EPiC Series in Education Science

Volume 5, 2023, Pages 104–115

Proceedings of NEMISA Digital Skills Conference 2023: Scaling Data Skills For Multidisciplinary Impact

Towards improving learning experiences in self-paced online learning courses

Sibukele Gumbo[1] and Hossana Twinomurinzi[1]

[1] University of Johannesburg, Johannesburg, South Africa

222239719@student.uj.ac.za, hossanat@uj.ac.za

Abstract

One of the key goals of higher education institutions (HEIs) is student career success. In HEIs, students are given the necessary subject-specific information, skills, and experience to achieve this. In South Africa, which has historically been one of the technological leaders in Africa, there are currently insufficient Fourth Industrial Revolution (4IR) specialists graduating from higher education to meet the rising need for a skilled workforce in those fields. The completion rate of supplementary online courses to expand this 4IR skills base is also a major concern. In this paper, we report on the strategy implemented to improve the course completion rate of a self-selected sample of students who attended face-to-face Data Science introduction workshops in the Eastern Cape province of South Africa. We achieved a 76% completion rate of two MOOCs courses in these workshops. We examined the contribution of the participant's personal factors and background contextual factors. We also listed any other factors suggested by the students, which collectively contribute to the learning experience construct of the social cognitive career theory (SCCT) and serve as a practical means of improving 4IR skills exposure and the outcome expectations. Therefore, this paper shows the mechanism which can be used to offer 4IR programmes, in HEIs, to raise students' self-efficacy and outcome expectations in fields such as Data Science. In addition, the learning experience inputs can be utilized to increase students' interest in majoring in 4IR courses.

1 Introduction

Massive Open Online Courses (MOOCs) are online programs intended for open (free) enrollment to a large audience. MOOCs offer a cost-effective and flexible method to master new skills, enhance careers, and deliver high-quality educational experiences on a large scale. MOOCs are an example of a self-paced learning strategy, which can be used for career advancement and supplemental learning. Other learning strategies include the use of internet resources, academic databases and discussion forums. The first MOOC was founded by Dave Cormier and was based on the connectivist theory of learning which emphasizes communal over individual learning (Mooc.org, 2023). Sustainable

Development Goal 4 (SDG4) of the United Nations (UN), which calls for "inclusive and equitable quality education. for everyone," is in accordance with the idea of MOOCs (United Nations, 2023). Given the high young unemployment rate, which positively correlates with limited education, SDG4 should be adopted by developing nations like South Africa (SA). Even though MOOCs have the potential to enhance educational results, countries still struggle to fully accept, integrate, and utilize them. For example, one publication states that the MOOCs current completion rate is 7–13% (Nesterowicz et al., 2022).

South Africa has twenty six public universities distributed throughout its nine provinces (SA Department of Basic Education, 2021). Four of them are located in the Eastern Cape province. None of the universities in the Eastern Cape provide Data Science courses as part of their regular curricula, according to (Twinomurinzi et al., 2022). Data Science is an interdisciplinary field with an emphasis on information extraction from data sets, which are frequently very large in size. The field includes analysis, preparing data for analysis, and presenting results to support organizational choices [5]. We see MOOCs as a viable alternative for building Data Scientists in the Eastern Cape. MOOCs can provide free access to online courses in HEI, in a variety of topic areas, including Data Science, for as many people as is practical. The principle that "information should be given freely and the desire to learn should be satisfied without demographic, economic, and geographical limits" is the foundation of MOOCs (Mooc.org, 2023).

We opted to use the learning experience construct of the social cognitive career theory (SCCT) model to assess the adoption of the Data Science programmes in a pilot workshop conducted at an Eastern Cape HEI. Robert W. Lent, Steven D. Brown, and Gail Hackett developed the SCCT model shown in Figure 1 (Medugorac et al., 2020), in 1994. It is a career development theory that explains how individuals make career choices by considering their personal and environmental factors, including their personal beliefs and attitudes, their self-efficacy, their outcome expectations, and their situational factors. It provides a systematic explanation for career development, respond to development of times as stated by (Buthelezi et al., 2009) and helps on focusing on special group like the youth on 4IR career The theory proposes that individuals make career decisions based on a dynamic interaction between their internal cognitive and affective processes and external environmental factors. SCCT provides a framework for understanding the complex process of career decision making and can be used to guide career counseling and support individuals in making informed career choices. It has gained appeal as a theoretical framework for examining career-related decisions across a range of groups, including jobs in Science, Technology, Engineering, and Mathematics (STEM). It is based on Albert Bandura's social cognition theory. Over the past 20 years, the SCCT model has been applied in numerous contexts to examine both professional and academic behavior. Its premise is that the best explanation for career development and decisions connected to careers is the consequence of a complicated web of interrelated elements.

The theory also offers opportunities for social support by instilling expectations, self-efficacy and through the use of observational learning as well as some reinforcements in order to achieve change in behavior. It further assist on personal judgment as to how well an individual can be able to execute a course of action which they are required to encounter with contemporary situations like the era of 4IR. (Buthelezi et al., 2009) specified that SCCT helps to construct a three factor interaction of careers which is crucial on this aspect too which comprise of

1. Self efficacy (can I do this),
2. Outcome expectations (what will happen if one doe it),
3. Personal goal (this regards to as how much one wants to do it).

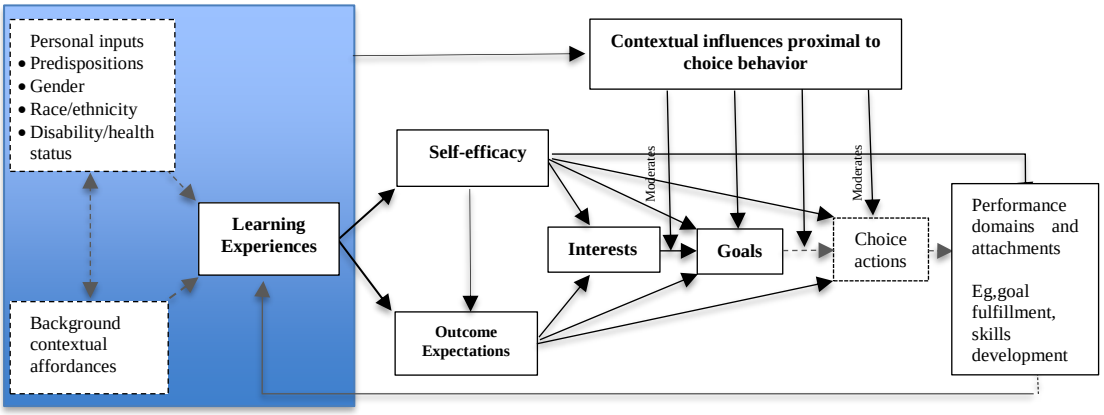

Figure 1: SCCT model (Medugorac et al., 2020) adapted as follows: we study factors highlighted in blue and relationships are indicated by dashed lines, while the remaining variables and factors are represented by solid lines.

According to SCCT (Figure 1), the career-relevant learning experience construct is influenced by individual inputs (e.g. gender and predispositions) and background contextual affordances. Learning experiences also includes performance successes, vicarious learning, social persuasion, and physiological arousal [6]. In order to understand the potential influences on students in HEI and their acceptance of MOOCs, a pilot survey was conducted to identity whether personal and background characteristics influenced students' MOOC learning experiences. We examined these learning experience precursors (highlighted in blue in Figure 1), to determine the extent in which they influenced students' completion of MOOCs courses. The respondents also shed some light on additional factors that affect MOOCs completion and those that could improve these online programs in developing countries. We also discuss 1) the aspects of the Data Science MOOCs that the respondents consider useful or valuable and 2) suggestions by the respondents on how to improve the courses. The following research questions were addressed:

• RQ1: What are the personal and background contextual characteristics of the self-group of HEI students undertaking MOOCs in the Data Science domain?

• RQ2: What other factors can improve MOOCs completion?

The rest of the paper is structured as follows: Section 2 describes the method used to collect data, Section 3 is the results and discussions and, Section 5 is the conclusion.

2 Method

In (Class Central, n.d.), IBM is listed as a provider of MOOCs. The choice to use its platform was based on the availability of free, industry-related courses to increase student competence. We invited first and second year Information Technology (IT) students based at a HEI in the Eastern Cape to a face-to-face workshop to introduce two Data Science courses offered on the IBM online platform. The purpose of the workshop was to introduce students to self-learning technologies and educational opportunities so they update their knowledge, skills, and understanding. IBM offers each participant free access to content, course material, lab exercises, and a certificate. Text, videos, online access to labs, and an examination at the end of each course comprise the content.

The courses chosen were Data Science 101 (IBM, n.d.-a)and Data Science Tools (IBM, n.d.-b). The students were required to bring their laptops and headphones to the workshops. We made use of wireless internet-equipped university facilities. At the end of the workshop, we used an online instrument to conduct a survey where willing workshop participants gave feedback on the following elements:

1. Personal inputs such as gender, age, and ethnicity;
2. Background contextual affordances, including hometown, highest qualification earned, grades in math and English in the matriculation, prior online training, and successful completion of online training;
3. Aspects of the workshop and courses that were considered useful or could be improved.

We also observed all the participants throughout the session and observed whether respondents proceeded with additional online courses after a workshop session. The results from the workshops are described in the next sections.

3 Results and discussions

3.1 Workshop participation

Three Data Science pilot face-to-face workshops took place in October 2022. Table 1 shows the number of students who 1) expressed their interest in the workshops, 2) attended the workshops and 3) completed the online survey. In summary, 320 students expressed their interest in the workshops, 2) 201 students (or participants) attended and 3) 72 students (or respondents) completed the online survey to evaluate their learning experience. The three venues used for the workshops could to accommodate 40, 70 and 100 students, respectively. Due to space restrictions, additional students who had not rsvped were turned away. None of the participants had previously undertaken any IBM online courses; new online IBM platform user accounts were created for all of them.

	Expressed interest	Participated in workshop	Completed online survey
Venue 1	37	26	14
Venue 2	113	73	50
Venue 3	170	102	8
TOTAL	320	201	72

Table 1: Expressed interest vs Participated in workshop vs Completed online survey

3.2 Survey respondents' personal inputs

The survey respondents personal inputs such as gender, age, and ethnicity are summarized in this section. These responses represent key pieces of their demographic data.

39 respondents (or 54.2) were female and 33 respondents (or 45.8%) were male as illustrated in Figure 2. This demonstrates that there is a rise in female students who are interested in technology-related courses at HEIs. The interest in females in technology courses complements the survey by Deloitte Worldwide which states that 'leading global technology organizations will, on average, have nearly 33% of women working for them in 2022, up little over 2 percentage points from 2019'(Hupter et al., 2021).

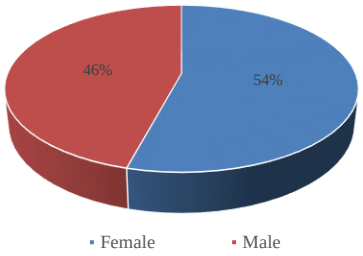

Figure 2: Gender of respondents

Figure 3 illustrates that the respondents' ages ranged from 18 to 34, with a mean age of 22. In the South African perspective, all respondents were considered youth. The introduction of Data Science courses to them contributes towards the advancement and development of a youthful 4IR workforce.

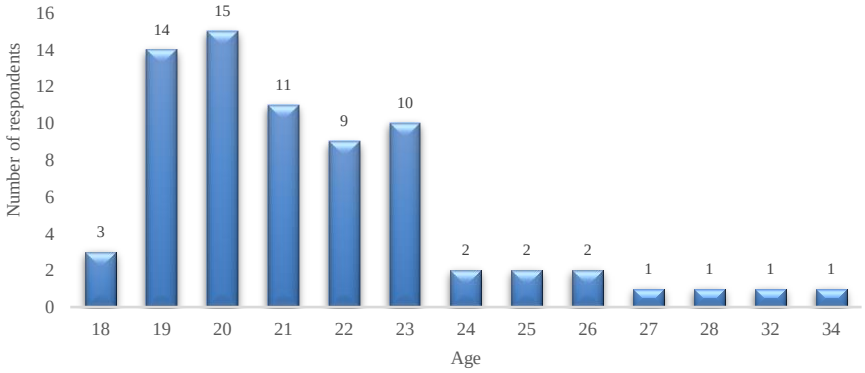

Figure 3: Age distribution of respondents

The four recognized ethnic groups in South Africa are Blacks (or Africans), Whites, Coloreds, and Indians. According to the 2011 census, there are 76.4% of Black South Africans, 9.1% of White South Africans, 8.9% of Colored South Africans, 2.5% of Indian South Africans, and 0.5% of Other/Unspecified South Africans ("Ethnic Groups in South Africa," 2022). Figure 4 displays the distribution of Data Science survey respondents' ethnic groups. One person was Colored, and the rest were African. Therefore, this study was based on non-native English-speaking MOOCs survey respondents.

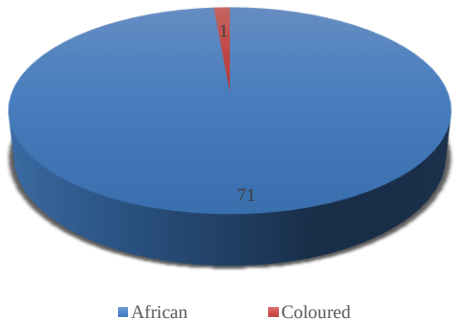

Figure 4: Distribution of ethnic groups

3.3 Survey respondents' background contextual affordances

The environmental background elements for the respondents are outlined in this section. We summarised the respondents' home areas, their highest degree held at the time they attended the Data Science sessions, their matriculation Math and English final results, and whether they had previously taken or completed a MOOC.

The distribution of the survey respondents' home areas are illustrated in Figure 5 and Figure 6. The majority of the students came from the Eastern Cape (68%) province where the university is located. 10% of the respondents were from Gauteng and 14% of the respondents came from Kwa-Zulu Natal. The remaining respondents were from the following geographic regions: Limpopo (6%) and 1.4% from Free State, Mpumalanga and North West. There were no Western Cape and Northern Cape respondents.

Figure 6 shows that 70% of the respondents were from urban areas, while 30% were from rural areas. This information demonstrates that the university enrols students from different regions of South Africa. The diversity of provinces and rural complement numbers is encouraging because empowering students from disadvantaged communities in STEM related careers is a significant step towards breaking the cycle of poverty, crime, and unemployment.

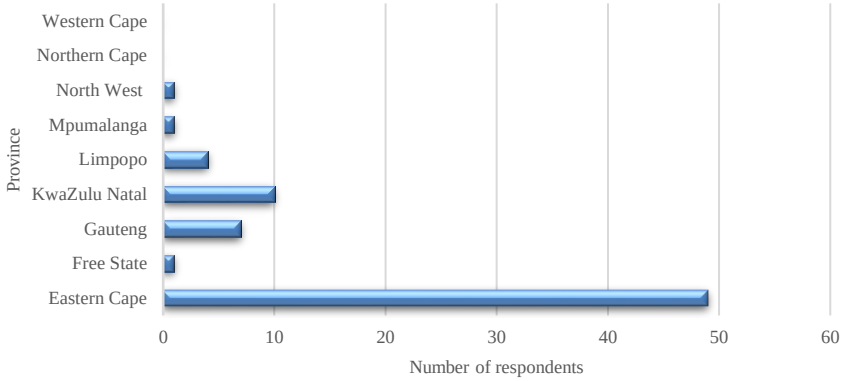

Figure 5: Distribution of respondents' home province

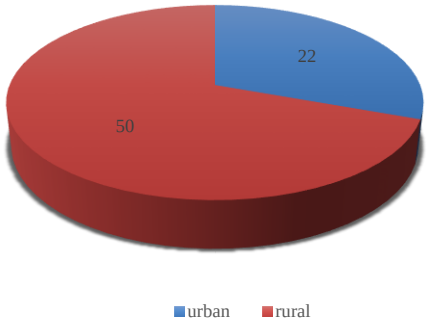

Figure 6: Distribution of respondents' home area

Amongst the respondents, 62 (or 86.3%) noted that Matric (National Senior Certificate) was their highest qualification. Matric, which is the qualification earned after completing the Grade 10 to Grade 12 period of further education and training, is equivalent to NQF level 4. Some of the respondents had alternative HEI credentials such as Bachelor, BTech and Diplomas. The distribution of the respondent's highest qualifications are illustrated in Figure 7.

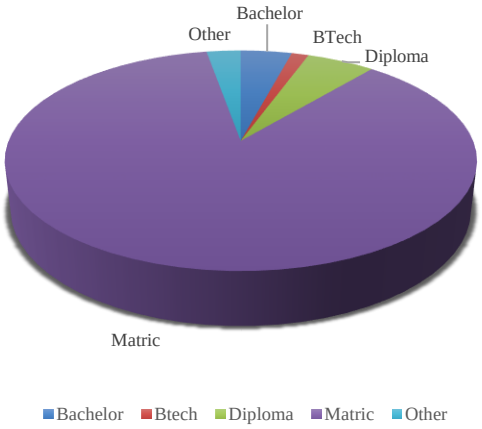

Figure 7: Distribution of highest qualifications

There are numerous concerns about the state of South Africa's educational system, particularly in Science and Mathematics subjects. All subjects have a standard pass rate of 30%, which sets a low pass bar. To make matters worse, more students than ever are choosing to take basic Mathematics Literacy exams, which disqualifies them from pursuing the courses that require Mathematics, at universities. We asked the respondents to state whether they had undertaken Mathematics Literacy or Mathematics examinations at Matric level.

Since none of the respondents were native English speakers (and the MOOCs chosen were taught in English), we also requested their level of English skill. Their English, Mathematics or Mathematics Literacy results are illustrated in Figure 8.

The respondents' English scores ranged from ‘moderate achievement’ (40–49%) to ‘outstanding achievement’ (80–100%). The respondents' average English grade was 60-69% which is considered a ‘significant accomplishment’ according to (Nethononda, 2022). We concluded that the respondents had a good understanding of the language used throughout the IBM course content.

16 respondents reported on their Mathematics Literacy results and 56 respondents reported on their Mathematics results. The fact that more respondents reported on Mathematics results, which is held to a higher level, was encouraging. The Mathematics results of the respondents ranged from ‘moderate achievement’ (30-39%) to ‘outstanding achievement’ (80-100%), while the Mathematics Literacy results ranged from ‘adequate achievement’ (50-59%) to ‘outstanding achievement’ (80-100%). This means that the Mathematic average of 40 – 49% was lower than that of Mathematics Literacy (60-69%).

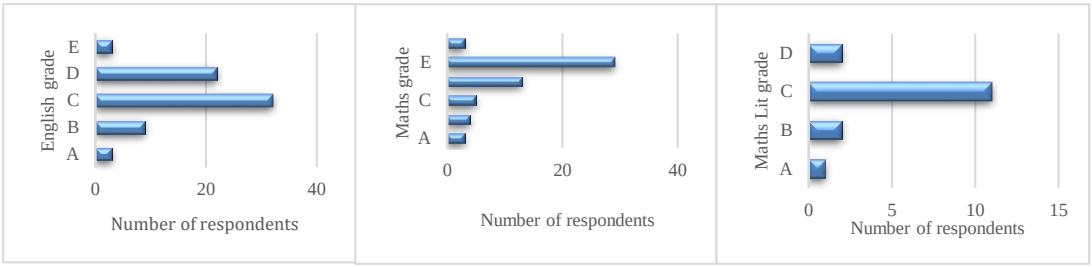

Figure 8: English, Mathematics and Mathematics Literacy matric results, respectively

The respondents were asked if they have even taken any MOOC/s prior to these face-to-face introduction sessions. 41 (or 57%) responded that they have taken a MOOC before, while 31 respondents (or 43%) had never done so. These results are illustrated in Figure 9. Although the participants did not specify which MOOC they had undertaken, they verbally acknowledged that none of the courses they had taken were IBM courses.

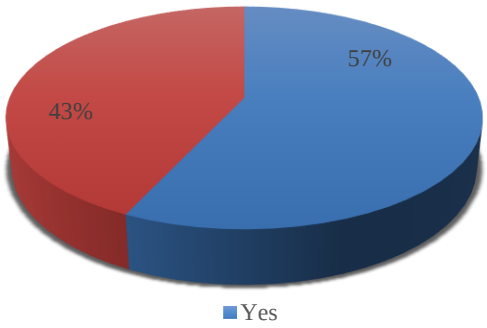

Figure 9: Undertaken online MOOC/s before

As previously stated, the face-to-face sessions were conducted to introduce two IBM courses to the participants. The first course, Data Science Foundations – Level 1, covered aspects such as ‘What is Data Science?’, ‘Why Data Science?’ and ‘Where to start?’. The second course, ‘Data Science Tools’ was a theoretic introduction to Data Science Tools such as Python and R. We observed that all

participants were excited to be part of the sessions because there are limited opportunities to study Data Science in the Eastern Cape province.

All 72 survey respondents completed at least one of the two MOOCs over the course of the workshop day. Specifically, only one of the 72 respondents completed one course, while the respondents completed at least two courses.

We also observed how many workshop participants completed at least one course during the workshop session. 141 (or 70%) of the 201 workshop participants submitted at least one verified IBM course completion certificate by the end of the workshop. This MOOC completion rate was substantially higher than the MOOC completion rate noted in (Nesterowicz et al., 2022).

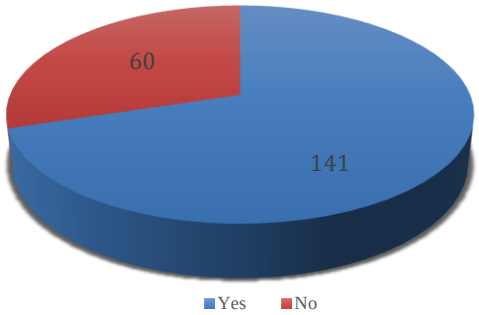

Figure 10: Completed at least one IBM MOOC by the end of the workshop session

We monitored the workshop participants for a month to ascertain whether they would complete additional IBM courses on their own. The total number of certificates received by the end of the month are illustrated in Figure 11. The highest number of IBM course completions was 12, while the lowest number was 0. 12 (or 22%) of the workshop participants were unable to finish any IBM course and 36 participant's finished just one course. This means that 24% of the participants did not fulfill the workshop requirements, while 76% completed the two IBM courses.

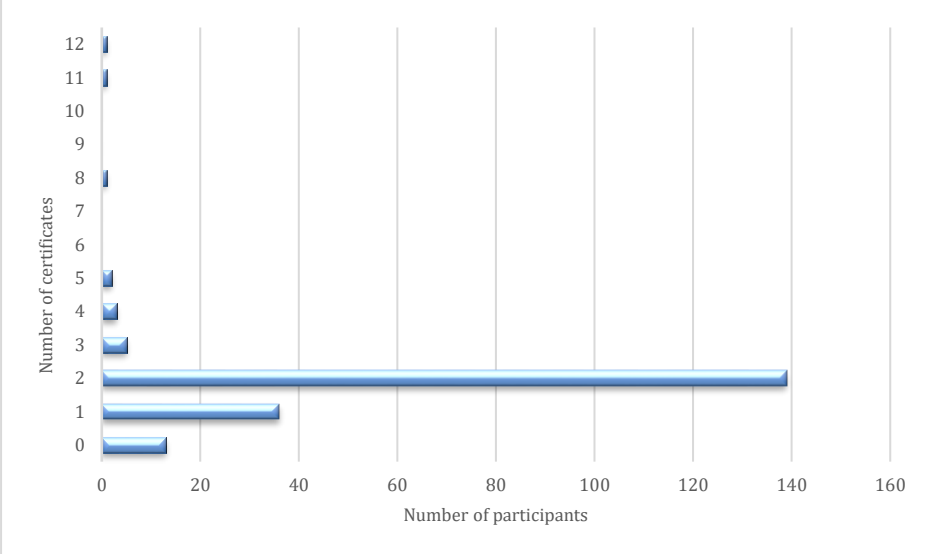

Figure 11: Certificates received between October and November 2022 from all workshop participants

3.4 Other survey feedback

The online survey respondents were asked what aspects of the workshop and course content they found valuable. The respondents appreciated exposure to Data Science concepts, basics and tools and the fact that the content was relevant to students who were not in the IT field. They felt confident that they could do more with the introductory skills which they had learnt.

In terms of the course layout, some respondents noted that it was useful that the MOOCs were in the form of video content which featured interviews of Data Scientist's personal experiences. The introduction of programming languages such as R and Python and the prospect of practically acquiring additional programming skills made them verbally commit to exploring more courses on the IBM platform. They also appreciated that there were quizzes and an examination to test their knowledge, although they would have preferred class discussions. The respondents felt that the course completion certificate was the most valuable component of the workshop as it would assist them in applying for jobs in the future.

The survey respondents provided feedback on what aspects of the workshop, course content and general learning can be improved. The respondents' main point was that coaching sessions, such as this workshop series, were critical in increasing the exposure of Data Science courses in the Eastern Cape. They felt that it was the responsibility of HEI to conduct these coaching sessions. This is in line with the social persuasion component of learning experiences, in the SCCT model. Through coaching, the participants were taught to believe that they could effectively obtain Data Science skills. They further suggested that the workshops and content should be more accessible and be introduced to learners in high school. If possible, these workshops should be incentivized. The face-to-face workshops, through coaching, were an opportunity for the physiological arousal of individuals' about MOOCs.

The respondents noted that more class discussions sessions should be allowed in order for MOOCs courses to be easier. The workshop session only allowed limited interaction between participants. We found that the concept of vicarious learning influenced the completion rate as the participants saw others completing the MOOCs successfully. However, a few respondents felt that learning as individuals was more fruitful, therefore the workshops should be offered online only. Performance success based on the completion of the first MOOC encouraged most participants to complete the second MOOC with more ease. While some respondents stated that more advanced MOOCs content and more practical work (instead of theory) should be introduced in the future; the majority of the respondents were content with the workshop and MOOCs courses chosen as an introduction.

The workshop participants were fortunate to have laptops, smartphones and a connection to the internet. However, we found that due to electricity challenges, the internet connection was unstable, and therefore affected the completion rate of the MOOCs as the online video content could not be viewed or downloaded. In a country like South Africa, where at present electricity is a challenge, it might be necessary to present offline MOOCs to improve the learning experience. There were additional suggestions that MOOC online content should also be printed as this would aid further research and studies and to promote continuity.

4 Conclusions

This paper reports on a MOOCs workshop series to introduce the concept of Data Science to HEI students in the Eastern Cape province of South Africa. We examined traits that are precursors of the SCCT's learning experience construct and provide us with a useful way to increase 4IR skill exposure and outcome expectancies in IT students enrolled at this particular university. The study showed that there is a strong interest in courses such as Data Science in HEI as shown by the number of students

who expressed their interest and the course completion rate of 76%. We noted the lack of Data Science programmes in the Eastern Cape province, therefore the findings demonstrated that the skill-base in this field can be increased if more effort is put by HEIs into establishing 4IR training programmes through MOOCs. Encouraging youth to develop skills is one of the first stages in creating the 4IR workforce and can aid in achieving SDGs. The concept of formally integrating MOOCs into HEI curriculum is a mechanism which can be used globally to increase the skills development of students.

Before a conclusion can be drawn regarding their learning experiences effect on student development, more research measurements and analyses are necessary on other constructs such as outcome expectations and self-efficacy in the SCCT model. Our analysis invites further exploration into other constructs in the SCCT model. Future studies should examine the MOOCs results and make use of hypothesis testing to determine how effective 4IR programmes are.

One of the study's limitations was the non-normality of the data in this sample; there were very few variances especially in the respondents demographic data. According to the features of the sample, students who took the Data Science introduction had high levels of interest before data collection, which probably affected the predictive value of these dimensions in subsequent analyses. This study suffers from problems related to voluntary responding, as certain groups are more likely to participate in a research, due to their personal characteristics. The participation of female respondents to male respondents was almost equal.

Students can improve 4IR skills and personal performance standards through ongoing physical exposure to online courses. People are likely to set objectives for maintaining or growing their involvement in an activity as soon as they become interested in it. People are more likely to develop a long-lasting interest in a task when they believe they are capable of executing it and when they anticipate the task will result in positive outcomes. The implementation of finely planned support of community and institutions for career choices in sustainability may be useful in giving rise to careers supporting the transition to more sustainable forms of economy.

References

Buthelezi, T., Alexander, D., & Seabi, J. (2009). Adolescents' perceived career challenges and needs in a disadvantaged context in South Africa from a social cognitive career theoretical perspective. *South African Journal of Higher Education*, *23*(3), Article 3. https://doi.org/10.4314/sajhe.v23i3.51033

Class Central. (n.d.). *IBM Courses*. Retrieved January 15, 2023, from https://www.classcentral.com/institution/ibm

Ethnic groups in South Africa. (2022). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Ethnic_groups_in_South_Africa&oldid=1122246833

Hupter, S., Mazumder, S., Bucaille, A., & Crossan, G. (2021). *Women in the tech industry: Gaining ground, but facing new headwinds*. Deloitte Insights. https://www2.deloitte.com/za/en/insights/industry/technology/technology-media-and-telecom-predictions/2022/statistics-show-women-in-technology-are-facing-new-headwinds.html

IBM. (n.d.-a). *Data Science 101*. Retrieved February 14, 2023, from https://cognitiveclass.ai/courses/data-science-101

IBM. (n.d.-b). *Data Science Tools*. Retrieved February 14, 2023, from https://cognitiveclass.ai/badges/data-science-tools

Medugorac, V., Šverko, I., & Babarovic, T. (2020). Careers in Sustainability: An Application of Social Cognitive Career Theory. *International Journal for Educational and Vocational Guidance*, *20*(3), 477–499. https://doi.org/10.1007/s10775-019-09413-3

Mooc.org. (2023). *Massive Open Online Courses*. https://www.mooc.org

Nesterowicz, K., Bayramova, U., Fereshtehnejad, S.-M., Scarlat, A., Ash, A., Augustyn, A. M., & Szádeczky, T. (2022). Gamification Increases Completion Rates in Massive Open Online Courses. *International Journal of Information and Communication Technology Education*, *18*(1). https://doi.org/10.4018/IJICTE.294447

Nethononda, M. (2022). *Matric symbols and level*. Matric College. https://www.matric.co.za/matric-symbols-and-level/

SA Department of Basic Education. (2021). *List of South African Universities*. https://www.education.gov.za/FurtherStudies/Universities/tabid/393/Default.aspx

Twinomurinzi, H., Mhlongo, S., Bwalya, K., Bokaba, T., & Mbeya, S. (2022). Multidisciplinarity in Data Science Curricula. *African Conference on Information Systems and Technology*. https://digitalcommons.kennesaw.edu/acist/2022/presentations/21

United Nations. (2023). *Sustainable Development Goals (SDG 4) Goal 4: Quality Education*. United Nations Western Europe. https://unric.org/en/sdg-4/