exercises below?

Select the exercise that you would give to Nancy by clicking on the button on the right of your chosen exercise.

Exercise	Example	Difficulty	Select
Addition of 3 balls representing 1 digit numbers. The totals in the baskets are 1 digit numbers. For example: **9, 2 and 7**. These are an easier type of exercise than the ones she did before.	9 2 7	Difficulty Level 1	○
Addition of 4 balls representing 1 digit numbers. The totals in the baskets are 1 digit numbers. For example: **3, 1, 5 and 9** These are the same type of exercise that she did before.	3 1 5 9	Difficulty Level 2	○
Addition of 4 balls representing 1 digit numbers. The totals in the baskets are 2 digit numbers. For example: **7, 5, 4 and 6** These exercises are more difficult than those she did before.	7 5 4 6	Difficulty Level 3	○
Addition of 5 balls representing 1 digit numbers. The totals in the baskets are 1 digit numbers. For example: **6, 2, 3, 4 and 3** These exercises are considerably more difficult than those she did before.	6 2 3 4 3	Difficulty Level 4	○
Addition of 5 balls representing 1 digit numbers. The totals in the baskets are 2 digit numbers. For example: **2, 8, 6, 9 and 3** These exercises are the most difficult of all exercises.	2 8 6 9 3	Difficulty Level 5	○

Figure 2: Screenshot of the exercise selection stage of the study

6 IMPACT OF PERSONALITY AND MENTAL EFFORT ON EXERCISE SELECTION

Using the set of exercises of validated difficulty levels and validated mental effort statements, we investigate what impact personality and mental effort have in selecting the next exercise for learners. In this study, we focus on learners who have 'just passed' as performance.

6.1 Methodology

We used the User-as-Wizard [21] methodology for our study where the role of the system is played by participants in selecting exercises for a learner to do next. This study builds on previous research investigating the impact of personality and performance in selecting the next exercise for learners [25] by including invested mental effort and using validated difficulty levels of exercises. Following the methodology used by [8], we use crowd-sourcing first to inspire the algorithm, and will validate results with teachers later, followed by investigating the actual impact on learners.

6.2 Participants

201 participants successfully completed the study (107 males, 94 females, 0 undisclosed; 26 aged 18 - 25, 114 aged 26 - 40, 59 aged 41 - 65, 1 aged 65 and 1 undisclosed; 27 were students, 9 were teachers and 165 other). The study was administered through crowd-sourcing on Amazon's Mechanical Turk (MTurk) [22]. To be eligible to be part of the study, participants had to pass an English Cloze test. Participants were also required to successfully complete 5 short exercises (of difficulty levels 1 to 5) similar to the ones that they could select for the learners to do next. In addition, participants had to have a 90% acceptance rate meaning that 90% of the work they do on MTurk is accepted by other requesters as being of good quality.

6.3 Materials

Validated stories conveyed the self-esteem of a fictional learner 'Nancy' (see Table 4). A sentence was added to indicate her past performance, and one of the validated mental effort statements was shown from the bold items in Table 3. The statements were selected to ensure a good spread of difficulties, and had good inter-rater agreement. However, we decided to exclude "no effort" and "maximum effort" as they were used in the explanation of the scale that participants saw in the validation experiment, and we excluded "average effort" as this could be affected by the learning domain. In addition, a set of exercises with validated difficulty levels (see Table 1) were presented to participants to select the one Nancy should do next.

6.4 Procedure

Participants began by completing five short exercises just like the ones that the learners would do so they gained an understanding of the different difficulty levels. The order of exercises presented to participants was from easiest to most difficult. In a between-subjects design, participants were then asked to select the exercise Nancy should do next given her self-esteem, past performance and invested mental effort, as shown in Figure 2. Participants were informed that it was their opinion that counted and as such there were no right or wrong answers.

6.5 Variables

The dependent variable for this study is the difficulty level of the exercise selected for Nancy to do next. Participants could choose between 5 difficulty levels (level 1 to 5). They were told that Nancy had done 10 exercises of difficulty level 2 before, and that she *just passed*. So, participants could select exercises of the same difficulty (level 2), or an easier difficulty (level 1) or of varying degrees of more difficulty (levels 3-5).

The independent variables used for this study are: learner self-esteem, high and low and their invested mental effort in solving the exercises (*minimal effort, little effort, moderate effort, much effort and all possible effort*).

6.6 Hypotheses

- **H1**: Overall, participants will select a more challenging exercise for High SE learners than for Low SE learners.
- **H2**: Overall, participants will select a more challenging exercise for learners with a lower mental effort than high mental effort
- **H3**: Participants will select a different difficulty level for the exercise depending on the combination of SE level and mental effort.

6.7 Results

Table 5 shows the results for each condition. We ran a 2-way ANOVA of the independent variables *self-esteem × effort* for *difficulty*. This was significant for both *effort* ($F(4, 200) = 4.12$, $p < 0.004$) and *self-esteem* ($F(1, 200) = 14.04$, $p < 0.001$), however, the interaction of *effort × self-esteem* was not significant.

Figure 3 shows the results overall for effort and self esteem. Looking at effort alone, it can be seen that a difficulty of 3 (*slightly harder than before*) is the most popular choice for all levels of effort. However, when effort is minimal or little, a higher percentage of participants recommend a higher level of difficulty than this (the majority choose a difficulty of 4 or 5). H2 is thus confirmed, participants did choose an exercise of higher difficulty for learners who required little effort to complete the exercise.

For self-esteem, Figure 3 shows that although an exercise of slightly harder difficulty (level 3) remains the most popular choice in both conditions, for low self-esteem, participants pick an exercise of lower or the same difficulty more often than in the high condition. Thus there is support for hypothesis H1.

With respect to hypothesis H3, we have evidence that both Self-Esteem and mental effort mattered to participants when selecting exercises overall, but as there is no interaction effect, we do not have strong enough evidence to factor in both mental effort and self-esteem at the same time, meaning this is not well supported.

7 CONCLUSION AND FUTURE WORK

From this study we conclude that learner personality and mental effort are important considerations for exercise selection. We now have a set of exercises with validated difficulty levels, validated mental effort levels and validated personality trait stories. From

Table 5: Exercise Selection Study Results

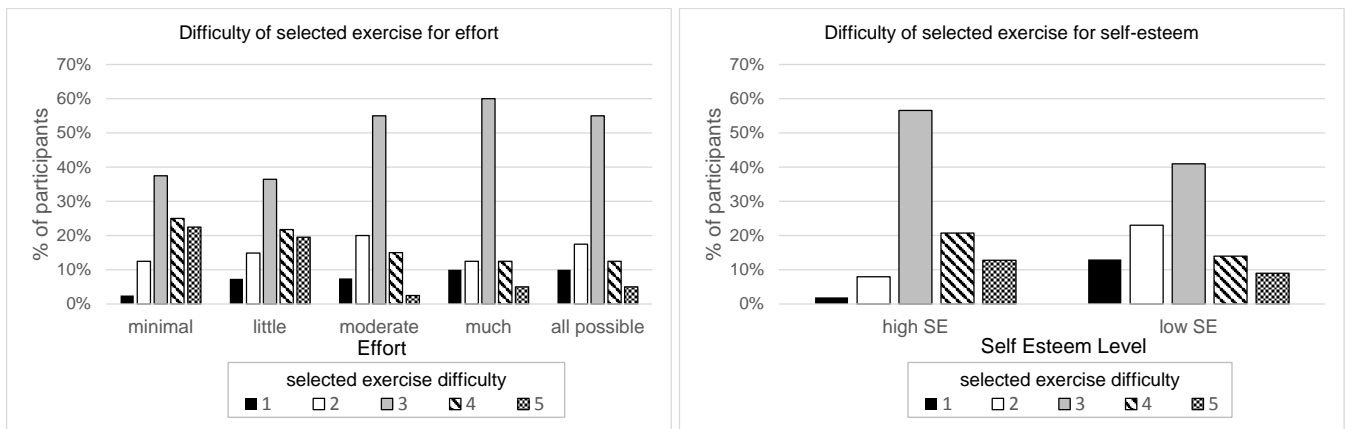| | Selected difficulty level (% of participants) | | | | | | | | | |
| | 1 | | 2 | | 3 | | 4 | | 5 | |
Effort	High SE	Low SE	High SE	Low SE	High SE	Low SE	High SE	Low SE	High SE	Low SE
minimal effort	0	5	0	25	40	35	25	25	35	10
little effort	5	10	5	25	43	30	29	15	19	20
moderate effort	5	10	10	30	60	50	20	10	5	0
much effort	0	20	10	15	80	40	10	15	0	10
all possible effort	0	20	15	20	60	50	20	5	5	5

Figure 3: Chosen exercise difficulty for effort (left) and self-esteem (right).

our current findings, we now have an indication of how exercises can be selected to achieve better learning outcomes. We know that overall, according to our participants, more challenging exercises should be selected for those learners who completed the task with less effort. A further study is needed to investigate the actual effect on learners of doing so.

We also know that if learner self-esteem is low, participants selected an exercise of the same or easier difficulty more often. However, as there was no clear interaction effects discovered, further investigations are required to determine how best to proceed with a strategy for adaptation to both characteristics at the same time. A future study could investigate whether learners with low self-esteem who have 'just passed' really would prefer easier exercises than those with high self-esteem in a controlled setting.

In this paper, we only considered one learner performance (*just passed*). Obviously, learner performance will affect the difficulty of subsequent exercise selection, however it may be that mental effort and personality will trigger different adaptation strategies for exercise selection at these performance levels. Future studies based on the methodology outlined in this paper will investigate this. Furthermore, we can investigate the effect of other personality traits using existing *Personality trait stories* to allow an ITS to adapt to other facets of learner personality. Based on previous research [7], we expect learner conscientiousness and neuroticism from the five-factor model to be relevant traits.

Due to the need for large amounts of data, this paper utilized crowd-sourcing to recruit participants. Any resulting adaptations require the input of experts in the learning domain for further refinement and to verify that the adaptations are appropriate. We can then incorporate these findings into an algorithm for an intelligent tutoring system to make use of these adaptations.

REFERENCES

[1] Gino Camp, Fred Paas, Remy Rikers, and Jeroen van Merrienboer. 2001. Dynamic problem selection in air traffic control training: A comparison between performance, mental effort and mental efficiency. *Comput Hum Behav* 17, 5 (2001), 575–595.
[2] Madeline M Carrig, Gregory G Kolden, and Timothy J Strauman. 2009. Using functional magnetic resonance imaging in psychotherapy research: A brief introduction to concepts, methods, and task selection. *Psychotherapy Research* 19, 4-5 (2009), 409–417.
[3] Gemma Corbalan, Liesbeth Kester, and Jeroen JG Van Merriënboer. 2006. Towards a personalized task selection model with shared instructional control. *Instructional Science* 34, 5 (2006), 399–422.
[4] Gemma Corbalan, Liesbeth Kester, and Jeroen JG van Merriënboer. 2008. Selecting learning tasks: Effects of adaptation and shared control on learning efficiency and task involvement. *CEP* 33, 4 (2008), 733–756.
[5] Lyn Corno and Mary Rohrkemper. 1985. The intrinsic motivation to learn in classrooms. *Research on motivation in education* 2 (1985), 53–90.
[6] Matt Dennis, Judith Masthoff, and Chris Mellish. 2012. The quest for validated personality trait stories. In *Proceedings of IUI 2012*. ACM, ACM, 273–276.
[7] Matt Dennis, Judith Masthoff, and Chris Mellish. 2015. Adapting Progress Feedback and Emotional Support to Learner Personality. *IJAIED* (2015), 1–55.
[8] Matt Dennis, Judith Masthoff, and Chris Mellish. 2016. Adapting progress feedback and emotional support to learner personality. *International Journal of Artificial Intelligence in Education* 26, 3 (2016), 877–931.

[9] Paula J Durlach and Jessica M Ray. 2011. *Designing adaptive instructional environments: Insights from empirical evidence.* Technical Report. DTIC Document.

[10] Paula J Durlach and Randall D Spain. 2014. *Framework for Instructional Technology: Methods of Implementing Adaptive Training and Education.* Technical Report. DTIC Document.

[11] Eduardo Backhoff Escudero, Norma Larrazolo Reyna, and Martín Rosas Morales. 2000. The level of difficulty and discrimination power of the Basic Knowledge and Skills Examination (EXHCOBA). *Revista Electrónica de Investigación Educativa* 2, 1 (2000), 2.

[12] LR Goldberg. 1980. Some ruminations about the structure of individual differences: Developing a common lexicon for the major characteristics of human personality. In *Invited paper, Convention of the Western Psychological Association, Honolulu, Hawaii.*

[13] Foteini Grivokostopoulou, Isidoros Perikos, and Ioannis Hatzilygeroudis. 2015. Estimating the Difficulty of Exercises on Search Algorithms Using a Neuro-fuzzy Approach. In *Tools with Artificial Intelligence (ICTAI), 2015 IEEE 27th International Conference on.* IEEE, 866–872.

[14] Foteini Grivokostopoulou, Isidoros Perikos, and Ioannis Hatzilygeroudis. 2017. Difficulty Estimation of Exercises on Tree-Based Search Algorithms Using Neuro-Fuzzy and Neuro-Symbolic Approaches. In *Advances in Combining Intelligent Methods.* Springer, 75–91.

[15] Todd F Heatherton and Janet Polivy. 1991. Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social psychology* 60, 6 (1991), 895.

[16] Anita Woolfolk Hoy. 2013. A reflection on the place of emotion in teaching and teacher education. *Advances in Research on Teaching* 18 (2013), 255–270.

[17] Timothy A Judge, Amir Erez, Joyce E Bono, and Carl J Thoresen. 2002. Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? *Journal of personality and social psychology* 83, 3 (2002), 693.

[18] Danny Kostons, Tamara van Gog, and Fred Paas. 2010. Self-assessment and task selection in learner-controlled instruction: Differences between effective and ineffective learners. *Computers & Education* 54, 4 (2010), 932–940.

[19] Chenghua Lin, Dong Liu, Wei Pang, and Zhe Wang. 2015. Sherlock: A Semi-automatic Framework for Quiz Generation Using a Hybrid Semantic Similarity Measure. *Cognitive computation* 7, 6 (2015), 667–679.

[20] Abraham H Maslow. 1973. *On dominance, self-esteem, and self-actualization.* Maurice Bassett.

[21] Judith Masthoff. 2006. The user as wizard: A method for early involvement in the design and evaluation of adaptive systems. In *5th Workshop on User-centred Design and Adaptive Systems.* 460–469.

[22] MT. 2012. Amazon Mechanical Turk. http://www.mturk.com. (2012).

[23] Juliet Okpo. 2016. Adaptive exercise selection for an intelligent tutoring system. In *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization.* ACM, 313–316.

[24] Juliet Okpo, Matt Dennis, Judith Masthoff, Kirsten A Smith, and Nigel Beacham. 2016. Exploring Requirements for an Adaptive Exercise Selection System. In *Proceedings of the 6th Workshop on Personalization Approaches in Learning Environments (PALE 2016).* 24th conference on User Modeling, Adaptation, and Personalization (UMAP 2016), CEUR workshop proceedings, this volume.

[25] Juliet Okpo, Matt Dennis, Kirsten Smith, Judith Masthoff, and Nigel Beacham. 2016. Adapting exercise selection to learner self-esteem and performance. In *Intelligent Tutoring Systems.* Springer, 517–518.

[26] Paul R Pintrich and Elisabeth V De Groot. 1990. Motivational and self-regulated learning components of classroom academic performance. *Journal of educational psychology* 82, 1 (1990), 33.

[27] Morris Rosenberg. 1986. *Conceiving the self.* RE Krieger.

[28] Morris Rosenberg, Carmi Schooler, Carrie Schoenbach, and Florence Rosenberg. 1995. Global self-esteem and specific self-esteem: Different concepts, different outcomes. *ASR* (1995), 141–156.

[29] Nabiollah Sadeghi, Zalina Mohd Kasim, Bee Hoon Tan, and Faiz Sathi Abdullah. 2012. Learning styles, personality types and reading comprehension performance. *English Language Teaching* 5, 4 (2012), 116.

[30] Ron JCM Salden, Fred Paas, and Jeroen JG Van Merriënboer. 2006. Personalised adaptive task selection in air traffic control: Effects on training efficiency and transfer. *Learning and Instruction* 16, 4 (2006), 350–362.

[31] Ana C Stephens, Eric J Knuth, Maria L Blanton, Isil Isler, Angela Murphy Gardiner, and Tim Marum. 2013. Equation structure and the meaning of the equal sign: The impact of task selection in eliciting elementary studentsfi understandings. *JMB* 32, 2 (2013), 173–182.

[32] Tamara Van Gog, K Anders Ericsson, Remy MJP Rikers, and Fred Paas. 2005. Instructional design for advanced learners: Establishing connections between the theoretical frameworks of cognitive load and deliberate practice. *Educational Technology Research and Development* 53, 3 (2005), 73–81.

[33] Tamara Van Gog, Liesbeth Kester, and Fred Paas. 2011. Effects of concurrent monitoring on cognitive load and performance as a function of task complexity. *Applied Cognitive Psychology* 25, 4 (2011), 584–587.

[34] M Vandad Sharifi, Ahmad Hajebi, and Reza Radgoodarzi. 2015. Twelve-month prevalence and correlates of psychiatric disorders in Iran: The Iranian Mental Health Survey, 2011. *Archives of Iranian medicine* 18, 2 (2015), 76.

[35] Gang Wu and Irene Cheng. 2007. An interactive 3D environment for computer based education. In *Multimedia and Expo, 2007 IEEE International Conference on.* IEEE, 1834–1837.