



**University of Fort Hare**  
*Together in Excellence*

**A MODEL FOR SECURE AND USABLE PASSPHRASES FOR  
MULTILINGUAL USERS**



**University of Fort Hare**  
*Together in Excellence*

**Pardon Blessings Maoneke**

2019

**A MODEL FOR SECURE AND USABLE PASSPHRASES FOR  
MULTILINGUAL USERS**

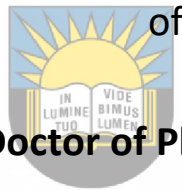
by

**Pardon Blessings Maoneke**

Student number: 201301198

**Thesis**

Submitted in fulfilment of the requirements for the degree



of  
**Doctor of Philosophy**

**University of Fort Hare**  
*Together in Excellence*

**Department of Information Systems**

at the

**University of Fort Hare**

**Promoter: Prof. Stephen Flowerday**

**Co-promoter: Dr Naomi Isabirye**

2019

## ABSTRACT

Research on more than 100 million passwords that have been leaked to the public domain has uncovered various security limitations associated with user-generated short passwords. Long passwords (passphrases) are considered an alternative solution that could provide a balance between security and usability. However, the literature shows a lack of consistency in the security and usability contributions of passphrases. For example, studies that investigated passphrase security focusing on structural dependencies at character level found passphrases to be secure. Inversely, other research findings suggest that passphrase security could be compromised by the use of predictable grammatical rules, popular words in a natural language and keyboard patterns. This is further exacerbated by research on passphrases that is focused on the Global North. This is a huge concern given that results from inter-cultural studies suggest that local languages do influence password structure and to some extent, password usability and security.

To address these gaps in the literature, this study used socio-technical theory which emphasised both the social and technical aspects of the phenomenon under study. Psychological studies show that the memory has limited capacity, something that threatens password usability; hence, the need to utilise information that is already known during password generation. Socio-cultural theory suggests that the information that is already known by users is contextually informed, hence socio-cultural theory was applied to understand the contextual factors that could be used to enhance passphrase security and usability. With reference to the Southern African context, this study argues that system designers should take advantage of a multilingual user group and encourage the generation of passphrases that are based on substrings from different languages. This study went on to promote the use of multilingual passphrases instead of emphasising multi-character class passwords.

This study was guided by design science research. Participants were invited to take part in a short password and multilingual passphrase generation and recall experiment that was made available using a web-based application. These passwords were generated by participants under pre-specified conditions. Quantitative and qualitative data was gathered. The study findings showed the use of both African and Indo-European languages in multilingual passphrases and short passwords. English oriented passwords and substrings dominated the multilingual passphrase and short password corpora. In addition, some of the short passwords and substrings in the multilingual passphrase corpora were found among the most common passwords of 2016, 2017 and 2018. Usability tests showed that multilingual passphrases are usable, even though they were not easy to create and recall when compared to short passwords. A high rate of password reuse during short password generation by participants might have worked in favour of short passwords. Nonetheless, participants appear to reflect better

usability with multilingual passphrases over time due to repeated use. Females struggled to recall short passwords and multilingual passphrases when compared to their male counterparts. Security tests using the Probabilistic Context-Free Grammar suggest that short passwords are weaker, with just more than 50% of the short passwords being guessed, while none of the multilingual passphrases were guessed. Further analysis showed that short passwords that were oriented towards an Indo-European language were more easily guessed than African language-oriented short passwords. As such, this study encourages orienting passwords towards African languages while the use of multilingual passphrases is expected to offer more security. The use of African languages and multilingual passphrases by a user group that is biased towards English-oriented passwords could enhance security by increasing the search space.

**Key words:** long passwords, passwords, passphrase, usability, security, password guessing.



University of Fort Hare  
*Together in Excellence*

## ACKNOWLEDGEMENTS

My sincere and deepest gratitude goes to my promoter, Professor Stephen Flowerday, for his intellectual, financial and moral support. You were always a source of profound knowledge and I will forever be grateful for your guidance. Secondly, I would like to thank my co-promoter, Dr Naomi Isabirye, for her assistance during data collection, with financial support and with critical thinking.

A special thank you to the Carnegie Mellon University Password Research Group which used their Password Guessability algorithms to test the strength of passphrases and short passwords for this study.

Another special thank you to Munyaradzi Katuruza for helping with data collection and thanks also to all those who took part in the experiment.

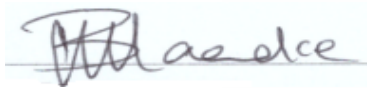


University of Fort Hare  
*Together in Excellence*

## DECLARATION

I, Pardon Blessings Maoneke, hereby declare that:

- I am fully aware of the University of Fort Hare's policy on plagiarism and I have taken every precaution to comply with the regulations.
- I am fully aware of the University of Fort Hare's policy on research ethics and I have taken every precaution to comply with the regulations. I have obtained an ethical clearance certificate from the University of Fort Hare's Research Ethics Committee and my reference number is the following: FLO081SMAO01.
- This thesis has not previously been submitted in full or partial fulfilment of the requirements for an equivalent or higher qualification.



**Pardon Blessings Maoneke**



**University of Fort Hare**  
*Together in Excellence*

09/09/2019  
(date)

## PUBLICATIONS

### Conference:

- 1) P. B. Maoneke & S. Flowerday. 2019. Password Policies Adopted by South African Organizations: Influential Factors and Weaknesses. In: Venter H., Looek M., Coetzee M., Eloff M., Eloff J. (eds) Information Security. ISSA 2018. Communications in Computer and Information Science, vol 973. Springer, Cham
- 2) P.B. Maoneke, S. Flowerday & N. Isabirye. 2018. The Influence of Native Language on Password Composition and Security: A Socioculture Theoretical View. *Paper Accepted for Presentation and Publication at the 33rd Proceedings of IFIP TC-11 SEC, 18-21 September 2018 Poznań, Poland.*
- 3) P.B. Maoneke, S. Flowerday & N. Isabirye. 2017. A Model for Usable and Secure Passphrases by Multilingual User Groups. *Proceedings of the Annual Security Conference/ISOneWorld Conference Information Institute Conference, Las Vegas, USA.*



### Journals:

University of Fort Hare  
Together in Excellence

- 1) P.B. Maoneke, S. Flowerday & N. Isabirye. 2020. Evaluating the Strength of A Multilingual Passphrase Policy (Computers and Security).
- 2) P.B. Maoneke, S. Flowerday & N. Isabirye. 2019. The Usability of Multilingual Passphrases. Computers in Human Behavior (Under review).

## DEDICATION

In loving memory of my grandmother: the late Mrs Eveldah Shumba

You will always be missed and loved!



University of Fort Hare  
*Together in Excellence*



# Table of Contents

CHAPTER 1: THE INTRODUCTION .....	1
1.0 Background.....	1
1.1 Statement of the problem.....	3
1.2 The research question .....	4
1.2.1 Research sub-questions.....	5
1.3 Research objectives.....	8
1.4 Theoretical foundation.....	8
1.5 Preliminary literature review .....	9
1.5.1 The socio sub-stream.....	9
1.5.2 The technical sub-stream .....	10
1.6 The significance of the study.....	11
1.7 Research methodology.....	11
1.7.1 Research paradigm and philosophical view .....	12
1.7.2 Research approach .....	12
1.7.3 Research method.....	12
1.7.3.1 Experiment design and data collection.....	12
1.7.3.2 Data analysis .....	13
1.8 Delimitation of the study.....	13
1.9 Ethical considerations.....	14
1.10 Study contribution.....	14
1.11 Outline of the study.....	17
CHAPTER 2: THE RESEARCH DESIGN AND METHODOLOGY .....	19
2.0 Introduction.....	19
2.1 Philosophical assumptions and research paradigm .....	19
2.1.1 Overview of research philosophies .....	21
2.1.2 Philosophical stance of this study.....	24
2.1.3 Design science research.....	25
2.1.4 The reasoning approach of the study.....	26
2.2 Design science research guidelines and process models.....	27
2.2.1 The application of Peffers et al.'s (2004) process model to this study .....	29
2.3 Research methods .....	34
2.4 Data collection.....	35
2.4.1 Secondary data collection.....	35
2.4.2 The experiment design and primary data collection.....	36
2.4.2.1 The questionnaire .....	38
2.4.2.2 The keystrokes .....	39
2.4.3 Target population and sampling method .....	40
2.4.4 The experiment and data collection administration .....	40
2.5 Data analysis.....	41

2.5.1 Password and passphrases characteristics .....	41
2.5.2 Short password and multilingual passphrase security .....	42
2.5.2.1. Hybrid password cracking algorithm .....	42
2.5.3 Short password and multilingual passphrase usability .....	43
2.6 Validity and reliability .....	43
2.7 Ethical considerations .....	45
CHAPTER 3: THEORETICAL FOUNDATION .....	47
3.0 Introduction .....	47
3.1 Socio-technical theory .....	47
3.2 Human memory functionality .....	50
3.2.1 Memory decay and password usability .....	52
3.2.2 Short-term memory capacity and password usability .....	53
3.3 Socio-cultural theory .....	54
3.3.1 The generic law of development. ....	56
3.3.2 Mediation .....	57
3.3.3 Generic domains .....	58
3.4 An overview of the social context and language development .....	59
3.4.1 Africa's language landscape .....	59
3.4.2 The practice of code switching in text messages .....	61
3.4.2.1 Functional code-switching .....	61
3.4.2.2 Principled code-switching .....	63
3.4.2.3 Meaningful code-switching .....	63
3.5 Chapter summary .....	64
CHAPTER 4: PASSWORD THREATS AND POLICIES IN USE .....	65
4.0 Introduction .....	65
4.1 Categories of authentication mechanisms .....	65
4.2 Password threats and attacks .....	66
4.2.1 Online password attack .....	67
4.2.2 Offline password attack .....	68
4.3 Password strength measurement .....	74
4.3.1 Entropy .....	75
4.3.2 Guess number .....	75
4.3.3 An overview of password strength measures .....	76
4.4 Factors of password strength and usability .....	76
4.4.1 Frameworks for password guidelines and best practices .....	77
4.4.2 Password policies .....	79
4.4.3 An overview of users and password management behaviours .....	88
4.4.3.1 Password generation .....	88
4.4.3.2 Keeping track of passwords, authenticating and changing passwords. ....	89
4.5 Chapter summary .....	91
CHAPTER 5: A MODEL OF USABLE AND SECURE PASSPHRASES .....	93

5.0 Introduction .....	93
5.1 Definition of a model .....	93
5.2 The rationale behind the proposed research model .....	94
5.3 Theoretical foundation .....	95
5.3.1 Multilingual passphrase security .....	97
5.3.1.1 Factors for passphrase strength .....	98
5.3.2 Password usability .....	102
5.3.2.1. Factors of passphrase usability .....	102
5.4 The proposed model .....	106
5.5 Chapter summary .....	108
CHAPTER 6: RESEARCH FINDINGS .....	110
6.0 Introduction .....	110
6.1 Sample profile .....	110
6.2 The data collection instrument .....	112
6.3 Social context overview .....	118
6.3.1 The generic law of development .....	118
6.3.2 Mediating symbolic tools and password characteristics .....	119
6.3.2.1 Short password characteristics .....	119
6.3.2.2 Multilingual passphrase characteristics .....	124
6.4 Instrument validity and reliability .....	130
6.4.1 Instrument validity .....	131
6.4.2 Instrument reliability .....	133
6.4.2.1 Instrument reliability – short passwords .....	133
6.4.2.2 Instrument reliability – multilingual passphrases .....	134
6.5 Password usability .....	134
6.5.1 Short password requirements usability .....	135
6.5.1.1 Short password generation usability .....	135
6.5.1.2 Short password recall usability .....	136
6.5.1.3 Short password usability differences and correlation analysis .....	139
6.5.2 Multilingual passphrase usability .....	140
6.5.2.1 Multilingual passphrase generation usability .....	140
6.5.2.2 Multilingual passphrase recall usability .....	141
6.5.2.3 Multilingual passphrase usability differences and correlation analysis .....	144
6.5.3 Comparison of short passwords and multilingual passphrases usability .....	145
6.6 Password strength .....	147
6.6.1 Short password strength .....	148
6.6.2 Multilingual passphrase strength .....	150
6.7 Chapter summary .....	151
CHAPTER 7: RESEARCH FINDINGS AND DISCUSSION .....	153
7.0 Introduction .....	153
7.1 Socio-cultural theory and password characteristics .....	153

7.1.1 A discussion of the findings on the generic law of development.....	154
7.1.2 A discussion of the findings on mediation.....	154
7.1.3 A discussion of findings on the generic domains .....	158
7.2 Password recall strategies .....	159
7.3 Factors of multilingual passphrase security and usability.....	159
7.3.1 Multilingual passphrase security .....	160
7.3.2 Multilingual passphrase usability .....	162
7.3.2.1 Multilingual passphrase effectiveness.....	162
7.3.2.2 Multilingual passphrase efficiency.....	165
7.3.2.3 Multilingual passphrase user satisfaction.....	167
7.4 Multilingual passphrase length and security.....	169
7.5 Chapter summary .....	169
CHAPTER 8: RESEARCH CONTRIBUTIONS AND RECOMMENDATIONS.....	171
8.0 Introduction .....	171
8.1 A synopsis of the study.....	171
8.1.1 The theoretical foundation .....	173
8.1.1.1 The social subsystem.....	174
8.1.1.2 The technical subsystem .....	175
8.1.1.3 A proposed model for secure and usable multilingual passphrases .....	176
8.2 The evaluation framework .....	177
8.2.1 The evaluation framework for this study .....	179
8.3 Evaluating a model for secure and usable multilingual passphrases.....	180
8.3.1 Evaluating multilingual passphrase security.....	181
8.3.2 Multilingual passphrase usability .....	183
8.3.2.1 Multilingual passphrase generation usability .....	184
8.3.2.2 Multilingual passphrase recall usability .....	186
8.4 Chapter summary .....	190
CHAPTER 9: CONCLUSION .....	191
9.0 Introduction .....	191
9.1 An overview of this study .....	191
9.2 Research questions and findings.....	194
9.3 Research contributions.....	199
9.4 Limitations of the study.....	201
9.5 Direction for future research.....	202
9.6 Conclusion .....	203
REFERENCES.....	205
APPENDIX A: THE EXPERIMENT DESIGN AND OVERVIEW .....	222
APPENDIX B: THE QUESTIONNAIRE .....	230
APPENDIX C: THE INFORMED CONSENT FORM.....	235
APPENDIX D: ETHICAL CLEARANCE CERTIFICATE .....	236

## LIST OF FIGURES

Figure 1. Security contributions: the Multilingual Passphrase Security Model .....	15
Figure 2. Passphrase usability contributions: a model for usable multilingual passphrases .....	16
Figure 3. Design science research methodology .....	28
Figure 4. Design science knowledge contribution framework.....	30
Figure 5. The research process of this study informed by Peffers et al. (2008) .....	33
Figure 6. Password generation experimental activities and data gathering.....	37
Figure 7. Capturing timestamps from keystrokes .....	39
Figure 8. The structure of the memory system.....	50
Figure 9. The perceived joint influence of social and technical subsystems on passphrases .....	96
Figure 10. A passphrase security model for a multilingual user group.....	102
Figure 11. A proposed model of usable and secure passphrases for a multilingual user group .....	108
Figure 12. Participants that took part in the short password and multilingual passphrase experiments.....	111
Figure 13. The distribution of gender across password experiments activities.....	111
Figure 14. Age group distribution across the sample.....	112
Figure 15. Activities of the study experiment and data collection .....	113
Figure 16. The short password generation platform. ....	120
Figure 17. Short password generation strategy.....	120
Figure 18. Password length of short passwords.....	123
Figure 19. Multilingual passphrase generation platform.....	125
Figure 20. Multilingual passphrase generation strategy.....	126
Figure 21. Common characters used to generate multilingual passphrases .....	128
Figure 22. Multilingual passphrase length according to the number of characters ...	129
Figure 23. Short password generation effectiveness and user satisfaction .....	136
Figure 24. Short password memorability strategy .....	136
Figure 25. Short password recall effectiveness and user satisfaction .....	137
Figure 26. The percentage of short password attributes (of all short password recall failure) that were often forgotten by users .....	138
Figure 27. Multilingual passphrase generation effectiveness and user satisfaction ..	141
Figure 28. Multilingual passphrase recalling strategy .....	142
Figure 29. Multilingual passphrase recall effectiveness and user satisfaction .....	143
Figure 30. The percentage of passphrase attributes (of all the passphrase recall failure) that were often forgotten by participants.....	144
Figure 31. Mean responses of theoretical variables on short passwords vs. passphrase creation, recall strategy and recalling .....	146
Figure 32. Short password guessing results using PCFG. ....	148

Figure 33. A comparison of guessing resistance between African language and English orientated short passwords .....	149
Figure 34. Short password length and resistance to password guessing .....	150
Figure 35. FEDS dimensions.....	178
Figure 36. Multilingual Passphrase Security Model .....	183
Figure 37. A model for Usable Multilingual Passphrases .....	189



University of Fort Hare  
*Together in Excellence*

## LIST OF TABLES

Table 1. Philosophical assumptions of the four paradigms aligned to the three philosophies discussed. ....	23
Table 2. An overview of password guessing algorithms .....	73
Table 3. Factors of passphrase usability considered for this study.....	104
Table 4. Questions for gathering data on PCE.....	114
Table 5. Password generation strategies that could reflect failure to effectively generate a short password or multilingual passphrase .....	115
Table 6. Questions for gathering data on PRE.....	115
Table 7. Password recall strategies that reflect failure to effectively recall a short password or multilingual passphrase .....	116
Table 8. Questions for gathering data on password creation user satisfaction (PCUS) .....	117
Table 9. Questions for gathering data on password recall user satisfaction (PRUS) ..	118
Table 10. Observed semantics used in user-generated short passwords.....	121
Table 11. Common password structures observed in the short password corpora ...	121
Table 12. Short password generation techniques assumed by participants .....	123
Table 13. Common substrings that were found among the most common passwords of 2016, 2017 and 2018.....	124
Table 14. Characteristics of observed multilingual passphrase structures.....	129
Table 15. Common substrings (in bold) that were found among the most common passwords of 2016, 2017 and 2018.....	130
Table 16. Rotated component matrix – password generation .....	132
Table 17. Rotated component matrix – password recall .....	133
Table 18. Reliability analysis – short passwords.....	134
Table 19. Reliability analysis – multilingual passphrases .....	134
Table 20. Multilingual passphrase vs. short password usability (time indicated in seconds).....	146
Table 21. T-tests for mean password length of usability on effectiveness and user satisfaction .....	147
Table 22. Data collection techniques and research sub-questions of this study.....	194

## LIST OF ABBREVIATIONS

DD: Key Down-Down

EFA: Exploratory factor analysis

FEDS: Framework for Evaluation in Design Science

FIPS: Federal Information Processing Standard

H: Hold

ISACA: Information Systems Audit and Control Association

ISO: International Organisation for Standardisation

LDS: L: Alphabetic letter, D: Digits and S: Symbol

LUDS: Lower-case, upper-case, digits and symbols

NIST: National Institute of Standards and Technology

PCA: Principal component Analysis

PCE: Password creation effectiveness

PCFG: Probabilistic Context-Free Grammar

PCUS: Password creation user satisfaction

PRE: Password recall effectiveness

PRUS: Password recall user satisfaction

PSM: Password strength meter

SP: Special publication

SPSS: Statistical Package for the Social Sciences

UD: Up-down

US: United States

USE: Usefulness, satisfaction and ease of use



University of Fort Hare  
*Together in Excellence*



## CHAPTER 1: THE INTRODUCTION

### 1.0 Background

The use of passwords as authentication mechanisms has remained popular since their inception (Denning, 1992, in Andersson & Saedén, 2013; Woods & Siponen, 2019). The ever-increasing need to protect information assets has prompted the need for further research on passwords, as user-designed passwords are considered weak (AlSabah, Oligeri, & Riley, 2018; Guo, Zhang, & Guo, 2019; Harris & Maymí, 2019; Inglesant & Sasse, 2010; Wang, Cheng, Gu, & Wang, 2015). Institutions are using various measures and policies to help users generate stronger passwords. These include rule-based password policies that encourage users to use different character sets, increase the password length and use password strength meters or adopt system-generated passwords (Houshmand & Aggarwal, 2012; Keith, Shao, & Steinbart, 2009). Such policies are, according to Keith, Shao, and Steinbart (2007), rooted in instruments like the United States (US) Federal Information Processing Standard (FIPS) of 1985. This standard is seen as one of the earliest sets of guidelines for creating strong passwords (Keith et al., 2007). In addition, the US National Institute of Standards and Technology's (NIST) Electronic Authentication Guideline is widely regarded as having served as the basis for most rule-based password policies (Grassi, Garcia, & Fenton, 2017; Houshmand & Aggarwal, 2012; Inglesant & Sasse, 2010; Ur et al., 2012; Wang, Cheng, et al., 2015) together with best practices and tools proposed by leading institutions such as the Information Systems Audit and Control Association (ISACA).

The literature suggests that system-generated and rule-based passwords are not always user-friendly and can easily be exploited by different user behaviours (AlSabah et al., 2018; Grassi et al., 2017; Guo et al., 2019; Komanduri et al., 2011; Renaud, Otondo, & Warkentin, 2019; Wang, Cheng, et al., 2015; Weir, Aggarwal, Collins, & Stern, 2010). For instance, users find rule-based and system-generated passwords difficult to memorise and often compensate for the shortcoming of memorising complex passwords by writing them down, storing them in insecure places and, in some cases, repeatedly using the same password across multiple domains because they are difficult to create and recall (Choong, Theofanos, & Lui, 2014; Inglesant & Sasse, 2010; Keith et

al., 2009, 2007; Melicher et al., 2016; Pilar, Jaeger, Gomes, & Stein, 2012; Renaud et al., 2019; Shay et al., 2016; Ur et al., 2012; Woods & Siponen, 2018). Users spend 1.5 to 2.25 business days each year generating new passwords (Choong et al., 2014; Shay et al., 2014). Houshmand and Aggarwal (2012) go on to state that hackers take advantage of password reuse and simply target password databases of weak domains whose exposed passwords are displayed and used to train probabilistic-based password-cracking algorithms. Knowledge from password databases in the public domain is then used to guess targeted passwords using algorithms such as the Probabilistic Context-Free Grammar (PCFG) and Markov-chain-based attacks (Houshmand & Aggarwal, 2012; Wang, Cheng, et al., 2015).

While researchers have been focused on balancing password strength and usability, Choong et al. (2014) found that users prioritise password memorability, with little or no interest in password strength. The “interrelationship between text password security and memory theory has long been recognised” (Zhang, Luo, Akkaladevi, & Ziegelmayer, 2009, p. 1). Hence, there is a need to understand human memory if password usability is to be improved (Woods & Siponen, 2018). Human memory can only recall three to four items when presented with a set of totally random items (Cowan, 2000). As such, if users are to generate memorable passwords, password policies should seek to exploit the lexical and logical semantics in the long-term memory (Cowan, 2000; Gruszka & Orzechowski, 2016). This has motivated the idea of using long phrases (passphrases) for authentication purposes (Andersson & Saedén, 2013; Grassi et al., 2017; ISACA, 2015; Keith et al., 2009; Melicher et al., 2016; Shay et al., 2016). It is believed that a user-defined passphrase can easily relate to a user’s long-term memory for memorability purposes, while their length promotes randomness. ISACA (2015) states that a passphrase is generally accepted as a more secure password. Study findings by Shay et al. (2014), Melicher et al. (2016) and Shay et al. (2016) suggest that passphrases offer better security even when facing resourceful PCFG.

The recognition of passphrases as an alternative measure for attaining password security and usability has not gone without caution. Keith et al. (2007) found that passphrases are prone to typographic errors. As such, Keith et al. (2009) recommend

the use of passphrases made up of sentences that follow grammatical rules or based on known words to address typographic errors. Contrary to the propositions by Keith et al. (2009), Rao, Jha, and Kini (2013) found that the use of predictable grammatical structures in passphrases can pose security limitations. In addition, Shay et al. (2016) and Bonneau and Shutova (2012) found that users base their passphrases on a few popular words in a language something that compromise security. It is possible that these findings can be explained by the fact that the resultant passphrases were generated using monolingual. Wang, Cheng et al. (2015) gave a plausible explanation on how language influence passphrases as they noted that, “even though generated and used in vastly diversified web services, passwords among the same language group have quite similar letter distributions” (p. 5). In addition, the success of the Markov Chain password guessing algorithm is relying on its ability to imitate character distribution in a language. Accordingly, this study proposes the use multilingual passphrases for African computer users who are characterised by a multilingual user group. Multilingual passphrases have the potential to increase the size of the passphrase search space thereby enhancing security (Rao et al., 2013).



### **1.1 Statement of the problem**

Research findings from more than 100 million passwords that were leaked to the public domain shows that password generation policies anchored on Lower-case, Upper-case, Digits and Symbols (LUDS) fail to encourage users to generate usable and strong passwords. The literature suggests that the use of passphrases is one of the alternative solutions for enhancing the security and usability of passwords (Blanchard, Malaingre, & Selker, 2018; Bonneau & Shutova, 2012; Braunstein, 2015; Grassi et al., 2017; Juang & Greenstein, 2018; Shay et al., 2016). However, studies based on different evaluation methods present mixed findings on the security and usability contributions of passphrases (Kelley et al., 2012; Shay et al., 2014, 2016; Bonneau & Shutova, 2012; Rao et al., 2013; Veras et al., 2014). For instance, passphrases are more likely to expose users to typographical errors when compared to short passwords (Keith et al., 2007). In addition, users are more likely to base their passphrases on a few popular words and grammatical rules in a language something that compromise security (Bonneau & Shutova, 2012; Rao et al., 2013; Shay et al., 2016). This is further exacerbated by

research that has been focused on passphrases of the Global North where English is often the first language. This is a huge concern given that inter-culture password studies have shown that the password structure may differ according to local languages and culture something that reflects on security (AlSabah et al., 2018; Wang, Cheng, et al., 2015). Given these findings in the literature, the problem statement of this study is as follows:

*Generating and using passphrases in a manner that meets security and usability requirements remains a challenge.*

## **1.2 The research question**

To address the problem statement highlighted in Section 1.1, the following research question was formulated:

*How can local languages be exploited in order to improve the security and usability of passphrases?*



The literature suggests that users' local languages influence the characteristics of the passwords they generate (AlSabah et al., 2018; Bonneau & Xu, 2012; Narayanan & Shmatikov, 2005; Wang, Cheng, et al., 2015). For example, the Chinese computer users are more likely to base their passwords on digits (more than 50%) when compared to the English computer users (15.77%) (Li et al., 2014; Wang, Cheng, et al., 2015). Furthermore, some of the Chinese passwords are likely to be based on Pinyin names and the selection of some of the digits in passwords portray the pronunciation of specific phrases in Mandarin Chinese (Li, Han & Xu, 2014; Wang, Cheng, et al., 2015; Yang, Hung, & Lin, 2013). As a result, a significant number of Chinese passwords are concentrated around a few keyboard character-keys something that makes them prone to online password guessing when compared to passwords generated by the English computer users (Wang, Cheng, et al., 2015). Nevertheless, a significant number of Chinese passwords are resistant to resourceful offline password guessing attacks when compared to English passwords (Wang, Cheng, et al., 2015). In addition, AlSabah et al. (2018) notes that passwords that were generated by different nationalities (Arabs,

Indians, Pakistanis, Philippines and English speaking nationalities) following the same password policy could be differentiated according to culture. Hence, understanding African language landscape can play a pivotal role in influencing users to generate secure and usable passwords. In contrast to the Chinese and some Arabic contexts reported in the literature (AlSabah et al., 2018; Wang, Cheng, et al., 2015); the majority of African languages, are based on the Latin alphabet. This, coupled with a multilingual user group points to the fact that computer users from the African context might present unique password generation behaviours. In terms of language, Africa portrays a unique context where English is the dominant language of instruction and first written language in literacy, while African languages are the first spoken languages (Deumert & Lexander, 2013; Dyers & Davids, 2015; Lexander, 2011; Ndlovu, 2016).

### 1.2.1 Research sub-questions

The main research question was broken down into the following four sub-questions:

- *What are the different password policies in use?*

This research sub-question explores different password policies in use. Chapter 4 of this study found that password policies could be split into password composition policy, system-generated passwords, password strength meters, and system and user-generated password policies. This study went on to investigate a password composition policy focusing on passphrases. The study thus promotes the use of multilingual passphrases in order to attain security and usability.

- *What are the language characteristics that could be considered to enhance the security of user-generated passphrases?*

In order to address this sub-research question, this study explored password threats. Understanding password threats was critical as it created an anticipation of language characteristics that could be considered when enhancing passphrase security. Chapter 4 of this study identified online and offline password threats as the security threat model for this study. As such, a strong passphrase and short password is one that can resist online and offline password threat. Users have been shown to base their passphrases on certain linguistic attributes within a language (AlSabah et al., 2018;

Melicher et al., 2016; Veras, Collins, & Thorpe, 2014). Rao et al. (2013) used Parts-of-Speech (POS) tagging to model grammatical structures in a passphrase and successfully demonstrated how grammatical rules in a language could be exploited by password hackers (online and offline password threats). Furthermore, some users have a tendency to base their passphrases on a few popular words in a language (AlSabah et al., 2018; Komanduri, 2016; Shay et al., 2016). Thus, this study explored the use of multilingual passphrases in addressing passphrase strength. While the NIST adopted Shannon's entropy and used it to motivate the generation of passwords based on different character classes, this study adopted a similar approach, but promoted the generation of passphrases based on multilingual instead of multi-character classes. Each of the study participants was asked to generate a passphrase based on at least two substrings from different languages. The same group of participants was also asked to generate a short password of at least eight characters long, based on different character classes. A PCFG password guessing algorithm was used to guess user-generated passphrases and short passwords in an offline password attack. Results from the password guessing algorithm showed that short passwords were weaker than passphrases based on substrings from different languages. Short passwords oriented towards the English language were found to be weaker when compared to short passwords oriented towards African languages.

- *What are the factors affecting the usability of passphrases?*

This sub-question evaluated factors that contribute to the usability of multilingual user-generated passphrases. The sub-question attempted to explain how knowledge that was gained and stored in the near permanent memory could, unconsciously, be considered when generating usable passphrases. The focus was on the use of multilingual passphrases. The International Organisation for Standardisation (ISO) 9241-11 standard was used to evaluate the usability of multilingual passphrases in this study. The ISO 9241-11 standard evaluates usability in terms of effectiveness, efficiency and user satisfaction. Chapter 5, Section 5.3.2.1, explains the meaning of the usability factors of effectiveness, efficiency and user satisfaction within the context of passphrases. Usability was evaluated during passphrase generation and recall. The results of this study show that passphrases are usable, though not as usable as short

passwords, which was unexpected. A high rate of short password reuse could explain this finding. Furthermore, effectiveness, efficiency and user satisfaction were found important during passphrase generation, while only passphrase effectiveness and efficiency were found important during passphrase recall. User satisfaction was not influential during passphrase recall.

- *What are the password characteristics of a multilingual user group?*

To address this research question, this study investigated password characteristics that include the use of semantics, password structures and the use of popular passwords by participants during password generation. Users adopt these password characteristics in order to enhance password usability. However, these password characteristics may have an impact on password security. Li, Wang, and Sun (2016) showed how password attackers could exploit the use of semantic information to enhance password guessing. Similarly, Weir, Aggarwal, De Medeiros and Glodek (2009) showed that password structures could be exploited during password guessing with damning effect. Narayanan and Shmatikov (2005) exploited the orientation of passwords in a user's language to generate an effective password guessing algorithm. Furthermore, password attackers often adopt an optimal approach to password guessing where the most popular passwords are targeted first during password guessing (Shay et al., 2016). Findings from this study suggest that users adapted semantic information such as names and existing passwords. English language-oriented words and phrases dominated the short password corpora, which also included short passwords with words and names oriented to African languages. The use of different languages in short password generation reflected the multilingual characteristics of the participants. Furthermore, there was wide use of alphabetic letters, digits and symbols. Three symbols that were widely used included the @, # and \$ symbols. The majority of participants preferred to combine a word or phrase with a number and symbol, in that order, during short password generation. In addition, 3% of the short passwords were found among the popular passwords of 2016, 2017 and 2018 according to SplashData (2016, 2017, 2018).

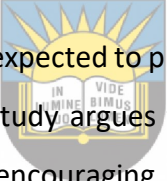


Passphrase characteristics were found to be comparable to short password characteristics to some extent. The majority of participants adopted names and English language-oriented words which dominated the majority of substrings used in passphrases. User-generated passphrases mainly constituted two or more substrings. A significant proportion of passphrases had a substring that resembled a short password in the case of password reuse. In addition, 4% of the passphrases had substrings that were found among the popular passwords of 2016, 2017 and 2018, according to SplashData (2016, 2017, 2018).

### **1.3 Research objectives**

The objective of this study is to produce:

*A model for secure and usable multilingual passphrases for a multilingual user group.*



The proposed model is expected to play a pivotal role in informing requirements for passphrase policies. This study argues that understanding the social context can help derive constructs for encouraging the generation of secure and usable passphrases. Socio-cultural theory was used to understand the research context of this study and the informing factors that could be considered when enhancing passphrase security and usability. Rao et al. (2013) found that the use of grammatical structures may compromise the security of passphrases. In addition, the dominance of a few selected popular words in a language in user-generated passphrases could potentially weaken passphrases (Komanduri, 2016; Shay et al., 2016). The proposed model for this study moves away from orienting passphrases towards a single language by promoting the use of multilingualism in user-generated passphrases. By so doing, this study's model is expected to enhance passphrase strength by increasing the passphrase search space. Increasing the search space was attained by orienting towards passphrases in multiple languages, instead of a single language.

### **1.4 Theoretical foundation**

This study was grounded in socio-technical theory, which gives equal importance to the socio- and technical subsystems when addressing a problem. By so



doing, this study abandoned the traditional way of solving an Information Systems problem by focusing on technical aspects (technical subsystem) in the hope that the context would adapt to the solution (Durkin, Mulholland, & McCartan, 2015; Shin, 2014). To understand the socio sub-stream, this study starts with an overview of the functionality of the human memory using Atkinson and Shiffrin's (1968) stages of memory theory, a popular theory in psychology and studies on password memorability. Given the influence of one's language in password generation (AlSabah et al., 2018; Bonneau & Xu, 2012; Narayanan & Shmatikov, 2005; Wang, Cheng, et al., 2015), this study used socio-cultural theory to understand the language terrain (socio subsystem) of the researched context. Section 1.5 gives a brief overview of the preliminary literature review indicating the subjects that were discussed within the socio- and technical subsystems.

### **1.5 Preliminary literature review**

This preliminary literature review focuses on the socio- and technical sub-streams.



#### **1.5.1 The socio sub-stream**

This sub-stream focused on understanding the functionality of a human memory and establishing the social context of the study. This is critical when exploring ways of improving passphrase usability (AlSabah et al., 2018; Woods & Siponen, 2019). Atkinson and Shiffrin (1968) argued that the memory can be divided into the sensory, short-term and long-term memory. The short-term memory which plays a critical role in password generation is limited in capacity (Atkinson & Shiffrin, 1968; Miller, 1956; Woods & Siponen, 2019). To increase the capacity of the short-term memory, users can make use of information in the long-term memory when performing cognitive tasks (Keith et al., 2007; Miller, 1956). In addition, socio-cultural theory proposes that cognition development is inspired by developments within the context of the subject or user. This suggests that, if a user is to generate a password, the resultant password would reflect contextual characteristics of the user. For example, AlSabah et al. (2018) show the influence of culture in password generation. According to socio-cultural theory, the influence of context in cognition development and its subsequent influence on

passwords could be explained by using three principles, namely, the generic law of development, mediation and the generic domains. Section 3.3 of this study shows how contextual factors affect password generation as purported by socio-cultural theory.

### **1.5.2 The technical sub-stream**

The technical sub-stream of this study explained different authentication mechanisms, password threats, available policies for secure and usable passwords as well as constructs for ascertaining password security and usability. In particular, this study evaluated the authentication mechanisms in use. This was done with the aim of positioning the use of passwords when compared to other authentication mechanisms. Section 4.1 in Chapter 4 gives an overview of different password authentication mechanisms and justifies the use of passwords. Once the use of passwords was justified, this study went on to discuss common password threats that could compromise security. Section 4.2 in Chapter 4 identifies the security threat model for this study. It is indicated that this study focuses on securing passwords against online and offline password security threats. The need to address online and offline password threats was elevated by the leakage of more than 100 million passwords into the public domain. In addition, the password policies in use are also discussed in Chapter 4. Due to the security and usability limitations of short passwords (Keith et al., 2007; Melicher et al., 2016; Shay et al., 2016), this study recommends the use of passphrases.

Chapter 5 presents the security and usability constructs for passphrases in this study. Factors of password security proposed by the study include the use of juxtaposed substrings from different languages, using a dictionary check and increasing passphrase length. The problem statement highlighted limitations in the available passphrase policies. A review of the socio sub-stream showed that the context of this study is characterised by users who could generate a passphrase based on multilingualism. The resistance of user-generated passwords to password guessing was used as a measure of strength or security, while the ISO 9241-11 standard on usability was used to evaluate the usability of user-generated passphrases based on multilingual substrings. The ISO 9241-11 standard defines usability as “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction

in a specified context of use”. Based on this definition, this study conceded that the usability of an authentication mechanism is the extent to which it is easy for a user to generate, remember and correctly enter a passphrase into the login prompt (Keith et al., 2007; Shay et al., 2014).

### **1.6 The significance of the study**

Studies on passphrases suggest that this approach may offer a better balance between security and usability, although, to date, the results on the subject remain inconclusive (Blanchard et al., 2018; Melicher et al., 2016; Shay et al., 2016, 2014). This study contributed to the body of password security and usability knowledge by exploring the use of multiple languages in user-generated passphrases. The study demonstrated the feasibility of using passphrases composed of substrings oriented in different languages that a user already knows. Furthermore, the few available studies on passphrases that have been conducted were based on western countries with limited studies from developing countries (Blanchard et al., 2018; Bonneau & Shutova, 2012; Melicher et al., 2016; Pilar et al., 2012; Rao et al., 2013; Shay et al., 2016; Veras et al., 2014). This is critical given that different cultures adapt and conceptualise the use of information and technologies uniquely (Winschiers-Theophilus & Bidwell, 2013). For example, AlSabah et al. (2018), Wang, Cheng, et al. (2015) and Li et al. (2014) note that the character composition of passwords owned by users from diverse cultures is significantly different. These findings are in line with propositions in socio-cultural theory that argue that contextual factors help shape the development of a human being. Similarly, Deumert and Masinyana (2008) and Lexander (2011) used a text message service to show that users from different contexts orient their messages in different ways according to the languages in the context. Accordingly, this study motivates the use of multilingual passphrases.

### **1.7 Research methodology**

This study focused on proposing a model for secure and usable passphrases. This section gives an overview of the research methodology for the study.

### **1.7.1 Research paradigm and philosophical view**

Section 2.1.1 in Chapter 2 discusses different research philosophies and paradigms. This study assumed a pragmatist research paradigm; hence a design science research approach was adopted which falls within the pragmatist paradigm. Design science is a problem-solving research approach that entails knowledge creation through building and evaluating human-made artefacts (Peppers, Tuunanen, Rothenberger, & Chatterjee, 2008; Hevner, March, Park, & Ram, 2004), something that is consistent with the focus of this study. This study focused on producing a model for secure and usable multilingual passphrases.

### **1.7.2 Research approach**

Section 2.1.4 in Chapter 2 identifies and discusses different reasoning approaches including inductive, deductive and abductive reasoning. This study assumed an abductive reasoning approach. The design of an artefact in the form of a model in Chapter 5 is in line with the propositions in the abductive research reasoning. The study went on to use a deductive reasoning approach during the evaluation of the proposed artefact using empirical evidence.



### **1.7.3 Research method**

This study adopted the guideline for design science proposed by Peppers et al. (2008). Hevner et al. (2004) and Vaishnavi and Kuechler (2015) also propose separate design science guidelines and process models; however, Peppers et al.'s (2008) design science guideline was arrived at after consolidating ideas from other process models in the literature, Hevner et al.'s (2004) included. Hence, their guidelines are comprehensive. In addition, this study used mixed methods for data collection and analysis, thus collecting both qualitative and quantitative data. The use of both quantitative and qualitative research methods is consistent with the literature (Keith et al., 2007; Melicher et al., 2016; Shay et al., 2016).

#### **1.7.3.1 Experiment design and data collection**

Data was gathered to evaluate the strength and usability of passphrases and short passwords. An experimental framework for gathering passwords using different

policies used by Shay et al. (2016) and Komanduri (2016) was adapted for this study. The experiment for this study was based on an online web application. Participants were asked to generate a password following two different password generation rules, namely, short passwords and passphrases. Short password generation was guided by a comprehensive eight-character (Comp8) password policy that required participants to generate an eight-character password with different character classes. The proposed model in Chapter 5 was translated into a passphrase policy that requires users to generate a passphrase based on juxtaposed substrings from different languages. An online questionnaire was used for data collection. In addition, raw passwords together with key logs generated during the experiment were gathered. Section 2.4.2 gives a detailed discussion of the experiment design and data collection procedures.

#### **1.7.3.2 Data analysis**

This study used statistical analysis techniques on the data that was gathered using a questionnaire and raw data generated by key logs. These statistical analysis techniques were used to establish short password and passphrase usability as explained in Section 2.5. Short password and passphrase resistance to password guessing was used as a measure of password strength; the use of password guessing as a measure of strength is justified Chapter 4. A PCFG proposed by Komanduri (2016) was used for password guessing. This PCFG is regarded as being among the best password guessing algorithms that can be used on short passwords and passphrases. Content analysis was used to analyse raw passwords in order to establish password characteristics guided by Section 2.5.1. Once data was analysed, Chapter 8 went on to evaluate the proposed model.

#### **1.8 Delimitation of the study**

The study extended the evaluation and validation of passphrases to the Southern African context. These passphrases were based on substrings oriented to different languages. The study used participants from selected universities in Namibia and South Africa, focusing on how users could generate secure and usable passphrases using the knowledge they already had. Other authentication mechanisms such as the use of (what you are) biometrics, keystroke analysis and (what you have) tokens were

not considered in this study. Furthermore, the security threat model considered for this study was online and offline password threats. Non-technical and passive online password attacks are beyond the scope of this research project. Section 4.2 in Chapter 4 further explains the password threat delimitation in this study.

## 1.9 Ethical considerations

This study adhered to ethical requirements suggested in the literature. An ethical clearance certificate was issued by the University Research Ethics Committee with the reference number FLO081SMAO01. Requirements for non-maleficence, fairness in selecting participants, putting in place measures to ensure privacy and anonymity of participants were met. Chapter 2 gives more details on the ethical considerations for this study.

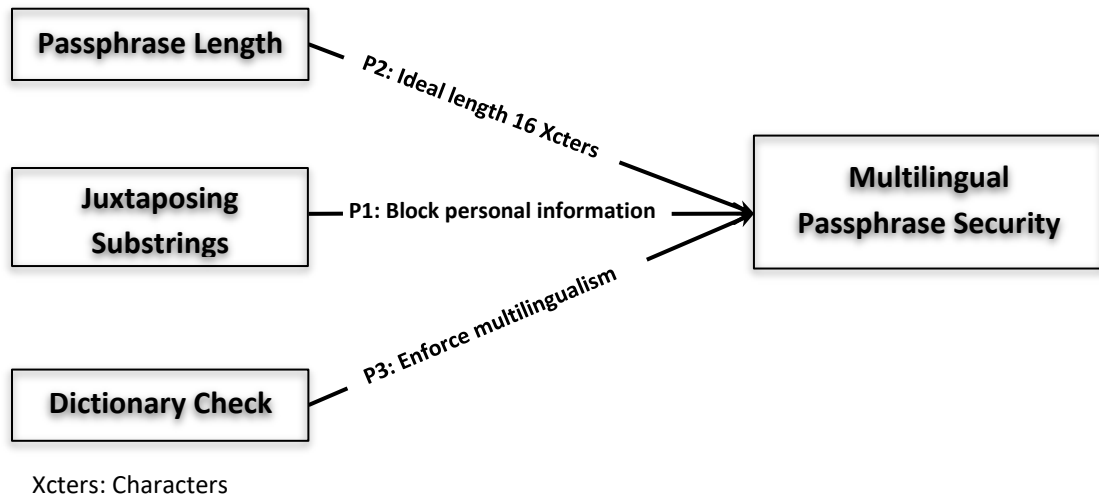
## 1.10 Study contribution

This study developed a model of secure and usable multilingual passphrases. The security and usability contributions on passphrases made by this study are as follows:



University of Fort Hare  
*Together in Excellence*

**Security contributions.** This study shows that it is possible for a multilingual user to generate stronger multilingual passphrases when compared to short passwords based on different character classes. A multilingual passphrase of at least 16 characters proved to be secure when compared to short passwords. Thus, the study showed that, while the dominance of English language within the context may influence password choices, adopting multilingual passphrases can give users an opportunity to generate more secure passphrases. A context-informed dictionary check (a word list with words based on languages in the context) can be used to enforce the use of multilingual passphrases. Figure 1, the Multilingual Passphrase Security Model, summarises the security contributions of this study. This is followed by a list of security propositions made by the study.



**Figure 1. Security contributions: the Multilingual Passphrase Security Model**

The following security propositions were formulated:

- **Proposition P 1:** *Restricting the use of personal information in user generated passphrases by juxtaposing substrings from different languages can enhance security.*
- **Proposition P2:** *Increasing the length of a multilingual passphrase to at least 16 characters based on at least two substrings enhances passphrase security.*
- **Proposition P3:** *The use of dictionary checks can motivate users to base their passphrases on multiple languages.*

**Usability contributions.** This study proposed a model for guiding the usability of multilingual passphrases where effectiveness, efficiency and user satisfaction were found to be important. These usability findings are with reference to passphrase generation and recall, as shown in Figure 2.

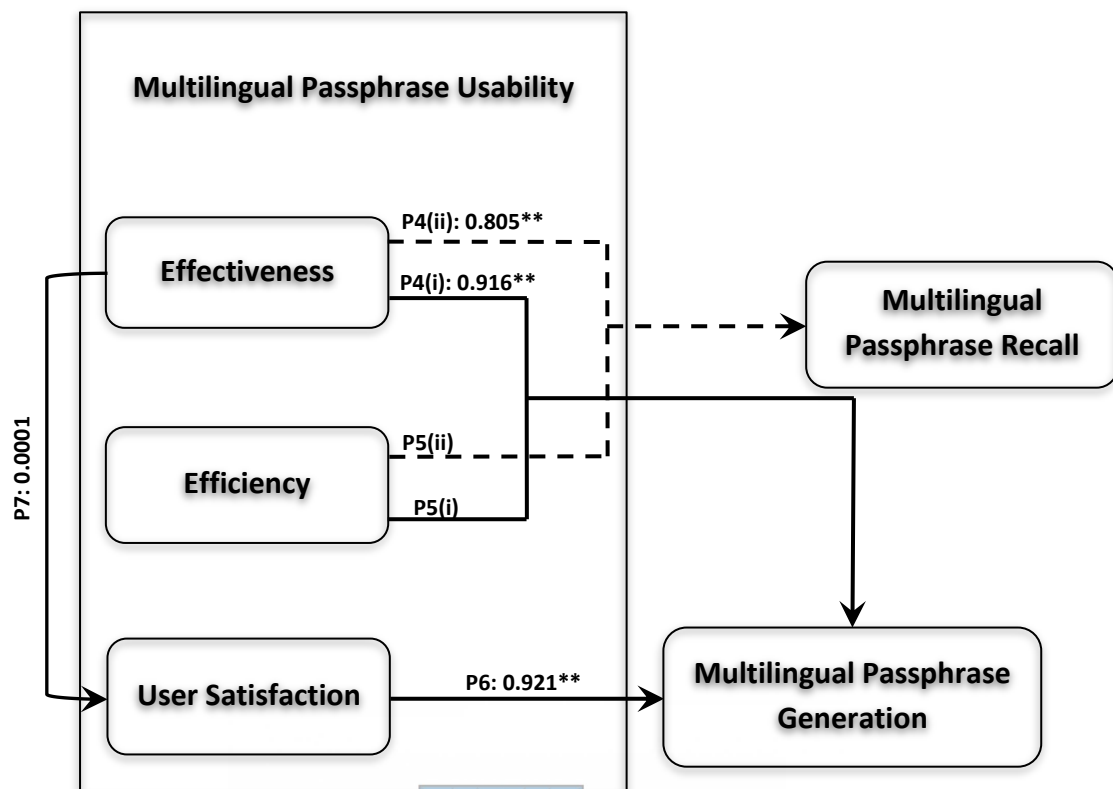


Figure 2. Passphrase usability contributions: a model for usable multilingual passphrases

The propositions summarised in Figure 2 on passphrase generation and usability with respect to factors of effectiveness, efficiency and user satisfaction are as follows:

#### Effectiveness:

- **Proposition P4(i):** The ability to effectively generate a multilingual passphrase without experiencing negative consequences positively influences passphrase usability.
- **Proposition P4(ii):** Effectively recalling a multilingual passphrase without experiencing negative consequences leads to passphrase usability.
- **Proposition P7:** Effective multilingual passphrase generation positively influences user satisfaction with the passphrase policy.



#### **Efficiency:**

- **Proposition P5(i):** Efficacy during multilingual passphrase generation positively influences passphrase usability.
- **Proposition P5(ii):** The repeated use of a multilingual passphrase over time positively influences the usability of passphrases.

#### **User satisfaction:**

- **Proposition P6:** User satisfaction with a multilingual passphrase policy during passphrase generation leads to passphrase usability.

### **1.11 Outline of the study**

**Chapter 1** introduced the research, discussing the research context, the problem area, objectives, literature review and an overview of the research methodology. **Chapter 2** continues by articulating the research methodology used for this study. This chapter explains the use of design science research in this study. **Chapter 3** explains socio sub-stream of this study, as envisaged by socio-technical theory. The functionality of the memory is also explained. Further, socio-cultural theory is used to understand the contextual setting of the study. **Chapter 4** goes on to explain technical sub-stream of the study according to socio-cultural theory. The focus falls on positioning text-based authentications among other authentication mechanisms, discussing password threats and security measures. In addition, password guidelines, best practices and policies are evaluated. Chapter 4 concludes with an evaluation of user behaviour during password generation and recall. **Chapter 5** demonstrates the use of the socio sub-stream explained in Chapter 3 and the technical sub-stream explained in Chapter 4 to develop a proposed model. This is commensurate with research paradigm of the study which encourages the design of research outputs where a model is one of the possible outputs. The proposed model identifies the position of the study regarding passphrase security and usability.

**Chapter 6** presents the study findings. Findings on passwords are presented according to the password policies pertaining to short passwords and passphrases investigated. **Chapter 7** continues by discussing the research findings from Chapter 6.

The chapter establishes the utility, efficacy and quality of the passphrase policy by comparing findings on passphrases against those on short passwords. The findings in Chapter 7 are also compared with those in the literature. This leads to conclusions on the study findings. In addition, **Chapter 8** uses conclusions drawn in Chapter 7 to outline the research contributions and recommendations of this study. In addition, the proposed model is evaluated to outline the study's contributions. Subsequently, the model in Chapter 5 is modified to reflect the study findings. **Chapter 9** concludes the study, indicating what has been done to address the research questions and meet the research objectives. The contributions and limitations of the study are also delineated and possible future research areas are explained.

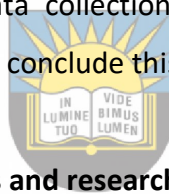


University of Fort Hare  
*Together in Excellence*

## CHAPTER 2: THE RESEARCH DESIGN AND METHODOLOGY

### 2.0 Introduction

Philliber, Schwab, and Samsloss (1980, in Yin, 2003) define research design as a blueprint for research that deals with at least four problems: what questions to study, what data is relevant, what data to collect and how to analyse the data. This definition concurs with a view by Eisenhardt (1989, in Iacono, Brown, & Holtham, 2011), who argues that a research process starts with defining the research question followed by identifying constructs from the literature, and then comparing and contrasting the themes emerging from the fieldwork with the literature. Accordingly, the previous chapter delineated the study. This chapter presents the research design and methodology followed in this study, expounds on the philosophical commitments, and explains the research reasoning approach and the application of the design science research. Techniques for data collection and analysis are also outlined. Ethical considerations and a summary conclude this chapter.



### 2.1 Philosophical assumptions and research paradigm

A research philosophy “refers to a system of beliefs and assumptions about the development of knowledge” (Saunders, Lewis, & Thornhill, 2016, p. 124). It reflects how a researcher views the world during the development of knowledge. This implies that different philosophies have different ways of viewing and understanding the world. However, authors on research methodology across different research disciplines have motivated the idea that a research paradigm is a framework that reflects a particular research philosophy (Collis & Hussey, 2013; Khan, 2014; Neuman, 2014). In other words, a research paradigm is considered to be a “philosophical framework that guides how scientific research should be conducted” (Collis & Hussey, 2013, p. 43). Neuman (2014) adds that a paradigm is a whole system of thinking that is defined as “a general organising framework for theory and research that includes basic assumptions, key issues, models of quality research, and methods for seeking answers” (p. 96). Thomas Kuhn, the researcher who first introduced the term and ideas around a research paradigm, defines a research paradigm as “a set of values and techniques which is shared by members of a scientific community, which act as a guide or map, dictating

the kinds of problems scientists should address and the types of explanations that are acceptable to them” (Kuhn, 1970, p. 175, in Khan, 2014, p. 225).

While there is a degree of consistency and understanding of what a philosophy and paradigm is, it should be noted that the literature shows a lack of consistency in its descriptions, categorisations and classifications of research philosophies (Collis & Hussey, 2013; Mkansi & Acheampong, 2012). This poses a major challenge to researchers as they try to identify and select philosophical assumptions that suit their studies (Mkansi & Acheampong, 2012). For instance, Collis and Hussey (2013) suggested that there are two main categories of research philosophies, namely, realism that emanates from the natural science discipline and idealism that emanates from the social science research discipline. Collis and Hussey (2013) went on to identify positivism and interpretivism as research paradigms that broadly fall within the realism and idealism philosophy respectively. They also acknowledge the existence of other research paradigms, depending on the extent of orientation towards positivism or interpretivism. On the contrary, Saunders et al. (2016) suggest five research philosophies, namely, positivism, critical realism, interpretivism, postmodernism and pragmatism. Collis and Hussey (2013) and Saunders et al. (2016) further suggest that the paradigmatic framework of these philosophies can be considered according to three perspectives, namely, ontology, epistemology and axiology. However, Collis and Hussey (2013) add two perspectives to their paradigmatic framework, namely, rhetorical and methodological assumptions.

Ritchie and Lewis (2003 in Mkansi & Acheampong, 2012) suggest a slightly different philosophical framework which they argue comprises an ontological and epistemological perspective, respectively. The ontological perspective could be viewed with respect to realism, materialism, critical realism, idealism and relativism, while the epistemological perspective includes positivism and interpretivism. This framework is close to Burrell and Morgan (1978) in Iviri and Venable's (2009) philosophical framework with two extreme ontological categories, namely, realism and nominalism; two epistemological positions, namely, positivism and anti-positivism; methodology and ethics. Drawing from these views on research philosophy and paradigm, it is clear

that authors have different perceptions on what constitutes a philosophical framework. This suggests the importance of identifying the characteristics of a philosophical framework assumed by a study. This study assumed the following attributes of a philosophical framework that are commonly used in the literature for reviewing philosophical and paradigmatic views:

- Ontology – relates to the nature of reality.
- Epistemology – relates to the relationship between the researcher and the phenomena under study. It looks at how knowledge is acquired.
- Methodology – assumes the processes and method through which the researcher acquires knowledge (Khan, 2014).
- Axiology – Vaishnavi and Kuechler (2015) indicate that axiology is a Greek word. Accordingly, this study adopts a view by Vaishnavi and Kuechler (2015) that conforms to the Greek definition of axiology where axiology is seen as the assumption of that which is of value from a research study or what is worth researching.



Understanding the attributes of a philosophical framework is important as this helps to define the philosophical view of a study. The next section uses the above attributes of a philosophical framework to explain different philosophies and to identify the philosophy that was used in this study.

### **2.1.1 Overview of research philosophies**

This section discusses three philosophical stances, namely, realism, idealism and pragmatism. The study acknowledges the existence of other research philosophies as indicated in the literature (Mkansi & Acheampong, 2012) and Section 2.1. The discussion of realism in this study was motivated by the fact that it is the oldest and first philosophy identified and was followed by idealism (Collis & Hussey, 2013). Collis and Hussey (2013) further state that realism and idealism gave rise to the positivist and interpretivist paradigms respectively. In addition to realism and idealism, Information Systems and business management researchers identified a third philosophy, namely, pragmatism (Ågerfalk, 2010; Hevner et al., 2004; livari, 2007; livari & Venable, 2009;

Mkansi & Acheampong, 2012; Saunders et al., 2016; Vaishnavi & Kuechler, 2015). Together, these three are the most commonly used philosophies in Information Systems (Mkansi & Acheampong, 2012). Hence the need for researchers to determine their philosophical orientation before designing a research project (Collis & Hussey, 2013). This section demonstrates the depth of understanding of the difference between different philosophies as a way of justifying philosophical choices for this study (Saunders et al., 2016), where the focus is on the realism, idealism and pragmatist philosophies respectively:

**Realism:** The philosophy of realism is arguably what Saunders et al. (2016) consider to be positivism. Realism is a philosophical stance of natural scientists whose research is mainly characterised by physical objects or phenomena. The realism is founded on the premise that there is only one single reality for every phenomenon under study. This single reality is considered independent from the researcher and can be measured and scientifically verified to produce law-like generalisations. A study conducted following the positivist paradigm often adopts a deductive reasoning approach and quantitative research methods that lead to the truth and theoretical explanations.



University of Fort Hare  
Together in Excellence

**Idealism:** The philosophy of idealism, as referenced by Collis and Hussey (2013), is arguably what Saunders et al. (2016) refer to as interpretive philosophy. Idealism is common to social science research and was developed to address the limitations of realism. Idealism posits that, unlike objects usually researched by natural scientists, humans have multiple realities which are subject to the interpretations of the researcher. Studies conducted following the philosophy of idealism are often characterised by an inductive reasoning approach in which theories emerge from data. Idealism also uses qualitative research methods.

**Pragmatism:** The philosophy of pragmatism aims to reconcile both objectivism and subjectivism (Saunders et al., 2016). The pragmatist philosophy is associated with applied research that aims to make practical and theoretical contributions. This philosophy is distinguished from other philosophies by its quest to make tangible useful

artefacts that solve a particular problem. The literature suggests action research and design science as two paradigms aligned to principles in the pragmatist philosophy (Ågerfalk, 2010; Hevner et al., 2004; livari, 2007). There is ongoing debate whether action research and design science should be seen as research paradigms (Baskerville, Kaul, & Storey, 2015; Gregor & Hevner, 2013; Hevner et al., 2004; Peffers et al., 2008; Vaishnavi & Kuechler, 2015). livari and Venable (2009) are of the view that design science is not a paradigm but a research orientation that is “based on more or less ‘positivistic’ or ‘interpretivist’ assumptions” (p. 7). In other words, design science is considered a research orientation that is shaped by the extent of one’s orientation to positivist or interpretivist paradigmatic assumptions (livari & Venable, 2009). In light of this ongoing debate, this study views design science as a research approach or orientation that falls under the philosophy of pragmatism as shown in Table 1. Mixed methods are often used in association with pragmatism.

Table 1 uses the philosophical framework in Section 2.1 to summarise the characteristics of paradigms that fall under the philosophies of realism, idealism and pragmatism.



University of Fort Hare  
Together in Excellence

**Table 1. Philosophical assumptions of the four paradigms aligned to the three philosophies discussed (livari & Venable, 2009; Vaishnavi & Kuechler, 2015).**

Philosophical assumption	Research paradigm/philosophy		
	Realism (Positivism)	Idealism (Interpretivism)	Pragmatism
Ontology	Single reality Universal Probabilistic	Multiple realities Socially constructed	Design science Multiple, contextually situated alternative world-states Socio-technology enabled
Epistemology	Objective; detached observer of truth	Subjective	Knowing through making; objectively constrained construction within context; iterative circumscription
Methodology	Observations Quantitative Statistical	Participation; qualitative Hermeneutical; dialectical	Developmental Measure impact of the artefact
Axiology	Truth; predictions	Understanding	Control; creation; utility; understanding

### 2.1.2 Philosophical stance of this study

Information Systems is a multi-paradigm, multidisciplinary research discipline that thrives on applied research (Niehaves, 2007; Peffers et al., 2008). This makes Information Systems research unique when compared to established research disciplines that Vaishnavi and Kuechler (2015) consider to be “paradigmatic” disciplines, where research can be done without reference to any philosophical stance, as noted by Mkansi and Acheampong (2012). Information Systems remains a young research discipline with a limited theoretical base and methodological background; for these reasons it relies on vast theories and methodological frameworks from other research disciplines. This calls for Information Systems researchers to clarify their philosophical commitments as these have an impact on how “we understand what it is we are investigating” (Johnson & Clark, 2006, in Saunders, Lewis, & Thornhill, 2009, p. 108).

This study subscribed to the pragmatist philosophy. Pragmatism is one of the three philosophical views that is usually considered for Information Systems including critical social theory (interpretivism) and positivism (Mkansi & Acheampong, 2012). Pragmatism was considered for this study owing to its emphasis on applied research and practical implications (Ågerfalk, 2010). The study was guided by propositions in the behavioural science and design science research disciplines in its quest to develop a model for usable and secure multilingual passphrases. It should be noted that users find it difficult to generate and recall passwords. Accordingly, the study aimed at addressing a socio-technical problem that arose from the intersection between the technical nature of Information Systems and the behavioural responses to technology by users (Niederman & March, 2012). The overall aim of the study was to develop a model that could be used in crafting passphrase policies that offer better utility, quality and efficacy in terms of security and usability. The pragmatist philosophy guided the study towards the proposition of an alternative password authentication policy from which users’ behavioural changes were evaluated to measure the extent of usefulness and utility.

Given that this study focused on proposing a new innovative artefact for enhancing the usability and security of passphrase authentication policies, this study



adopted the design science research guidelines. As shown in Table 1, design science is a research approach that falls under the pragmatist philosophy. Iivari and Venable (2009) argued that design science research involves building new innovative artefacts for solving a defined problem. Vaishnavi and Kuechler (2015) cemented this definition by stating that design science research “changes the state-of-the-world through the introduction of novel artefacts” (p. 31). This is in stark contrast to action research; an approach that also falls under pragmatism, which aims at understanding the existing reality (Iivari & Venable, 2009), something that was not the primary aim of this study. This study sought to create a new reality, in the form of a model, whose utility, efficacy and quality were evaluated in a given contextual environment by theorising the context.

### **2.1.3 Design science research**

Design science research became a subject of interest in Information Systems following Hevner et al.'s paper publication titled, “Design Science in Information Systems Research” in 2004 (Goes, 2014; Gregor & Hevner, 2013; Prat, Comyn-Wattiau, & Akoka, 2015). Both the design science research and the Information Systems discipline are young (Gregor & Hevner, 2013); hence, the building blocks of design science in Information Systems research are still maturing and its theoretical base is not yet stabilised (Prat et al., 2015; Vaishnavi & Kuechler, 2015). As a result, design science is often seen as a methodological “hodgepodge” that does not fit squarely into the “existing research pigeonholes but, from appearances, might fit partly into all of them” (Baskerville, Kaul, & Storey, 2015, p. 542). It is therefore important to clarify the position of this study in its conceptualisation and use of design science research.

Section 2.1.1 acknowledged some sections of the literature that identify design science as a research paradigm. Given that design science is a young subject, this study concluded that it falls under the pragmatist paradigm or philosophy. The study adopted the focus of Baskerville et al. (2015) on research outputs that result from the building (design) and evaluation (science) of artefacts, and acknowledged that design science has the potential of being multi-paradigmatic and can produce innovative artefacts. Section 2.2 discusses the application of design science in this study.

#### **2.1.4 The reasoning approach of the study**

Now that the philosophical stance and paradigmatic views of this study have been clarified, it is important to explain its reasoning approach and demonstrate how the approach is commensurate with the adopted philosophical framework. Saunders et al. (2016) identified three reasoning approaches to theory development, namely, induction, deduction and abduction. Lee, Pries-Heje, and Baskerville (2011) are of the view that induction and deduction are the dominant reasoning approaches to theory building in many research disciplines. Inductive reasoning involves building theory from cases or data, while deduction focuses on evaluating or testing theories (Eisenhardt & Graebner, 2007). Deduction has its roots in the natural sciences, while induction is aligned to the social sciences (Saunders et al., 2016). Unlike induction and deduction that involve the creation of knowledge in some sort of mechanistic way or following a logical sequence, abduction involves a researcher's creativity in the creation of new knowledge (Kovács & Spens, 2005; Lee et al., 2011). Kovács and Spens (2005) suggest abduction is used when the problem at hand is unique and cannot be solved by the available theories. Through an iterative creative process, the researcher has to create a new theory or framework (artefact) that offers propositions set to address an identified problem. Abduction often works with deduction as the newly created framework has to be evaluated within a context and final propositions deduced, based on empirical evidence (Vaishnavi & Kuechler, 2015; Kovács & Spens, 2005).

This study aimed at contributing to theory development and proposing an artefact. These contributions are consistent with expected outcomes of design science research (Baskerville et al., 2015; Gregor & Hevner, 2013; Lee et al., 2011; Prat et al., 2015). For instance, Prat et al. (2015) stated that design science should be developing theories in addition to building and evaluating artefacts. Accordingly, the conducting of this study assumed an abductive reasoning approach that promotes the creation of new design science theory and artefacts. Baskerville et al. (2015) argue that design science is a duality of design and science. A researcher abductively uses the available literature to design and create a new world or phenomenon-artefact. The science component goes on to deductively evaluate a newly created world to ascertain the utility of the artefact. This is consistent with the literature that suggests that design science involves

two major activities, namely, designing or building and evaluating innovative artefacts (Hevner et al., 2004; Iivari & Venable, 2009; Niederman & March, 2012; Peffers et al., 2007; Prat et al., 2015; Vaishnavi & Kuechler, 2015).

## **2.2 Design science research guidelines and process models**

The literature presents different but related design science research guidelines and process models that depict how knowledge and an artefact are developed (Hevner et al., 2004; Peffers et al., 2008; Vaishnavi & Kuechler, 2015). Hevner et al. (2004) propose seven guidelines that should be considered when conducting and evaluating good design science research. Using the available literature, Peffers et al. (2008) propose six guidelines for design science research, while Vaishnavi and Kuechler (2015) fairly recently developed a five-step design science research process model that elaborates the knowledge usage and knowledge-building process within a cycle of each design science research. Vaishnavi and Kuechler (2015) went on to concede that their design science process model is similar to that of both Hevner et al. (2004) and Peffers et al. (2008) among others. However, Vaishnavi and Kuechler (2015) argue that their model focuses more on generating design science knowledge.

The conducting of this study subscribed to design science research guidelines and the process model developed by Peffers et al. (2008). Their process model has six activities, namely, problem identification and motivation; definition of the objectives for a solution; design and development; demonstration; evaluation, and communication. Figure 3 summarises the process model proposed by Peffers et al. (2008). These design science research guidelines by Peffers et al. (2008) were developed following a consensus method using design science process models in the literature. As such, this process model is compatible with propositions in the leading design science guidelines by Hevner et al. (2004), among other leading design science process models. Hence, propositions by Peffers et al. (2008) are considered as one of the excellent design science research process models (Vaishnavi & Kuechler, 2015). Recent work by Gregor and Hevner (2013) has shown how the process model designed by Peffers et al. (2008) could be used to demonstrate the knowledge contributions of design science research.

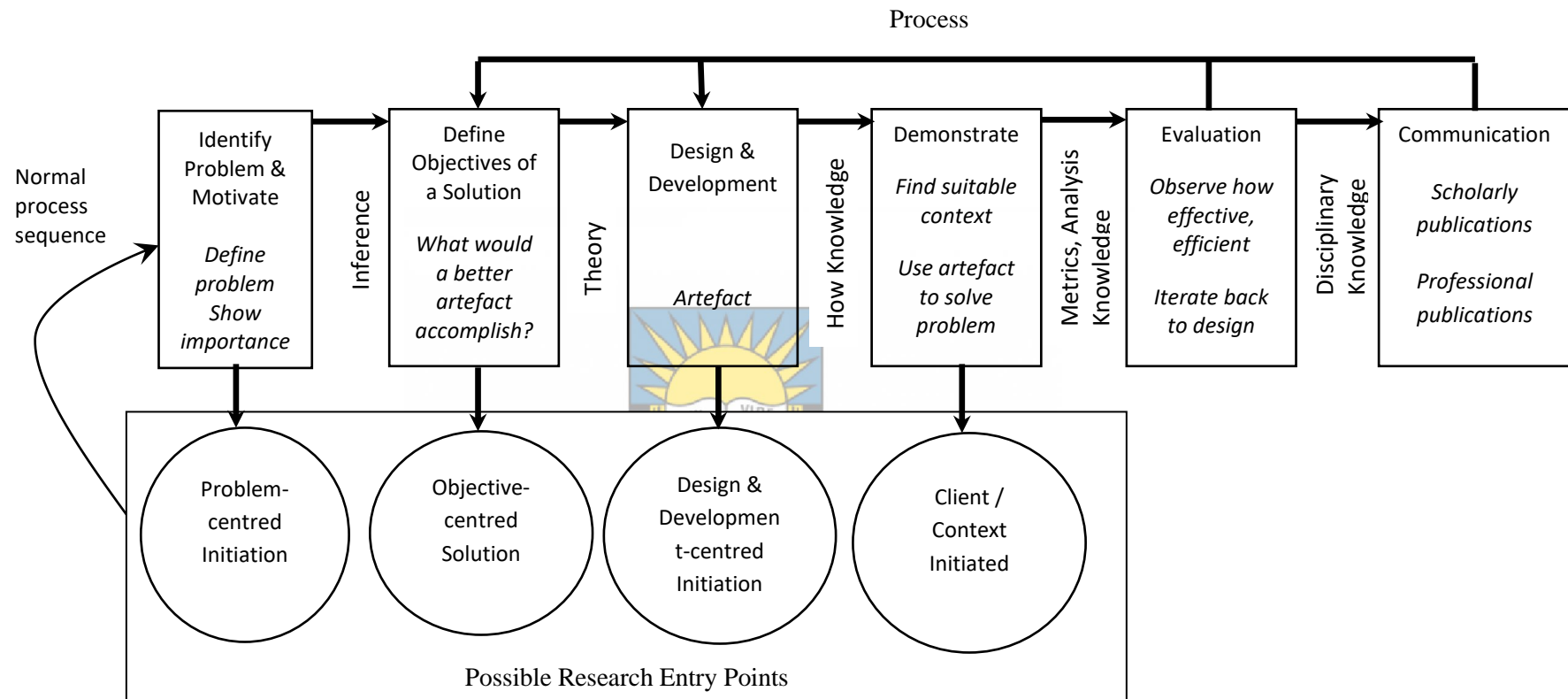


Figure 3. Design science research methodology (adopted from Peffers et al., 2008, p. 54)

### 2.2.1 The application of Peffers et al.'s (2004) process model to this study

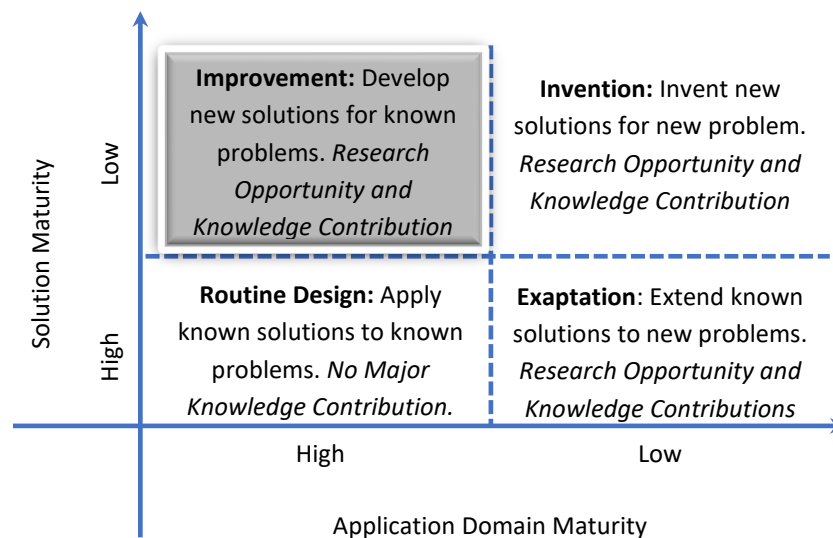
This section revisits the activities in Peffers et al.'s (2008) design science process model thereby explaining what was done when this study adopted this model. Peffers et al. (2008) concede that it is perfectly normal for a design science study to start from any activity other than activity one. This is in line with suggestions by Hevner et al. (2004), who also concede that there is no one fixed layout for addressing the design science research guidelines proposed in their framework as long as every guideline is addressed. This study is problem centred; as such, it started with **Activity One (Problem identification and motivation)**, shown in Figure 3, which saw the identification of the research problem to motivate the importance of the study. The research problem was established from a literature review. Chapter 1 of this study used “kernel theories” in the literature to motivate the argument that the research problem of the study is of a socio-technical nature. The research problem, as stated in Chapter 1, is specified below:

*Generating and using passphrases in a manner that meets security and usability requirements remains a challenge.*



**Activity Two: Define the objectives of a solution.** This activity focuses on the inference of an objective or the deliverable of the research; that is, a feasible solution to the identified problem. The utility of a solution to the identified problem should be scientifically proven using qualitative or quantitative methods (Peffers et al., 2008). In light of the research problem, this study sought to develop a model for generating secure and usable multilingual passphrases. The study drew from the understanding of short password and passphrase challenges and aimed at proposing an improved artefact as suggested in the design science knowledge contribution framework by Gregor and Hevner (2013), shown in Figure 4. The model developed in this study could be used to inform the design of secure and usable passphrase authentication policies. The literature confirms that a model is an example of an artefact that could be produced by design science research (Hevner et al., 2004). The model proposed for this study is a product artefact that is socio-technical in nature (Venable, Pries-Heje, & Baskerville, 2012). Venable et al. (2012) state that a socio-technical artefact is an artefact whose

utility is a result of its interaction with humans. Thus, such artefacts are not purely technical where human use is not required.



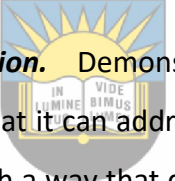
**Figure 4. Design science knowledge contribution framework (adapted from Gregor & Hevner, 2013, p. 245)**



**Activity Three: Design and development.** This activity focuses on designing and developing an artefact. The desired functionality and architecture of the artefact are defined during this activity. The available theory – often known as the kernel theory or justificatory knowledge or reference theory – provides guidelines and principles for grounding the development of the artefact and justifying its expected functionality (Gregor & Hevner, 2013; Hevner et al., 2004; Peffers et al., 2008). Such theories can emanate from research disciplines other than the Information Systems field. Chapters 3, 4 and 5 of this study address Activity Three: design and development. The study followed suggestions to search through the available literature and use creativeness (abduction) to develop an artefact for addressing the research problem. To propose a social-technical artefact, the study was grounded in socio-technical theory, which emphasises a “joint optimisation of the technical and socio subsystems, rather than the optimisation of the technical subsystem and the adaption of the social subsystem around it” (Cherns, 1987, in Durkin et al., 2015, p. 948). This was motivated by the fact that when an information system is applied to a contextual environment, its success depends on social rather than technological constructs (Doherty, 2014). Accordingly,

the following chapters in this study contribute to the design and development activity as follows:

- Chapter 3 discusses and synthesises the social component of socio-technical theory, which are in turn expounded by socio-cultural theory. Language and the psychological development of human cognition are used to build the foundation of the argument for secure and usable multilingual passphrases in the study.
- Chapter 4 discusses and provides a synthesis of the technological components of socio-technical theory. The password security threat model for this study is conceptualised and measures for attaining password strength are explained. Available password policies, guidelines and best practices are discussed, exposing their limitations.
- Chapter 5 proposes a model for secure and usable passphrases based on the findings of Chapters 3 and 4. Abduction is used to arrive at the proposed model.



**Activity Four: Demonstration.** Demonstration involves the deployment of the newly developed artefact so that it can address the identified problem. The idea is to put the artefact into use in such a way that data can be gathered in order to ascertain its utility in addressing the identified problem using Activity Five. In particular, the factors of passphrase security and usability in the proposed model were used to guide the design of a passphrase policy. The passphrase policy was used to guide password generation on a web-based application by participants. Using web-based applications in an experimental demonstration of password policies is consistent with related studies (Komanduri et al., 2011; Shay et al., 2016). The demonstration saw participants generating their own short passwords and passphrases which were used for authentication purposes during the experiment. Section 2.4.2 and Appendix A give full details on the way the artefact for this study was demonstrated.

**Activity Five: Evaluation.** Data collected following a triangulation of different techniques was analysed to establish the usability and security contributions of the two selected password policies that were demonstrated. Chapter 6 presents findings from data collection. Chapter 7 interprets research findings and Chapter 8 continues by expounding this study's research contributions. The artefact was adjusted according to

findings from data collection and Chapter 9 gives a conclusion to the study. In addition, Chapter 9 paves the way for future research based on study findings. Section 2.5 explains the data analysis techniques for this study.



University of Fort Hare  
*Together in Excellence*



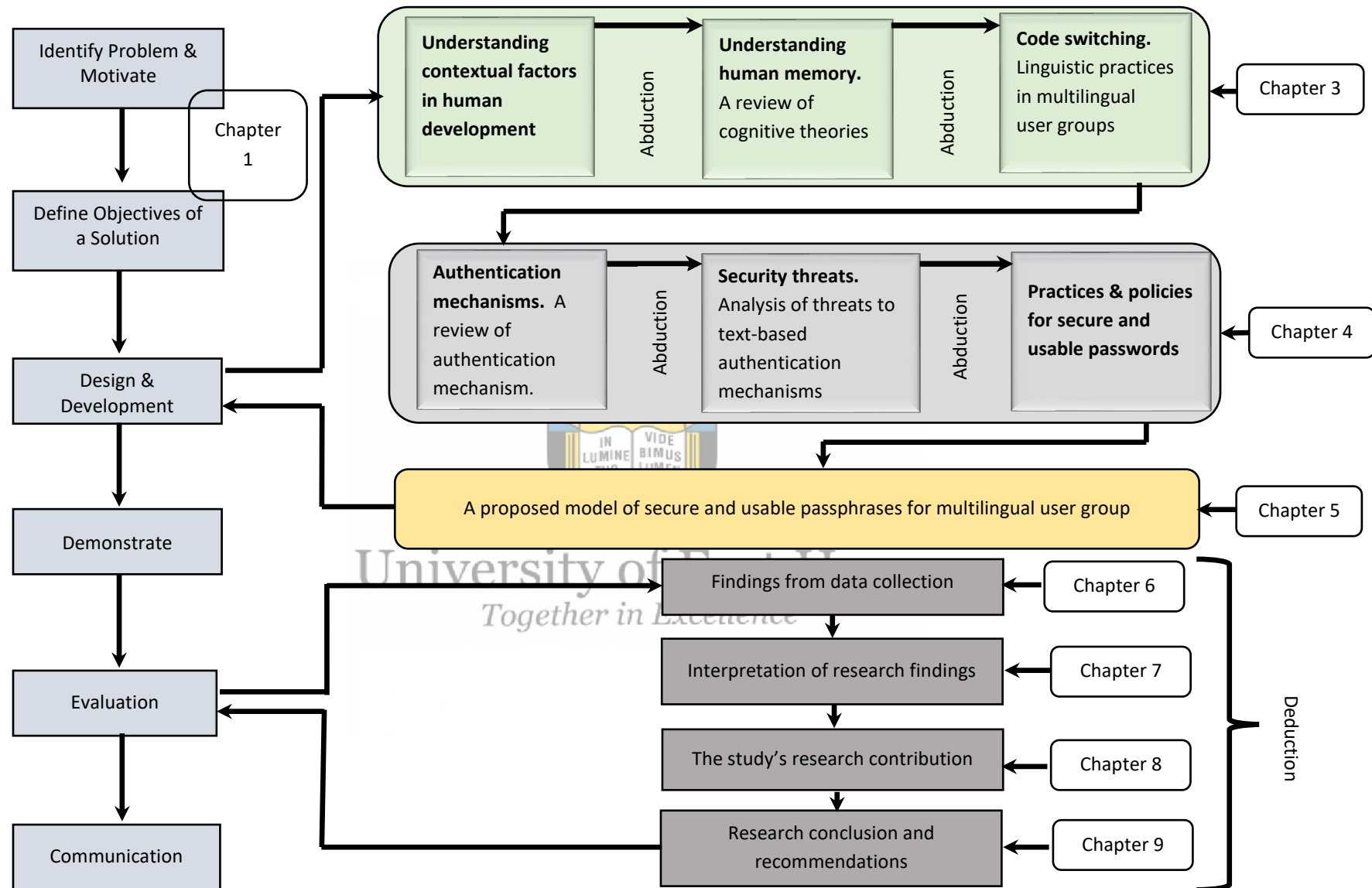


Figure 5. The research process of this study informed by Peffers et al. (2008)

**Activity Six: Communication.** The study findings will be communicated through publications at conferences and in research journals. However, the proposed model for secure and usable multilingual passphrases was presented at a security conference, Information Institute Conferences, Las Vegas, USA in 2017. The conference was used to share ideas with security experts on the model proposed by this study. In addition, part of the research findings was presented at the 33rd IFIP TC 11 International Conference, SEC 2018 held at the 24th IFIP World Computer Congress, WCC 2018 Poznan, Poland. At least two more publications are being worked on as part of communicating the complete findings from this study to the research community. On completion, the thesis will also be made available through the university. Figure 5 summarises the adaptation of Peffers et al.'s (2008) process model in this study.


### 2.3 Research methods

The use of design science research implies that this study focused on building and evaluating artefacts. Mixed methods were used for data collection and analysis. Design science could use qualitative and quantitative research methods at the same time (Iivari & Venable, 2009). Venkatesh, Brown, and Bala (2013) observed that a study may employ multiple research methods by conforming to both the qualitative and quantitative worldviews. This is perfectly consistent with the pragmatist research paradigm. The use of mixed methods in this study was motivated by a need to understand the phenomenon under investigation holistically (Venkatesh et al., 2013). This is critical given that Chapter 1 of this study showed that research findings on passphrase security and usability remain fragmented and inconclusive. In addition, Chapter 3 contextualises the view of this study on passphrases, something that supports a holistic view of the study problem.

A literature review by Venkatesh et al. (2013) showed that the use of qualitative and quantitative methods in mixed methods can be done concurrently or sequentially. Using qualitative and quantitative research methods concurrently implies that methods in a mixed research would be used independently of each other in order to understand the phenomenon under study – in this case, the security and usability of passphrases. By contrast, using qualitative and quantitative methods sequentially suggests that

findings from the two methods would be analysed together in order to corroborate each other. This study assumed a sequential use of mixed methods. It is assumed that findings from qualitative or quantitative research methods can inform each other. This is consistent with the literature on password usability and security (Keith et al., 2009; Komanduri et al., 2011; Melicher et al., 2016; Shay et al., 2016). For instance, Keith et al. (2009) state that when a user fails to log in successfully, the system keeps a record of log in failure while the user generates a perception of system usage. Hence, log in trails can be accessed to gather data on log in failure, while a questionnaire survey can be used to gather data on users' perceptions of an authentication mechanism. Data generated from these different sources would then be used sequentially to understand passphrase usability. The next sections explain how data collection and analysis were conducted using mixed methods in this study. In addition, the measures that were taken to ensure methodological validity and reliability are explained.

## **2.4 Data collection**



The literature on design science research suggests that the data collection process has to be done following two separate phases: secondary and primary data collection. Hevner et al. (2004) propose an information systems research framework that indicates researchers having to use secondary data creatively to justify the shape and design of the proposed artefacts. Further data would be collected in the form of primary data to evaluate the utility of the artefact. This study used an experiment to facilitate primary data collection. The sections below discuss the activities of secondary and primary data collection in this study.

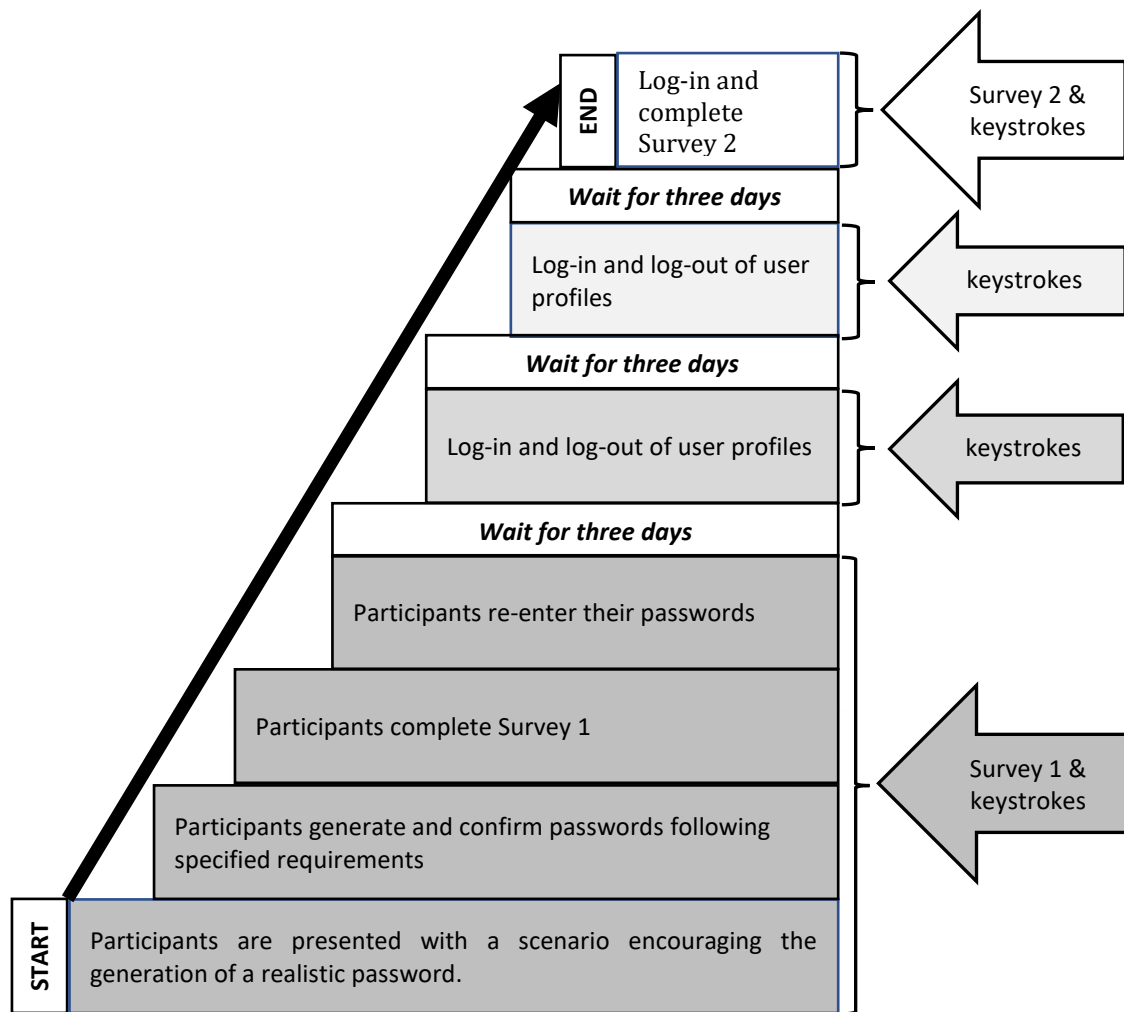
### **2.4.1 Secondary data collection**

Figure 5 shows that Chapter 1 of this study used a literature review to identify research gaps. In addition, human cognition theories were used to enhance the understanding of human memory functionality. Understanding human memory functionality is critical considering that memorability is one of the main influential factors when deciding on passphrase and password choices (Choong et al., 2014). Practices in multiple languages by a multilingual user group were evaluated to establish the feasibility of extending such practices to text-based authentication. Socio-cultural

theory is used in Chapter 3 to justify the possibility of using multilingual based passphrases. In addition, previous research findings on passphrase usability and security are reviewed in Chapter 4. As shown in Figure 5, findings from the literature review of Chapter 3 and 4 are used, through abduction, to propose an artefact for this study. The use of the literature review in this study is consistent with suggestions in the literature, as Hevner et al. (2004) posit that design “is a search process” among available theories “to discover an effective solution to a problem” (p. 88). Similarly, Yin (2012) arguably suggests that the available theories can be used for coming up with the initial design of an artefact. Coming up with theoretical justifiable arguments that explain the expected functionality of the artefact is important, given that design science problems often cannot be solved by the available theories (Vaishnavi & Kuechler, 2015).

#### **2.4.2 The experiment design and primary data collection**

The primary data was gathered by using an experiment and two sets of questionnaires. Experiments and existing leaked password corpora are the commonly used data sources of passwords (Li et al., 2014; Shay et al., 2016; Wang, Cheng, et al., 2015; Yang, Hung, & Lin, 2013). This study focused on proposing a unique passphrase policy that promotes the use of multilingualism; as such, an experiment was deemed suitable as there was no password corpora in the public domain that could have been used to address the research questions of the study. An experimental framework for password generation and use that was developed by Shay et al. (2016) and Komanduri (2016) was adapted for this study. This experimental framework has been in use since 2011. Shay et al. (2016) and Komanduri's (2016) experimental framework allow participants to generate a password following specified conditions and kept the data for each participant organised.



**Figure 6. Password generation experimental activities and data gathering**

The experiment in this study was based on a web application built specifically for the purpose of this experiment. Upon opening the password generation platform, participants were presented with a scenario encouraging the generation of a realistic password as purported by Shay et al. (2016) and Komanduri (2016). Each participant was required to generate a short password and use it over a period of fourteen days, following steps shown in Figure 6. The short password policy is a popular policy that was designed following a guideline by the NIST (Shay et al., 2016). Once done (as indicated by completing Survey 2) with short passwords, the same group of participants was invited to return to the experiment and generate a passphrase based on multilingualism as specified in the proposed model in Chapter 5. Similarly, upon completion of passphrase generation, participants were asked to complete Survey 1 and re-enter their passphrase as shown in Figure 6. Re-entering a password (passphrase

or short password) was meant to demonstrate the ability to recall a password after a cognitive burdening exercise of completing the questionnaire. This practice of using a distractor task in the protocol of an experiment is a common approach for testing short-term memory (Blanchard et al., 2018). Participants were asked to visit the experiment platform after three days and log into their profiles to demonstrate their ability to recall the passphrase. Waiting for three days prior to asking participants to log into their profiles is the standard time period that was used by related studies for testing password memorability (Shay et al., 2016; Komanduri, 2016).

Figure 6 shows the activities of password generation, completing Survey 1, password recall and completing Survey 2 at the end of fourteen days. Data was gathered through questionnaires (Surveys 1 and 2) at intervals shown in Figure 6. In addition, raw passwords and key strokes were captured during password generation and logging into user profiles. The gathered data was used for evaluating password policy usability and security. Appendix A gives a detailed overview of the experimental activities and data collection procedures.

#### **2.4.2.1 The questionnaire**

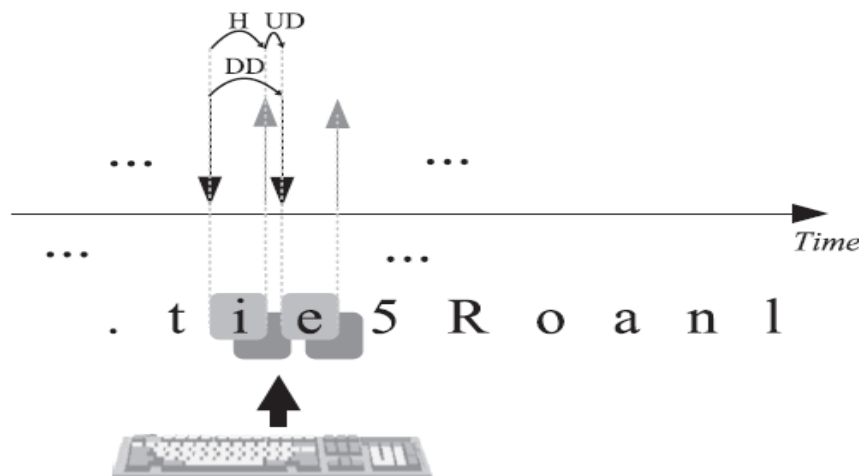
Two structured questionnaires were used for data collection. Questionnaire 1 (used in Survey 1) was used to gather demographic data and data on password (short password and multilingual passphrases) generation strategy, as well as participants' perceptions of password generation usability. This questionnaire was completed on the first day of the experiment. Questionnaire 2 (used in Survey 2) was used to gather data on password recall strategies that were assumed by participants during the fourteen days they were logging in and out of their profiles, as well as data on participants' perceptions of the usability of password recall. These questionnaires automatically appeared online on the experiment platform at designated time periods, as shown in Figure 6. Appendix C shows the questionnaires that were used in this study. Chapter 6 shows how questions in the questionnaire are linked to constructs in the proposed model.

#### 2.4.2.2 The keystrokes

The process of password creation and use generates a number of data sources that can be referred to for ascertaining usability factors (Keith et al., 2009; Melicher et al., 2016) as outlined in Chapter 5. This study used method triangulation in some instances to gather user and system generated data for evaluating password usability during primary data collection. The use of triangulation is consistent with the literature as explained in Section 2.3. This is common to socio-technical artefacts that include the human aspect. Raw data on timestamps for each key event allowed for the computation of the following timestamps that were used for keystroke data analysis:

- Key Down-Down (DD): the time that elapses between pressing two consecutive keys
- Hold (H): time interval between pressing and releasing a key
- Up-Down (UD): the time that elapses between releasing one key and pressing the next key.

Keystroke dynamics that make use of the above key-press timestamps appear to yield better accuracy when estimating password usability (Montalvão, Freire, Bezerra, & Garcia, 2015). Figure 7 shows a visual representation of the way in which this study used keystrokes DD, H and UD to capture timestamps for a password like tie5Roan1.



**Figure 7. Capturing timestamps from keystrokes (adapted from Montalvão et al., 2015).**

#### **2.4.3 Target population and sampling method**

Participants (students) from selected South African and Namibian universities were enrolled to participate in the experiment. Within these universities, South African, Namibian and Zimbabwean nationals were targeted. There is a significant number of Zimbabweans studying in South African and Namibian universities. Alomari and Thorpe (2019) justify the proxy of using university students with their finding that the attitude and behaviour of university students towards passwords is comparable to that of the general population. In addition, von Zezschwitz, De Luca and Hussmann (2013) observed that, on average, computer users generate their first password at the age of 15 and these passwords often remain unchanged or experience minimal changes as users adapt their first passwords for different accounts. As such, it was expected that the targeted university students were still going through their early encounters with passwords. Participants who took part in the study were purposefully selected.

#### **2.4.4 The experiment and data collection administration**

Potential participants were invited to take part in the experiment through in-class announcements and the distribution of posters at the targeted universities. Interested participants were given an overview of the experiment and asked to sign an informed consent form. Contact details of those who had joined the experiment (mobile phone number or electronic mail address) were secured and these were used for sending out reminders during periodic logging in and completion of the Questionnaires. The logging in activities of each participant were monitored to make sure that reminders were sent to those who were behind on the experiment activities. Participants who missed a logging in session by more than four days were considered drop outs and they were eliminated from the experiment. The whole experiment lasted for four weeks: two weeks for short passwords and another two weeks for passphrases. A participant was considered to have completed either a short password or passphrase experiment by completing Questionnaires 1, 2 and logging into their profiles at least three times, as shown in Figure 6. Incentives in the form of airtime, memory sticks and smart phones were used so as to encourage participation. At the end of the first two weeks of the experiment, all participants were given airtime worth R40 (approximately US\$3.50). Completing the passphrase experiment would see a participant receiving a



16gb memory stick plus airtime worth R30. Those who completed both the short password and the passphrase experiment stood a chance to win a smart phone. Subsequently, four smart phones were won by four participants through an open draw.

## **2.5 Data analysis**

Primary data was analysed to establish the utility of using a multilingual passphrase policy over a short password. The next sections present the data analysis techniques that were applied in this study in order to determine password characteristics, security and usability.

### **2.5.1 Password and passphrases characteristics**

This study used short password and multilingual passphrase characteristics to reflect the social context as explained in Chapter 3. The password characteristics that were analysed include language orientation, adopted password structures, password length, the adoption of global passwords and the use of substrings (Komanduri, 2016; Shay et al., 2016; Ur et al., 2016; Wang, et al., 2015; Weir et al., 2009). Language orientation in user-generated passwords was used to establish the influence of contextual factors. Lantolf, Thorne, and Poehner (2015) state that “language in all its forms is the most pervasive and powerful cultural artefact that humans possess to mediate their connection to the world, to each other, and to themselves” (p. 5). This could explain why user-generated passwords can be differentiated on the basis of a user’s language and culture (AlSabah et al., 2018). Content analysis was used to identify the use of languages in user-generated passwords. Levenshtein’s edit distance, explained in Chapter 4, was used to measure the distance between passwords and dictionary words (Campbell, Ma, & Kleeman, 2011; Ur et al., 2015). Two language experts were engaged to identify passwords oriented towards African languages. The engaged experts consisted of a Namibian and a South African national. In addition, this study analysed password structures reflecting the use of different character classes as reported by Weir et al. (2009). Section 4.3 in Chapter 4 gives a detailed analysis of different password structures. Levenshtein’s edit distance (von Zezschwitz et al., 2013) was also used to measure the distance between the user-generated passwords of this study against popular passwords of 2016, 2017 and 2018 that were released by

SplashData (2016, 2017, 2018). The analysis of password characteristics was reported using descriptive statistics to reflect observed magnitudes.

### **2.5.2 Short password and multilingual passphrase security**

This study proposed the use of passphrases based on multilingualism in order to enhance security, making use of guess numbers to estimate the strength of the passphrases and short passwords that were gathered using the experiment outlined in Section 2.4.2 and Appendix A. Kelley et al. (2012) define a guess number as a measure of the number of guesses needed by a short password or passphrase cracking algorithm to accurately and completely estimate a given short password or passphrase. In a way, a guess number shows the resistance of a short password or passphrase to password guessing. As such, short password and passphrase resistance to guessing was used to measure strength. Section 4.3 in Chapter 4 justifies the use of guess numbers as a measure of password strength. Kelley et al. (2012) suggest that guess numbers allow for password policy and inter-cultural comparison by computing the percentage of passwords that can be cracked by an algorithm as well as those that can be cracked at a given number of guesses. Accordingly, comparisons were done to establish the difference in strength between short passwords and passphrases.

#### **2.5.2.1. Hybrid password cracking algorithm**

This study adopted and used a password cracking algorithm proposed by Komanduri (2016) for guess number estimations. Komanduri (2016) modified Weir et al.'s (2009) PCFG and proposed a hybrid password cracking algorithm. This algorithm thrives on its capability to learn the likely passwords to be selected by users, given a password policy. The algorithm works with a token table that contains a sample of passwords generated using a given password policy. The token table keeps a record of the actual passwords and the frequency of password occurrence in the corpus. In addition, passwords with multiple words (passphrases) are conjoined together, removing spaces in between words, and the frequencies of the resultant strings are also computed and kept in the token table. Using data in the token table, Komanduri's (2016) password cracking algorithm is capable of self-generating passwords that were not seen in the password corpus. Approximately 19.4 million entries of passwords and

dictionaries in the public domain were used as training data for the password guessing algorithm (Ur, Segreti, et al., 2015). The sources of the training data include leaked passwords from RockYou (leaked in 2009), MySpace (leaked in 2006), Yahoo! (leaked in 2012) and public dictionaries namely the Google Web Corpus Dictionary, Web2 Dictionary and Inflection Dictionary (Ur, Segreti, et al., 2015). Access to the PCFG algorithm was facilitated by the Carnegie Mellon University Password Research Group's Password Guessability Service through their website: <https://pgs.ece.cmu.edu/>. Chapter 4 discusses how different versions of the PCFG could be used for guessing short passwords and passphrases.

### **2.5.3 Short password and multilingual passphrase usability**

The study used statistical techniques to compare the usability attributes of short passwords and multilingual passphrases. Data that was gathered from the questionnaires was analysed using a t-test to establish the equality of means for the usability constructs. A paired sample t-test was used to compare participants' perceptions of usability constructs between short passwords and multilingual passphrases. A correlational analysis was used to access the linear relationship between usability constructs or factors of the study.

It should be noted that data for evaluating usability for this study constituted two different data sets: quantitative and categorical data. Different data sets were a result of using two different data gathering techniques, that is, the web-based experiment platform through key logs and questionnaires. As such, data analysis of the two separate data sets was done in sequence as explained in Section 2.3. This study used Wilcoxon non-parametric tests to establish the difference in time taken to generate and key in a passphrase and short password.

## **2.6 Validity and reliability**

Validity measures the extent to which the gathered data accurately measures that which it is meant to measure; in this case, constructs in the proposed model. However, Roberts, Priest, and Traynor (2006) are of the view that validity can be split into external and internal validity. External validity ensures that study findings can be

generalised or extended to other research contexts. Accordingly, this study targeted different population groups in terms of gender and tribes of participants. As such, the study ensured a fair representation of the populace, a move that promoted external validity. Roberts et al. (2006) further suggest that internal validity is measured by content, criterion and construct validity. The literature shows that there are different ways of measuring validity that could be considered in a research study (Bell, Bryman, & Harley, 2015). This study focused on construct validity. Construct validity evaluates the extent to which the questionnaire measured the constructs in the proposed model. As such, construct validity ensured that the research conclusions could “be made from the operationalisations of a study to the theoretical constructs on which operationalisations are based” (Yilmaz, 2013, p. 318).

This study used factor analysis to measure construct validity. Henson and Roberts (2006) state that “factor analysis can be used to determine what theoretical constructs underlie a given data set and the extent to which these constructs represent the original variables” (p. 396). An exploratory factor analysis (EFA) was used to establish the content validity of the subsections of the questionnaire that gathered data on passphrases and short passwords usability. Williams, Brown, and Onsman (2012) indicate that EFA is applicable to samples of at least 100 participants. However, the bigger the sample size, the more effective EFA is in measuring construct validity. The computation of factor analysis suggests that factors that represent error or noise have to be eliminated from the sample in such a way that only those factors that contribute to the solution are retained. This study used the eigenvalue rule (Henson & Roberts, 2006). If the eigenvalue is greater than one (eigenvalue > 1 rule), all factors are said to be substantially contributing to the solution; hence, all the considered factors should be retained (Henson & Roberts, 2006). However, any factor with a greatest loading of less than 0.45% should be excluded from the study in line with recommendations by Reise, Waller, and Comrey (2000). These guidelines were adopted as evidenced by the use of the EFA in this study. Chapter 6 presents the findings of the EFA. It also shows the constructs that were evaluated using EFA, indicating how these constructs link with the proposed model in Chapter 5 and the questionnaire.

Reliability is defined as the “degree to which a research instrument measures a given variable consistently every time it is used under the same condition with the same subjects” (Yilmaz, 2013, p. 317). Reliability is often measured by computing the internal consistency in participants’ responses. Internal consistency reflects the extent to which responses from participants relate to each other; that is, are the indicators collectively scoring high or low for the whole data set under review? Cronbach’s alpha coefficient was used to measure internal consistency or reliability (Roberts et al., 2006). Computation of Cronbach’s alpha coefficient results in a figure that ranges between 0 (zero), meaning a lack of internal consistency in the dataset, and 1 (one), meaning a perfect correlation with complete internal consistency in the dataset (Bell et al., 2015). Results of 0.8 and above reflect an acceptable level of internal consistency, although results of 0.7 could be considered acceptable. This implies that a coefficient result of 0.6 and below in the dataset under review shows poor internal consistency. These guidelines were adopted in the evaluation of internal reliability in this study.

## 2.7 Ethical considerations

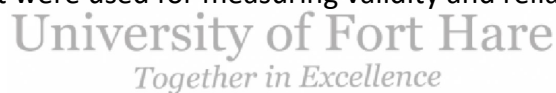
A summary of the research methodology and an overview of the experiment for this study were sent for ethical clearance. A research ethical clearance certificate was obtained from the University of Fort Hare Research Ethics Committee. This suggests that the study complied with ethical principles, namely, autonomy and respect for the dignity of persons; beneficence; non-maleficence and justice (Wassenaar, 2006). Further to that, participants were asked to complete an informed consent form that assured participants of their confidentiality and anonymity. Ethical considerations were critical in this study for proper guidance and transparency, as Choong et al. (2014) note that users can be sensitive to inquiries related to passwords. Therefore, data collected in this study was used for research purposes only. The study also committed to reporting aggregated statistical information of the gathered data during the results presentation with individual account details kept confidential (Wang, Cheng, et al., 2015). Data access controls were put in place to minimise the risk of password database access.



University of Fort Hare  
Together in Excellence

## 2.8 Chapter summary

This chapter explained the research methodology for this study. The methodology explained in this chapter guided the study towards the achievement of the study objectives. In particular, the chapter gave an account of different philosophical assumptions and paradigmatic frameworks. The study is aligned to the pragmatist paradigm, a philosophy that emphasises applied research with practical implications, something that correlates with the focus of this study. A design science research approach was assumed, which saw the study following an abduction reasoning approach in which the literature was used to come up with the initial “design” of an artefact. The chapter went on to show that the study used an experiment to demonstrate the utility of the designed artefact. A web-based experiment platform was used to generate short passwords and passphrases, and for the gathering of responses from questionnaire surveys. Mixed methods were used that saw the collection of data using system logs and completed questionnaires. Statistical analysis was used in a deductive manner during the evaluation of the usability of the artefact, while guess numbers were used to estimate password strength and factor analysis and Cronbach’s alpha coefficient were used for measuring validity and reliability.



The next chapter explains the first stage of data collection for this study which involved a literature review.

## CHAPTER 3: THEORETICAL FOUNDATION

### 3.0 Introduction

This study proposes a model for generating secure and usable multilingual passphrases. An Information Systems research framework by Hevner et al. (2004) points to the assumption that one of the ways for attaining rigour in design science research is by grounding a study in existing theories that are related to the research domain. Information Systems is an applied research discipline that can be conducted in a multidisciplinary and multicultural context (Niehaves, 2007; Peffers et al., 2008). In that regard, this study followed propositions of socio-technical theory for addressing a problem by giving attention to both the social and technical aspects. Accordingly, this chapter gives an overview of socio-technical theory and explains the biological functionality of the human memory. It goes on to discuss contextual factors with a primary focus on language. A socio-cultural theory was used to explain how human mental development occurs within a social context – including language learning. The chapter also reflects on password characteristics aligning to the principles of socio-cultural theory. Studies on text messages were used to portray the influence of the social context on language development as purported in socio-cultural theory. In particular, the practice of code-switching in computer-mediated communication was used to support language development within the targeted research context. This study argues that, if users are to generate secure and usable passwords, the design of password policies should be informed by users' linguistic practices and the way the human memory operates.

### 3.1 Socio-technical theory

There are growing calls for adapting and using theories that cut across different research disciplines when addressing Information Systems problems (Shin, 2014). Socio-technical theory is one such theory that moves away from a traditionally narrow approach of focusing on technological subsystems with an assumption that the social subsystem will adapt to technical requirements (Durkin et al., 2015; Shin, 2014). Instead, socio-technical theory gives equal importance to both the social and technical subsystems. It argues that when an information system is deployed in an organisational

context, its success is socially constructed instead of being determined by technical capabilities (Doherty, 2014). This suggests that socio-technical theory can be used to inform the design of an information system that is compatible with its environment of use. This study used socio-technical theory to inform the design of usable and secure text-based password policies. The study advocates for what Shin (2014) termed a human-centred approach to designing usable and secure password policies. In other words, this study made use of a social lens to inform the design of usable and secure password policies (Durkin et al., 2015).

The United Kingdom's London-based Tavistock Institute introduced socio-technical theory towards the late 1950s, following a labour study finding that focusing on mechanisation alone may not translate to improved productivity (Durkin et al., 2015). Since then, interest in socio-technical theory has grown from the social sciences to other research disciplines including Information Systems (Eason, 2008 in Davis, Challenger, Jayewardene, & Clegg, 2014). The fact that socio-technical theory is a cross-disciplinary theory has led to a lack of consensus on an interdisciplinary definition of the theory (Wu, Fookes, Pitchforth, & Mengersen, 2015). Wu et al. (2015) suggest that elements of socio-technical theory subsystems are diverse and differ in relation to the problem domain. These differences can be easily traced in studies that adopted socio-technical theory. For instance, Durkin et al. (2015) suggest that the technical subsystem is composed of the "technology, machinery, processes, procedures and the physical environment", while the social subsystem is composed of "structure, people and their attitudes, behaviours and relationships" (p. 948). On the other hand, a study by Shin (2010) limited its view of socio-technical theory to what they termed technology issues, government, industry and social plus cultural issues. Further to that, Mumford proposed a system development methodology, which drew upon socio-technical theory principles, called ETHICS- "Effective Technical and Human Implementation of Computer-based Systems" (Sawyer & Jarrahi, 2013, p. 9). Wu et al. (2015) used a literature review in an attempt to bring together disparate views on elements and dimensions of socio-technical theory. They proposed a unified hierarchical structure in which they argued that socio-technical theory can be broken down into three subsystems, namely, social, technical and natural environment. Their study went on to

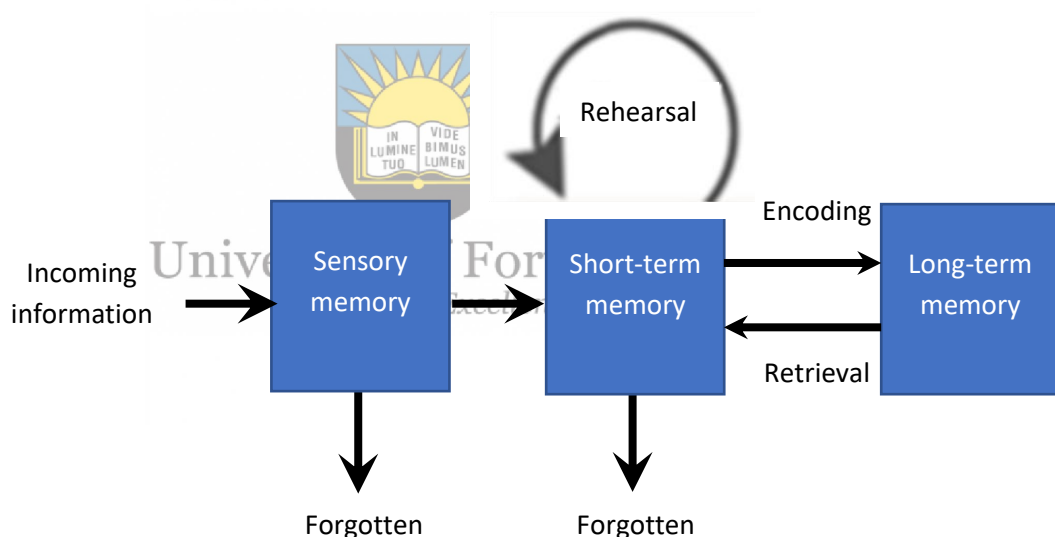


identify subsystems under each element in a hierarchy. This study conceded that the subsystems of socio-technical theory include the social and technical subsystems. A stakeholder analysis was used to guide the identification of elements that make up the social and technical subsystems as done by Lin, Paragas, Goh, and Bautista (2016). Stakeholder analysis allowed for the identification of individuals and groups that are affected by or affect the generation and use of passwords. In addition, stakeholder claims, resources used and stakeholder power, legitimacy and interests should be ascertained (Shin & Song, 2012).

The literature, with support from empirical evidence, suggests password users and system designers are major password stakeholders. It is well documented that users find it difficult to generate strong and usable passwords (Andersson & Saedén, 2013; Bonneau & Shutova, 2012; Keith et al., 2007). Furthermore, the literature on password generation suggests that human memory is the most important resource used by users when generating and recalling passwords (Zhang et al., 2009). It should be noted that socio-technical theory motivates a need to address Information System problems by understanding the context or by viewing them from a human-centred perspective (Shin, 2014). Accordingly, this chapter expounds the study's considerations of a social subsystem from a socio-cultural perspective focusing on language development and use, with the aim of understanding the context within which passwords are generated. The literature shows that user-generated passwords are linked to a user's context, courtesy of the spoken or written language(s) (Bonneau & Shutova, 2012; Rao et al., 2013; Veras et al., 2014). Hence, this chapter explains the human memory functionality and goes on to expound the use of socio-cultural theory in the study to explain human cognition development and use within a social context. It was expected that language learning and use within a context would help identify linguistic characteristics that could be exploited in order to improve the security and usability of passwords. Elements of the technical subsystem of socio-technical theory in this study are discussed in Chapter 4.

### 3.2 Human memory functionality

The use of passwords requires users to generate, learn, retain and recall passwords (Woods & Siponen, 2018, 2019). These activities rely on the memory which is responsible for the storage and retrieval of information that was acquired through the different senses. Therefore, information on how the memory operates can be used to understand how users can better generate, learn, retain and recall passwords. Atkinson and Shiffrin (1968) proposed a stages of memory theory that shows different components of the memory and how it works. This multi-store theory has been widely used in psychological studies and research on passwords (Al-Ameen, Wright, & Scielzo, 2015; Woods & Siponen, 2019; Zhang et al., 2009). According to Atkinson and Shiffrin (1968), the memory can be split into the sensory, short-term and long-term memory, as shown in Figure 8.



**Figure 8. The structure of the memory system (Atkinson & Shiffrin, 1968)**

Figure 8 illustrates that new information find its way into the memory through various senses. Information can easily decay, a natural process of forgetting that occurs over time, within milliseconds while in the sensory memory (Atkinson & Shiffrin, 1968). The sensory memory makes use of the short-term memory, also known as the working memory, to keep information memorable. However, the short-term memory can only hold new information for 30 seconds before it completely decays and is limited in capacity to  $7 \pm 2$  chunks of information (Atkinson & Shiffrin, 1968; Miller, 1956). The

short-term memory is critical because this is where cognitive activities relating to learning, thinking and problem-solving occur. Rehearsing new information is one of the methods for avoiding information decay in the short-term memory. Goldstein (2011, in Woods & Siponen, 2019) defines rehearsal as a process of repeating the same “information over and over” again or actively maintaining the information through repeated training and use (p. 63). In addition, newly generated information with cues that can easily be linked to the existing information in the long-term memory has a greater chance of being memorable (Al-Ameen, Wright et al., 2015; Baddeley, 2009b, in Woods & Siponen, 2019; Zhang et al., 2009). However, besides forgetting new information through decay, it is also possible that the newly acquired information interferes with the already existing information or vice-versa, thereby posing further memorability challenges to users (Zhang et al., 2009). According to Atkinson and Shiffrin (1968), the longer information stays in the short-term memory the more likely it will be copied or transferred to the long-term memory where permanent and memorable information resides waiting to be retrieved.



When Atkinson and Shiffrin's (1968) theory is applied to passwords, it implies that a new password will find its way into the memory through the sensory memory and will be learnt in the short-term memory before it is copied or transferred to the long-term memory where it will be stored permanently while waiting for retrieval during password recall (Woods & Siponen, 2019). However, complex passwords with more than seven characters or words that appear to be random to a user may not last for more than 30 seconds in a user's memory. This possibly explains why users struggle to memorise system-assigned passwords (Shay et al., 2012). The next section discusses theoretical views on forgetting/decay and the capacity of the short-term memory. The focus is on exploring ways of promoting password usability by overcoming password decay and the challenges associated with the limited capacity of the short-term memory.

### 3.2.1 Memory decay and password usability

Information in the short-term memory either ends up in the long-term memory or decays given the limited amount of time it can remain in the short-term memory (Cowan, 2014; Everitt, Bragin, Fogarty, & Kohno, 2009; Zhang et al., 2009), as highlighted in the previous section. Factors contributing to decay are interference and a lack of cues between new information and the information that resides in the long-term memory (Everitt et al., 2009; Zhang et al., 2009). The theory of interference posits that newly acquired information can interfere with the already existing near-similar information, thereby hampering information recall efforts (España, 2016; Zhang et al., 2009). España (2016) goes on to state that interference is one of the reasons why users avoid new password creation. The literature on passwords is awash with reports of password reuse (Bang, Lee, Bae, & Ahn, 2012; Hayashi & Hong, 2011; Helkala & Bakås, 2013; Stobert & Biddle, 2014; von Zezschwitz et al., 2013), something that clearly demonstrates the resistance of users to generating new passwords.

The literature on passwords proposes rehearsal or repeated use of the newly generated passwords in order to prolong their stay in the short-term memory in the hope that they will eventually be transferred to the long-term memory and avoid interference (Bonneau & Schechter, 2014; Woods & Siponen, 2019). A literature review by Woods and Siponen (2019) reports that mass or distributed repeated use of the password after generation improves memorability. These authors further demonstrate that increasing the number of password verifications improves memorability (Woods & Siponen, 2019). In addition, study findings by Bonneau and Schechter (2014) led them to the conclusion that dispersed password rehearsal or training improves password memorability. These findings are consistent with the views in the multi-store memory model proposed by Atkinson and Shiffrin (1968).

However, Simon (1974, in Zhang et al., 2009) argues that “information is not processed in single strands or discrete entities but as ‘chunks’ of similar or equivalent data” (p. 4). Zhang et al. (2009) went on to show that password interference remains a challenge in password recall when users are required to recall many passwords. Everitt et al. (2009) made a similar finding; for example, they observed that repeated use

helped participants to recall a single graphical password. They also noted that participants with four graphical passwords suffered at least ten times as many password recall failures owing to interference when compared to those with a single password (Everitt et al., 2009). These findings from Everitt et al.'s (2009) study are corroborated by study findings in Woods (2017). Woods (2017) investigated the effect of repeated use and the frequency of password recall. Ten passwords per participant were used. The results show that there was no relationship between the repeated use of passwords and memorability (Woods, 2017). Furthermore, the frequency of password recall did not seem to influence memorability. These findings call for further research on memorability and rehearsal of multiple passwords.

There are indications that interference is more pronounced when there is limited long-term knowledge about the subject (Cowan, 2014). This suggests that generating new information based on already existing information may reduce interference. Atkinson and Shiffrin (1968) state that the process of searching and retrieving information from the long-term memory depends on the availability of traces of information. “When the trace is strong and complete, related information is expected to be easier to locate and recall” (Zhang et al., 2009, p. 5). Thus, to better retain new information by avoiding interference, one could consider using the already existing meaningful cues in the long-term memory (Al-Ameen, Wright et al., 2015). Similarly, to enhance memorability, this study attempts to follow an approach that makes use of the already known information in password generation.

### **3.2.2 Short-term memory capacity and password usability**

Cognitive load theory can be used to explain how short-term memory capacity affects memorability (Paas & Ayres, 2014; Woods & Siponen, 2019). Cognitive load theory outlines the relationship between the effort needed to process information vis-à-vis the amount of information to be processed (Woods & Siponen, 2019). It has already been stated that the short-term memory is limited to  $7 \pm 2$  elements of information. Hence, as more information is loaded into the short-term memory for processing, so the effort to process the information increases. Inglesant and Sasse (2010) and Woods and Siponen (2019) state that users generate passwords under the

distraction of work-related tasks, password policies or people in the background, which negatively affects password learning. However Paas and Ayres (2014) are of the opinion that one can overcome the challenges resulting from limited short-term memory capacity by using information that is already stored in the long-term memory. Thus, handling complex tasks that exceed the capacity of the short-term memory can easily be achieved by making reference to information in the long-term memory. Chunking theory, which was developed by Miller (1956), postulates that meaningful information in the long-term memory can be grouped together into new information that is easy to memorise. Thus, “highly meaningful words are easier for a person to learn and remember than less meaningful words, with meaningful being defined by the person’s number of associations with the word” (Newell, Shaw, & Simon, 1961 in Carstens, Malone, & McCauley-Bell, 2006, p. 100). For instance, a passphrase is a typical example of chunks of information that were grouped together to formulate a single meaningful and memorable phrase (Keith et al., 2007; Woods & Siponen, 2019). Keith et al. (2007) found that passphrases are easy to remember even though they initially expose users to many typographical errors. Several studies have since shown that passphrases are more secure and usable than short passwords (Melicher et al., 2016; Shay et al., 2016). This study likewise makes use of the propositions in chunking theory, thus recommending the use of multilingual passphrases.

### **3.3 Socio-cultural theory**

Lev Semyonovich Vygotsky, a Russian psychologist, is widely credited for laying out the theoretical framework of socio-cultural theory during the 1930s (Mercer & Howe, 2012; Zuengler & Miller, 2006). Although the intellectual roots of socio-cultural theory date back to the 18th and 19th century (Lanolf, Thorne, & Poehner, 2015), it was not until the mid-1990s that this theory became popular as a theory for explaining psychological development within a social context. The growth in popularity of the theory gave an alternative view to the dominant cognitive views for explaining learning. The cognitive views argued that learning is individualistic, hence it is a mental process that occurs internally, independent of the social context (Zuengler & Miller, 2006). Socio-cultural theory, on the other hand, proposed that an individual’s mental functioning is related to participation in contextual social interactions (Scott & Palincsar,

2013). Thus, socio-cultural theory explains the relationship that exists between mental functioning and one's cultural and institutional setting. Although socio-cultural theory and cognitive views have different approaches to explaining and understanding psychological development, these theorists all aim to study human mental development and functionality. This is important to this study as it could explain a person's use of language(s) and linguistic attributes which could be exploited by password policy designers.

Vygotsky's grounding of socio-cultural theory drew inspiration from three principles of Marxist theory which advance the notion that

- “human consciousness is fundamentally social, rather than merely biological, in origin”
- human mental activity “is mediated by material artefacts – psychological and symbolic tools/signs” such as language, in order to organise or understand the world
- the understanding of human development “should be holistic in nature” (Lantolf et al., 2015, p. 2).



University of Fort Hare  
*Together in Excellence*

Socio-cultural theory further proposes three principles that explain psychological development in line with the principles of Marxist theory. These principles include the genetic law of development, mediation and genetic domains. Socio-cultural theory and its principles have received wide use in explaining high-level mental activities such as language learning, whose occurrence is, according to the theory, socially constructed instead of being biologically constructed alone (Lantolf, 2000; Lantolf et al., 2015; Mercer & Howe, 2012; Zuengler & Miller, 2006). Other high-level mental activities include voluntary attention, intentional memory, logical thinking and problem solving. Furthermore, socio-cultural theory has been found to be useful in explaining human mental functionality in a global environment (Marginson & Dang, 2017). Accordingly, this study concedes that password generation and use are higher mental activities that involve voluntary attention, intentional memory, logical thinking and problem solving, as purported by Vygotsky (Lantolf et al., 2015; Zhang et al., 2009). Hence, password generation and use need to be understood from a sociological perspective and help

inform the design of password policies that promote the generation of usable and secure passwords. The next section discusses the three principles of socio-cultural theory and goes on to suggest certain implications of these principles for user-generated passwords.

### **3.3.1 The generic law of development.**

Vygotsky argues that human psychological development is socially constructed. An individual's setting, as determined by culture, language, history, peer groups and institutional structures at school or in the workplace, plays a critical role in shaping the initial human mental development (Lantolf et al., 2015). Vygotsky points out that the "human psychological process does not pre-exist inside the head waiting to emerge at just the right maturational moment" (Lantolf, 2000, p. 14) and that mental development is not an inborn capacity that unfolds with time (Zuengler & Miller, 2006). Instead, human psychological development occurs across two levels; first at a social level as one interacts with those in one's social environment and then at an individual level. For instance, an individual first learns a language by receiving instructions from guardians. With time, the learner becomes acquainted with the subjects to the extent of regulating the subjects with own mental functions; this speaks to an activity that describes internal human development (individual level). Once the subjects are internalised, they become available as a cognitive resource.

The propositions in the generic law of development imply that the social environment in which a computer user resides has an influence on the password a user is likely to generate. An English-speaking individual is expected to generate an English language-oriented password. For instance, Voyiatzis, Fidas, Serpanos, and Avouris (2011) observed the use of native Greek language-oriented passwords by Greek computer users. Similarly, Bonneau and Xu (2012) observed passwords that reflected contexts of different computer users as reflected by different language orientations, including Hebrew, Spanish and Chinese. Another study by AlSabah et al. (2018) showed that even if passwords are generated following a similar password policy, the resulting passwords will still differ according to the culture of the user.



### 3.3.2 Mediation

Through his socio-cultural theory, Vygotsky reasoned that human mental functioning is mediated by cultural artefacts (tools) as humans move through social and individual levels of psychological development. “Just as physical tools serve as auxiliary means to enhance the ability to control and change the physical world, symbolic tools serve as an auxiliary means to control and reorganise our biologically endowed mental processes” (Lantolf et al., 2015, p. 5). Thus, physical tools are externally oriented and assist one to master one’s nature, while symbolic tools are internally oriented tools that promote mastery of oneself (Vygotsky, 1978, p. 55, in Marginson & Dang, 2017). Tools that are seen as cultural artefacts include language, diagrams, maps, mnemonic techniques, computers, calculators, paint brushes and so on (Vygotsky, 1981, in Scott & Palincsar, 2013). Language is considered a powerful symbolic tool that allows individuals to mediate social interactions and regulate the conducting of cognitive activities such as thinking and problem solving. As a result, the human mind is seen as a functional system that is culturally shaped (Lantolf, 2000).

Within the context of this study, the mediational domain suggests that computer users prefer contextually developed symbolic tools for generating passwords. For example, an analysis of more than 100 million publicly leaked English and Chinese passwords shows that close to 50% of Chinese passwords are purely digit-based when compared to English passwords that are mainly a concatenation of English words and digits or words in the English dictionary (25.88%) (Bonneau & Xu, 2012; Li et al., 2014; Wang et al., 2015; Yang et al., 2013). Furthermore, the symbolic tools preferred by the Chinese when generating passwords were found to be Pinyin names (one in every ninth password), while English users preferred to adapt an English name in their passwords (one in four passwords). Another interesting finding that could be related to the influence of the mediational domain is the use of digits in passwords that portray the pronunciation of phrases in Mandarin Chinese, such as 5201314 which translates to “I love you forever” (Li et al., 2014; Wang et al., 2015; Yang et al., 2013). All these findings reflect on the preferred contextual symbolic tools in password generation as argued by the mediational principle.

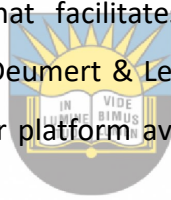
### 3.3.3 Generic domains

Socio-cultural theory substantiates the notion that higher mental functionality is always in motion and goes through continuous changes (Marginson & Dang, 2017). Vygotsky argued that a generation inherits cultural artefacts from previous generations and acts on the artefacts, resulting in modifications before passing on the artefacts to the next generation. Accordingly, such a phenomenon needs to be understood from a historic point of view. Vygotsky's socio-cultural theory advanced four generic domains from which the study of human mental development can be done. These include the phylogenetic, sociocultural, ontogenetic and microgenetic domain (Lantolf, 2000). The phylogenetic domain focuses on the natural biological development of a human as a specie, while the sociocultural domain looks at the changes occurring in the social environment where human species reside (Lantolf, 2000; Marginson & Dang, 2017). These include changes in the symbolic and physical tools that mediate human mental development. Furthermore, the ontogenetic domain focuses on how individuals acquire mediational tools, like language, as they move from being a novice in the subject to an expert. Cross (2006, in Marginson & Dang, 2017) suggests that the microgenetic domain focuses on "momentary fragments of development" (p. 119), such as learning a word or the grammatical features of a language or being trained to perform a task in an experiment (Lantolf, 2000).

When applied within the context of this study, the generic domain implies that user-generated passwords are likely to evolve with time. This may result from password reuse, as users adapt to different password policies. For instance, a longitudinal study by Von Zezschwitz et al. (2013) found that user passwords evolved over time as a result of changes in password security requirements. Such evolutions involved minor changes to existing passwords as users adapted old passwords in an attempt to comply with password requirements without compromising memorability. In addition, Jakobsson and Dhiman (2013) observed that users can evolve their passwords by making spelling mistakes, insertions, concatenating different character classes and replacing different character classes.

### **3.4 An overview of the social context and language development**

The previous sections discussed the occurrence of human mental development from a socio-cultural theoretical perspective and looked at ways in which the human memory operates. Socio-cultural theory advances the notion that cultural symbolic tools, like language, play a critical role in reorganising the biologically endowed mental processes. A review of the principles in socio-cultural theory in Section 3.3 showed how local languages influence the characteristics of user-generated passwords. It is therefore argued that if passwords are to be usable, password policies should encourage password generation and user requirements that align to that which a human mind is moulded to comprehend or understand. The next section explores the language landscape of the study research area. The aim is to establish the characteristics of the research context in terms of language use. Studies on text messages were used to give a picture of the extent of language understanding. The mobile phone technology that facilitates text messages has out-diffused prior technologies (Kalba, 2008 in Deumert & Lexander, 2013); hence, text messages have the potential to make a wider platform available for sourcing reference information relating to language use.



University of Fort Hare  
*Together in Excellence*

Morel, Bucher, Doehler, and Siebenhaar (2012) suggest that the practice of code-switching in communication demonstrates one's linguistic skills. Accordingly, Section 3.4.2 reflects on text message code-switching with the primary aim of explaining socio-contextual influences on mental development as delineated in socio-cultural theory. This study aimed at exploiting people's understanding of language(s) and used this to inform the design of password policies that promote the generation of usable and secure passwords.

#### **3.4.1 Africa's language landscape**

Socio-cultural theory argues that, cognition and/or human mental development is initially a feature of social contextual factors, such as the spoken or written language. Accordingly, this section analyses Africa's language landscape in order to anticipate its potential influence on text message code-switching practices, something that is argued to be a reflection of the would-be user-generated passwords.

In terms of spoken and written language, Africa is regarded as a multilingual society in which individuals speak and write at least two different languages. Deumert and Lexander (2013) sum up Africa's language landscape by stating that "in many African countries being 'literate', that is, educated, continues to refer not simply to the ability to read and write, but also to be able to do this in English, French or Portuguese (as well as Spanish in Equatorial Guinea)" (p. 525). The existence of different cultural tribes and colonisation are often considered to be major factors promoting multilingualism in Africa. In particular, the official languages of most African countries are those of the former colonial masters (Lexander, 2011). For instance, Namibia has multiple ethnic groups that speak more than eleven indigenous languages and three Indo-European languages (Peters, Winschiers-Theophilus, & Mennecke, 2015). English is Namibia's official language and the language of instruction starting from secondary education (Peters et al., 2015). The same applies to Senegal, a country with twenty-five to thirty recognised indigenous languages where French is the official language and language of instruction in public schools (Lexander, 2011). Likewise, other countries like Botswana, Côte d'Ivoire, Ghana, Nigeria and South Africa, just to name a few, portray a similar language landscape (Deumert & Lexander, 2013; Dyers & Davids, 2015; Ndlovu, 2016). As observed across Africa, quite often the language of instruction at learning institutions differs from the home language (Deumert & Lexander, 2013; Dyers & Davids, 2015; Lexander, 2011; Ndlovu, 2016). Thus, even though the orthography of the first spoken indigenous African language exists, these local languages are not always learnt at school and not used as the language of instruction, leaving one exposed to a multilingual social environment (Deumert & Lexander, 2013; Lexander, 2011).

In addition to the different languages that characterise Africa, there are numerous dialects of these languages. Some of the dialects do not have documented or known orthography (Deumert & Lexander, 2013; Lexander, 2011). It is therefore argued that human cognitive development in such a socio-cultural environment has great potential for promoting bilingualism or multilingualism, something that creates a platform for numerous text message variations that can contribute to uniqueness. It is believed that such linguistic characteristics could be exploited to come up with a password policy that maximises security and usability. The next section uses an

overview of findings on text message code-switching reported in the literature as evidence that demonstrates the social contextual factors in human cognitive development as explained by socio-cultural theory.

### **3.4.2 The practice of code switching in text messages**

The literature suggests that text message variations are diverse to such an extent that the definition of code-switching that limits the practice to juxtaposing at least two different languages (codes) in a message may not be adequate (Morel et al., 2012). In addition, Morel et al. (2012) note that if language is the only determinant of a code, words like “chaos” or “version” with a similar spelling in English, German and French would be difficult to categorise when used in written code-switching. Hence, this study concedes that code-switching variations in text messages are diverse and can also be determined by “speech style, specific vocabulary use and graphical cues” (Morel et al., 2012, p. 265). Accordingly, this study rather adopted Tagg, Baron, and Rayson's (2012) three factors that motivate text message variations, namely, functionality, principle and meaning. These factors are used to guide the outlay of evidence on text message code-switching practices noted in the literature. The factors influencing text message variations proposed by Tagg et al. (2012) are in sync with the principles of this study as they also take into account sociocultural influences on text message code-switching practices.

#### **3.4.2.1 Functional code-switching**

Code-switching in text messages is functional in that text variations are often a result of functional demands during an interaction, such as the need to make a quick response or the need to produce an understandable message. To some extent, functional code-switching can be explained by Thurlow and Brown's (2003) proposition that users practise code-switching in order to attain three maxims, namely, brevity and speed, paralinguistic restitution and phonological approximation. For instance, the need for brevity and speed would see users making use of abbreviations and letter-number homophones and less use of punctuation and spaces (Thurlow & Brown, 2003). Paralinguistic restitutions will be used to express emotions and emphasis. On the other

hand, phonological approximations would be used to emulate informal speech, for instance.

Studies from different contexts reflect the role of functional code-switching and the need to attain Thurlow and Brown's (2003) maxims. Keong, Gill, Noorezam, and Abdulrazaq (2012) observed that Malay university students used lexical reductions that reflected brevity and speed as well as paralinguistic restitution in their English text messages. Similarly, Deumert and Masinyana (2008) observed that isiXhosa-English speaking South Africans demonstrated code-switching practices in their English text messages that included the use of abbreviations, non-standard spellings, paralinguistic restitutions and phonological approximations. Furthermore, some text messages were written following Afro-American vernacular. Another study by Lexander (2011) showed that Senegalese are comfortable abbreviating and using creative spelling in French text messages. It is argued that these text message code-switching practices are a reflection of the underlying social context.



However, of particular interest is the finding that there is little evidence to support the fact that people aim to attain Thurlow and Brown's (2003) maxims when writing text messages using African languages (Deumert & Lexander, 2013; Deumert & Masinyana, 2008; Dyers & Davids, 2015; Lexander, 2011; Ndlovu, 2016). Lexander (2011) noted that most text messages written in native Senegalese languages were expressed following the standard language orthography. Deumert and Masinyana (2008) made similar findings that text messages written in isiXhosa were written following the standard orthography. Similar findings were made in other African countries such as Nigeria, Ghana and Côte d'Ivoire (Deumert & Lexander, 2013). Hence, Thurlow and Brown's (2003) maxims have to be applied with caution as they are not universal.

In addition, functional code-switching variations may result in the use of more than one language base, something that reflects contextual and social dimensions in an interaction (Morel et al., 2012). For example, Lexander (2011) noted text messages with two language bases, namely, French and Wolof in his corpus. This reflected the

immediate social context where Wolof is the dominant spoken language (spoken by 80–90% of the population) in Senegal, with French spoken by a minority but being their first written language (Lexander, 2011). Deumert and Masinyana (2008) made a similar finding of text messages that included two language bases, namely, English and isiXhosa. This finding shows the influence of the social context where English is the dominant language of instruction and first written language in literacy, while isiXhosa is a spoken language (Deumert & Lexander, 2013; Deumert & Masinyana, 2008; Lexander, 2011). Therefore, findings from these studies confirm principles of socio-cultural theory on human cognition development within a context.

#### **3.4.2.2 Principled code-switching**

Text message code switching is principled in the sense that the choices of spelling variations are informed by the orthographic principles of the language. This implies that the pattern of spelling variations for a particular language might be constant across different contexts. A literature review by Tagg et al. (2012) led to the observation that some spelling variations on the English language by isiXhosa-speaking people were comparable to those of US English speakers. Similarly, Chiluba (2008) observed common worldwide English language spelling variations in the Nigerian text message corpus. It can be argued that this reflects the generic domain principle of socio-cultural theory, where symbolic tools constantly change.

#### **3.4.2.3 Meaningful code-switching**

Tagg et al. (2012) suggest that some spelling variations are socially and contextually constructed as one attempt to portray oneself in a meaningful way. For instance, one might practise code-switching with the aim of reflecting a fun person or in an attempt to express a casual conversation. In addition, some spelling variations can reflect the society or social group to which one belongs. Text message corpora of studies conducted in Nigeria, Senegal and South Africa show a corroborative finding that text messages written in English have spelling variations that are unique to particular social contexts, thereby making meaningful code-switching in particular contexts (Chiluba, 2008; Deumert & Lexander, 2013; Deumert & Masinyana, 2008; Lexander, 2011). Based on these research findings in the literature, it is argued that

human cognitive development in a multilingual environment gives one the potential to practise meaningful code-switching that is reflective of the social environment as suggested by socio-cultural theory.

### **3.5 Chapter summary**

This chapter motivated a need to address the password security and usability predicament from a socio-technical view. The chapter started with an overview of the functionality of the memory. It was found that the memory is limited in capacity and can hold new information for a very short period of time only. Accordingly, this study sought to exploit information or knowledge in the long-term memory. The contextual factors as determined by socio-cultural theory were discussed. It is argued that the understanding of language development and its use is of paramount importance to the generation of secure and usable passwords because users' choice of passwords is often related to the words in a natural language one speaks (AlSabah et al., 2018; Bonneau & Shutova, 2012; Rao et al., 2013; Veras et al., 2014). In particular, this study proposes to extend the practice of text message code-switching to the generation of usable and secure passwords. It was noted that some of the users' practices of code-switching could lead to unique messages that could only be traced back to particular social contexts. Hence, the study looked at exploiting such code-switching practices in generating usable and secure passwords.

This chapter looked at the socio subsystem. The next chapter goes on to explore the technical subsystem. It is argued that, together, the socio and technical subsystems could best inform the design of secure and usable password policies.



## **CHAPTER 4: PASSWORD THREATS AND POLICIES IN USE**

### **4.0 Introduction**

The aim of this study was to develop a model of secure and usable multilingual passphrases. To assist in achieving this aim socio-technical theory was adopted. Accordingly, the previous chapter discussed the study perspective on the socio subsystem as envisaged by socio-technical theory, while socio-cultural theory was used to present the view of this study on the socio subsystem. This chapter contributes to the study by discussing constructs in the technical subsystem. Thus, the overall aim of the chapter is to provide an understanding of the security and usability concerns associated with the available password authentication designs. To meet the objectives of this chapter, an analysis of authentication mechanisms is done to establish the position of text-based authentications. The chapter goes on to discuss password threats, subsequently suggesting a commensurate password strength measure. This is followed by an overview of international password guidelines and best practices for enhancing password strength and usability. Password generation policies are also reviewed focusing on their security and usability contributions. By so doing, this chapter creates a basis for a model for generating secure and usable passphrases, which will be proposed in Chapter 5.

### **4.1 Categories of authentication mechanisms**


Authentication mechanisms can be broadly categorised into knowledge-based (what one knows), token-based (what one has) and biometric-based (what you are) mechanisms (ISACA, 2015). Knowledge-based authentication mechanisms include textual passwords, personal identification numbers, passphrases and graphical passwords (Taneski, Heričko, & Brumen, 2014). Token-based authentication makes use of physical tokens that carry identification details, for example a bank credit card. Biometric-based authentication makes use of unique biometric traits such as finger prints, faces and palm prints. Another biometric authentication mechanism is behavioural biometrics, which is based on one's unique characteristics or behaviour when performing certain activities. Common behavioural-biometrics include keystroke dynamics (Ciampa, 2013). All these authentication mechanisms afford users different

security and usability facilities. As such, they can be used in any combination to come up with multifactor authentication.

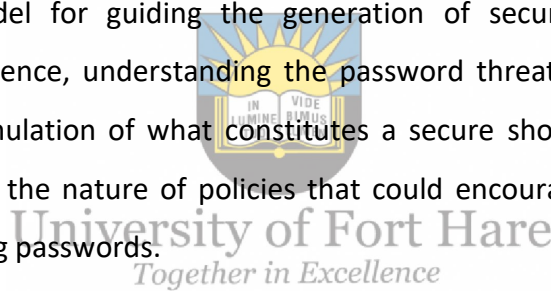
Of interest to this study is the knowledge-based authentication mechanism (text password) that remains a dominant access control mechanism, despite inherent limitations that were noted from as far back as the late 1970s (Dell'Amico, Michiardi, & Roudier, 2010; Li et al., 2014; Wang, Cheng, et al., 2015). Passwords were the first established authentication mechanism and are expected to remain dominant for the foreseeable future (Dell'Amico et al., 2010; Mazurek et al., 2013; Wang, Cheng, et al., 2015; Weber et al., 2008). From a system implementer's point of view, passwords are regarded as a low cost, easy to implement authentication mechanism.

#### **4.2 Password threats and attacks**

The literature identifies a wide range of attacks that could exploit weaknesses in text-based authentications (Campbell et al., 2011; Dell'Amico et al., 2010; Grassi et al., 2017; Harris & Maymí, 2019; Vaithyasubramanian & Christy, 2015). These threats can be broadly grouped into

- 
- University of Fort Hare  
*Together in Excellence*
- non-technical password attacks that include dumpster diving, over the shoulder surfing and social engineering
  - passive online password attacks that include wire sniffing, man in the middle and replay attacks
  - online password attacks such as brute force and password guessing attack
  - offline password attacks that rely on having access to a password file and guessing the passwords it contains. Passwords in a password file are usually stored in encrypted or hashed format that can be cracked using password guessing algorithms (Campbell et al., 2011; Dell'Amico et al., 2010; Grassi et al., 2017; Harris & Maymí, 2019; ISACA, 2015; Vaithyasubramanian & Christy, 2015).

There is a growing interest in researching online and offline password attacks following leaks of password databases that exposed millions of hashed passwords from popular sites, such as CSDN, Tianya, Duduniu, 7k7k, 178.com, RockYou, Facebook, Adobe, Middle East Bank, the South African traffic fine platform and Yahoo (AlSabah et al., 2018; Dell’Amico et al., 2010; Florêncio, Herley, & van Oorschot, 2014; Houshmand & Aggarwal, 2012; IOL, 2018; Shay et al., 2016; Wheeler, 2016). While online password attacks can be mitigated by setting a limit to login attempts, it is the offline password attacks that pose a significant threat given that the attacker has an unlimited number of password guessing attempts. Improved computer hardware and software means that one can make more than 350 billion password guesses in a matter of hours (Ciampa, 2013; Weir et al., 2010). This study is limited to online and offline password attacks that could be mitigated by generating complex passwords (Dell’Amico et al., 2010; Grassi et al., 2017). This is consistent with the focus of this study that seeks to propose a model for guiding the generation of secure and usable multilingual passphrases. Hence, understanding the password threats considered for this study guides the formulation of what constitutes a secure short password or multilingual passphrase and the nature of policies that could encourage users to generating the perceived strong passwords.



#### **4.2.1 Online password attack**

An online password attack involves an attacker repeatedly trying candidate passwords on a live system. Attackers using this approach often try popular passwords first or the victim’s personal information (Florêncio et al., 2014b). Such an attack can assume a brute force attack. A brute force attack tries every password combination until a match is found (Weir et al., 2009). Creating a strong password, setting a limit to log-in attempts and delaying future log-in attempts after every failed log-in attempt are some of the measures for curbing online password attacks (Dell’Amico et al., 2010; Florêncio et al., 2014; Ur et al., 2016). However, recent evidence shows that moving away from trawling password attacks towards targeted password attacks using personal information gives password attack perpetrators more leverage for overcoming measures against online password attacks (Wang & Wang, 2015; Wang, Zhang, Wang, Yan, & Huang, 2016).

#### 4.2.2 Offline password attack

An offline password attack occurs when a perpetrator undetectably gains access to an encrypted password file (Florêncio et al., 2014b). In an offline attack, the attacker has an unlimited number of attempts at guessing passwords. Given the right hardware and software, one can make more than 350 billion password guesses in a matter of hours (Ciampa, 2013; Weir et al., 2010). With more than 100 million passwords leaked to the public over the past ten years (Li et al., 2014), offline password attacks that are difficult to defend have become one of the most researched threats (Ciampa, 2013; Weir et al., 2010). Offline password attacks make use of a guessing attack, a technique that was used first by Morris and Thompson in their 1979 publication (Wheeler, 2016). While Morris and Thompson used a dictionary attack for password guessing, password guessing attacks have since evolved. For example, the introduction of John the Ripper in the 1990s and the recent probability-based guessing attacks, namely, the PCFG and Markov-chain, and non-probability based guessing attacks, such as the zxcvbn (Weir et al., 2009; Wheeler, 2016). A number of other probability-based password guessing attack varieties have since been proposed, inspired by ideas in the PCFG or the Markov-chain with the aims of improving efficiency and effectiveness in password guessing (Kelley et al., 2012; Komanduri, 2016; Shay et al., 2016; Wang & Wang, 2015). These password guessing attacks are discussed next so that those posing a great threat to passwords can be enlisted. This will help to inform the anticipated magnitude of password strength.

**4.2.2.1 John the Ripper (JTR).** The literature suggests Morris and Thompson's use of a dictionary attack in password guessing inspired the development of JTR (Bailey, Dürmuth, & Paar, 2014; Ur et al., 2015). JTR uses a password dictionary or wordlist (which can be a combination of a natural language dictionary and a publicly available password set) to learn different mangling rules to be followed when generating candidate passwords during a password guessing attack (Ur et al., 2015). Mangling rules exploit common vulnerabilities found in the password dictionary such as replacing an "a" with "@" or replacing "i" with "1". Hence, mangling involves character modifications and appending number(s) to a subtext, something that could generate new candidate passwords out of the passwords in the dictionary (Duermuth et al.,

2015). JTR is comparable to Hashcat, another password guessing attack technique (Ur et al., 2015). The only difference is, when referencing multiple password dictionaries to generate candidate passwords during password guessing, JTR uses a single password mangling rule at a time on each dictionary before moving to the next rule. On the other hand, Hashcat iterates all the mangling rules on a password dictionary before moving to the next dictionary (Ur et al., 2015).

There are different versions of JTR, each with unique effectiveness and efficiency in password guessing. In addition, JTR can be implemented in different modes with each mode having specific mangling rules applicable to it (Bailey et al., 2014). These modes include the following:

**(i) Wordlist mode.** Also known as the dictionary mode. A password dictionary is fed into JTR such that passwords in the dictionary will be used as candidate passwords when guessing particular passwords. During password guessing, different mangling rules are used, depending on what was learnt from the password dictionary.

**(ii) Incremental mode.** The incremental mode makes use of the Markov model in a guessing attack where all possible combinations for generating candidate passwords are attempted (Dürmuth et al., 2015). Statistical frequencies learnt from the password dictionary are used to prioritise different possible combinations (Ji et al., 2017).

Despite the different versions, empirical evidence suggests JTR performs best when cracking short passwords (less than nine characters), even though it may not always outperform PCFG password guessing technique (Dürmuth, Chaabane, Perito, & Castelluccia, 2013; Ji et al., 2017). However, different JTR and Hashcat configurations may yield password guessing results that “are frequently comparable to, and sometimes even more effective than the probabilistic approaches” (Ur et al., 2015, p. 464). The password dictionary size and its closeness to the password policy whose passwords are being guessed are some of the critical success factors of JTR (Dürmuth et al., 2013).

**4.2.2.2 Probabilistic Context-Free Grammar (PCFG).** Weir et al. (2009) used context-free grammar to develop a probabilistic based algorithm for password guessing. Their algorithm is based on identifying password structures which they argue have different probabilities of occurrence. The identified password structures are arranged in descending order of occurrence such that the most common passwords structures are tested first (Li et al., 2016; Weir et al., 2009). The PCFG makes use of a password dictionary to learn common password structures that are denoted by an “S” for symbol, “D” for digit, “L” for letter and associated probabilities. Thus, the password structure is based on the character class tokenisation or tagging. In a way, these structures illustrate various mangling rules practised by users during password generation (Wang, Cheng, et al., 2015).

The PCFG has proven effective in password guessing even when applied to both real world and experimentally generated passwords, following different password policies based on different languages (Dell’Amico et al., 2010; Houshmand & Aggarwal, 2012; Ji et al., 2017; Li et al., 2014; Wang, Cheng et al., 2015; Wang & Wang, 2015; Ur et al., 2015). In addition, Dell’Amico et al. (2010) found PCFG more effective in password guessing compared to JTR. By contrast, Ji et al. (2017) and Ur et al. (2015) concluded that the success of an algorithm in password guessing is subject to configuration settings. They, however, conceded that PCFG is stable and improves its guessability with more training data, even when compared to best algorithms like the Markov chain (Ji et al., 2017; Ur et al., 2015).

Further to that, since its inception in 2009, PCFG has had numerous extensions and modifications to exploit different semantic structures followed by users when generating passwords (Houshmand, Aggarwal, & Flood, 2015; Komanduri, 2016; Li et al., 2016; Rao, Jha, & Kini, 2013; Shay et al., 2012; Veras et al., 2014; Wang et al., 2016). For example, Li et al. (2016) modified the original PCFG and proposed Personal-PCFG, an algorithm that could study and guess passwords whose structures are derived from personal information such as name, date of birth, identification number, mobile phone number and email address. When applied to a Chinese dataset, their Personal-PCFG improved password guessing by up to between 309% and 634% more than the original

PCFG (Li et al., 2016). Wang et al. (2016) also extended PCFG, proposing a personal identification information-type guess attack (TarGuess). Their study found that password guessing could be improved by 20.26% when the attacker is in possession of a victim's personal information that includes an email address, account name, name, birthday, phone number and national identification number. The literature suggests the effectiveness of password guessing is mainly attributed to Chinese users' high use of personal information in their passwords with 60.1% of passwords containing personal information (Li et al., 2016; Wang et al., 2016).

Houshmand et al. (2015) defined different keyboard patterns and went on to modify the original PCFG in such a way that it could learn and tag different keyboard patterns. By focusing on keyboard patterns, their algorithm could guess 22% more passwords than the original PCFG. Further to that, Veras et al.'s (2014) semantically informed PCFG guessed 67% more LinkedIn passwords than the original PCFG. Their algorithm also cracked 32% more Myspace passwords than the original PCFG. When the context-free grammar technique is used to tag long passwords based on grammatical structures (Parts of Speech), results show that efficiency in password guessing is improved to 18.7% when compared to the original PCFG (4.8%) (Rao et al., 2013). A study by Komanduri (2016) also modified the original PCFG and proposed a hybrid PCFG that could guess long passwords. He shifted from character class tokenisation to n-gram tokenisation (word-based tagging), a move that subsequently improved guessing of passphrases. These findings suggest the versatility of the PCFG and its capability to exploit various user password generation behaviours.

**4.2.2.3 Markov chain process.** Narayanan and Shmatikov (2005) used the Markov model to develop a password cracking algorithm. They observed that a user's choice of password characters is not randomly distributed. Instead, the distribution of password characters selected when generating passwords is similar to that of the character distribution in a user's language, something that reduces the search space (Narayanan & Shmatikov, 2005). Accordingly, their proposed Markov chain makes use of wordlist(s) and aa password dictionary to learn letter distribution and transition probabilities (Dell'Amico et al., 2010; Dürmuth et al., 2013). When guessing passwords, the learnt



probability distribution determines password length (Dell’Amico et al., 2010). In guessing a short password, characters of the password are guessed one after the other, according to transition probability. Thus, Markov chains use details of the predecessor character to guess the most likely character to come next, depending on the transition probability learnt from analysing wordlists and password dictionaries.

The Markov chain has since experienced different modifications with the aim of exploiting users’ password generation behaviours. For example, Kelley et al. (2012) adopted Markov chain and modified it to improve its efficiency in password guessing. Wang and Wang (2015) recently modified the determination of weights on the probability of “name-related letter segments” to enhance the effectiveness of the Markov chain algorithm when guessing passwords based on a user’s names (p. 7). Similarly, Dürmuth et al. (2013) demonstrated that the effectiveness of Markov chains can be increased by 5% when the algorithm is configured in such a way that it exploits passwords generated using personal information that includes the first name, surname and date of birth. Dell’Amico et al. (2010) found that Markov chain can guess strong passwords better than the JTR.



University of Fort Hare  
*Together in Excellence*

**4.2.2.4 Dropbox’s zxcvbn.** Zxcvbn is an open source algorithm for password guessing. It was introduced in 2012 and has seen various modifications to enhance guessing performance (Wheeler, 2016). Unlike PCFG and Markov chain that use probability, zxcvbn uses heuristics to guess passwords. Hence, zxcvbn is a low-cost password guessing algorithm that works with small password samples and does not require powerful computers as is the case with resourceful probabilistic algorithms, like PCFG and Markov chains. Furthermore, zxcvbn seems to be a better password guessing algorithm than the currently commercialised measures and algorithms for guiding users to generate secure passwords (de Carnavalet & Mannan, 2014). When compared to leading password guessing algorithms such as the PCFG, zxcvbn is comparable up to  $10^5$  password guessing attempts (Wheeler, 2016). To guess a password, zxcvbn makes use of its understanding of a password pattern. The algorithmic design of password guessing using zxcvbn is composed of three phases, namely, pattern matching, estimating and searching, and it makes use of password dictionaries to learn popular



passwords. During pattern matching, zxcvbn makes use of eight pre-defined pattern matching categories, namely, token (such as an English word), attempting the words in reverse, using common sequence, attempting repeating sequence, keyboard pattern, date and brute force should the previous seven patterns not have been established. Any possible mangling rules, for example, L33T are investigated during pattern matching. Once a match has been found, an estimation of the number of guesses needed to guess a password is done depending on the rank or frequency of the matched password substring in the password dictionaries. If the guessing attempts required to crack the password are few, it suggests the password is weak. Wheeler (2016) gives a detailed overview of the zxcvbn.

**4.2.2.5 Overview of password guessing algorithms.** Table 2 provides a summary of the reviewed password guessing algorithms. It shows the capabilities of these algorithms and their limitations. However, Table 2 suggests that basing passwords on predictable mangling rules, personal information and keyboard patterns may not be enough to protect the password against offline password attacks.

**Table 2. An overview of password guessing algorithms**

Password guessing algorithm	Use of a password dictionary	Password guessing approach	Applicability	Guessed samples	Computer resources
JTR	Yes	Probabilistic (target substring mangling rules)	Short passwords (< 10 characters long)	Simple mangling rules	Mild to high use
PCFG	Yes	Probabilistic (target password structures and according to their probabilities)	Short passwords Passphrases	<ul style="list-style-type: none"> <li>Keyboard patterns</li> <li>Personal information</li> <li>Can guess passwords not in the password dictionary</li> </ul>	Mild to high use
Markov chain	Yes	Probabilistic (target character distribution and transformation probability)	Short passwords Passphrases	<ul style="list-style-type: none"> <li>Personal information</li> <li>Simple mangling rules</li> <li>Can guess passwords not in the password dictionary</li> </ul>	Resource intensive
Zxcvbn	Yes	Heuristics	Short passwords (< 10 characters long)	<ul style="list-style-type: none"> <li>Simple mangling rules</li> <li>Keyboard patterns</li> </ul>	Not resource intensive

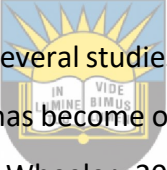
### 4.3 Password strength measurement

This section explores different measures of password strength and identifies a password strength measure that is commensurate with the model of password threats (offline and online password threats) in this study. Accordingly, this section defines the meaning of password strength or security for this study. This is critical as it helps one to understand the confines of passphrase security in sub-question 2 of this study.

Research on password strength has been a subject of interest for some time. The literature presents a number of measures for ascertaining password strength (Kelly et al., 2012; Shay et al., 2016; Weir et al., 2010). These include entropy, guess numbers, Levenshtein's edit distance and Zipf's Law (AlSabah et al., 2018; Blanchard et al., 2018; Campbell et al., 2011; Guo et al., 2019; Kelley et al., 2012; Komanduri, 2016; Malone & Maher, 2012; Melicher et al., 2016; Shay et al., 2016; Weir et al., 2010). Levenshtein's edit distance focuses on the number of characters that need to be altered as one converts a password to a dictionary word (von Zezschwitz et al., 2013). This technique assumes that users adapt dictionary words and apply minor changes when generating passwords. A small edit distance reflects a weak password. The previous chapter motivated the use of multilingualism in promoting password randomness, something that could require numerous dictionaries to ascertain Levenshtein's edit distance. In addition, multilingualism is expected to reduce password skewedness which limits the applicability of Zipf's law. Besides, the study focuses on increasing the search space and argues that it is the randomness of a phrase that contributes to password strength, irrespective of language. Although Levenshtein's edit distance and Zipf's law can help understand password distribution, both these techniques eventually require a password strength measure such as guess numbers to confirm whether passwords are indeed weak or not. In particular with regard to the Zipf's law, it is less effective on small and evenly distributed password samples (Malone & Maher, 2012). In addition, focusing on online and offline password threats implies that a commensurate password strength measure for this study should be able to establish a password's susceptibility to password guessing. Based on these arguments, this study focuses on the most commonly used measures for estimating password guessing, namely, entropy and guess number (Kelly et al., 2012).

#### 4.3.1 Entropy

Entropy is a widely conceptualised theoretical view on password strength. It draws on Claude Shannon's measure of unknown information (Weir et al., 2010). With its roots in passwords dating back to 1985, entropy was popularised by the NIST in 2006 through the Special Publication (SP) 800-63 Electronic Authentication Guideline, with the latest edition released in 2017 (Grassi et al., 2017). The documentation by NIST has since become very influential in guiding the design of password policies when generating secure passwords (Weir et al., 2010). To enhance password strength, entropy focuses on maximising the password search space by encouraging the use of all possible character sets available on the ASCII character code, for instance, when generating a password. It is assumed that users would randomly select character keys within the different classes (upper-case, lower-case, numbers and symbols) when generating passwords, something that would translate to random passwords.



It should be noted that several studies backed by empirical evidence have shown that entropy is ineffective and has become obsolete as a measure of password strength (Dell'Amico & Filippone, 2015; Wheeler, 2016). More than twenty million real-world passwords guessed so far show that there is no correlation between entropy and a password's resistance to guessing (Kelly et al., 2012; Komanduri, 2016; Weir et al., 2010; Shay, et al., 2016). In addition, entropy can label common passwords such as "P@55word" strong despite the fact that these could easily be guessed by probabilistic password guessing algorithms (Blocki, Komanduri, Procaccia, & Sheffet, 2013; Mazurek et al., 2013). These findings, among other factors, prompted NIST to abandon its tradition of encouraging different character sets as a measure for enhancing password strength (Grassi et al., 2017).

#### 4.3.2 Guess number

Dell'Amico and Filippone (2015) define password strength as a function of the number of guessing attempts needed to guess a particular password by any given password guessing algorithm – preferably probabilistic guessing algorithms. Different researchers concur with this proposition (Mazurek et al., 2013; Shay et al., 2015; Ur et

al., 2016; Wang & Wang, 2015; Wheeler, 2016). The more guessing attempts required to guess a password, the stronger the password. This approach is based on the assumption that a password perpetrator uses an optimal strategy, where passwords with few guessing attempts are targeted first (Houshmand & Aggarwal, 2012). However, the success of guess number depends on the amount and quality of data used to train the guessing algorithm (Mazurek et al., 2013), which implies that its effectiveness as a depicter of password strength is subject to available resources. This study gathered plain text passwords, on which guess numbers were used to estimate strength. Appendix A explains the study protocol for gathering raw passwords. Guess number is a better measure of password strength, considering the password threats discussed in Section 4.2.

#### **4.3.3 An overview of password strength measures**

There are different measures for estimating password strength in the literature. These are entropy, guess number, Levenshtein's edit distance, statistical techniques and Zipf's Law. Guess number was considered a suitable password strength measure for this study, as it aligns with the password threat model of online and offline threats proposed by the study. Using guess number allows this study to establish the resistance of passwords to password guessing in an offline and online password attack. The next section makes use of the literature to explore measures for promoting the generation of secure passwords. The focus is on establishing the security and usability contributions of password guidelines, best practices and policies in the literature.

#### **4.4 Factors of password strength and usability**

A number of measures aimed at promoting the generation of secure and usable passwords have been studied (Ur et al., 2012). These measures attempt to address password threats discussed in Section 4.2. However, as research is progressing, there appears to be a bias towards password security with little attention given to password usability (Babb, Keith, & Steinbart, 2016). This is so despite the adverse effect password usability has on password security. As a result, this study advanced a notion that it is impossible to maximise password strength without paying attention to password usability. Hence, the next section explores password guidelines and best practices, as

well as password policies focusing on password security and usability. The section addresses the first research sub-question that sought to establish password policies in use.

#### **4.4.1 Frameworks for password guidelines and best practices**

AlFayyadh, Thorsheim, Jøsang, and Klevjer (2012) identified four internationally recognised password authentication frameworks that could be considered as a source of guidance when implementing text-based password authentications. These include the USA's Electronic Authentication Guideline (US NIST SP800-63B) for Federal Agencies; the eID Interoperability for Pan European Electronic Government Services for the European Union; the Framework for Authentication and Non-Repudiation in Electronic Communication for the Norwegian public sector, and the National e-Authentication Framework for the Australian government (AlFayyadh et al., 2012; Grassi et al., 2017). These frameworks propose that a service provider should evaluate the risks that might have an impact on their online service provision and ascertain a commensurate authentication assurance level. A proportionate restrictive password policy would then be implemented depending on the authentication assurance level (AlFayyadh et al., 2012). Password authentication principles advanced by these best practices include

- a definition of password strength and usability
- storage of the password file in encrypted format and
- the enforcement of regular password changing (AlFayyadh et al., 2012; ISACA, 2015).

While these principles are considered valuable, their usability remain debatable. These principles are reviewed next:

**4.4.1.1 A definition of password strength and usability.** There is a belief that password strength should be determined by the character set (upper- and lower-case letters, symbols and numbers) used when generating a password (ISACA, 2015). The complexity of passwords as depicted by the number of character sets and length are determined by the sensitivity of the online service under consideration (ISACA, 2015;

AlFayyadh et al., 2012). Accordingly, entropy, discussed in Section 4.3, informs the character set and length requirements for password policies. However, the major advocate for measuring password strength using entropy, NIST, recently shifted from this stance. The NIST SP 800-63B, a subcomponent of the SP 800-63-3 suite of 2017, promotes the use of a blacklist to restrict the use of common passwords and dictionary words during password generation (Grassi et al., 2017). The SP 800-63B also discourages the use of keyboard patterns and personal information in user-generated passwords. Passwords should be at least eight characters long. The NIST SP 800-63B is of the view that by observing these requirements, the resulting passwords would be secure. However, to attain usability, password policy designers are required to give users feedback and guidance during password generation (Grassi et al., 2017). According to Furnell, Khern-am-nuai, and Esmael (2018), password feedback provides ratings that estimate password strength, while password generation guidance goes on to provide “more explanatory detail about how well the resulting password would serve” (p. 5) users.



**4.4.1.2 Password file storage in encrypted format.** The NIST SP800-63 suggest passwords should be stored in an encrypted format, irrespective of the extent of sensitivity of the account. Increasing occurrences of password file leaks point to the importance of storing the password file in an encrypted format. Gaining access to raw passwords can give attackers an easier way of knowing password structures, information that could be used in further password attacks. Service providers can use different one-way hashing algorithms (cryptography) and store passwords in a hash format. Hashing passwords is expected to burden password attackers as they try to extract the original passwords. In addition, passwords can be salted, a move that further improves password security while in storage. A salt is a user-specific string that is added to each password before hashing (Shay et al., 2016). Hence, a salting string can differ even for users sharing the same password.

However, system users cannot rely on hashing and salting for password strength as these measures do not always compensate for weak passwords. For example, there are systems that either use a poor hashing algorithm or store passwords in plain text

(Bauman, Lu, & Lin, 2015; Florêncio et al., 2014b). In addition, the characteristics of most common passwords among users are already known, following leaks of more than 100 million passwords. These details on their own could aid the successful guessing of more passwords despite a service provider's use of a hash function and salting algorithm. These findings go on to suggest a need for password policies that promote the generation of unique and secure passwords.

**4.4.1.3. Forced regular password changes.** Forced regular password changes are also known as password expiration policies. ISACA (2015) makes it categorically clear that high level accounts should have their passwords changed on a regular basis. The idea behind this requirement is to encourage users to generate strong passwords over time (Houshmand & Aggarwal, 2012). However, this measure is rarely implemented by text-based authentication designers (Florêncio et al., 2014b). Where it has been implemented, results have not been encouraging. For example, Zhang, Monroe, and Reiter (2010) analysed passwords extracted from a real-life operational system to establish security contributions of regular password changing enforced by a password expiry policy. They found that users change their old passwords very little as more than 41% future passwords were guessed using information of preceding passwords (Zhang et al., 2010). A study by Chiasson and van Oorschot (2015) concurs with these findings, as it was noted that forced password changes through regular password expiry leads to new passwords that are related to the previous ones. From a usability point of view, users (80%) are of the opinion that password expiry should not be enforced (Rinn, Summers, Rhodes, Virothaisakun, & Chisnell, 2015). Users prefer to be given room to change their passwords when they deem it necessary (Rinn et al., 2015). Similarly, US federal employees studied by Choong et al. (2014) found forced regular password changes time consuming and unassailable.

#### **4.4.2 Password policies**

A password policy is broad and presents guidelines on password generation requirements, password expiry, password reuse, log-in attempts and password recovery procedures. Some of these guidelines were discussed in Section 4.4.1. This section focuses on password generation requirements. Such policies outline a



predefinition of acceptable and unacceptable passwords (Blocki et al., 2015). Calls for secure passwords in light of weak but memorable user-generated passwords in the 1990s motivated a need for password policies (Shay et al., 2016). These password generation policies make different security and usability contributions with some policies struggling to find a balance between the two (Braunstein, 2015; Inglesant & Sasse, 2010; Komanduri et al., 2011; Ur et al., 2012).

The literature suggests that password generation policies include the password composition policy, system assigned password policy, system and user-generated password policy, and password strength meters (Houshmand & Aggarwal, 2012; Ur et al., 2012; Wang & Wang, 2015; Weir et al., 2010). Of these policies, password strength meters are recommended as others force users to adhere to specific guidelines (von Zezschwitz et al., 2013). These password generation policies are discussed next, identifying their security and usability contributions. The usability of these password policies is evaluated, paying attention to the four stages of the password life cycle – password generation, keeping track of passwords, authenticating and changing passwords (Choong et al., 2014; Stobert & Biddle, 2014).



**4.4.2.1. Password composition policy.** The password composition policy is arguably the brainchild of the NIST old version of the SP 800-63 Electronic Authentication Guideline and is widely used (AlFayyadh et al., 2012; Florêncio & Herley, 2010; Houshmand & Aggarwal, 2012). Also known as an “explicit password creation policy”, this policy defines that which constitutes an acceptable password (Weir et al. 2010, p. 171) in terms of the minimum length and use of symbols, numbers, uppercase and lowercase letters. In other words, the password composition policy promotes the generation of passwords the security of which is anchored in lower-case, upper-case, digits and symbols (LUDS). The policy works best with proactive check mechanisms to monitor the use of different character classes during password generation (Shay, et al., 2012). In addition, a dictionary check and blacklist could be used to forestall users from adopting common passwords (Komanduri et al., 2011; Ur et al., 2012; Shay et al., 2016; Weir et al. 2010). Below is an overview of the security and usability limitations of the password composition policy.



**Security limitations.** A ground-breaking study by Weir et al. (2009) found that password composition policies may not always lead to strong passwords as generally assumed. Findings from Weir et al.'s study has been corroborated by several other studies (Komanduri et al., 2011; Ur et al., 2012; Weir et al. 2010). Among the security limitations of the password composition policy is a finding that users fulfil password requirements in predictable ways (Grassi et al., 2017). However, the literature motivates the use of a blacklist to restrict the use of popular passwords and words (Kelley et al., 2012; Komanduri et al., 2016; Grassi et al., 2017). The next section on usability discusses some of the user behavioural practices that compromise the security of passwords generated under the password composition policy.

**Usability limitations.** When a password composition policy is implemented, users find password generation and learning frustrating, and often fulfil password requirements in a predictable manner (Komanduri et al., 2011; Ur et al., 2012). Hence, a complex password composition policy leads to user frustration during password generation, password reuse, basing passwords on semantic information, failing to memorise passwords and being faced with authentication challenges. Some of these usability concerns result in password security ramifications which are discussed next:

- i. **Frustration during password generation.** Users find it difficult to meet the character class requirement and sometimes the required password length. This was demonstrated by users who required more password generation attempts (2 to 3.35 attempts) to generate a password that is “at least 8 characters” long with a “lowercase English letter, uppercase English letter, digit and symbol” when compared to generating a password of at least twelve characters in length (Komanduri et al., 2011; Shay et al., 2016, p. 12). Melicher et al. (2016) observed a number of deletions during password generation as a result of unusable password composition policies. They found that users may end up failing to meet password requirements which can lead to subsequent quitting of the password generation process (Komanduri et al., 2011; Shay et al., 2016).
- ii. **Password reuse.** From a usability point of view, password reuse helps users reduce the burden of memorising many passwords given the ever-increasing number of password-driven accounts, ranging from 6 to 199 per computer user

and averaging 25 per user (Bang et al., 2012; Hayashi & Hong, 2011; Helkala & Bakås, 2013; Stobert & Biddle, 2014; Von Zezschwitz et al., 2013). However, from a security point of view, a study from South Korea demonstrated how 150 000 accounts on a secure website were successfully guessed using information on 2.3 million log-in credentials secured from sites with weak password authentication designs (Bang et al., 2012). Studies have investigated how users reuse passwords. According to Choong et al. (2014), users are more likely to make minor changes to existing passwords or use existing passwords as is or use old passwords when they find password requirements complex and burdensome. If a user decides to modify existing passwords, the resultant passwords are on average arrived at after deleting or inserting two to three characters “at the beginning or end or at both ends of a string” (Das, Bonneau, Caesar, Borisov, & Wang, 2014, p. 5). These changes include adding a number or symbol at the front or capitalising the first letter depending on password requirements (Das et al., 2014; Rinn et al., 2015; Ur et al., 2016; Ur et al., 2015; Von Zezschwitz et al., 2013). Eventually, these usable and predictive ways of fulfilling the password composition policy compromise the security contributions made by the policy.

- iii. **Using semantic information in passwords.** Users are more likely to adapt semantic information when faced with a restrictive password composition policy (Shay et al., 2010). There are also instances where users make use of personal information, even if password requirements are not so restrictive. Personal information often adapted during password generation includes dictionary words, website information or personal information such as hobbies, their names or names of loved ones, date of birth, address, phone numbers and identification numbers. The use of words in a language is clearly shown by different character distribution across the password corpora from different contexts (AlSabah et al., 2018; Bonneau & Xu, 2012; Jakobsson & Dhiman, 2013; Maoneke, Flowerday, & Isabirye, 2018; Wang et al., 2015; Yang et al., 2013). For example, “English users use raw English words as a basis for passwords, while few Chinese users chose raw Pinyin words to build passwords, yet they prefer Pinyin names, especially full names” (Li et al., 2014; Wang et al., 2015, p. 7).

Furthermore, passwords based on keyboard patterns, digits and/or personal information are also more pronounced in Chinese datasets and those from the Middle East (AlSabah et al., 2018; Li et al., 2014; Wang et al., 2015). Section 4.2.2 explained how passwords generated following these strategies could easily be guessed by a password guessing algorithm.

- iv. **Failure to memorise passwords.** The memorability of a password is the most important factor users consider when generating a password (Choong et al., 2014). Similarly, Ur et al. (2015) found that users (35%) consider password memorability a primary concern. Users often opt to memorise a passwords if they “perceive the benefit” of memorising a password outweighs the cost of writing down the password in the event of “a security breach” (Duggan, Johnson, & Grawemeyer, 2012, p. 416). Where users find it difficult to memorise passwords, they resort to writing down passwords, reusing passwords and using password managers (Choong et al., 2014; Komanduri et al., 2011; Stobert & Biddle, 2014; Shay et al., 2016). These password tracking techniques are more pronounced in complex and burdensome password requirements (Choong et al., 2014; Shay et al., 2010; Shay et al., 2016). For example, the majority of users who wrote down passwords were associated with a password requirement that users include at least three-character classes in their passwords (Komanduri et al., 2011; Shay et al., 2016). Furthermore, a number of users save passwords in a web browser (12–81%) instead of using formal password managers (4%) (Stobert & Biddle, 2014; Ur et al., 2015).
- v. **Authenticating challenges.** Users report mistyping (re-entry) or typographical errors when logging in, while others resort to copying and pasting passwords. This is more pronounced when using passphrases, mobile phones and/or where password policies are considered cumbersome (Choong et al., 2014; Keith et al., 2009; Melicher et al., 2016). Typographical errors occur when a user mistakenly strikes a nearby key or transposes the correct characters or “enters too few or too many keystrokes” (Keith et al., 2009). Typographical errors are not a result of memory loss, but a mistake during the execution stage as one is keying in a password (Keith et al., 2009). Nevertheless, typographical errors contribute to user frustration during logging in.

In light of the usability challenges, businesses in the private sector seldom adopt restrictive password composition policies. For example, large, high value accounts that are often attacked such as PayPal, Amazon, eBay, Facebook and Gmail have relatively weak password composition policy designs aimed at maximising usability (AlFayyadh et al., 2012; Florêncio & Herley, 2010; Furnell, 2016; Wang & Wang, 2015). This is despite the fact that leading frameworks on password policy guidelines recommending strict password composition policy design, as determined by the importance of the account (AlFayyadh et al., 2012; Florêncio & Herley, 2010).

Nevertheless, the literature reports a number of efforts aimed at enhancing the usability and security of password composition policies. Shay et al.'s (2015) multi-step interactive password creation process and real-time password requirements communication improved password generation usability. Mazurek et al. (2013) recommend that users avoid beginning passwords with uppercase letters and ending passwords with digits and symbols. In addition, a number of studies are motivating for the use of long passwords (Blanchard et al., 2018; Bonneau & Shutova, 2012; Braunstein, 2015; Komanduri et al., 2011; Shay et al., 2016; Ur et al., 2012). NIST appears to support the idea as they have now discouraged the use of different character sets in preference for passwords that are at least eight characters long, generated under the guidance of a blacklist and real-time feedback (Grassi et al., 2017). However, research on the security and usability contributions made by passphrases is still ongoing (Blanchard et al., 2018; Juang & Greenstein, 2018; Rao et al., 2013). Accordingly, this study aims to make its contribution by proposing a model that could be used when designing a password composition policy that leads to secure and usable multilingual passphrases.

**4.4.2.2. System assigned passwords.** This policy is centred on maximising password strength by avoiding password reuse and the generation of weak passwords by users (Houshmand & Aggarwal, 2010; Shay et al., 2012; Ur et al., 2012). Password authentication frameworks recommend system-assigned passwords for sensitive user accounts and these can be implemented through one-time passwords (AlFayyadh et al., 2012). However, this policy is less usable. For example, Shay et al. (2012) experimented

with system-assigned passphrases and passwords. They observed that users faced memorability challenges to the extent that they had to resort to writing down passphrases and passwords (Shay et al., 2012). Furthermore, users found it more difficult to type system-assigned passphrases accurately than system assigned passwords (Shay et al., 2012). Al-Ameen, Fatema, Wright, and Scielzo (2015) researched visual, verbal and spatial cues with the aim of enhancing the memorability of system-assigned passwords. They argued that different cues give users an opportunity to adopt a preferred password cue. Al-Ameen et al. (2015) reported a 100% memorability of passwords among 37 participants even after a week from the day of initially receiving the assigned password. Nevertheless, their approach had long login durations and was evaluated using a relatively small sample, something that limits the generalisability of their findings.

**4.4.2.3. System and user-generation password policy.** Following usability and security concerns associated with system assigned passwords and password composition policies, there are studies exploring the security and usability contributions of passwords that are core-generated by both the system and user (Blanchard et al., 2018; Furnell et al., 2018; Houshmand & Aggarwal, 2012; Komanduri, Shay, Cranor, Herley, & Schechter, 2014; Weir et al., 2010). For instance, a user-generated password is evaluated for strength using a probabilistic context-free algorithm. Should the user-generated password fail to meet the minimum strength threshold, the system would adjust the password by effecting a single character or more at any position of the password thereby improving password randomness (Blanchard et al., 2018; Houshmand & Aggarwal, 2012; Weir et al., 2010). Limiting password modifications to a few characters ensures memorability of the password is maintained (Houshmand & Aggarwal, 2012).

In addition, Komanduri, Shay, Cranor, Herley, and Schechter (2014) proposed an algorithm that could, given preceding character(s), predict the next character a user would be likely to type and give a warning depending on predicted security repercussions. The algorithm was named Telepathwords (Komanduri et al., 2014). For effective prediction and assistance during password generation, Telepathwords is first

trained on various password predictors such as common character sequences learnt from public passwords and language models, keyboard patterns, repeating string sequences and interleaving strings. The idea is to help users avoid generating passwords that are associated with common password predictors, something that compromises security. Data collection and analysis shows that Telepathwords improved security and memorability when compared to selected password composition policies (Komanduri et al., 2014). However, users complained of the annoying password generation process. Another study by Blanchard et al. (2018) proposed a passphrase generation technique in which users were guided to select words from an array of random words. Their technique improved passphrase strength and reduced the use of popular words and grammatical rules in user-generated passphrases.

System and user-generation password policy is considered less user-friendly (Weir et al., 2010). In addition, the overall implication of implementing such a presumably costly policy remains to be seen, given that the popularity of password use is down to low costs, easy and simple implementation compared to token- and biometric-based authentications (Bauman et al., 2015; Weber et al., 2008). However, such a policy can be useful in guiding password generation for high-value password accounts, such as those for financial services (Bailey et al., 2014; Mazurek et al., 2013).

**4.4.2.4. Password strength meters (PSMs).** PSMs are used to nudge users into generating strong passwords. Ur et al. (2012) defines a PSM as “a visual representation of password strength, often presented as a coloured bar on screen” (p. 1). Wang and Wang (2015) concur with this definition. To gauge password strength, PSMs can make use of a password guessing algorithm or entropy (password character composition) or detect weak and common keyboard patterns (Castelluccia, Dürmuth, & Perito, 2012; De Carnavalet & Mannan, 2014). Although PSMs are not as dominant as some password policies, their presence is quite notable on different websites including those offering electronic commerce services (Wang & Wang, 2015). Empirical evidence suggests the presence of a PSM does influence users into generating strong passwords (Babb et al., 2016; Castelluccia et al., 2012; Ciampa, 2013; De Carnavalet & Mannan, 2014; Egelman, Sotirakopoulos, Muslukhov, Beznosov, & Herley, 2013; Ur et al., 2012; Vance, Eargle,

Ouimet, & Straub, 2013). However, this finding is challenged by a number of factors. When implemented stringently, PSMs may frustrate users to the extent that they may not be motivated to improve their password strength ratings (Ur et al., 2012). In addition, there are suggestions that PSMs may not be influential when generating passwords for low value accounts given that users put more effort into creating strong passwords where the account is considered to be of high value (Egelman et al., 2013). This is supported by a finding that the “perceived threat, perceived password effectiveness” and one’s capability of efficiently meeting the password requirements together influence the intent to comply with the requirements (Mwagwabi, McGill, & Dixon, 2014, p. 3194). Vance et al. (2013) add that a simple PSM “does not significantly improve password strength” when implemented without interactive fear appeals to warn users of the possible threats (p. 2996).

Moreover, the literature suggests that the PSMs in use on most websites give inaccurate password strength estimates as they are not aggressive enough to encourage the generation of strong passwords (Castelluccia et al., 2012; Ciampa, 2013; Furnell, 2016; Maoneke & Flowerday, 2018; Ur et al., 2012; Ur et al., 2016). This is so despite the fact that PSMs can influence users to improve password strength by changing password length and character composition (Ciampa, 2013; Ur et al., 2012). This could be explained by the fact that private sector website owners focus more on usable password policies for their clients (AlFayyadh et al., 2012; Florêncio & Herley, 2010; Wang & Wang, 2015). Another explanation could be that there is no one standard approach for evaluating password strength given that a password considered weak by one PSM is considered strong by another (Ciampa, 2013; Wang & Wang, 2015).

However, the most perturbing observation is that despite progress on researching password guessing algorithms, there is limited use of these guessing algorithms in PSMs to measure password strength, with the majority of current approaches centred on password character composition (LUDS) or entropy (Babb et al., 2016; Ciampa, 2013; Egelman et al., 2013; Vance et al., 2013; Ur et al., 2012). Only Castelluccia et al. (2012) and Wheeler (2016) demonstrated the implementation of a password guessing algorithm on PSMs. For example, Castelluccia et al. (2012)



demonstrated the use of a Markov-based algorithm to assess password strength on a PSM. Wheeler (2016) explained the use of Dropbox's zxcvbn guessing algorithm when guiding users to generate secure and usable passwords.

#### **4.4.3 An overview of users and password management behaviours**

Users are an important stakeholder in the password ecosystem (Taneski et al., 2014) as they play a leading role throughout the password life cycle. Password life cycle is made up of four stages, namely, password generation; keeping track of passwords; authenticating and changing passwords (Choong et al., 2014; Stobert & Biddle, 2014). It is important to understand users' password tendencies throughout the password life cycle in light of password threats, memorability and cognition development, password guidelines and best practices that were discussed in this and the previous chapter. The knowledge about users' password behaviours could be used to inform the design of password policies or the creation of additional supportive measures to help users settle when using a particular password policy. User password tendencies are a direct reflection of the trade-offs between password security and usability requirements (Ur et al., 2016) or security and convenience (Tam, Glassman, & Vandenwauver, 2010; Woods & Siponen, 2019), something Duggan et al. (2012) refer to as a matter of weighing costs and benefits as users go through different stages of the password life cycle.

##### **4.4.3.1 Password generation**

Ur et al. (2015) and Renaud et al. (2019) found that users have predefined password generation strategies and routines they are not likely to abandon. Concatenation of different character classes is the mainly used password generation strategy followed by replacement (e.g. L33T) and making spelling mistakes (Jakobsson & Dhiman, 2013). This is done by applying minor changes to existing passwords or else users simply reuse existing passwords or recycle old ones (Choong et al., 2014; Von Zezschwitz et al., 2013). It should be noted that during password generation, users prioritise the generation of memorable passwords (Choong et al., 2014). Reuse of passwords or sections thereof and the adherence to semantic information are the overarching password generation strategies (Duggan et al., 2012) as users find their way



around password requirements and device constraints. These password generation strategies are not always used mutually exclusively.

Password rules that require users to include different character sets in their passwords are seen as cumbersome (Choong et al., 2014; Komanduri et al., 2011; Shay et al., 2016; Shay et al., 2010). A closer look at the problem shows that users very often struggle to meet the character class requirement and sometimes the required password length. Further to that, users fail twice as much to create passwords on mobile phones due to typing mistakes when compared to using traditional devices such as computers and laptops (Melicher et al., 2016). Contrary to this finding, there are suggestions that users (66%) do not find device constraints a concern when selecting passwords (Stobert & Biddle, 2014). It appears users have become used to generating passwords on mobile phones, despite the usability concern (Melicher et al., 2016). However, using a mobile phone for generating passwords results in users opting for all lower-case letter passwords (Melicher et al., 2016; Yang, Lindqvist, & Oulasvirta, 2014). Making passwords visible during password generation on mobile phones can significantly reduce typing errors and increase password strength (Melicher et al., 2016). Studies have shown that users can generate equally strong passwords using mobile phones as with computers (Yang et al., 2014), especially when generating long passwords or passphrases (Bonneau & Shutova, 2012; Shay et al., 2016; Melicher et al., 2016). This finding supports the aims of this study of generating secure and usable multilingual passphrases.

#### **4.4.3.2 Keeping track of passwords, authenticating and changing passwords.**

Users often rely on memorising passwords, writing down passwords, password reuse and using a password manager in order to keep track of passwords. Choong et al. (2014) observed that the main (69%) password tracking technique for frequently used passwords was memorising. This high percentage can be explained by the fact that repeated use and possibly password reuse improves memorability (Stobert & Biddle, 2014). However, few users (38%) could memorise less frequently used passwords (Choong et al., 2014). When users find it difficult to memorise passwords, they resort to writing down them down on different media or saving them in password

managers (Choong et al., 2014; Komanduri, et al., 2011; Stobert & Biddle, 2014; Shay et al., 2016). There are different explanations why users choose to write down passwords. These are the following

- i. complex and burdensome password policies that require three- or four-character classes
- ii. strategically done when a user perceives that “the benefit for not having to remember a password to outweigh the possibility of, and costs associated with, an increased likelihood of a security breach” (Duggan et al., 2012, p. 416)
- iii. strategically done to assist in recalling certain passwords. This is done until the password is memorised (Stobert & Biddle, 2014).
- iv. perceived to be a safe way of storing passwords (Ur et al., 2016).

The literature portrays mixed reactions to the security and usability contributions of writing down passwords. Ur et al. (2016) suggest that writing down passwords is not secure, even though computer users think otherwise. But, Tam et al. (2010) argue that computer users should be encouraged to write down passwords as long as they can keep their pads or notes with written passwords in a safe place. Some users keep media with written passwords safely while others keep them in unsafe places (Choong et al., 2014; Stobert & Biddle, 2014). This study is of the view that writing down passwords is a bad practice that should be avoided. Awareness campaigns could be considered so that users can make informed decisions on when to write down passwords and where to keep the media with written passwords. The use of password managers or password vaults to aid memorability and password tracking could be considered (Chatterjee, Bonneau, Juels, & Ristenpart, 2015). By using a password manager, users will only be required to memorise the master password that will be protecting all the other passwords. This study proposes that a strong and usable master password could enhance the security of password managers. Hence, the need for a secure and usable passphrase remains important.

Section 4.4.2 identified authentication challenges faced by users and recommendations in the literature. However, password changing as a component of the password life cycle is rarely practised. Users rarely change their passwords, as 70%

of those studied in Von Zezschwitz et al. (2013) were found to be using passwords they created during registration. Similarly, Stobert and Biddle (2014) observed that users rarely change passwords unless the password has been forgotten. Even if users do change their passwords, they often make sure that the resulting passwords are in synch with existing passwords (Stobert & Biddle, 2014).

#### **4.5 Chapter summary**

Passwords have been in use for a very long time, dating back to the ancient Roman times (Adeka, Shepherd, & Abd-alhameed, 2013). This trend is expected to continue into the foreseeable future. Accordingly, this chapter has reviewed different password security threats. The ever-increasing incidences of password file breaches have seen offline password attacks become a dominant password threat. Probabilistic algorithms are the most commonly used techniques for guessing passwords in an offline password attack. These password guessing algorithms are versatile and could exploit the password generation strategies often followed by computer users. The popularity of offline password threats and improved password guessing algorithms resulted in a theoretical shift in defining what constitutes a strong password, as various researchers prefer to use guess numbers over Shannon's entropy. However, some of the password strength guidelines and best practices are still using principles in Shannon's entropy as a measure of password strength, despite the fact that such principles have been found to be less effective.

Given password security threats posed by online and offline attacks, this chapter observed that recommended password best practices and password policies often fail to find a balance between security and usability. For instance, password expiry policies that were expected to encourage the generation of unique and secure passwords have been found to be ineffective and not user-friendly. Computer users often find ways to adapt and reuse their old passwords, something that is against the spirit of the password expiry policy. Storing passwords in encrypted format can easily be weakened by some service providers who use cheap and weak encryption algorithms or the generation of easy-to-guess passwords and password reuse that span secure and insecure sites. Again, restrictive password policies that are expected to enhance

security have been found unusable and users often fulfil these password requirements in predictable ways. As a result, the private sector remains sceptical about adopting restrictive password policies as this may frustrate clients, leading to a loss of business.

In light of these findings on password guidelines, best practices and policies vis-à-vis offline password threats and user behaviour, the next chapter proposes an alternative approach for addressing the password security and usability challenges identified in this chapter. Chapter 5 is guided by findings in both Chapter 3, the socio subsystem, and Chapter 4, the technical subsystem, to propose a model of secure and usable passphrases.

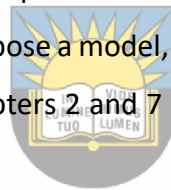


University of Fort Hare  
*Together in Excellence*

## CHAPTER 5: A MODEL OF USABLE AND SECURE PASSPHRASES

### 5.0 Introduction

There is growing interest in research aimed at improving the usability and security of text-based authentication mechanisms, with passwords being a particular focus (Blanchard et al., 2018; Choong et al., 2014; Grassi et al., 2017; Kelley et al., 2012; Shay et al., 2016; Wang, Cheng, et al., 2015). The previous chapter explained the different password guidelines, best practices and policies that are in use and identified the security and usability challenges associated with text-based authentication mechanisms. This chapter adds to the study by proposing a tentative model of secure and usable multilingual passphrases. This marks a critical step in this study, which aimed at proposing a model of secure and usable multilingual passphrases. Accordingly, the chapter defines a model and goes on to explain the rationale for proposing a model for this study. This chapter explains the use of the socio (Chapter 3) and technical (Chapter 4) subsystems to propose a model, which is demonstrated and evaluated using criteria explained in both Chapters 2 and 7 in order to address the problem statement of the study.



University of Fort Hare

*Together in Excellence*

### 5.1 Definition of a model

March and Smith (1995) define a model as “a set of propositions or statements expressing relationships among constructs” (p. 6). Hevner et al. (2004) add that a model is an abstraction that shows how things are through the use of constructs. March and Smith (1995) define constructs as “concepts from the vocabulary of a domain” that are used to describe a problem or solution. A model is one of the various artefactual outcomes that are expected from design science research. Thus, the ultimate goal of design science research is to create an outcome (artefact) in the form of a model or instantiation or methods that reflect the reality of the phenomenon under investigation. This is in sharp contrast to natural science or social science research which seeks to understand reality by testing a model or theories (Peffer et al., 2008). Nevertheless, the proposition of constructs in design science research models should be informed by reliable theories within the domain. A model is a suitable outcome for this study, as it will show the relationship between the key constructs that influence the

generation and use of secure and usable multilingual passphrases. While the available frameworks, guidelines and best practices are important, they are not prescriptive, something that paves the way for abuse by password policy designers as they fail to interpret what is expected.

## **5.2 The rationale behind the proposed research model**

This study acknowledges other researchers who have investigated passphrase security and usability (Kelley et al., 2012; Komanduri, 2016; Rao et al., 2013; Shay et al., 2016). However, it is quite intriguing to note that rarely do the findings on passphrase security and usability from these studies complement each other. Instead, the literature presents mixed views that portray inconclusive findings on the security and usability contributions of passphrases. For instance, Kelley et al. (2012) and Shay et al. (2016) posit that passphrases are secure and usable when compared to short passwords. Conversely, Rao et al. (2013), Bonneau and Shutova (2012) and Veras et al. (2014) question the security contributions of passphrases, arguing that users often create passphrases following linguistic patterns that can be exploited and that compromise passphrase strength.



In short, studies that investigated passphrase security focusing on structural dependencies at character level found passphrases to be secure (Kelley et al., 2012; Shay et al., 2016; Shay et al., 2014), while studies that researched passphrase security based on linguistic properties such as grammatical structures, popular words in a language and keyboard patterns found those trying to crack passphrases prone to guessing (Rao et al., 2013; Bonneau & Shutova, 2012; Veras et al., 2014). This is because attackers can simply reduce the passphrase search space by exploiting inherent semantic patterns that users follow when generating passphrases. Furthermore, the current crop of studies on passphrases is dominated by English computer users. This is despite research findings by Wang, Cheng, et al. (2015) that suggest the structure of user-generated passphrases is influenced “by native languages (and culture background)” (p. 5). Furthermore, the success of the Markov chain password guessing algorithm is down to its ability to imitate character distribution in a language. Arguably, these findings suggest the importance of a user’s language when investigating

passphrase security and usability. Given that user-generated passphrases confined to a particular language are likely to be weak and easy to guess (Rao et al., 2013; Bonneau & Shutova, 2012; Veras et al., 2014), this chapter explores the use of multiple languages (multilingual phrases) when creating a user-generated passphrase. This is done with the idea of increasing the size of the passphrase's search space (Rao et al., 2013), without compromising the security or usability of passphrases.

Considering the above arguments, this chapter proposes a model of multilingual passphrase security and usability that is suitable for multilingual user groups. The chapter reflects on socio-cultural theory, passphrase security and usability and then proposes a tentative model. In particular, this chapter reports on findings from an abductive use of kernel theory or justification knowledge in the literature. Thus, the proposed model is an output of the design and development activity according to Peffers et al. (2008). The study argues that the current state of research findings on passphrase security and usability, coupled with conflicting findings, presents a research opportunity to explore alternative approaches to addressing the security and usability dilemma.



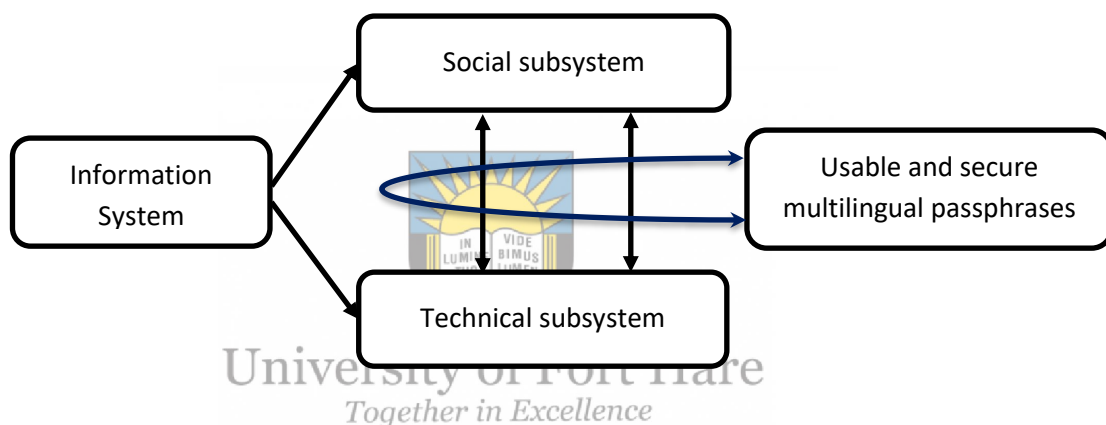
University of Fort Hare  
*Together in Excellence*

### **5.3 Theoretical foundation**

This study is grounded in socio-technical theory (Durkin et al., 2015). By adopting socio-technical theory, the study is of the view that a joint optimisation of the socio and technical subsystems enhances passphrase security and usability. Chapter 3 explored the view of this study of the social subsystem, while Chapter 4 provided a broad overview of the perception on the technical subsystem. In particular with regard to the social subsystem, this study used multi-store theory to explain the functionality of human memory and went on to use socio-cultural theory to explain how users acquire languages. Password security and usability has long been linked to the language and culture of users (Li et al., 2014; Wang, Cheng, et al. 2015; Bonneau & Shutova, 2012). For instance, Wang, Cheng, et al. (2015) and Li et al. (2014) established that user-generated passwords from a single language group exhibit consistency in terms of character distribution. However, there is a significant difference in character

distribution between passwords generated by different language user groups (Wang, Cheng, et al., 2015).

In addition, considerations regarding the technical subsystems in this study were limited to the processes and procedures that shape activities of password generation in light of password threats. These include password policies and the use of a blacklist. Technical subsystems relating to PSMs and real-time password generation feedback are beyond the scope of this study. In short, this study focused on the proposition of password policies whose implementation is informed by social subsystems as shown in Figure 9.



**Figure 9. The perceived joint influence of social and technical subsystems on passphrases**

For instance, Chapter 3 used socio-cultural theory to explain how a computer user in the context of this research is likely to acquire multiple languages. Chapter 3 continued by giving an account of the practice of code-switching in text messages that reflected the extent of multilingualism among computer users. Accordingly, this study argues in favour of extending the practices of code switching to user-generated passphrases, if usability and security is to be attained. For example, it is argued that individuals who are comfortable in both Chinese and English could be influenced through the use of relevant password policies in such a way that they generate passphrases that are based on words or substrings from these two languages. This study advanced the NIST's (SP 800-63) use of principles in Claude Shannon's entropy as explained in Chapter 4. However, instead of encouraging the use of all possible



character sets available on the ASCII character code, this study focused on promoting the use of substrings orientated to the different languages a computer user knows. The next sections on multilingual passphrase security and usability give further details on the implementation of multilingual passphrases in this study.

### **5.3.1 Multilingual passphrase security**

This study focused on the online and offline security threats explained in Section 4.2. Reports of database attacks exposing millions of hashed passwords increased the interest in researching solutions for online and offline password attacks. It should be noted that passwords are not always stored as plain text in a service provider's database, but in a hashed file. Every password in the hashed password file will have its own hash pattern generated by a hashing function. Upon gaining access to the hashed password file, security perpetrators generate candidate passwords together with their respective hash patterns. These hash patterns are then compared to those in the hashed password file to see if there is a match. A match of hash patterns suggests that a corresponding password has been successfully guessed. Hence, service providers can use slow hash schemes to reduce the likelihood of offline password guessing. However, Section 4.4.1.2 explained why users should not rely on hashing and salting algorithms for password security. Instead, users are encouraged to generate strong passphrases to protect themselves from online and offline password threats (Melicher et al., 2016; Shay et al., 2016).

Section 4.3 in Chapter 4 identified Shannon's entropy and guess number as the most common measures for estimating passphrase strength. This study assumed guess number, explained in Section 4.3.2, as the criterion for measuring password strength. Guess number has gained popularity following a research finding that Shannon's entropy misleads when it comes to estimating password strength (Shay et al., 2016; Wang, Cheng, et al., 2015; Weir et al., 2009). The analysis of guess number in Section 4.3.2 suggested that, from a probability theoretical point of view, a guessing algorithm would need a few attempts to guess popular passwords. Hence, this study adopted Rao et al.'s (2013) view that the size of the passphrase search space should be increased in order to maximise passphrase strength. The search space was defined as the set of all

possible unique values or words in different languages that could be used to generate passphrases (Rao et al., 2013). Accordingly, this study took advantage of multilingual users in its socio subsystem and recommended a passphrase policy that encourages the use of words from different languages. There are suggestions that the use of multiple languages in user-generated passphrases and short passwords promotes strength (Ur et al., 2016; Voyiatzis et al., 2011).

#### **5.3.1.1 Factors for passphrase strength**

Chapter 1 set out the second sub-question of this study as follows:

*What are the language characteristics that could be considered to enhance the security of user-generated passphrases?*

This section addressed the second research sub-question by identifying factors of password strength for this study. The literature motivated various factors that could be considered to enhance security. Shay et al.'s (2016) study analysed the structural dependencies of passphrases at character level. They found that passphrases were secure on condition that a blacklist is used to deter users from generating predictable phrases. In contrast, Rao et al.'s (2013) study focused on linguistic properties of passphrases and found them to be weak. Rao et al. (2013) observed that memorability demands force users to generate passphrases following a few selected grammatical rules. As such, resultant passphrases could easily be guessed by algorithms that exploit grammatical rules. In addition, Bonneau and Shutova (2012) observed the majority of user-generated passphrases are composed of words that are skewed towards popular words in a natural language. Bonneau and Shutova (2012) and Rao et al. (2013) therefore recommend the use of random words, if the resultant passphrases are to be strong.

This study argues that the research findings by Shay et al. (2016), Rao et al. (2013) and Bonneau and Shutova (2012) could be explained by the fact that the respective passphrases were generated using a single language. As such, this study sought to take advantage of a multilingual user group and encourage the generation of

passphrases based on words from different languages, which is expected to spread the character distribution of the passphrase corpora. This principle is somewhat related to views relating to NIST's use of Shannon's entropy, although this study focused on using a wide range of languages instead of character classes. In addition, Florêncio et al., (2014) is of the view that while a blacklist with  $10^6$  common words or passwords may be reasonable to deter users from using common words/passwords, it may not protect against all offline password guessing attacks. Besides, bigger lists with roughly  $10^{14}$  common words or passwords may compromise usability. This is further exacerbated by the fact that common passwords may vary from one context to the other depending on the language in use to such an extent that a passphrase policy that adopts a public list of blacklisted passwords may not guarantee security (AlSabah et al., 2018; Blocki et al., 2013).

The next section lays out constructs for ascertaining passphrase security that are deemed suitable for the multilingual user group in this study. This study has adopted the orientation of user-generated passphrases to different languages. Hence, juxtaposing substrings, passphrase length and dictionary checks are considered important constructs for enhancing passphrase security.

**Juxtaposing substrings.** Juxtaposing substrings from different languages as reflected in code switching has the potential to increase the search space and make the resultant passphrase random and difficult to guess. Passwords from different languages have different character distributions (Wang, Cheng, et al., 2015). As noted by Wang, Cheng, et al. (2015), if one tests the security strength of a native Chinese password together with English-based passwords, the Chinese password is likely to be considered stronger and more difficult to guess, even though it might be weaker when tested together with other Chinese passwords. As a result, this study suggested juxtaposing substrings from at least two different languages in order to enhance passphrase security.

While juxtaposing substrings from different languages appear analytically demonstrable in theory; the literature shows that other presumed assumptions such as

basing passwords on LUDS, a move that was widely expected to enhance password strength, did not yield the anticipated benefits as users found ways to manipulate the password generation requirements (AlSabah et al., 2018; Grassi et al., 2017; Guo et al., 2019; Komanduri et al., 2011; Wang, Cheng, et al., 2015; Weir et al., 2010). Furthermore, African languages are considered less attractive when compared to Indo-European languages as they are seen as complex and often have long words (Deumert & Masinyana, 2008). Hence, the need for a user study to demonstrate the use of African languages in passphrase generation and the influence of juxtaposing passphrases on security. In addition, juxtaposed substrings in a passphrase should be separated by blank spaces, a practice that was found to significantly enhance passphrase strength (Melicher et al., 2016; Shay et al., 2016). It is therefore proposed that:

*Proposition P1: Passphrases generated by juxtaposing substrings from different languages are more secure.*



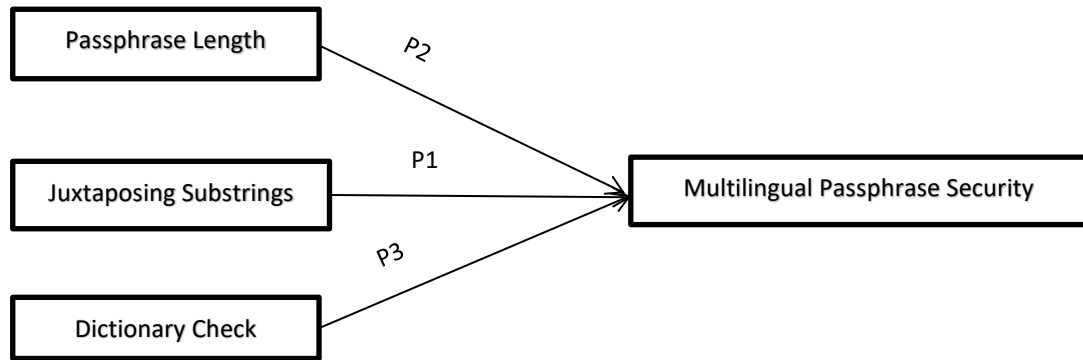
**Passphrase length.** Studies have shown that moving from short passwords to long passwords, herein referred to as passphrases, has the potential to enhance the overall security of passwords (Blanchard et al., 2018; Kelley et al., 2012; Komanduri et al., 2011; Melicher et al., 2016; Shay et al., 2016). For example, all the password policies recommended by Shay et al. (2016) and Melicher et al. (2016) on security grounds are strictly for generating passphrases. The recommended password policies encourage the generation of passphrases that range from twelve to more than sixteen characters long. However, observations on reviewed passphrases showed that passphrase length alone may not be adequate to compensate for the shortcomings of generating passphrases based on a few popular words (Bonneau & Shutova, 2012; Shay et al., 2016). In addition, Rao et al. (2013) used parts of speech tagging to demonstrate that passphrase strength is not a direct function of length. Underlying password structures together with length play a pivotal role towards a passphrase's strength (Shay et al., 2016; Rao et al., 2013). This study proposed the juxtaposing of substrings from different languages as the determining factor of the underlying structures of passphrases. Thus, length is considered a function of the juxtaposed substrings that make up the passphrase structures in this study. It is therefore proposed that:

*Proposition P2: Passphrase length and underlying passphrase structures can enhance passphrase security.*

**Dictionary check.** Chapter 4 observed that a blacklist or dictionary check is one of the recommended best practices for restricting users from basing their passphrases on common words and passwords. However, this study went further to use dictionary checks for making sure that user-generated passphrases are not based on a single language. Thus, dictionary check was used to promote the occurrence of juxtaposed substrings in a passphrase. Consequently, the use of a dictionary check was expected to enhance passphrase security by restricting users to generating passphrases based on substrings from multiple languages. This, therefore, enforced the view in this study that of increasing passphrase search space by encouraging the use of multilingual phrases. This study acknowledged that users may generate passphrases based on non-standard spelling or following phonological approximations – something that is common in code-switching among multilingual user groups (Carrier & Benitez, 2010; Deumert & Masinyana, 2008). Such substrings may not be detected by dictionary checks. However, findings from other research studies suggest that such practices have the potential to enhance passphrases security (Shay et al., 2016; Melicher et al., 2016). Appendix A explains and justifies the dictionaries that were used for the experiment in this study. It was therefore proposed that:

*Proposition P3: The use of dictionary checks can motivate users to base their passphrases on multiple languages.*

Based on the propositions above, a passphrase security model for a multilingual environment is shown in Figure 10.



**Figure 10. A passphrase security model for a multilingual user group**

### 5.3.2 Password usability

Section 4.4 of this study discussed how a bias towards passphrase security affects usability. Section 4.4.2 went on to expound the way in which some practices, assumed to be usable by users, have ricochet effects on passphrase security. As such, this study concedes that usability and security are equally important. The next section discusses factors of passphrase usability. By so doing, the section addresses the third research sub-question of this study which was formulated as follows:

  
 University of Fort Hare  
*What are the factors affecting the usability of passphrases?*

#### 5.3.2.1. Factors of passphrase usability

The definition and factors of usability in the ISO 9241-11 standard were adopted and used to guide passphrase usability in this study. The ISO 9241-11 is a tried and tested standard both in the industry and academia; hence, it is expected to provide balanced and complete measures for evaluating passphrase usability (Bevan, Carter, & Harker, 2015). The ISO 9241-11 standard defines usability as “the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use” (Bevan, Carter, Earthy, Geis & Harker, 2016, p. 269). This study refers to the latest definition of usability that is applicable to services. Old versions of the ISO 9241-11 suggest the standard was once limited to products. As required by the definition of usability, this study concedes that all users of password authentication mechanisms with multilingual skills are its

system users. Further, this study agrees with research findings by Choong et al. (2014) that irrespective of the importance of passphrase security to the systems administrator, generating a memorable passphrase that could be typed in easily, is a primary goal of system users.

It, therefore, was considered that users should be able to generate memorable passphrases that they can type in with effectiveness, efficiency and user satisfaction if they are to attain passphrase usability. Section 4.4 discussed different items in the literature that can be used to ascertain effectiveness, efficiency and user satisfaction. These included time taken to create a password; the use of semantic information during passphrase creation; password reuse; copy and paste; number of deletions during passphrase generation; passphrase storage on other media; the number of passphrase creation attempts; passphrase re-entry; attempts needed to correctly type in a passphrase; passphrase creation failure; failure to recall the passphrase and time spent entering passphrases (Keith et al., 2007; Melicher et al., 2016; Shay et al., 2016). Bevan et al. (2016) and Lund (2001) also define additional items that could be adapted and used as measures of usability. Some of these usability measures have received thorough validation as they have been tested on different password policies and different devices over a considerable amount of time (Melicher et al., 2016; Shay et al., 2016). These measures were combined and some were consolidated to arrive at items for ascertaining effectiveness, efficiency and user satisfaction for this study. Table 3 shows the usability factors for this study in accordance with the revised ISO 9241-11 standard.

**Table 3. Factors of passphrase usability considered for this study**

Factors	Items	Source
Effectiveness	Accurately recalling passphrases	Choong et al., (2014); Keith et al. (2007); Shay et al. (2016)
	Accurately generating a passphrase	
	Accurately typing in a passphrase	
	Meeting passphrase requirements	Shay et al. (2016)
	Passphrase creation failure: failure to meet requirements and passphrase mismatch	Melicher et al. (2016); Shay et al. (2016)
	The use of a passphrase reminder	Shay et al. (2016)
	Storage of passphrases on other media.	Choong et al. (2014); Melicher et al. (2016); Shay et al. (2016); Ur et al. (2016)
	Use of semantic information in passphrases.	AlSabah et al. (2018); Bonneau and Xu (2012); Shay et al. (2010); Wang et al. (2015)
	Passphrase reuse	Helkala & Bakås (2014); von Zezschwitz et al. (2013)
	Failure to recall passphrases	Stobert & Biddle (2014); Shay et al. (2016)
	Easy to use	Lund (2001)
	Simple to use	Lund (2001)
	Requires as few steps as possible to accomplish the task	Lund (2001)
	Flexible	Lund (2001)
	Effortless	Lund (2001)
	One can recover from mistakes quickly and easily	Lund (2001)
Efficiency	Time taken to type in a passphrase	Melicher et al. (2016); Shay et al. (2016)
	Passphrase creation attempts	Melicher et al. (2016); Shay et al. (2016)
	Passphrase recall attempts	Melicher et al. (2016); Shay et al. (2016)
	Time taken to generate a passphrase	Choong et al. (2014); Melicher et al. (2016)
User satisfaction	Overall satisfaction	Bevan et al. (2016); Lund (2001)
	Satisfaction with features	Bevan et al. (2016)
	It is pleasant to use	Bevan et al. (2016); Lund (2001)
	Feeling comfortable about the policy	Bevan et al. (2016)
	Recommending a friend	Bevan et al. (2016); Lund (2001)
	Finding the process fun	Lund (2001)
	Works according to user expectations	Lund (2001)
	The process is wonderful	Lund (2001)
	Willingness to utilise the password policy	Lund (2001)

**Effectiveness.** According to the ISO 9241-11 standard, effectiveness relates to the accuracy and completeness with which users achieve their goals, without experiencing negative consequences. Hence, a user's ability to accurately and completely generate, memorise and type in a multilingual passphrase involves items contributing to usability effectiveness. Users are also expected to meet all the multilingual passphrase requirements in terms of length, spacing substrings and using words from different languages. Potential negative consequences expected from the



generation, memorising and typing of multilingual passphrases include failure to create passphrases; the use of semantic information: adapting a date of birth, using a phone number, using native words, adapting a name, using English words, adapting an address, spelling abbreviations, using multilingualism, using non-standard spelling; passphrase reuse; failure to recall a passphrase; using a passphrase reminder; passphrase storage on other media: writing down passphrases, sharing with a friend, saving in computer and saving on a browser.

A review of the literature on psychological studies in Chapter 3 suggested that poorly designed passphrase policies make it difficult for users to easily exploit their short-term memory when generating and recalling passphrases. This study encourages users to generate multilingual passphrases based on their daily skills of code-switching in order to maximise passphrase generation, memorability and typing effectiveness with the aim of reducing any potentially negative consequences. The Usefulness, Satisfaction and Ease of use (USE) questionnaire proposed by Lund (2001) recommends different items for evaluating ease of use. Some of these items were adapted in this study as part of effectiveness. This allowed the study to determine how users could access their memory easily and effectively during password generation and recall. User convenience has been found critical during password generation and recall (Woods & Siponen, 2019). The items that were adapted from Lund (2001) are: “it is easy of use, it is simple to use, it requires as few steps possible to accomplish what I want to do with it, it is flexible, using it is effortless, I can recover from mistakes quickly and easily”. Table 3 shows items for evaluating effectiveness. It is therefore proposed that:

*Proposition P4: The ability to effectively generate, memorise and type in a multilingual passphrase without experiencing negative consequences enhances passphrase usability.*

**Efficiency.** The ISO 9241-11 standard defines efficiency as the amount of resources used to achieve an objective or goal. Within the context of this study, these resources include the time taken to generate a passphrase, the number of deletions during passphrase generation, the number of passphrase creation attempts, time spent

entering a passphrase and re-entries needed to correctly type in a passphrase as shown in Table 3. It is therefore proposed that:

*Proposition P5: Efficacy in multilingual passphrase generation, recall and typing in leads to passphrase usability.*

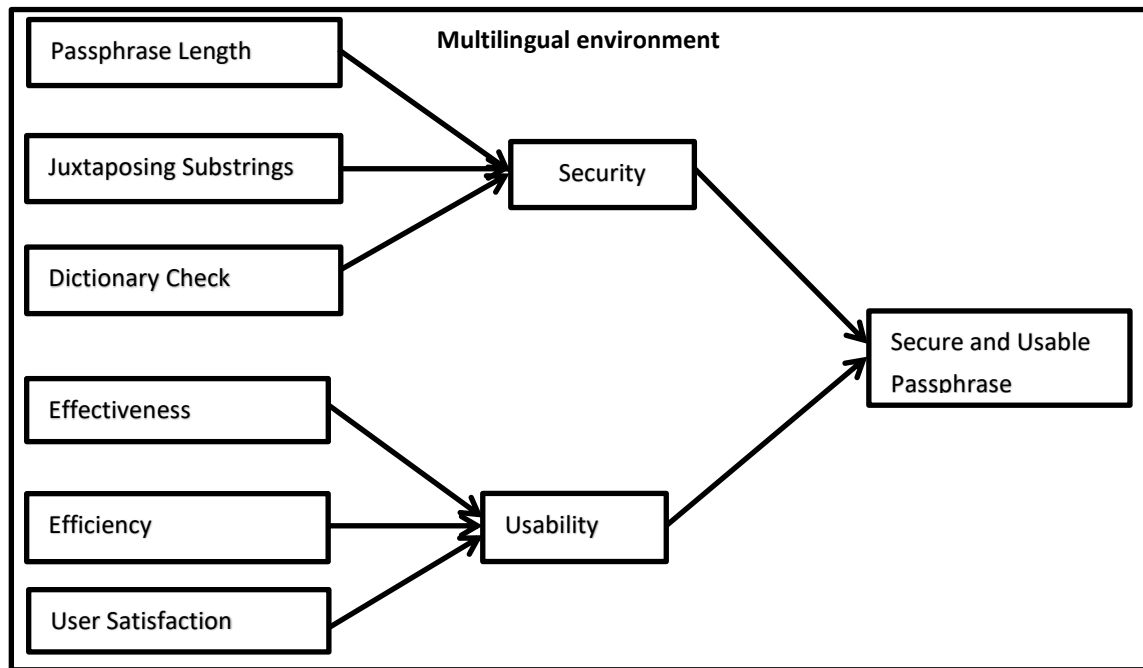
**User satisfaction.** Bevan et al. (2015) suggest that user satisfaction is the “positive attitudes, emotions and/or comfort resulting from” generating, recalling and typing in passphrases following a passphrase policy (p. 3). To our knowledge, very few if any studies have used the ISO 9241-11 standard to evaluate password usability. As a result, the majority of usability items in the literature are biased towards evaluating effectiveness and efficiency with little or no focus on items for evaluating user satisfaction. Some of the user satisfaction items in Bevan et al. (2016) and the USE questionnaire were adapted for this study. These include users’ overall satisfaction, satisfaction with a password policy features, whether the policy is pleasant to use, comfortable with the policy, a willingness to recommend the policy to a friend, finding the policy fun, the policy works the way users expect it to work, finding the policy wonderful and the willingness to utilise the password policy (Bevan et al. 2016; Lund, 2001). These items are summarised in Table 3. Participants who are frustrated by a complex password policy are expected to show a negative attitude towards the policy and a lack of interest in it on other platforms that use text-based authentications. Hence, showing a positive attitude towards items for evaluating password policy user satisfaction reflects on usability. Based on these arguments, it is therefore proposed that:

*Proposition P6: User satisfaction with a multilingual passphrase policy leads to passphrase usability.*

#### **5.4 The proposed model**

The usability and security propositions in this chapter led to a model shown in Figure 11. The past decades have seen a number of policy propositions for increasing the size of the search space from which unique passphrases could be drawn. For

example, “enforcing the inclusion of numbers and special characters, requiring both upper- and lower-case letters, and increasing minimum password lengths” (White, Shaw, Monroe, & Moreton, 2014, p. 1), while others recommend less use of predictable grammatical structures in passphrases (Rao et al., 2013), to mention but a few. Despite all these efforts, the security and usability of short passwords and passphrases remain questionable (Andersson & Saedén, 2013; Bonneau & Shutova, 2012; Kelley et al., 2012; Rao et al., 2012). As such, this chapter motivated the use multilingual passphrases. The study adopted principles in Shannon’s entropy as used by the NIST and proposed basing passphrases on multilingual substrings instead of multi-character sets. The use of multilingual substrings is expected to increase the size of the search space from which multilingual passphrases can be drawn, without compromising usability and security. Such policies are applicable in sub-Saharan Africa, for instance, where the majority of people do not use their first spoken language (African languages) as the official language – a move that has created a multilingual environment (Lexander, 2011). In addition, Kang's (2015) admission that globalisation is fast promoting a multilingual environment and his subsequent research findings suggest that the success in terms of usability and security of a text-based authentication mechanism may lie in policies that are driven by contextualised factors such as language. Figure 11 shows constructs of usability and security in the proposed model. Such a model can be used to inform the designing of passphrase policies for multilingual societies.



**Figure 11. A proposed model of usable and secure passphrases for a multilingual user group**



## 5.5 Chapter summary

This chapter identified a research gap in the literature on short password and passphrase policies. It was noted that the security and usability contributions of passphrases are yet to be thoroughly validated. As such, the chapter proposed the use of multilingual passphrases with the aim of enhancing passphrase security without compromising usability. The ISO 9241-11 standard was used to depict factors of multilingual passphrase usability. In particular regarding passphrase security, this chapter motivated the need to increase the size of the passphrase search space by using more than one language during the generation of passphrases. Socio-cultural theory suggests that users can learn and understand different languages which can be used simultaneously as revealed by the practice of code-switching. As such, this chapter presented its argument on how multilingual passphrases can address challenges that are herein considered a factor in generating passphrases using a single language. The chapter proposed dictionary checks, the use of juxtaposed substrings and passphrase length as major factors that could be looked at in order to enhance passphrases security in a multilingual environment. Furthermore, the ISO 9241-11 standard on usability was

used to provide guidance in identifying factors of usability. This standard has been tried and tested; hence, it can provide a complete set of factors for ascertaining the usability of multilingual passphrases. These factors include effectiveness, efficiency and user satisfaction. Extending a common practice of code-switching in a multilingual environment to authentication is expected to enhance the usability of passphrases. In conclusion, it is argued that juxtaposing during passphrase generation can be done with effectiveness and efficiency, thereby resulting in user satisfaction.

The proposed model in this chapter was demonstrated using an experiment and was subsequently evaluated to establish its utility over the popular short password policy. Appendix A presents the experiment protocol that was used to demonstrate the proposed model. The next chapter presents findings from the demonstration of the proposed model.



University of Fort Hare  
*Together in Excellence*

## CHAPTER 6: RESEARCH FINDINGS

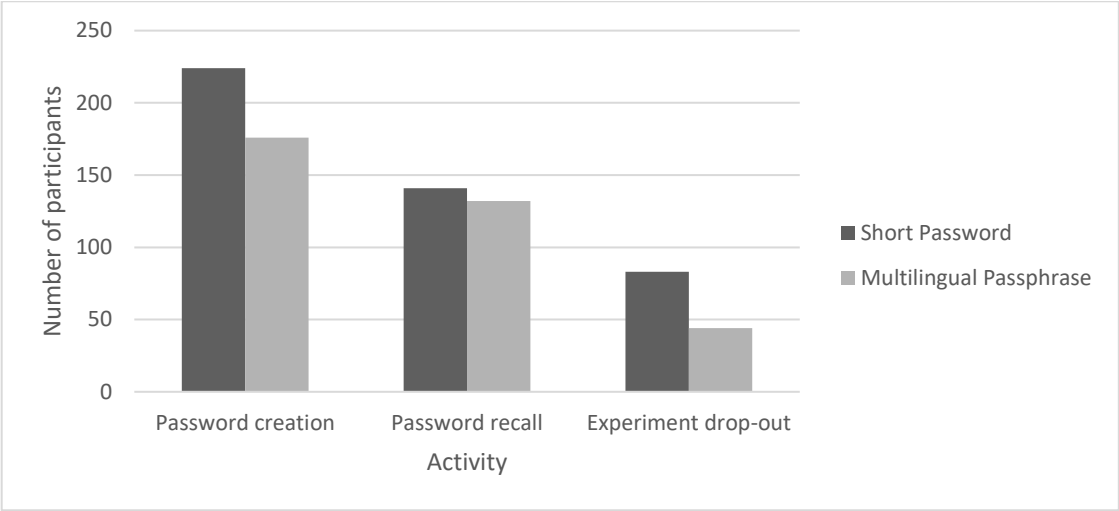
### 6.0 Introduction

The previous chapter presented a model for usable and secure multilingual passphrases. This chapter adds to the study by presenting the results of the data collection. The findings presented in this chapter are used to evaluate the model proposed in Chapter 5, which argues in favour of multilingual passphrases that supplant short passwords as they are seen to be less secure and not user-friendly. Nevertheless, this study gathered data on short passwords and multilingual passphrases with the aim of making security and usability comparisons between the two samples (short passwords and multilingual passphrases). The next section of this chapter presents the sample profile. The data collection instrument used in this study is outlined. The chapter goes on to give an account of contextual findings that reflect the propositions in the socio-cultural theory. Findings on contextual factors are angled at exposing the influence of language characteristics on password generation. An evaluation of the reliability and validity of the data collection instrument follows. This activity then validates the data collection instrument that was used to gather data for assessing the usability of short passwords and multilingual passphrases. An outlay of short passwords and multilingual passphrase usability and security concludes the chapter.

### 6.1 Sample profile

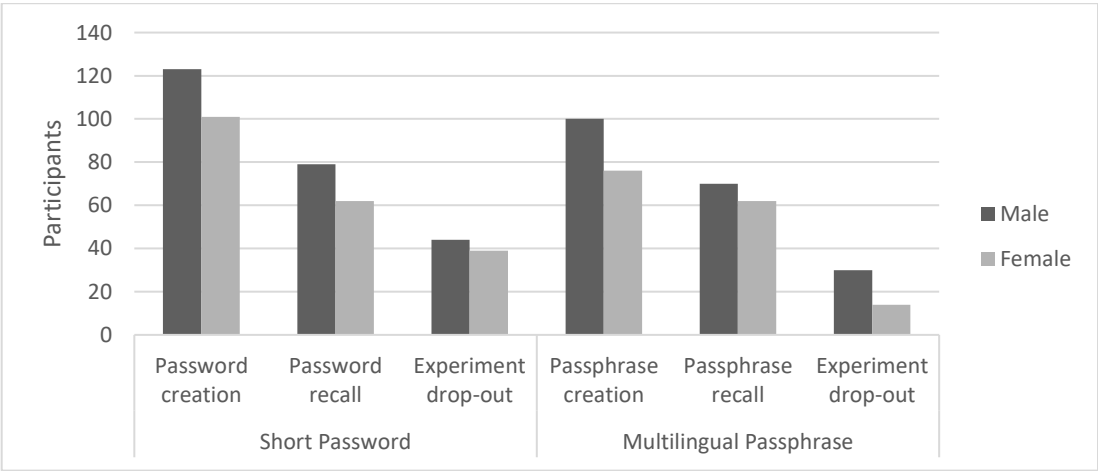
This section reports on the demographics of participants. Two hundred and twenty-four (224) students participated in the experiment but only 112 completed both sets of the experiment: short passwords and multilingual passphrases. This participation rate suggests that 50% of the participants dropped out during the course of the experiment. Shay et al. (2016) experienced a fairly similar dropout rate in their password experiment, as 41% of their participants failed to complete the experiment. Closer analysis of participants within the experiment categories of short passwords and multilingual passphrases shows that 224 participants created a short password, but only one 141 of these went on to participate in short password recall. This suggests that 37% of the participants dropped out during short password creation and recall. In the case of multilingual passphrases, 176 participants created a passphrase and only 132 of these

took part in the passphrase recall exercises. Figure 12 summarises these results. It should be noted that some participants participated in either one or both of the short password and multilingual passphrase experiments.



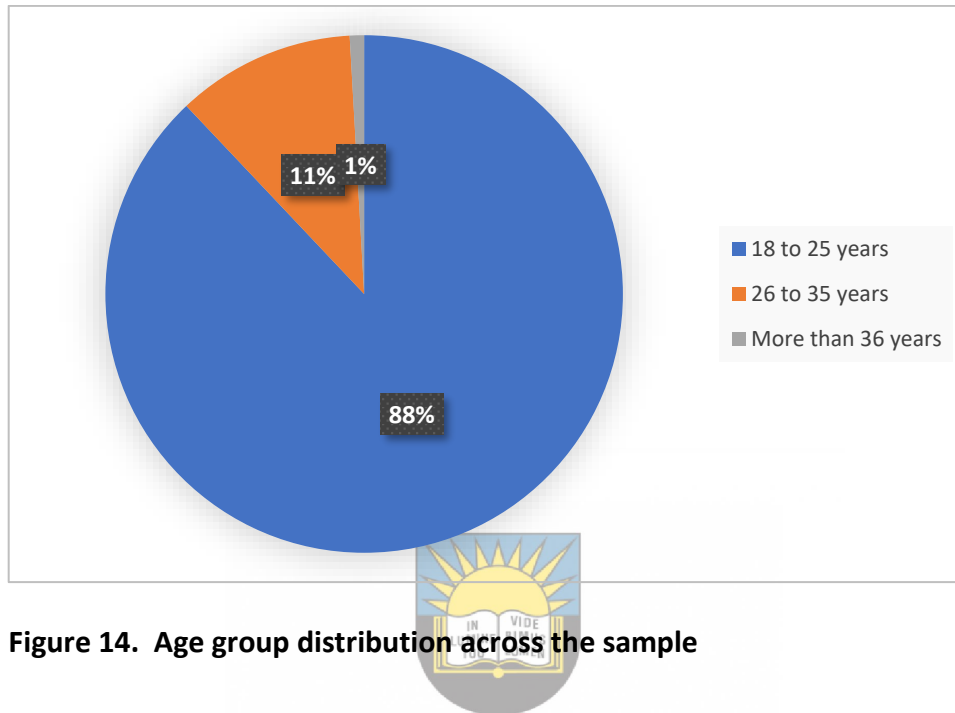
**Figure 12. Participants that took part in the short password and multilingual passphrase experiments**

An analysis of gender distribution across participants shows that 54.85% were male participants while 44.15% were female. Figure 13 shows the distribution of participants according to gender throughout the experiment. It can be seen that male participants dominated across short passwords and multilingual passphrase, creation and recall.



**Figure 13. The distribution of gender across password experiments activities**

Data about the age distribution of participants was gathered. Results show that the majority (88%) of participants were within the age group of 18 to 25 years old, as shown in Figure 14. This statistic can be explained by the fact that only university students were engaged during the experiment.



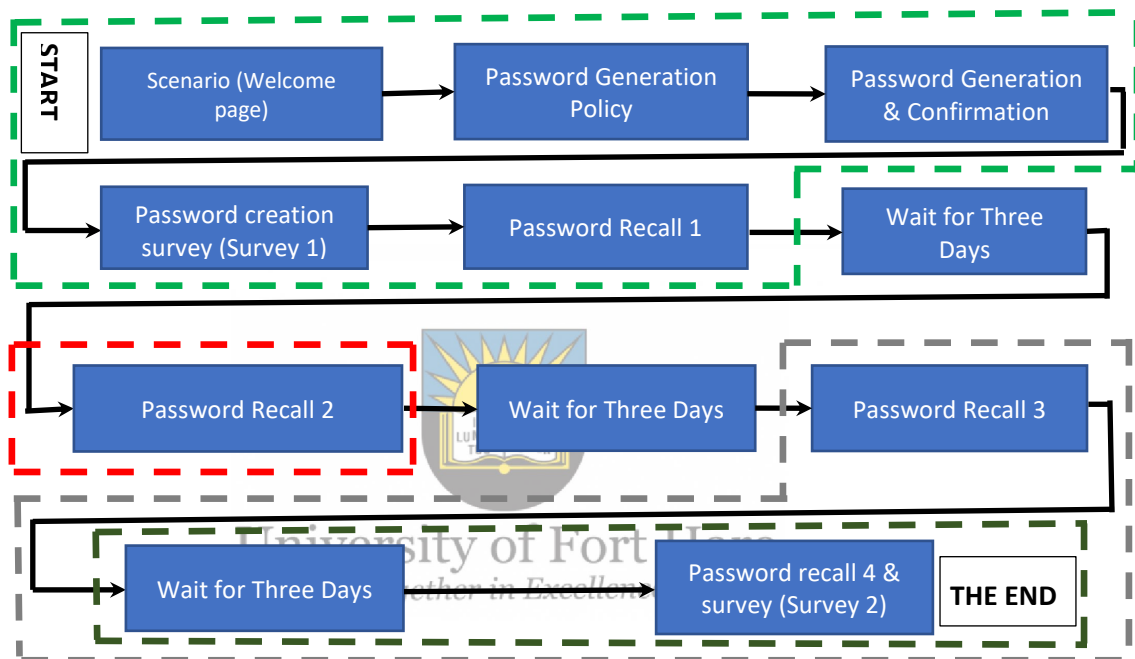
**Figure 14. Age group distribution across the sample**

## 6.2 The data collection instrument

Chapter 5 proposed a model for secure and usable multilingual passphrases. In addition, Chapter 2 indicated that a mixed research method was used for data collection. This involved a triangulation of data collection methods for gathering quantitative and qualitative data. Related studies on passwords make use of a similar approach for data collection where questionnaire and system generated data is gathered (Melicher et al., 2016; Shay et al., 2016). Data was gathered by means of an experiment which was designed in such a way that a participant would generate a password following specified password requirements, complete a questionnaire (Survey 1) and login into their profiles (Password Recall 1) on the first day of the experiment. Participants were asked to wait for three days before being invited to make a second login (Password Recall 2) which was followed by waiting for another three days before making the next login attempt (Password Recall 3), as shown in Figure 15. The end of the experiment was marked by logging in for a fourth time and completing a questionnaire for evaluating password recall usability (Survey 2). The idea was to gather



data for testing how user-friendly it was for participants to generate and recall short passwords and multilingual passphrases over a period of two weeks. Short passwords and multilingual passphrases were generated and used in succession, thus participants generated short passwords and used them over a period of two weeks following the steps outlined in Figure 15. The following two weeks saw the same participants generating multilingual passphrases and repeating the activities outlined in Figure 15. Appendix A provides detailed scenarios of activities summarised in Figure 15.



**Figure 15. Activities of the study experiment and data collection**

The proposed model in Chapter 5 contains three factors of usability, namely, effectiveness, efficiency and user satisfaction. This section shows how the items under these usability factors are linked to questions in the questionnaire and to qualitative data collection methods used in this study.

### 6.2.1 Data collection methods for evaluating effectiveness

Table 3 in Chapter 5 identified items for password (short passwords and multilingual passphrases) effectiveness usability. The updated ISO 9241-11 standard goes on to suggest that failure to generate and recall passwords effectively leads to negative consequences that should be evaluated. In particular to this study, the

negative consequences include password reuse; the use of semantic information in passwords (adapting a date of birth, using a phone number, using native words, adapting a name, using English words, adapting an address, using spelling abbreviations, using multilingualism, using non-standard spelling); failing to accurately and completely recall passwords, and storage of passwords on other media (writing down passwords, copy and paste, and automatic text entry). These items were used to construct questions for gathering data to evaluate usability in terms of password creation and recall effectiveness.

**Password creation effectiveness (PCE):** Table 4 shows the link between the questions in the questionnaire and items of password creation effectiveness that were identified in Chapter 5.

**Table 4. Questions for gathering data on PCE**

Item	Question	Item labelling
Easy to use	Creating a password for this study was easy.	PCE1
Simple to use	Creating a password for this study was simple.	PCE2
Meeting passphrases requirements	Password creation requirements for this study are user-friendly.	PCE3
Requires as few steps as possible to accomplish the task	It required few steps for me to create a password for this study compared to when I use other policies.	PCE4
Flexible	Password creation requirements for this study were flexible.	PCE5
Effortless	Creating a password for this study was effortless.	PCE6
One can recover from mistakes quickly and easily	I could quickly and easily recover from the mistakes I made during password creation.	PCE7
Accurately generate passphrase	I can successfully create a password using requirements specified for this study every time.	PCE8

In addition, data for evaluating the failure to effectively create a password was gathered using questions in Table 5. The gathered data focused on incidences of password reuse and the use of semantic information in passwords.

**Table 5. Password generation strategies that could reflect failure to effectively generate a short password or multilingual passphrase**

Strategy	Question
Password reuse	I created my password based another password I already know.
Adapting a date of birth	The password I created is based on a date of birth.
Used a phone number	I created my password based on a phone number.
Adapted native words	I created my password based on words in my mother tongue.
Adapted a name	I created my password based on the name of someone or something I know.
Used English words	I created my password based one or more words in English.
Adapted an address	I created my password based on an address that I know of.
Used spelling abbreviations	I created my password using spelling abbreviations (slang or colloquial terms).
Used multilingualism	The password I created is based on words written using different languages.
Used non-standard spellings	I created my password using non-standard spelling.

**Password recall effectiveness (PRE):** Table 6 shows the link between the questions in the questionnaire and items of password recall effectiveness. Focus is on questions used to gather data for evaluating one's ability to accurately recall a short password and multilingual passphrase. It should be noted that system-generated data was gathered and used to evaluate the ability of participants to accurately type in a password. For every logging in attempt, the web application platform gathered user-generated data that included the actual password that was used in a login attempt, timestamps for each keyboard key presses (DD, H, DU) and an indication of whether logging in was successful or failed.

**Table 6. Questions for gathering data on PRE**

Item	Question	Item labelling
Easy to use.	Recalling a password for this study was easy.	PRE1
Simple to use.	Remembering a password for this study was simple.	PRE2
Requires as few steps as possible to accomplish the task	It required a few steps for me to remember a password for this study when compared to when I create a password using other policies.	PRE3
Effortless	Remembering a password for this study was effortless.	PRE4
One can recover from mistakes quickly and easily	I could quickly and easily recover from the mistakes I made during logging in.	PRE5
Accurately recalling passphrases	I could successfully remember the password I was using for this study every time.	PRE6

In addition, data for evaluating the consequence of failing to effectively recall passwords was gathered. This is in line with propositions in the updated ISO 9241-11 standard. The focus was on gathering data that reflected failure to accurately and completely recall short passwords and multilingual passphrases, and storing these on other media. Participants who required more than four login attempts were considered to have failed to accurately and completely recall their passwords (Shay et al., 2016). Data for evaluating failure to accurately and completely recall passwords was gathered using the key logs generated as participants logged into their profiles. In addition, data for evaluating the use of other storage media by participants during the experiment was gathered using the questions in Table 7.

**Table 7. Password recall strategies that reflect failure to effectively recall a short password or multilingual passphrase**

Password recall strategy	Question
Memorised	I managed to memorise and remember the password I was using for this study.
Shared with friend	I had to share my password for this study with a colleague in case I forgot it.
Saved on the computer	I wrote and saved my password somewhere on the computer because I could not remember it.
Saved on the browser	I saved the password I used for this study on the internet browser because I could not memorise it.
Saved on the phone	I had to save my password for this study on my mobile phone in case I forgot it.
Wrote down	I had to write down my password on a piece of paper in case I failed to remember my password.

### 6.2.2 Data collection methods for evaluating efficiency

Table 3 in Chapter 5 identified the following as items for password efficiency usability:

- Password creation attempts
- Time taken to create a password
- Time taken to type in a password
- Password recalling attempts

Data for assessing the efficiency usability was gathered by means of the use of keystrokes. As indicated earlier, data on timestamps for each keypress during password

generation and recall was gathered. This implies that the raw candidate passwords that were attempted during generation and recall were captured. This data was used to derive password creation and recall attempts. The data for evaluating the time taken to create and type in a password was also gathered using keystrokes as participants went about the activities of password generation and logging in. Data on the timestamps (DD, H, DU), explained in Section 2.4.2.2; was used to arrive at the time taken to create and type in a password.

### 6.2.3 Data collection methods for evaluating user satisfaction

Table 3 in Chapter 5 of this study identifies items for measuring user satisfaction. These items were used to derive questions for gathering data that was used to assess user satisfaction with password creation and recall. Tables 8 and 9 shows the link between user satisfaction items and the respective questions.

**Table 8. Questions for gathering data on password creation user satisfaction (PCUS)**

Item	Question	Item labelling
Overall satisfaction	I was satisfied with the password creation process for this study.	PCUS1
Recommending it to a friend	I would recommend the password creation process for this study to a friend.	PCUS2
Finding the process fun	Creating a password for this study was fun.	PCUS3
Works according to user expectations	Password creation process for this study works the ways I want it to work.	PCUS4
The process was wonderful	Creating a password for this study was wonderful.	PCUS5
Willingness to utilise the password policy	I would prefer to use this study's way of creating a password on Facebook and my email.	PCUS6
It is pleasant to use	It was pleasant to create passwords for this study	PCUS7

**Table 9. Questions for gathering data on password recall user satisfaction (PRUS)**

Item	Question	Item labelling
Recommend it to a friend	I would recommend the password creation process for this study to a friend because it helps one create a password that is easy to remember.	PRUS1
Overall satisfaction	I was satisfied with recalling passwords for this study.	PRUS2
Works according to user expectations	The process of remembering the password I created for this study occurred the way I wanted it to occur	PRUS3
Willingness to utilise the password policy	I would not need to write down or store my passwords as much if I always used a password format like the one I used for this study.	PRUS4
Comfortable about the policy	I could remember more passwords at once if I always used a password format like the one I used for this study.	PRUS5
Satisfaction with features	The password format I used for this study helped me create a password that was easy to remember.	PRUS6

### 6.3 Social context overview

Data was gathered on demographics to establish characteristics of the social context of this study. The majority of participants who created passwords were South Africans (41%) followed by Namibians (31%) and Zimbabweans (26%). Understanding of the social context was used to evaluate the principles of socio-cultural theory, namely, the generic law of development, mediation and generic domains. This study argued that contextual factors influence psychological development as suggested by socio-cultural theory. Hence, computer users are expected to reflect contextual factors by orienting user-generated passwords to languages and following cultural practices that are common within their contexts. Findings on socio-cultural theory are presented next.

#### 6.3.1 The generic law of development

To evaluate the generic law of development, this study used data on users' computer skills, first language, second language and ethnic grouping. Participants were asked to rate themselves on a scale of 1 to 5 to determine their level of computer skills. Exactly 41% indicated that they were novice users, 38% suggested they were intermediate, while 11% suggested that they were advanced computer users. The remaining participants (10%) indicated that they were below average to novice computer users. These results can be explained by the fact that the experiment

targeted information systems, computer science and informatics students who had access to computers while at the university.

In terms of spoken and written language, 36% of those who created short passwords indicated that their first language was isiXhosa. IsiXhosa is a dominant language in the targeted province of South Africa where the experiment was carried out. Only 25% indicated that they speak Shona while 18% selected Oshiwambo as their first spoken language. Shona is a dominant Zimbabwean language while Oshiwambo is a popular language among the Vambo tribe, a dominant Namibian tribe. The remaining participants (21%) represented different native Zimbabwean, Namibian, South African and Indo-European languages in small percentages. Of all the participants analysed, the dominant second language was shown to be English (94%). These findings suggest participants are multilingual users with a greater chance of speaking and writing African language or an Indo-European language. The dominance of English as a second language can be explained by the fact it is the first written and official language of the targeted countries (Namibia, South Africa, Zimbabwe).



### **6.3.2 Mediating symbolic tools and password characteristics**

This section reflects on short password and passphrase characteristics that portray the preferred mediational symbolic tools within the researched context. The evaluation of the mediational symbolic tools used is reported according to password type: short passwords and multilingual passphrases.

#### **6.3.2.1 Short password characteristics**

The data to evaluate the short password characteristics was gathered using questions in Table 5. In addition, raw passwords were also gathered for further characteristics analysis. Accordingly, short password characteristics were evaluated in terms of the use of semantic information, password structures, password length and the use of popular passwords. Figure 16 is a visual representation of the web application platform where participants generated short passwords following specified requirements.

Home Join Sign in

Create a new password following the password requirements

New Password\*

\*New Password

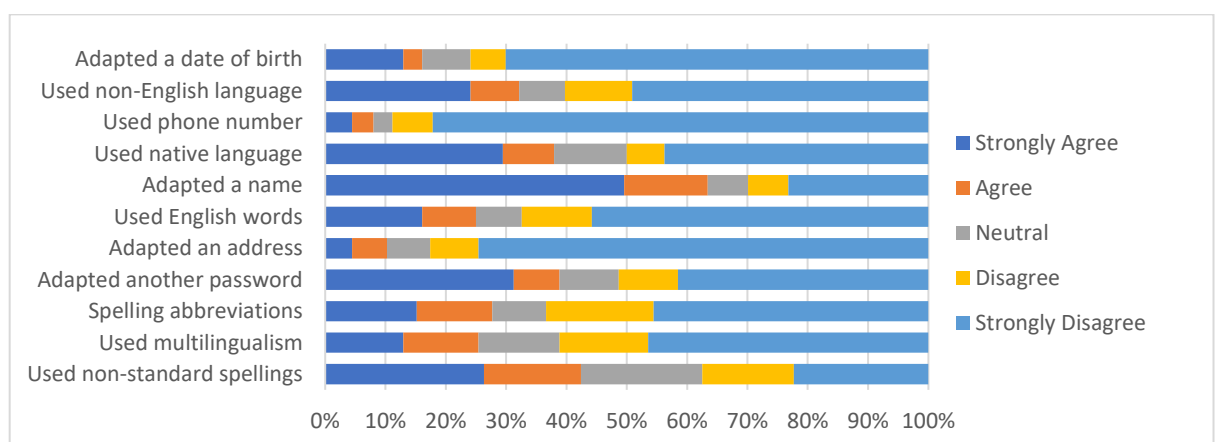
- Must be at least 8 characters long, without any spaces
- Must contain at least one capital letter, one lowercase letter, one number and one special character (symbols like &, \$, @, #, ! and \*)
- Must not contain your username or personal details

Confirm Password

Sign up Clear

**Figure 16. The short password generation platform.**

**The use of semantics.** A questionnaire and raw passwords were analysed to establish the use of semantic information in password generation. Results from a questionnaire presented in Figure 17 show that close to 50% of the participants strongly agreed that they adapted names when generating short passwords, while just above 30% adapted an existing short password. For every new short password, there was approximately a 30% chance that the resulting short password would be oriented towards an African language. Given the prevalence of adapting names as short passwords, it can be argued that the use of an African language was a result of users adapting African names instead of using words in African language. Furthermore, just above 25% of the participants claimed to have used non-standard spellings in short password generation.



**Figure 17. Short password generation strategy**



A content analysis on the short password corpora supported the above findings on the use of semantic information when generating short passwords. Strings that appeared as names and words were found popular among semantics used by participants during short password generation as shown in Table 10.

**Table 10. Observed semantics used in user-generated short passwords**

Semantics used	Total	Percentage
English_word	46	20
English_name	28	12
English_phrase	9	4
Random	30	13
isiXhosa_name	29	13
isiXhosa_word	13	6
Shona_name	17	7
Shona_word	6	3
Multilingual	12	5
Oshiwambo_name	9	4
Oshiwambo_word	2	1
KB_pattern	3	1
Place_name	3	1
Website_name	3	1
Other	14	6

Table 10 also shows the dominance of orienting passwords to English words. However, names oriented to African languages (isiXhosa, Oshiwambo and Shona) together contributed a larger proportion of all short passwords.

**Table 11. Common password structures observed in the short password corpora**

Password Structure	Total	Percentage
LDS	48	21
LSD	51	23
SLD	13	6
LD	38	17
Other	74	33

**Password structure.** User-generated short passwords were evaluated to establish common password structures. These password structures were classified according to character composition such as L: alphabetic letter, D: digits and S: symbol

(Weir et al., 2009). For example, password Favour@123 can be reduced to  $L_6S_1D_3$  which translates to an LSD password structure. Table 11 shows the dominant password structures that were found in the short password corpus.

A review of digits and symbols used in the short password corpora was done. The selection of digits in 27% of the reviewed short passwords appeared to be based on semantics. The majority (10%) of short passwords had digits that resembled the year of birth of the participant as they were written like 1995 or 95. Other participants (5%) included a two-digit number between 18 and 25, a figure that was assumed to resemble the participant's age. Our reasoning is based on the fact that most participants indicated that they were of the age group 18 to 25 years old. In addition, 6% included a four-digit number that resembled a year, for instance 2017, while 4% of the participants included a keyboard pattern with at least three digits following a particular pattern. Common keyboard patterns of digits included a 123 or 777. In addition, participants did not show random selection of symbols for their passwords. The majority (42%) of the participants included the "@" symbol followed by those who used a "#" (15%) in their short passwords. A significant number (7%) of participants used a "!" and another 7% used a "\$" as a symbol in their short passwords. These findings clearly demonstrate that users prefer a small pool of potential symbols when generating short passwords.

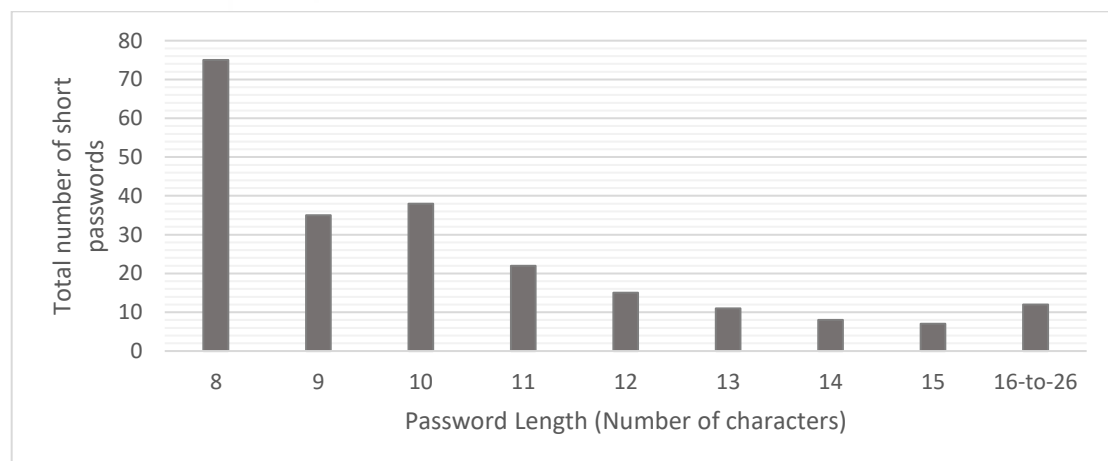
Furthermore, user-generated passwords were analysed to establish whether participants had arrived at their short passwords by making spelling mistakes, insertions, concatenating different character classes and replacing different character classes (Jakobsson et al., 2013). Spelling mistakes are difficult to identify and users often do this with or without intent. For instance, writing the word "password" as "passwrod". Insertion may involve placing a block of unique characters inside a string. For instance, inserting the number 77 into the name Christina to come up with a password like "Christi77na" (Jakobsson & Dhiman, 2013). Concatenation involves combining different strings together. For example, combining the strings "password" with "@" and "2019" in order to generate a password like "password@2019". Lastly, replacement involve users replacing different character classes, for example LEET could

be written as L33T (Jakobsson & Dhiman, 2013). These strategies can also be used to explore how complex a password is depending on the assumed technique when generating a short password. Results in Table 12 show that concatenation is the most dominant technique used by participants to arrive at a short password. Findings on the popularity of concatenating different character classes were corroborated by the popularity of LDS or LSD and SLD password structures in Table 11.

**Table 12. Short password generation techniques assumed by participants**

Password Generation Technique	Example	Frequency	Percentage
Concatenation	Favour@123	178	79
Insertion	M201401521s	13	6
Replacement	P@55w0rd	24	11
Random	QH0xaH619	9	4

**Password length.** An analysis of password length was done. While participants were required to generate an eight-character password, a significant number of participants generated short passwords that were more than eight characters long. Figure 18 provides a breakdown of the length of all short passwords analysed in this study.



**Figure 18. Password length of short passwords**

**The use of popular passwords.** Data was also gathered to establish whether short passwords were similar to global passwords. The focus was on establishing

whether a subset of at least four successive characters or the whole of the password appeared among the top-100 most common passwords of 2016, 2017 and 2018, according to SplashData (2016, 2017, 2018). Shay et al. (2014) used a similar approach to identify the most common substring within their password corpora. Close to 3% of the short password corpora was found to have substrings that appeared in the list of the most common passwords. These passwords are shown in Table 13 together with a popular password to which they are compared. Other short passwords in the password corpora had various mangling rules that distanced them from selected passwords in the list of most common passwords. For example, short passwords such as “P@55w0rd777”, “P@5sword” and “Pa\$\$word2” can be traced back to password or password1 which are in the 2016, 2017 and 2018 lists of most common passwords.

**Table 13. Common substrings that were found among the most common passwords of 2016, 2017 and 2018**

Substrings observed in the short password corpora	Related common passwords of 2017	Position on the most common passwords list
Qwerty123#	Qwerty	4
12345ABcd!@#\$	12345	5
Ilove!1995	Iloveyou	10
Fucku@1992	Fuckyou	52
fuCKuRobson5@	Fuckyou	52
Pewdiepiebananasoda123\$	Banana	61

#### 6.3.2.2 Multilingual passphrase characteristics

This section presents multilingual passphrase characteristics that reflect preferred symbolic tools within the context. The focus is on the use of semantics, passphrase structures, passphrase length and the use of common substrings. Figure 19 shows a visual representation of the web platform that was used to generate multilingual passphrases together with the specified requirements.

Home Join Sign in

Create a new password following the password requirements

New Password\*

\*New Password

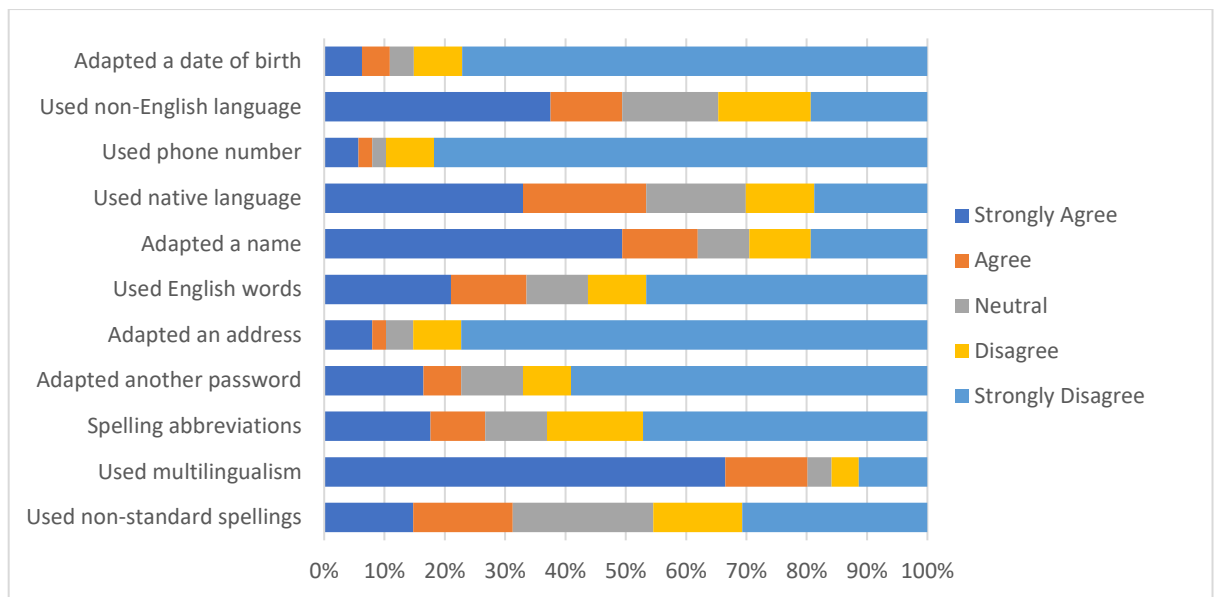
- Have at least two words.
- The words making a password should be from at least two different languages for example an English word and the other can be an Afrikaans word
- The words making a password should be separated by at least one space or non-letter sequence.
- The password should have at least sixteen characters

Confirm Password

Sign up Clear

**Figure 19. Multilingual passphrase generation platform**

***The use of semantics.*** After completing two weeks of the short password experiment, participants were asked to generate a multilingual passphrase basing their passphrases on substrings from at least two different languages. Participants responded positively to this requirement, as reflected by more than 65% who strongly agreed that they had based their passphrases on multilingualism. Close to 50% of the participants adapted names during multilingual passphrase generation. A similar observation was made on short passwords. However, there was a fair share of participants who oriented their multilingual passphrases towards non-English languages (more than 35%), while just above 20% of the participants adapted English words in their passphrases. Of interest is an observation that less users (15%) adapted existing passwords when compared to short passwords. This could be explained by a paucity of password policies encouraging users to generate passphrases. Figure 20 summarises these results.



**Figure 20. Multilingual passphrase generation strategy**

A further content analysis was done on the multilingual passphrase corpora. Our observations show that the majority of multilingual passphrases had a substring that is oriented to an English word (56%). Such a multilingual passphrase could be a combination of an English word and an Afrikaans word (6%) or an Afrikaans word and an English word (2%). In addition, some of the multilingual passphrases had African language-oriented word juxtaposed with an English language-oriented word (7%) or vice-versa (9%). It was also found that substrings that resembled a name (48%) represented a significant proportion in the multilingual passphrase corpora. These names could be a participant's full name (English name and an African surname), which was observed in 11% of the whole multilingual passphrase corpora, or the name of a known place (3%). Approximately 29% of the African names were observed in the multilingual passphrase corpora. Only 13% of the observed substrings in the multilingual passphrase corpora resembled Afrikaans words.

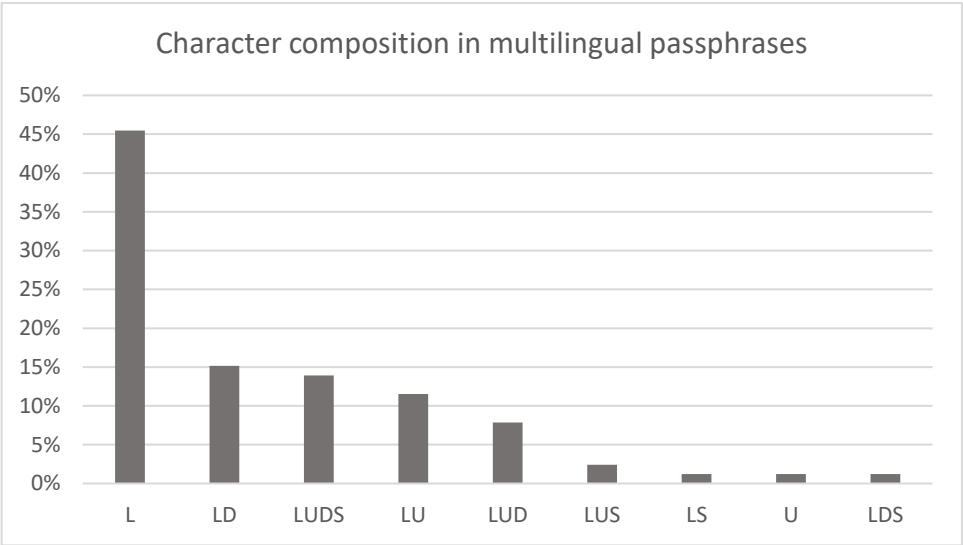
A comparison was done between a short password and a multilingual passphrase generated by the same participant to establish the practice of password reuse. A multilingual passphrase with a substring of at least four consecutive characters that matched those in a short password were seen as multilingual passphrases that were arrived at after password reuse. This is the standard used in the literature when

comparing the similarity of passwords (Shay et al., 2014). It was observed that 19% (33) of the multilingual passphrases had substrings that were similar to those found in short passwords. In cases where participants reused their short passwords, a substring on the far left of the resulting multilingual passphrase was often (72%) a result of adapting a short password.

The study went on to investigate the magnitude of transformation that was done to a short password as it was transformed into a substring of a multilingual passphrase. Levenshtein's edit distance was used to measure the distance between a substring in a multilingual passphrase and a short password (Das et al., 2014). The edit distance shows the number of characters that need to be changed in order to convert a substring in a multilingual passphrase to a short password. Results from Levenshtein's edit distance showed that 6% of the multilingual passphrase corpora had a substring that was completely identical to a short password. The two substrings of these multilingual passphrases were based entirely on the original short password that was keyed in twice, separated by a space to come up with a passphrase that had two substrings. Furthermore, 7% of the multilingual passphrase corpora had a substring that was one to three characters short of the original short password. The remaining 5% had substrings that were four to seven characters away from the original short password.

**Passphrase structure.** It was also important to investigate the use of different character classes in multilingual passphrases. The evaluation of character class use was done in terms of the use of L: lower-case alphabetic letters; U: upper-case alphabetic letters; D: digits and S: symbols. It was important to establish the use of these different character classes as they might have had an influence on multilingual passphrase strength and usability. The results showed that there is a dominant use of lower-case alphabetic letters (45%) followed by the use of combined lower-case alphabetic letters and digits (15%). Interestingly, participants generated multilingual passphrases based on LUDS (14%). This can be explained by the fact that 19% of all the multilingual passphrases had a substring that was extracted from a short password. This reflects a

precedence of password reuse. Figure 21 displays the findings on the use of different characters during multilingual passphrase generation.



**Figure 21. Common characters used to generate multilingual passphrases**

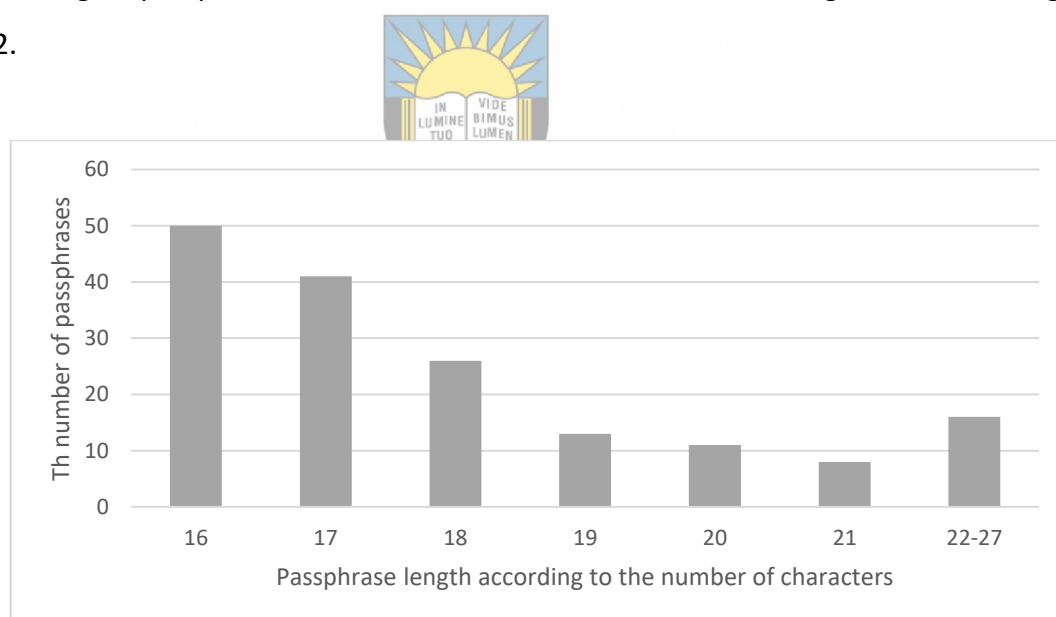
In addition, multilingual passphrases were analysed to establish the number of substrings that make up a passphrase. The aim was to observe the use of a non-alphabetic letter sequence (space) during multilingual passphrase generation. Table 14 shows that the majority of the user-generated multilingual passphrases had two substrings (78%), while a few had three substrings (7%) and a very few had more than three substrings. There was a fair use of keyboard patterns based on digits within multilingual passphrase substrings. These keyboard patterns were always found on the right end of the multilingual passphrase as either a stand-alone substring or extending another substring without a non-alphabetic character sequence separating the two, as shown in Table 14.



**Table 14. Characteristics of observed multilingual passphrase structures**

Characteristics of observed substrings	Example	Frequency	Percentage of Substrings
Two Substrings	Totsiens Goodbye	129	78
Three Substrings	house imba indlu	11	7
Four Substrings and more	werk by die see is easy	2	1
Two Substrings+KB_PATTERN	sikhosonke swaartbooi 12345	5	3
Two substrings_KB_PATTERN	Because dankie123	4	2

**Passphrase length.** An analysis of multilingual passphrase length was done. It was found that the majority of multilingual passphrases had a length of sixteen characters, as recommended by the passphrase policy. However, the majority of multilingual passphrases were more than sixteen characters long, as shown in Figure 22.

**Figure 22. Multilingual passphrase length according to the number of characters**

**The use of common substrings.** An analysis was done to establish whether multilingual passphrases have substrings that appear in the list of most common passwords of 2016, 2017 and 2018, according to SplashData (2016, 2017, 2018). Seven multilingual passphrases (4%) displayed a substring that is among the common passwords in 2016, 2017 and 2018, as shown in Table 15.

**Table 15. Common substrings (in bold) that were found among the most common passwords of 2016, 2017 and 2018**

Substrings observed in the passphrase corpora	Related common passwords of 2017	Position on the most common passwords list
Design @ <b>123456</b>	123456	1
Mpho Sleep <b>12345</b> AB	12345	5
<b>iloveyou</b> etumona	iloveyou	10
<b>Welcome</b> to kanyemba@97	welcome	12
M@engahama <b>Passw0rd</b>	passw0rd	19
<b>password</b> sinozuko	password1	29
Nust@ <b>computer</b> 17	computer	45

#### 6.4 Instrument validity and reliability

Testing for validity and reliability is the first step to analysing short password and multilingual passphrase generation and recall usability. A questionnaire survey was used to gather data for evaluating usability factors for generating and recalling short passwords and multilingual passphrases. Data for assessing usability (efficiency) was gathered using key logs as participants went about the activities of password generation and logging in. This section reports on validity and reliability tests that were done on the data that was gathered using a questionnaire. The gathered data corresponded to two factors, namely, effectiveness and user satisfaction, as shown in the proposed model in Chapter 5.

Explorative Factor Analysis (EFA) was used to test construct validity. Factor analysis allows for the identification of key constructs that explain the investigated factors of effectiveness and user satisfaction. SPSS version 25 was used for this analysis. All tests were carried out at the 95% level of significance. It should be noted that factors of usability (effectiveness and user satisfaction) were evaluated twice, thus during password generation and recall. The EFA was conducted using principal component analysis (PCA). Explorations of various factor solutions were conducted employing additional extraction and data rotation methods to find the most parsimonious set of factors. The most parsimonious result was achieved with two factors by employing

equamax rotation. The next section reports on the validity of usability factors during password generation and recall.

#### **6.4.1 Instrument validity**

The PCA was done on fifteen (15) items (combined items from Tables 4 and 8) for password generation and eighteen (18) items for password recall (combined items from Table 6 and 9). The cumulative variance for the two factors for password generation was 58.12%, which accounts for almost 60% of the total variability. This marks an acceptable threshold (Williams et al., 2012). Similarly, to determine factors of password recall, two usability factors, namely, effectiveness and user satisfaction, were analysed. An acceptable threshold of cumulative variance for the two factors was found to be 62.09%. On both occasions, password generation and recall, all the evaluated factors had an eigenvalue greater than one (1), as recommended by the Kaiser rule (Mertler & Vannatta, 2004).



In addition, this study used guidelines established by Reise et al. (2000) and found all fifteen (15) items for password generation (creation) loading higher (that is,  $\geq 0.45$ ) on the primary loadings of their respective components. A summary of the results of the rotated factor matrix is presented in Table 16. These items loaded higher than or equal to 0.461 on PCE. PCUS comprised seven items, which had high loadings (i.e. all  $\geq 0.718$ ) suggesting the strength of the empirical validity of the construct.

**Table 16. Rotated component matrix – password generation**

Item	PCE	PCUS
PCE1	0.812	
PCE2	0.811	
PCE3	0.658	
PCE4	0.554	
PCE5	0.461	
PCE6	0.781	
PCE7	0.552	
PCE8	0.595	
PCUS1		0.721
PCUS2		0.819
PCUS3		0.718
PCUS4		0.736
PCUS5		0.782
PCUS6		0.755
PCUS7		0.793

Note: Extraction Method: Principal Component Analysis. Rotation Method: Equamax with Kaiser Normalization.  
Rotation converged in 3 iterations.

Similarly, guidelines established by Reise et al. (2000) were used on twelve items for password recall and loaded high (that is,  $\geq 0.45$ ) on the primary loadings of their respective components. Items one to six on Table 17 contributed to PRE and these items resulted in high loadings (i.e. all  $\geq 0.570$ ), suggesting the strength of the empirical validity of the construct. Another six items loaded high on PRUS with all the items loading high (i.e.  $\geq 0.502$ ) on the primary loadings.

**Table 17. Rotated component matrix – password recall**

Item	PRE	PRUS
PRE1	0.570	
PRE2	0.634	
PRE3	0.658	
PRE4	0.596	
PRE5	0.742	
PRE6	0.696	
PRUS1		0.502
PRUS2		0.645
PRUS3		0.676
PRUS4		0.814
PRUS5		0.808
PRUS6		0.718

Note: Extraction Method: Principal Component Analysis. Rotation Method: Equamax with Kaiser Normalization.

Rotation converged in 5 iterations.



## 6.4.2 Instrument reliability

This section reports on the instrument reliability in relation to short passwords and multilingual passphrases. Internal consistency was ascertained using the Cronbach's alpha coefficient.

### 6.4.2.1 Instrument reliability —short passwords

Table 18 shows the reliability of each scale as it relates to the variables measured. The Cronbach's alpha for the scales ranged from 0.809 to 0.921 which shows a high reliability of coefficients for short passwords and constructs according to the criteria explained in Section 2.6.

**Table 18. Reliability analysis – short passwords**

Variable/s	Valid N	The number of questions	Cronbach's $\alpha$
<b>Short password creation</b>	222	15	0.921**
1. Effectiveness	223	8	0.853**
2. User satisfaction	222	7	0.912**
<b>Short password recall</b>	140	18	0.901**
1. Recall strategy	141	6	0.809**
2. Effectiveness	141	6	0.836**
3. User satisfaction	140	6	0.862**

\*\*Significantly acceptable reliability

#### 6.4.2.2 Instrument reliability – multilingual passphrases

Table 19 shows the reliability of each scale as it relates to the variables measured. The Cronbach's alpha for the scales ranged from 0.783 to 0.921, which shows a high reliability coefficient for the instrument used on multilingual passphrases and its constructs, as explained in Section 2.6.

**Table 19. Reliability analysis – multilingual passphrases**

Variable/s	Valid N	The number of questions	Cronbach's $\alpha$
<b>Passphrase creation</b>	173	15	0.916**
1. Effectiveness	175	8	0.834**
2. User satisfaction	174	7	0.921**
<b>Passphrase recall</b>	130	18	0.893**
1. Recall strategy	131	6	0.783**
2. Effectiveness	132	6	0.805**
3. User satisfaction	131	6	0.915**

\*\*Significantly acceptable reliability

#### 6.5 Password usability

This study gathered data to test the usability of short passwords and multilingual passphrases. This section presents the findings of the evaluation of factors of usability during short password/multilingual passphrase generation and recall. The section ends with a comparison, in terms of usability, between short passwords and multilingual passphrases.

### 6.5.1 Short password requirements usability

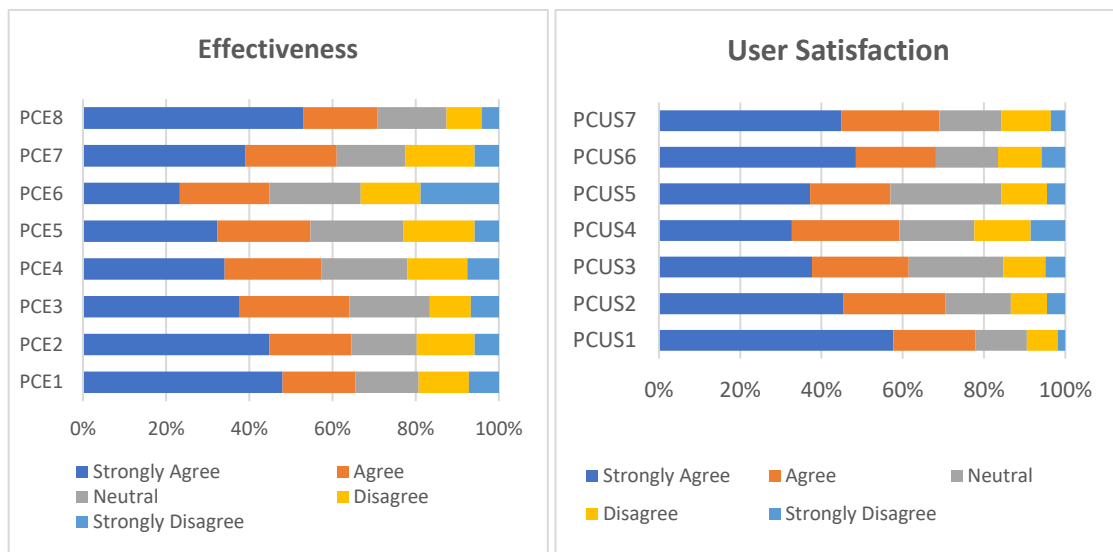
This section reports on the usability of short passwords during password generation and recall and focuses on reporting findings on the usability factors of effectiveness, user satisfaction and efficiency.

#### 6.5.1.1 Short password generation usability

Two hundred and twenty-four (224) participants generated short passwords as indicated earlier. Section 6.3.2.1 explored different short password generation strategies used by participants. As such, this section presents findings on short password generation usability. A one-sample test was used to evaluate a user's perception of how easy it was to generate a short password. Results from the one-sample test showed that participants rated effectiveness (mean = 3.7057, SD = 0.90548,  $t = 11.639$ ,  $p = < 0.0001$ ) and user satisfaction (mean = 3.8880, SD = 0.97012,  $t = 13.639$ ,  $p = < 0.0001$ ) significantly higher than 3 on a 5-point Likert scale. This implies that participants found short password generation user-friendly in terms of effectiveness and user satisfaction. A closer look at short password generation effectiveness showed that participants strongly agreed (47%) that generating short passwords was easy (PCE1), would always be easy (50%) and that it was a simple (45%) task to do (PCE2). Furthermore, 39% of the participants strongly agreed that they found it easy to recover (PCE7) from mistakes made during short password generation. However, the results also suggested that it required a measure of effort to generate a short password as shown in Figure 23.

In addition, Figure 23 shows that user satisfaction in short password generation was rated highly in terms of usability. Only a small number of participants felt that the system did not work the way they wanted it to work during short password generation. Data on keylogs was gathered to ascertain short password generation efficiency. This included data on the time taken to generate and confirm short passwords and the number of attempts required to generate a short password in accordance with specified short password requirements. The total time for generating short passwords was arrived at by adding together the initial short password generation and confirmation time. It took, on average, 82.2 seconds to generate a short password and 34.3 seconds

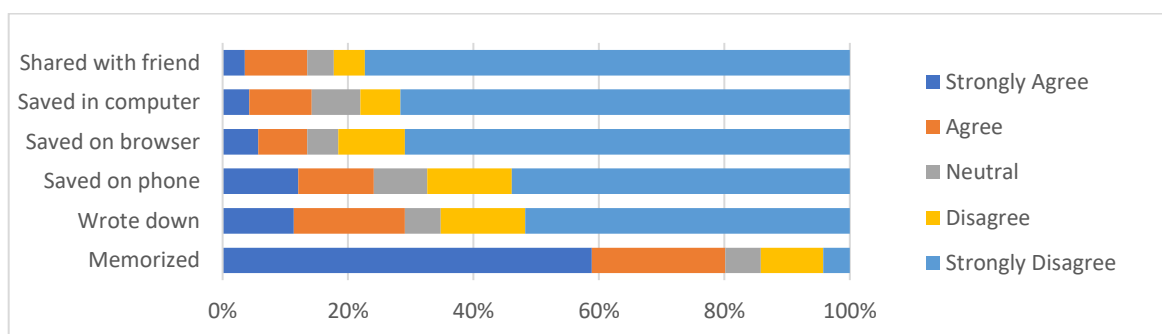
to confirm the generated short password. Hence, it took 116.5 seconds in total to generate a short password. Furthermore, participants required, on average, 2.1 short password generation attempts to completely and accurately generate a short password.



**Figure 23. Short password generation effectiveness and user satisfaction**

#### 6.5.1.2 Short password recall usability

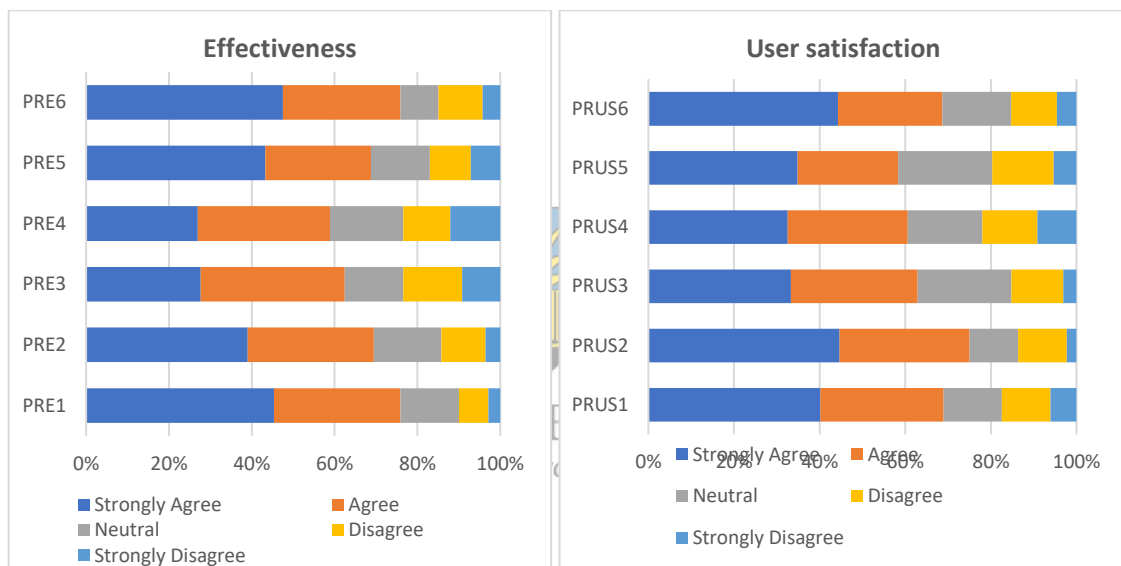
This section presents the findings on the short password recall strategies used by participants and expounds on password recall usability. One hundred and forty-one (141) participants who generated a short password took part in password recall. The findings on short password recall strategy show that close to 60% strongly agreed that they had managed to memorise their passwords. However, 12% of the participants strongly agreed to having saved their short passwords on their phones with 11% writing down their short passwords in case they forgot. Figure 24 summarises these findings.



**Figure 24. Short password memorability strategy**



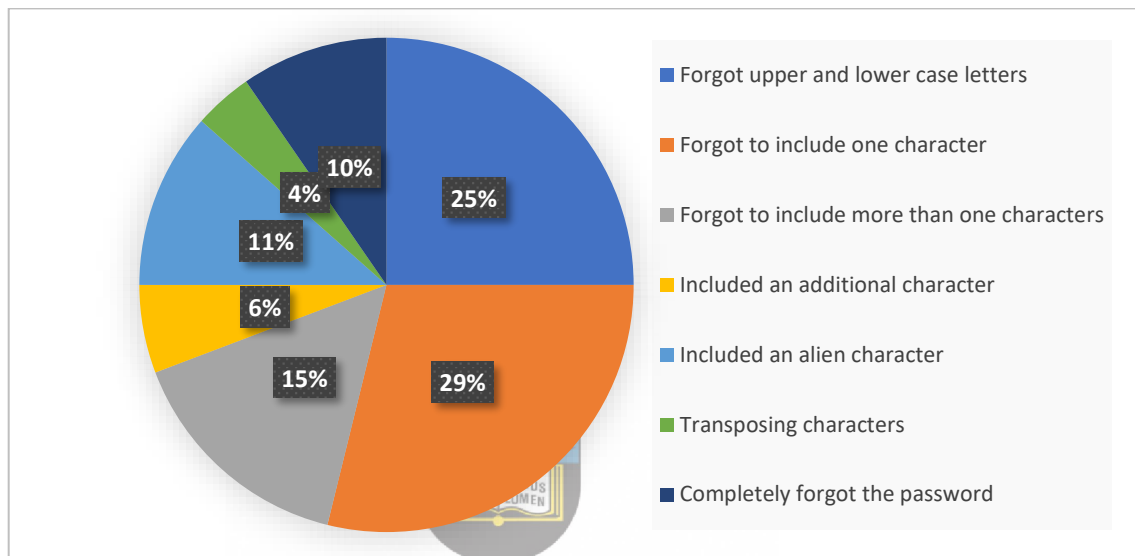
Data was gathered to evaluate how easy it was for participants to memorise short passwords, as demonstrated by their capability to recall a short password every three days over a period of two weeks. Results from a one-sample test showed that, on average, participants significantly rated short password recall higher than three on effectiveness (mean = 3.8322, SD = 0.89874,  $t = 10.995$ ,  $p = < 0.0001$ ). Furthermore, participants significantly rated user satisfaction with short password recall higher than three (mean = 3.8905, SD = 0.85774,  $t = 12.284$ ,  $p = < 0.0001$ ). These results suggest that, on average, participants could accurately and completely recall their short passwords with satisfaction. Figure 25 shows detailed findings on the evaluated attributes of short password recall effectiveness and user satisfaction.



**Figure 25. Short password recall effectiveness and user satisfaction**

Findings on password recall effectiveness were corroborated by results from key logs. Key logs were used to gather data on short password recall as demonstrated by participants' capability to accurately type in their passwords. The results showed that on all but one occasion, participants who had failed to accurately key in their short passwords were one or two characters away from the actual short password. Levenshtein's edit distance was used to compute the number of characters that needed to be changed in order to arrive at the actual short password from a wrongly keyed in password. Accordingly, the results revealed that all participants who had committed an error during logging in were 2.2 characters, on average, away from the actual short

password. Twenty-five per cent (25%) of all those who participated in short password recall failed to accurately and completely key in their short passwords. Ten per cent (10%) of the cases who failed to accurately key in a short password were as a result of typographical errors. Section 4.4.2.1 in this study defined typographical errors and went on to reason that typographical errors are not a result of memory loss but a mistake during the execution stage.



**Figure 26. The percentage of short password attributes (of all short password recall failure) that were often forgotten by users**

It was important to establish login errors that were due to failing to accurately and completely recall the short password. As such, login failure due to typographical errors was eliminated from the sample of reasons why participants had failed to successfully key in their passwords. It was therefore found that approximately 22% of the participants failed to accurately and completely recall their short passwords. Further investigation showed that the majority (29%) of participants who failed to recall their short passwords had forgotten one of the characters in the actual short password. Twenty-five per cent (25%) of short password recall failures were a result of failing to know which character(s) were to be written in upper or lowercase. Only 15% of the login failure attempts were a result of failing to recall and insert more than one character that was in the actual short password. Furthermore, failure to recall a short

password saw some of the participants including a character that was alien (11%) to the actual password. Figure 26 summarises these findings.

In addition, data gathered using key logs was used to evaluate efficiency according to log-in attempts and duration/time. The results show that participants needed 1.3 login attempts, on average, to accurately login into their profiles. Participants logged into their profiles on four separate occasions. Findings from the data gathered by key logs showed that it took 30 seconds, on average, to successfully login into a profile on one's first return; 33 seconds, on average, to successfully login into a profile on the second return; and this dropped to twenty-six (26) and twenty-one (21) seconds on the third and fourth login attempts, respectively. The constant drop in time taken to login can be explained by the fact that participants were learning and getting used to their short passwords overtime.

#### **6.5.1.3 Short password usability differences and correlation analysis**

An independent-samples t-test was done to compare the means between male and female ratings on short password generation and recall. Levene's test for homogeneity of variance (homoscedasticity) was used. Significant differences in mean ratings between males and females were only noticed on short password recall strategy. Males (mean = 4.3734; SD = 0.72413) showed a significantly higher mean rating on short password recall strategy than their female (mean = 3.8629; SD = 0.1.08245) counterparts ( $t = 3.195$ ;  $Pr > |t| = 0.002$ ). This could be explained by the fact that a number of female participants agreed to having written down their short passwords on a piece of paper or on their mobile phones or to having saved them somewhere or shared the short password with a colleague in case they forgot it.

An investigation for a linear relationship in the data, which allowed for a correlational analysis, was carried out. The non-parametric Spearman's rho bivariate correlation coefficient (two-tailed test) was used. The results showed that short password generation effectiveness had a strong, positive significant relationship with short password generation user satisfaction ( $r = 0.658$ ;  $p = < 0.0001$ ). Thus, short password generation effectiveness leads to user satisfaction. Similarly, short password

recall strategy had a strong, positive significant relationship with short password recall effectiveness ( $r = 0.489$ ;  $p = < 0.0001$ ) and short password recall user satisfaction ( $r = 0.449$ ;  $p = < 0.0001$ ). These results suggest that the ability to recall a short password promotes recall effectiveness and user satisfaction.

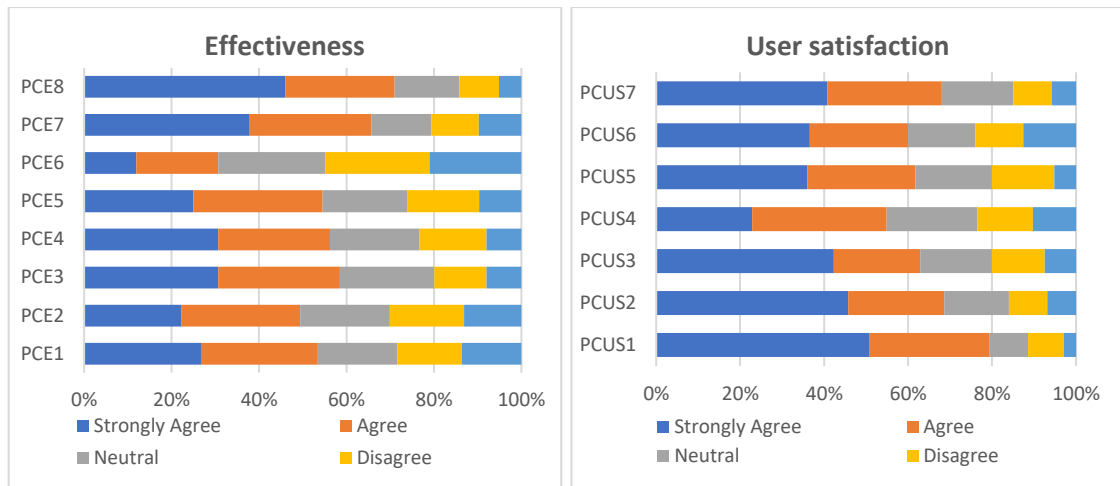
### **6.5.2 Multilingual passphrase usability**

This section reports on the usability of multilingual passphrases during passphrase generation and recall. Thus, findings on the usability factors of effectiveness, user satisfaction and efficiency are reported.

#### **6.5.2.1 Multilingual passphrase generation usability**

A total of 176 participants generated multilingual passphrases. Section 6.3.2.2 explored different multilingual passphrase generation strategies that were used by participants. This section reports on multilingual passphrase generation usability. A one-sample test was done to establish the overall perception of participants on multilingual passphrase generation usability, focusing on passphrase generation effectiveness and user satisfaction. The results showed that multilingual passphrase generation effectiveness (mean = 3.4743, SD = 0.87215,  $t = 7.194$ ,  $p = < 0.0001$ ) and user satisfaction (mean = 3.7874, SD = 0.99814,  $t = 10.405$ ,  $p = < 0.0001$ ) were significantly higher than 3. This implies that participants had found multilingual passphrase generation significantly usable as determined by effectiveness and user satisfaction.

Figure 27 shows that, even though multilingual passphrase generation was perceived as effective, participants had reservations about this password policy. For example, just above 30% thought it was effortless (PCE6) to generate multilingual passphrases. Close to 50% thought multilingual passphrases were simple (PCE2) to generate, while just above 50% thought that multilingual passphrases were easy (PCE1) to generate and that multilingual passphrase requirements were flexible. Furthermore, a detailed analysis of multilingual passphrase generation user satisfaction in Figure 27 showed that participants were satisfied with multilingual passphrase generation. All user satisfaction attributes that were evaluated showed that more than 50% of the participants agreed that the process had unfolded to their satisfaction.

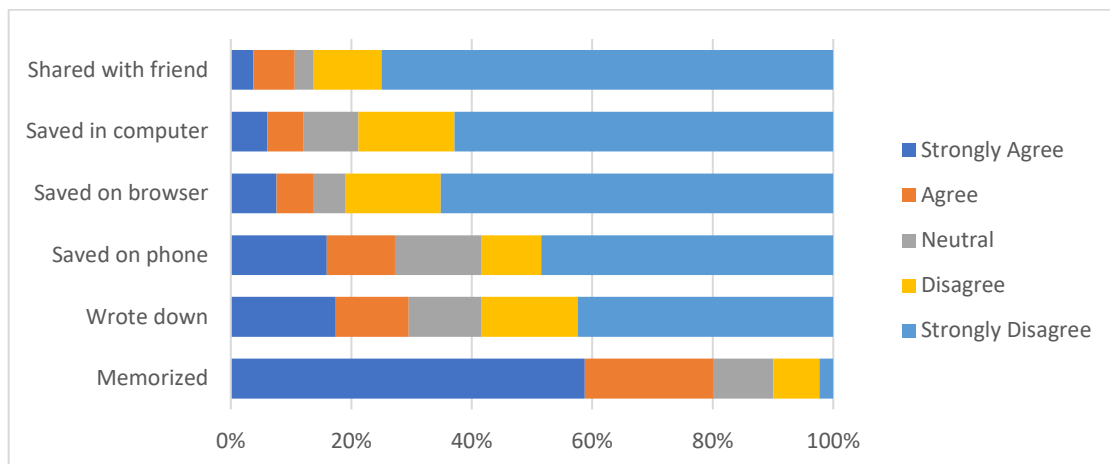


**Figure 27. Multilingual passphrase generation effectiveness and user satisfaction**

Key logs generated by the system during multilingual passphrase generation were analysed to establish passphrase generation efficiency. It was found that, on average, it took participants 246.9 seconds to initially generate a multilingual passphrase. Observations from key logs showed that participants often started multilingual passphrase creation by adopting a short password or name or a phrase. However, a couple of deletes and attempts to meet the multilingual passphrase requirement resulted in either a completely new phrase from the original thought or a phrase with substrings from a short password (19%). A limited amount of password reuse might explain the high average time needed to generate a multilingual passphrase. In addition, it took participants, on average, 74 seconds to confirm the generated multilingual passphrase. Hence, it required 320.9 seconds, on average, to completely generate a multilingual passphrase. This result is corroborated by a finding that participants required, on average, 4.5 attempts to create a multilingual passphrase.

#### 6.5.2.2 Multilingual passphrase recall usability

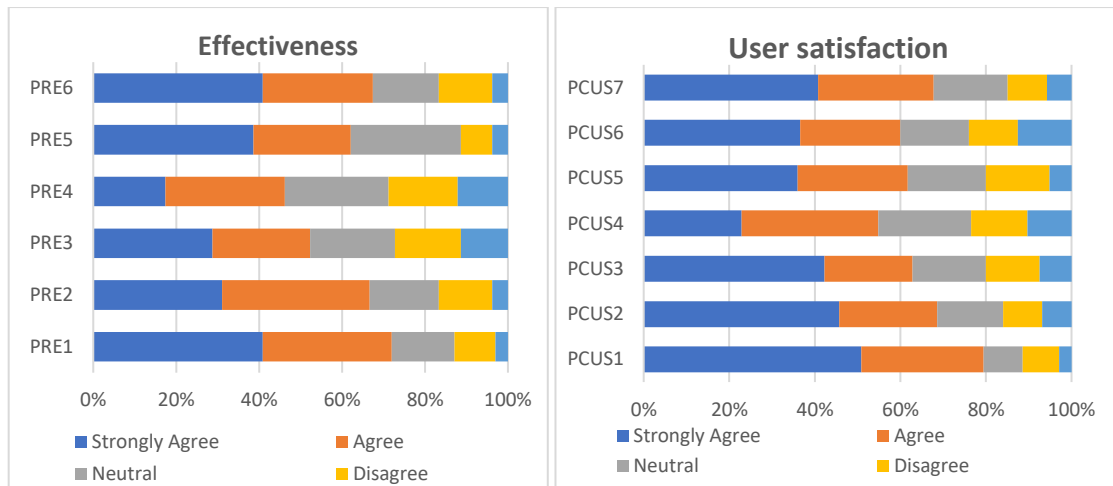
One hundred and thirty-one (131) participants took part in the multilingual passphrase recall exercise and completed a questionnaire on passphrase recall or memorability. This section reports on multilingual passphrase recall strategies used by participants prior to evaluating the usability of these strategies. Figure 28 shows that close to 60% of the participants strongly agreed that they had managed to memorise and recall their multilingual passphrase.



**Figure 28. Multilingual passphrase recalling strategy**

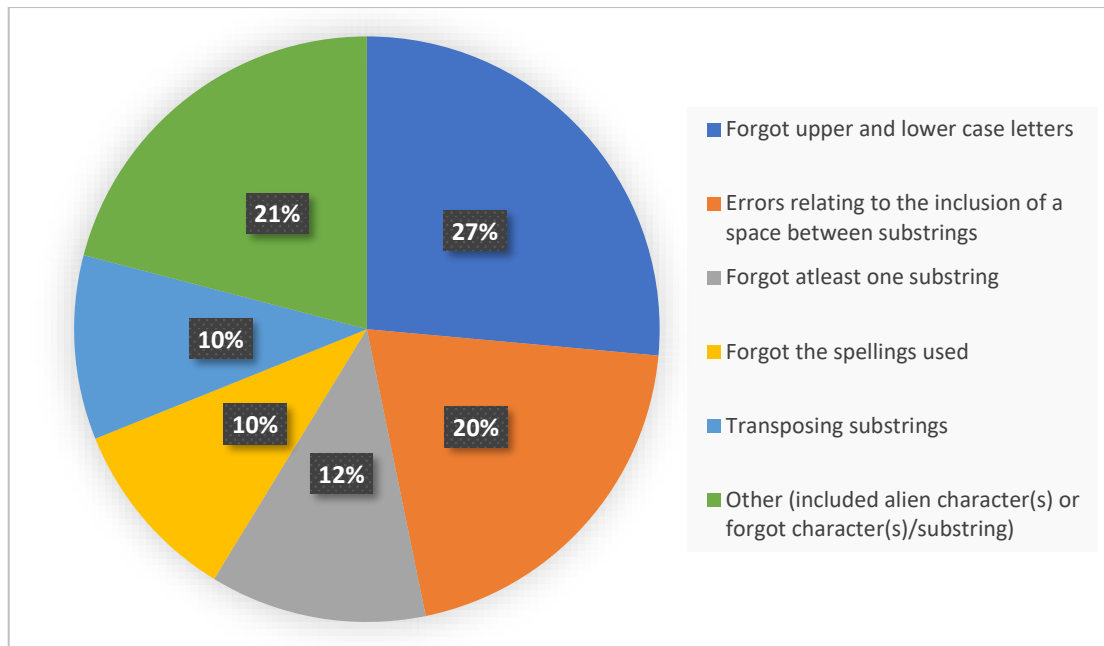
Furthermore, participants were asked to rate the usability of multilingual passphrase recall. A one-sample test that was conducted on multilingual passphrase recall showed that participants significantly rated passphrase recall effectiveness higher than 3 (mean = 3.6881, SD = 0.85516,  $t = 9.245$ ,  $p < 0.0001$ ). Furthermore, participants significantly rated user satisfaction with passphrase recall (mean = 3.8168, SD = 1.01229,  $t = 9.235$ ,  $p < 0.0001$ ) higher than 3. These findings suggest that, on average, participants could accurately and completely recall their multilingual passphrases with satisfaction.

A detailed analysis of all items assessed under multilingual passphrase recall effectiveness in Figure 29 showed that participants at least agreed that the process was effective on most attributes. However, participants felt it required a great deal of effort to recall multilingual passphrases, while just above 50% agreed that it required a few more steps to recall a multilingual passphrase (PRE3).



**Figure 29. Multilingual passphrase recall effectiveness and user satisfaction**

Similarly, Figure 29 shows that participants expressed satisfaction with multilingual passphrase recall. However, just below 40% of the participants strongly agreed that they would have struggled to remember more passphrases at once (PRUS5) if they were to adopt such a password policy. These findings were corroborated by findings from key logs, which showed that 40% of the participants had failed to accurately and completely type in their multilingual passphrases. Levenshtein's edit distance went on to show that all participants who had failed to accurately key in their multilingual passphrases had resulted in candidate passphrases that were 4.2 characters, on average, away from the actual passphrase. Twenty-six per cent (26%) of all the login failures on multilingual passphrases had been due to typographical errors, while the remaining 74% were accounted for by a failure to recall the multilingual passphrase. It was important to establish the causes of multilingual passphrase recall failure as reflected in the login errors. The majority (27%) of participants had struggled to accurately recall characters that were to be keyed in as upper- or lower-case letters. This was followed by a failure to recall and accurately include a space between substrings (20%). Figure 30 shows these errors which affected multilingual passphrase recall effectiveness.



**Figure 30. The percentage of passphrase attributes (of all the passphrase recall failure) that were often forgotten by participants**

Data on multilingual passphrase efficiency shows that it took participants 132.9 seconds on average to accurately login to their user profiles during their first logging in session. This finding suggests multilingual passphrases were not easy to recall completely and accurately. However, it took less time (60.3 and 28.2 seconds) on average to accurately key in a multilingual passphrase on participants' second and third return. This trend changed at the last login session as it took 48.7 seconds to accurately type in a multilingual passphrase. These findings suggest it will take time for participants to efficiently memorise and recall a multilingual passphrase. Lastly, participants required, on average, two login attempts to successfully login into their profiles.

#### **6.5.2.3 Multilingual passphrase usability differences and correlation analysis**

An independent-samples t-test was done to compare the mean ratings between male and female respondents focusing on multilingual passphrase generation and recall. Levene's test for homogeneity of variance revealed significant differences in mean ratings on multilingual passphrase recall strategy and recall effectiveness. In both cases, males had significantly higher mean ratings on multilingual passphrase recall



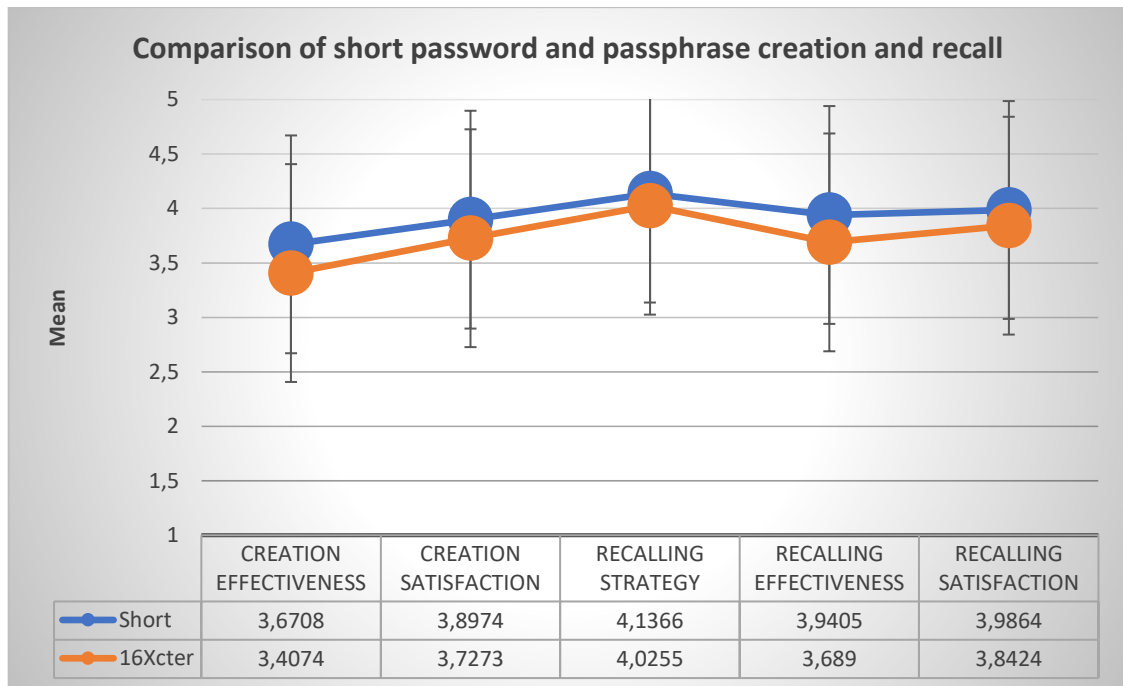
strategy (mean = 4.2524; SD = 0.78212;  $t = 2.523$ ;  $Pr > |t| = 0.013$ ) and recall effectiveness (mean = 3.8548; SD = 0.77028;  $t = 2.423$ ;  $Pr > |t| = 0.017$ ). A closer look at the data showed that more female participants wrote down their multilingual passphrases on a piece of paper and saved passphrases on their mobile phones in case they forgot. In addition, relatively more female participants indicated that recalling a multilingual passphrase for this study was not always simple and that it required a few more steps for them to recall their passphrase. However, multilingual passphrase generation effectiveness ( $t = 1.108$ ;  $Pr > |t| = 0.269$ ), user satisfaction ( $t = -0.079$ ;  $Pr > |t| = 0.937$ ) and recall user satisfaction ( $t = 0.369$ ;  $Pr > |t| = 0.713$ ) did not yield any significant difference among the investigated demographics.

In addition, a correlation analysis was done to establish potential relationships between evaluated constructs. The results showed that multilingual passphrase generation effectiveness had a strong, positive significant relationship with multilingual passphrase generation user satisfaction ( $r = 0.629$ ;  $p = < 0.0001$ ). Thus, effective multilingual passphrase generation promotes user satisfaction. On the other hand, multilingual passphrase recall strategy had a moderate, positive significant relationship with passphrase recall effectiveness ( $r = 0.364$ ;  $p = < 0.0001$ ) and passphrase recall user satisfaction ( $r = 0.398$ ;  $p = < 0.0001$ ). Thus, the ability to recall a multilingual passphrase mildly promoted recall effectiveness and user satisfaction.

### **6.5.3 Comparison of short passwords and multilingual passphrases usability**

This study aims to propose a password policy that could encourage users to generate and recall secure and user-friendly passwords, by motivating the use of multilingual passphrases. It is therefore important to understand how multilingual passphrases compare to short passwords in terms of usability. A paired samples t-test was conducted in order to make mean rating comparisons on short password and multilingual passphrase theoretical constructs, namely, creation effectiveness, creation user satisfaction, recall strategy, recall effectiveness and recall user satisfaction. In all cases, the short password policy had higher mean ratings, as shown in Figure 31. However, significant differences were noticed on password creation effectiveness (mean = 0.2633;  $df = 111$ ;  $t = 3.215$ ;  $p = 0.002$ ), creation user satisfaction (mean =

0.1701;  $df = 109$ ;  $t = 2.053$ ;  $p = 0.042$ ) and recall effectiveness (mean = 0.2514;  $df = 111$ ;  $t = 2.331$ ;  $p = 0.022$ ). There were no significant differences on mean ratings for recall strategy (mean = 0.1111;  $df = 110$ ;  $t = 1.273$ ;  $p = 0.206$ ) and recall user satisfaction (mean = 0.1439;  $df = 100$ ;  $t = 1.339$ ;  $p = 0.183$ ).



**Figure 31. Mean responses of theoretical variables on short passwords vs. passphrase creation, recall strategy and recalling**

Furthermore, Table 20 shows that it required twice as much effort to efficiently generate a passphrase, confirm and recall it. Login errors occur twice as much with multilingual passphrases than they do with short passwords. These findings corroborate the results in Figure 31 that shows multilingual passphrases are significantly less usable when compared to short passwords.

**Table 20. Multilingual passphrase vs. short password usability (time indicated in seconds)**

Password	Initial creation	Confirmation	Creation attempts	Logging in errors	First recall	Second recall	Third recall	Fourth recall
Passphrase	247	74	4.5	2	133	60	28	49
Short password	82	34	2.1	1.3	30	33	26	21

An investigation was done to establish the password length that appeared to be user-friendly to the participants. Figures 18 and 22 on password length show that participants went beyond the minimum recommended length for short passwords and multilingual passphrases. Therefore, it was interesting to establish the password length that appeared to be most user-friendly to participants when both short passwords and multilingual passphrases were considered basing on views from usability factors of effectiveness and user satisfaction. The usability of passwords whose length ranged from 8 to 21 characters long were considered. Eight characters are the minimum recommended characters for short passwords, while only a few passwords exceeded 21 characters as shown in Figure 22. Table 21 shows results from an independent-samples t-test which compared the means of the usability and non-usability of passwords in terms of their lengths. Usable password lengths are those with a mean that was greater or equal to 3 for effectiveness and user satisfaction. Levene's test for homogeneity of variance verified that this assumption of equal variances holds in both samples. The mean password length for usable passwords was almost 14 characters in both cases, i.e. mean = 13.70 for effectiveness and mean = 13.58 for satisfaction. However, there was no significant difference in means of usable and non-usable passwords. This lack of significance suggests that passphrases with slightly more than 14 characters are still user-friendly.

**Table 21. T-tests for mean password length of usability on effectiveness and user satisfaction**

Study Variable	Gender	Mean	SD	Levene's Test for Equality of Variances		t-test for Equality of Means				
				F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
<b>Creation Effectiveness</b>	Non-usable	13.69	4.739	0.915	0.339	-0.019	387	0.985	-0.011	0.556
	Usable	13.70	5.163							
<b>Creation Satisfaction</b>	Non-usable	13.68	4.862	0.236	0.627	0.166	383	0.868	0.097	0.586
	Usable	13.58	4.702							

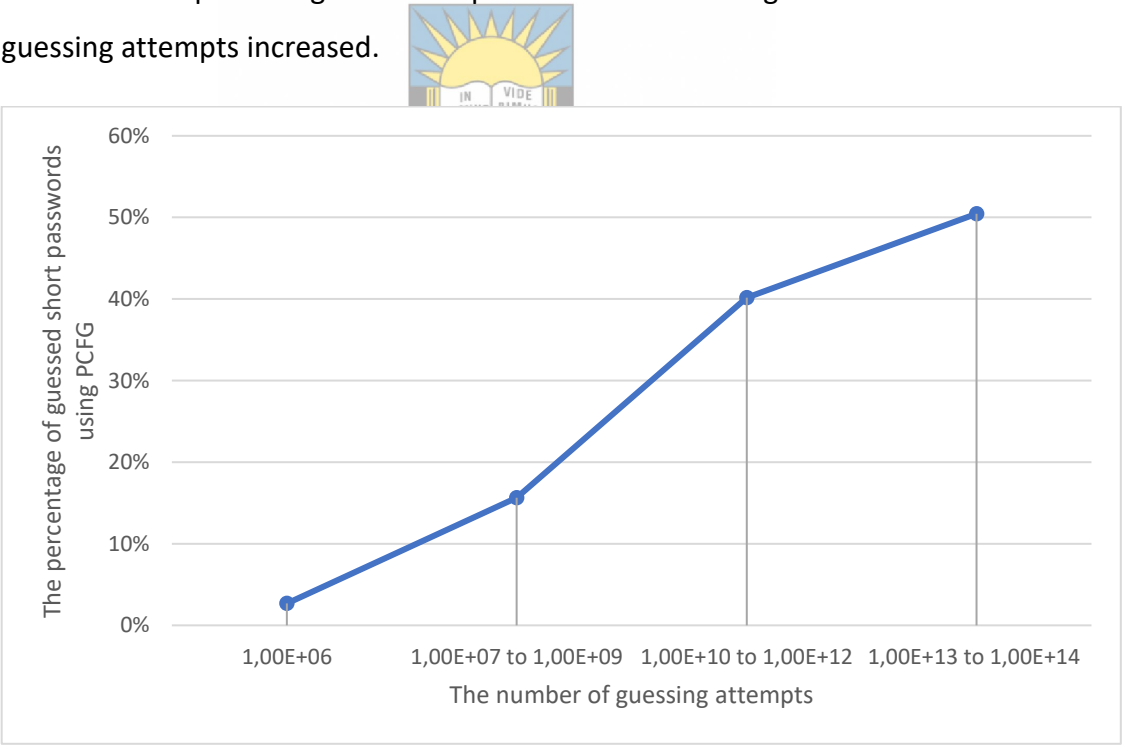
## 6.6 Password strength

Chapter 5 explained the perception of this study on short password or multilingual passphrase strength, conceding that password strength is determined by the resistance of the short password or multilingual passphrase to guessing. This study used a PCFG password guessing algorithm to evaluate the strength of short passwords

and multilingual passphrases (Ur, Segreti, et al., 2015). Section 2.6.2.1 explained the PCFG that was used in this study and Section 4.2.2.2 evaluated the PCFG. The following sections present findings on short password and multilingual passphrase strength.

**6.6.1 Short password strength**

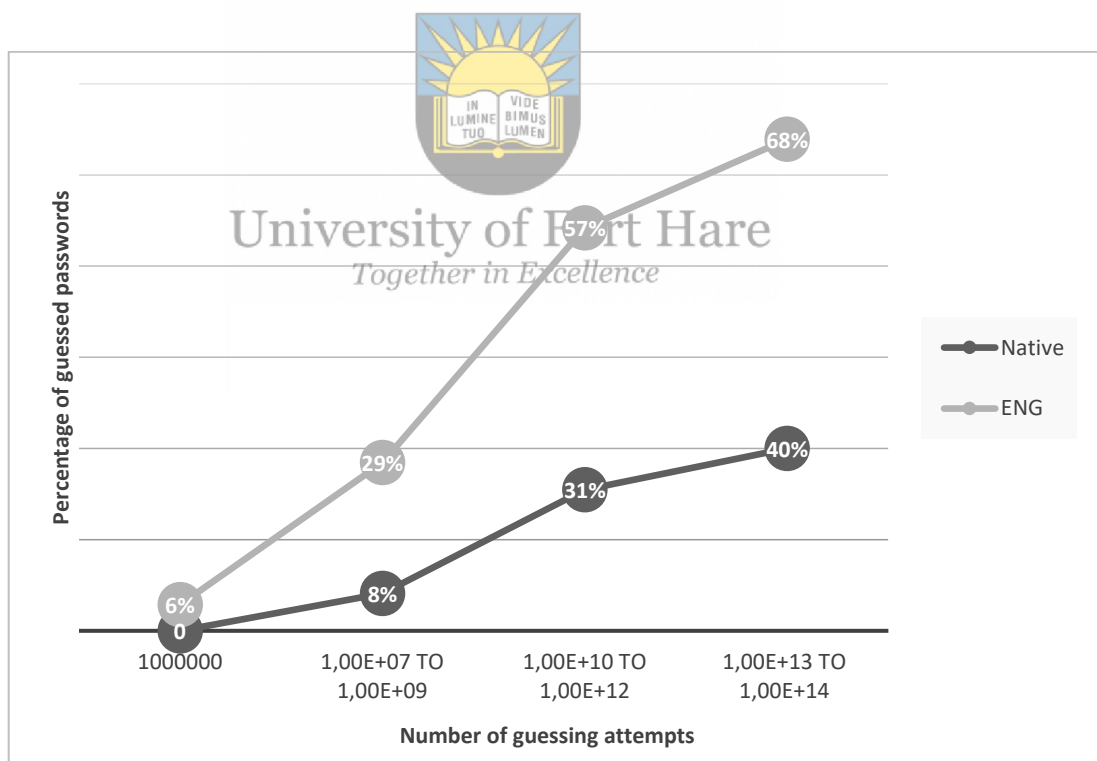
Of the 224 short passwords that were sent for password guessing using PCFG, 50.4% of these were guessed. It took less than a thousand attempts to guess the first short password. However, it required a lot of resources ( $10^{12}$  guessing attempts) to guess 40% of the short passwords and  $10^{14}$  attempts to guess 50% of the short passwords. The cut-off for guessing short passwords was set at  $10^{15}$  guessing attempts. It took at most two weeks for these passwords to be guessed and just over 50% of these were guessed. These findings suggest that a user-generated short password following LUDS has a less than 50% chance of resisting any guessing attempts by a PCFG. Figure 32 shows the percentage of short passwords that were guessed as the numbers of guessing attempts increased.



**Figure 32. Short password guessing results using PCFG.**

An analysis was carried out to establish the strength effects of orienting a short password to a particular language. The aim was to establish any security benefits of orienting passwords to African languages over Indo-European languages. Short

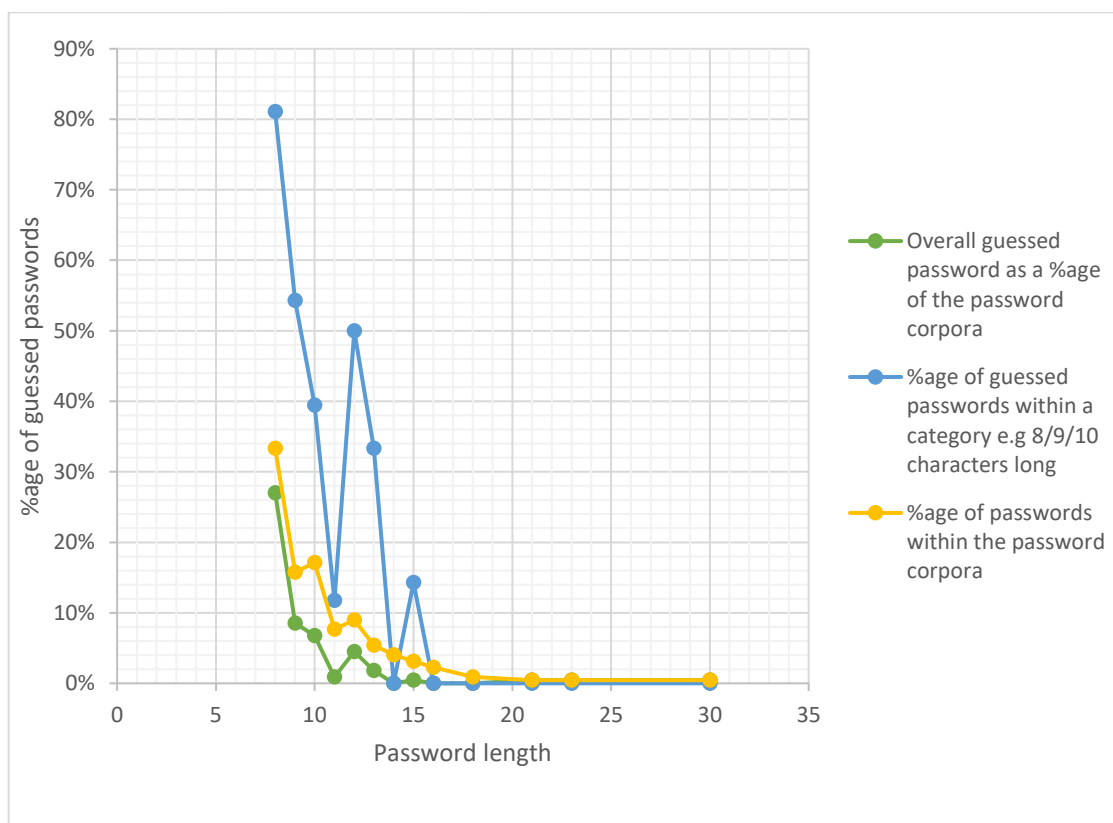
passwords oriented to African languages in the researched countries, namely, Namibia, South Africa and Zimbabwe, were collectively compared to those oriented to the English language. It should be noted that English is the first written language and is the official language of all three of the researched countries. Figure 33 shows the password strength benefits of orienting a short password to an African language, indicating that no short password that was oriented to an African language was guessed within the first one million guessing attempts, while 6% of the English language-oriented short passwords had been guessed at this stage. It took few guessing attempts (less than  $10^{12}$ ) to guess 40% of the short passwords, yet it took close to  $10^{14}$  guessing attempts to guess the same amount of African language-oriented short passwords. All in all, close to 70% of English-oriented short passwords were guessed compared to only 40% of African language-oriented ones that were guessed within the recommended  $10^{15}$  guessing attempts.



**Figure 33. A comparison of guessing resistance between African language and English orientated short passwords**

Further analysis was done to establish any relationships between short password length and their resistance to password guessing. Figure 18 shows that some

of the participants went beyond the short password minimum length requirement of eight characters. Figure 34 below shows that there might be a password strength benefit associated with generating a long password. For instance, 81% of all the short passwords that were eight-character long were guessed. This proportion (eight characters long) of the short passwords that was guessed contributed 27% to the whole short password corpora that was guessed. It should be noted that 33% of the short password corpora were eight characters long. The percentage of guessed passwords decreased as the password length increased as shown in Figure 34. None of the short passwords that were at least sixteen characters long were guessed within  $10^{15}$  guessing attempts. Passwords that were at least 16 characters long represented 2% of the short password corpora.



**Figure 34. Short password length and resistance to password guessing**

### 6.6.2 Multilingual passphrase strength

Multilingual passphrases generated by participants in this study were sent for password guessing using a PCFG. To guess multilingual passphrases, an option was

selected that could guess passwords that were at least sixteen characters long with at least two substrings. A multilingual passphrase guessing cut-off was set at  $10^{15}$  guessing attempts. After two weeks of guessing attempts, the PCFG could not guess any multilingual passphrase that had been generated using the password policy proposed in this study. This despite the presence of substrings that were among the popular passwords of 2016, 2017 and 2018 (as shown in Table 15) in some of the user-generated passphrases. It can be argued that the length of the passphrase and the incorporation of substrings oriented to different languages within a single passphrase reinforced the overall strength and resistance to guessing. This is supported by a finding in Figure 33 which showed that African language-oriented short passwords are slightly resistant to guessing compared to English language-oriented short passwords. The paucity of long passwords in the public domain might also have aided the resistance of passphrases.

## 6.7 Chapter summary

This chapter presented research findings from the analysis of the gathered data. Data gathered using an experiment was used to present findings on constructs/factors in the proposed model of Chapter 5. In other words, this chapter contributed to the use of design science research in this study by presenting data for evaluating proposed constructs that had been informed by theoretical insights in the literature. The study argued that understanding the social context can help system designers when designing usable and secure password policies. Hence, socio-cultural theory was assumed in guiding the choices of usable and secure password policies. Findings in this chapter showed that the social context does have an effect on password generation as purported in socio-cultural theory. Short passwords that were oriented to African or Indo-European languages or both reflected on the generic law of development. Participants showed a high use of names when generating short passwords and passphrases. This can be explained by memorability theories that propose that users may seek to use existing information in the long-term memory during password generation with the aims of reducing interference (España, 2016). Another line of thought is that users adapt names as passwords because they find it convenient or because it required little cognitive effort (Woods & Siponen, 2019). In addition, study



University of Fort Hare  
Together in Excellence

results show that some of the user-generated short passwords and passphrases had substrings that were found among the most common passwords of 2016, 2017 and 2018. Lastly, password reuse was more pronounced in short passwords when compared to passphrases.

Usability assessments in terms of efficiency, effectiveness and user satisfaction showed that short passwords were more usable than passphrases. While passphrases were seen as usable according to the t-tests conducted in this study, passphrases were significantly less usable when compared to short passwords. Accordingly, it required more than twice as much effort to generate and recall a passphrase. In addition, participants were twice as much more likely to forget a passphrase than a short password. However, high password reuse during short password generation might have influenced the perception of participants on short passwords usability. Chapter 3 suggests that the repeated use of the same information reduces the cognitive burden of generating and using a new password. Interestingly, continued rehearsal of passphrases courtesy of periodic logging in during the experiment appears to have aided memorability as participants were able to constantly reduce logging in time. Further, the results suggest that a password policy recommending the generation of passphrases that are 14 characters long is high likely to be user-friendly. In addition, results on security tests suggested that passphrases are far more secure than short passwords. None of the user-generated passphrases were guessed, yet just above 50.4% of the short passwords were successfully guessed by the PCFG. Results also showed that English language-oriented short passwords were guessed at a faster rate than those oriented to African languages.

The next chapter discusses the findings of this chapter in preparation for the evaluation of the proposed model.



## **CHAPTER 7: RESEARCH FINDINGS AND DISCUSSION**

### **7.0 Introduction**

The previous chapter presented the research findings from the data collection and analysis. This chapter discusses the research findings presented in Chapter 6 and makes comparisons with findings in the literature. The discussion of the study findings was split into two: an overview of password characteristics and of password security. Bonneau (2012) noted that this is the most common way of presenting findings on passwords since its use by Morris and Thompson in the late 1970s. Wang, Cheng, et al. (2015) went on to widely hypothesise that password characteristics are greatly influenced by African languages. As such, Chapter 3 of this study used socio-cultural theory to motivate the use of passwords based on multilingualism. Hence, the analysis of the password characteristics identified in this study is guided by the principles in socio-cultural theory. It is critical to understand password characteristics and their orientation to different languages as this has an effect on security (Jakobsson & Dhiman, 2013; Wang, Cheng, et al., 2015). The chapter goes on to discuss findings on password security and usability.

### **7.1 Socio-cultural theory and password characteristics**

This section discusses research findings on password characteristics according to the principles of socio-cultural theory. This theory was used to inform the theoretical foundation of this study and to explore the socio subsystem of this study. Chapter 3 of this study used socio-cultural theory to explain how contextual factors influence human mental development and functioning and its principles, such as the generic law of development, mediation and genetic domains, were used to explain human mental development within a context. It is argued that understanding language development can help readers to understand password characteristics and the exploration of opportunities to influence users to generate secure and usable multilingual passphrases. The literature shows that users often adopt common words in a language when generating passwords. The influence of language on this study's short password and multilingual passphrase generation is discussed next within the principles of the socio-cultural theory. By so doing, this section addresses the fourth research sub-

question of the study that was set out as follows: *What are the password characteristics of a multilingual user group?*

### **7.1.1 A discussion of the findings on the generic law of development**

The generic law of development proposes that mental development does not pre-exist, neither is it inborn; rather, it unfolds with time under the influence of the context. Hence, the context of a computer user is expected to influence the choices of passwords. Research findings in Chapter 6 reflect the different contextual environments in which the researched participants grew up. Various participants indicated that their first language was isiXhosa or Oshiwambo or Shona or another African language. However, nearly all participants (94%) indicated that their second language was English. This was reflected by the dominance of short passwords oriented to English (39%) and African (39%) languages, and multilingualism (5%). Thirteen per cent (13%) of the short passwords were considered random. Random passwords are those passwords with substrings that have no identifiable patterns and could not be oriented to a particular language (De Carnavalet & Mannan, 2014). A closer look at the short password corpora showed that some of the passwords resembled names of towns or institutions (3%) or names of websites (1%). These were noted to be names of local places or institutions or websites to which participants were affiliated. Nevertheless, these findings on orienting short passwords to languages within the context support propositions in the generic law of development. As such, it can be concluded that the researched context had a multilingual user group. The literature shows that African languages within the context do have an influence on the characteristics of short passwords (Bonneau & Xu, 2012; Wang, Cheng, et al., 2015). However, what is unique in the research findings of this study is a near balanced use of different languages by users during password generation when compared to the Chinese context where Pinyin words/names and numbers were dominant in the corpora (Bonneau & Xu, 2012).

### **7.1.2 A discussion of the findings on mediation**

Socio-cultural theory suggests that different cultural artefacts (symbolic tools), such as language, are used to mediate social interactions and regulate cognitive

activities of thinking and problem solving. These symbolic tools vary according to the context as suggested by the generic law of development. This study argues that these preferred symbolic tools reflect computer users' preferences in password generation as they orient their passwords to different languages and symbolic tools. The previous section showed that participants preferred using symbolic tools (African and Indo-European languages) within their context. The following sections explore these preferred symbolic tools and go on to offer possible explanations for their use within the context of short passwords and passphrases.

***Mediational symbolic tools common in short passwords.*** The results of this study on short passwords show that participants who oriented their passwords to the English language preferred to use English words (20% of the short password corpora) and phrases (4%) when compared to those who adapted English names (12%). The majority of those who oriented their passwords to an African language preferred to adapt African names (26%) over African words (10%). In addition, a small percentage of short passwords were comparable to the most popular short passwords of 2016, 2017 and 2018 according to SplashData. These findings are comparable to those in the literature. For example, Chinese computer users do not prefer to use native words (raw Pinyin words); instead they prefer adapting Pinyin names (Wang, Cheng, et al., 2015). One in four passwords of English users is based on an English name (Wang, Cheng, et al., 2015). However, the Chinese do not prefer to adapt English words; rather they prefer to base their passwords on digits (Bonneau & Xu, 2012; Wang, Cheng, et al., 2015), something that is in contrast to the findings of this study. Bonneau and Xu (2012) suggest that the high use of digits is because Chinese native languages, which are non-Latin-based languages, cannot be supported by the available character key code standards. In addition, this study recorded a low use of keyboard patterns (1%) as short passwords compared to Chinese (more than 8%) and English users (more than 2%) (Li et al., 2014).

This study is of the view that the dominance of English-oriented short passwords could be explained by the fact that English is the dominant language used on technological platforms, where respective mangling rules have since been established

(Deumert & Masinyana, 2008). Mangling rules are established alterations that could be done on substrings (Carrier & Benitez, 2010; Deumert & Masinyana, 2008; Morel et al., 2012). Using established mangling rules reduces the cognitive burden of adapting English words as passwords that require different character classes. African languages, on the other hand, are less used on technological platforms and have no developed mangling rules, something that makes adapting African words into passwords of a different character class less attractive. This argument is consistent with principles in theories of human cognition. According to the embedded process model, it would be easy for a computer user to access the already activated parts of the long-term memory that are composed of known English words and mangling rules during password generation (Schweppe & Rummer, 2014). As a result, when it comes to the use of an African language, it would be easy for users to adapt their names as passwords, something that would reduce the cognitive burden associated with adopting an African word. Based on these arguments, it is concluded that:



*Passwords oriented to African languages are dominated by local names.*

*Passwords oriented to Indo-European languages are dominated by words and phrases.*

Furthermore, close to a third of the participants showed that the inclusion of digits in short passwords does not occur randomly. Rather, the choice of what digits to include in the short password was informed by details in the user's context. Users preferred to include their year of birth or age as part of the resulting password. Some of the users included a four-number digit that resembled a year. In addition, users had a bias towards a small pool of symbols to include in their passwords. The "@" was by far the most used symbol by participants followed by a "#", "!" and a "\$". The "@" is used to concatenate a substring made of alphabetic letters to a segment of the password that is based on digits or used in mangling where a letter "a" is replaced by the "@" symbol. Similarly, the "#", "!" and a "\$" symbol were used in mangling where an alphabetic letter was replaced by a symbol. These practices are consistent with those reported in the literature (Dürmuth et al., 2013; Florêncio et al., 2014). For instance, the order and some of the preferred symbols by users from this study were

comparable to those preferred by the Arabic speaking users: “@”, “\_”, “!”, “\$” or Indians and Pakistanis: “@”, “\*”, “\$”, “#” or the Philippines: “@”, “\_”, “!”, “\*” (AlSabah et al., 2018). Interestingly, users in this study and English speaking users reported in AlSabah et al. (2018) appear to prefer the same top four symbols: “@”, “!”, “\$”, “#” during short password generation. These findings led to the following conclusion:

*The choice of symbols and digits in user-generated passwords does not follow a random distribution.*

**Mediational tools common in passphrases.** Participants used different languages to generate passphrases as required by the passphrase rules. The majority of passphrases had a substring that had the English word together with an African word or an Afrikaans word. African languages are the first spoken languages while English is the language of instruction and first written language (Deumert & Lexander, 2013; Lexander, 2011). It should be noted that the researched context is characterised by a population where users have an Indo-European name and African surname or vice-versa. This could possibly explain the reason why many participants, close to 50% of the corpora, adapted their full names as a suitable passphrase for this study. Bonneau and Shutova (2012) found that users adopted their personal names as passphrases. Generally, names in African languages are long (Deumert & Masinyana, 2008) and if adapted as part of the passphrase, could easily meet the length requirement of 16 characters that was specified for passphrases in this study. It is therefore concluded that:

*Users are likely to adapt their names during passphrase generation.*

In addition, passphrases recorded a drop in the use of digits and symbols. Forty-five per cent (45%) of the passphrases were based on alphabetical letters only. Adding different character classes in passphrases has long been considered less usable (Choong et al., 2014; Shay et al., 2014).

### 7.1.3 A discussion of findings on the generic domains

Socio-cultural theory argues that cultural artefacts (symbolic tools) evolve and change over time. These changes can be necessitated by various factors such as globalisation. Within the context of passwords, users often change existing passwords as they adapt to new password requirements. Von Zezschwitz et al. (2013) researched the way in which user-generated passwords change over time as users adapt existing passwords to new password requirements. This research study sought to establish user behaviour in password reuse by engaging the same participants during short password and multilingual passphrase generation. The results in Chapter 6 showed that substrings of selected multilingual passphrases (6% of the passphrase corpora) were similar in all respects to a short password. It was found that users often adopted a short password and then separated it from the second substring of the multilingual passphrase with a non-character sequence key (space).

In addition, it was observed that a number of participants (7%) removed one to three characters on the right end of their short passwords as they converted these to become part of multilingual passphrase substrings. This practice is in line with the practices found in the literature to be commonly executed (Das et al., 2014; Von Zezschwitz et al., 2013). Das et al. (2014) observed that the majority of users (43%) use identical passwords across different accounts while some users (19%) reuse a substring of other passwords. When reusing a substring of other passwords, resultant passwords are, on average, arrived at after deleting or inserting two to three characters “at the beginning or end or at both ends of a string” (Das et al., 2014, p. 5). The cognition burden associated with multilingual passphrase generation may explain users’ tendency of including a substring of their existing short passwords. Adapting a similar short password, removing one to three characters on the right side of the short password and frequently placing the reused components of the short password on the far-left side of the multilingual passphrase emphasises the magnitude of the cognitive burden faced by users during passphrase generation. Chapter 3 suggests that an already existing password is part of the working memory that can be easily accessed during password generation. This could possibly explain the inclusion of the reused substrings on the far left of the multilingual passphrase followed by a second substring of the multilingual

passphrase that is generated by exploiting the long-term memory. Though widely seen as usable, password reuse has the potential to compromise the strength of multilingual passphrases where an attacker is equipped with knowledge of a user's short passwords. It is, therefore, concluded that:

*Users are more inclined to reuse substrings from short passwords during multilingual passphrase generation.*

## **7.2 Password recall strategies**

Participants were asked to indicate their short password and multilingual passphrase recall strategies. The findings in Chapter 6 (Figure 31) of this study showed that users used similar password recall strategies for both the short password and the multilingual passphrase policy. On both occasions, users showed how they mainly relied on password memorisation. However, differences in password recall strategies were observed across the demographics that were considered in this study in both short password and multilingual passphrase recall strategies. More females than males significantly showed their reliance on writing down passwords or storing them on their phones or sharing passwords with a colleague in case they forgot them. This practice was common during short password and multilingual passphrase recall. This finding could be explained by the fact that males “trust their memory more than females” when recalling passwords (Helkala & Bakås, 2013, p. 352). However, it should be noted that “males reuse fewer passwords than females”, something that might have aided the ability of males to recall passwords (Helkala & Bakås, 2013, p. 352). Li et al. (2016) add that males are more likely to use personal information than females when generating passwords, something that helps password recall.

## **7.3 Factors of multilingual passphrase security and usability**

This section discusses the research findings of this study on factors of secure and usable user-generated multilingual passphrases. The factors of security and usability were analysed by making a comparison between findings on short passwords and multilingual passphrases. Studies advocating for passphrases used a similar approach in their attempts to justify the use of passphrases over short passwords (Melicher et al.,



2016; Shay et al., 2016, 2014). This study motivates the use of multilingual passphrases with the idea of increasing the passphrase search space. Passphrase length, juxtaposing substrings and a dictionary check were used to enhance passphrase strength in this study. Factors of usability that were analysed included effectiveness, efficiency and user satisfaction as defined by the ISO 9241-11 standard (Bevan et al., 2015). Findings on these factors of security and usability are discussed next.

### 7.3.1 Multilingual passphrase security

An overall picture of the research findings showed that multilingual passphrases are stronger than short passwords. This finding is discussed in detail in light of passphrase length, juxtaposing substrings and using a dictionary check.

**Passphrase length:** the study results in Chapter 6 showed that there are security benefits from generating a long password. Figure 34 showed that more short passwords resisted password guessing as their length had increased from the minimum recommended length of eight characters. Furthermore, all passphrases (had a minimum length of 16 characters) resisted password guessing. This finding concurred with a growing belief that passphrases are more secure than short passwords (Kelley et al., 2012; Komanduri et al., 2011; Melicher et al., 2016; Shay et al., 2016). This study used a PCFG that mainly relied on learning from existing passwords in order to guess another password corpus. It is possible that a lack of long passwords in the public domain might have worked in favour of long passwords. Nevertheless, findings from data analysis led to the following conclusion:

*Password length can enhance password security.*

**Juxtaposing substrings:** this study motivated juxtaposing substrings from different languages to enhance password security. This was done to eliminate users' reliance on a few selected words within a language. All participants indicated that they were multilingual; however, English was found to be the dominant second language. As such, potentially users might have oriented their passwords to a single dominant language and ended up basing their passwords on a few common words in that



language (Wang, Cheng, et al., 2015). This was proven by the dominance of English language-oriented short passwords that were easily guessed by the PCFG used in this study. African language-oriented passwords were not easily guessed when compared to English language-oriented passwords. As such, combining substrings from different language orientations during multilingual passphrase generation enhanced passphrase security as shown by their resistance to password guessing. This reasoning is reinforced by the fact that all of the guessed short passwords that were reused as one of the multilingual passphrase substrings did not end up compromising the security of the resulting passphrase. Thus, the strength of 6% of all the multilingual passphrases that were generated through password reuse was not compromised by the inclusion of a guessed short password. It may therefore be concluded that juxtaposing substrings from different language orientations enhanced overall multilingual passphrase security by increasing the passphrase search space. Rao et al. (2013) motivated the idea of increasing the passphrase search space if users were to overcome the use of selected semantics and enhance the strength of passphrases. As such, the following conclusion was made:



*Passphrases generated by juxtaposing substrings from different languages are more secure.*

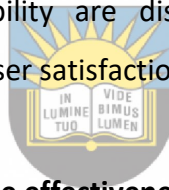
However, it was observed that there was an increased occurrence of adapting full names as multilingual passphrases. This could be explained by the fact that the majority of users' names have both African and English language-oriented substrings. The literature shows that the use of personal information in user-generated passwords has dire security consequences especially when faced with a targeted password attacker (Wang & Wang, 2015; Wang et al., 2016). A trawling attack was used in this study to evaluate multilingual passphrase strength in which none of the passphrases were found to be weak. This finding should be taken with caution given the magnitude of using personal information in user-generated passphrases.

**Dictionary check:** research findings from Chapter 6 suggest that the use of a dictionary check managed to influence users to generate multilingual passphrases

based on more than one language. English-oriented passwords dominated short passwords but this changed in passphrases where substrings were required to be of different languages. Still, English language-oriented substrings remained dominant as they were mixed with either an African language or Afrikaans, a common Indo-European language within the context. By forcing users to orient their substrings to different language orientations, the dictionary check ensured that juxtaposing of substrings was possible.

### **7.3.2 Multilingual passphrase usability**

Research findings showed that short passwords were more usable than multilingual passphrases. Even though multilingual passphrases were found usable, they were not as usable as short passwords, a finding that is consistent with sections of the literature (Keith et al., 2007; Melicher et al., 2016; Shay et al., 2016). Findings on multilingual passphrase usability are discussed next in detail with respect to effectiveness, efficiency and user satisfaction.



#### **7.3.2.1 Multilingual passphrase effectiveness**

Multilingual passphrase usability in terms of effectiveness was evaluated during passphrase generation and recall (memorability).

Multilingual passphrase generation effectiveness was evaluated using data that was gathered using the questions shown in Table 4 in Chapter 6. Subsequent results from the statistical analysis of the data gathered showed that multilingual passphrase generation effectiveness was significantly rated above the mean. This suggests that users found multilingual passphrase generation usable in terms of effectiveness. When compared to short passwords, multilingual passphrase generation was found to be significantly less effective. This study finding is in contrast to the finding by Shay et al. (2016) which showed no significant difference between passphrase and short password generation difficulties. These contrasting findings could be explained by the fact that Shay et al. (2016) used a dictionary check that restricted the use of common passwords on their short password policy, something that might have elevated the complexity of

generating short passwords in their study. In addition, short passwords in this study might have appeared easy to generate because participants simply adapted existing passwords and no blacklist was used. In fact, just over 30% of the participants adapted an already existing short password compared and approximately 15% adapted existing passwords during multilingual passphrase generation. Password reuse is a widely reported password generation strategy in the literature (Bang et al., 2012; Das et al., 2014; Rinn et al., 2015; von Zezschwitz et al., 2013; Woods & Siponen, 2019). Cognitive load theory suggest that password reuse reduces the burden of generating and recalling passwords (Paas & Ayres, 2014; Woods & Siponen, 2019) something that translates to usability.

In addition, nearly 50% of the participants adapted names during short password and multilingual passphrase generation. This might have assisted users during short password generation. However, in the case of multilingual passphrases, typing a long password might have resulted in more typographical errors, something that could potentially frustrate users during multilingual passphrase generation (Keith et al., 2007). Furthermore, some participants appeared to struggle to meet passphrase requirements, for example some participants had to use a sketchpad to generate a multilingual passphrase and double check if the length requirement of at least sixteen characters had been met before keying in their multilingual passphrases on the web application platform. These findings from password generation effectiveness led to the following conclusion:

*Users are more likely to experience negative consequences in their attempt to accurately and completely generate multilingual passphrases than they will when generating short passwords.*

In addition, multilingual passphrase recall effectiveness was evaluated. Table 6 in Chapter 6 showed the questions that were used to gather data for evaluating multilingual passphrase recall effectiveness. Data analysis indicated that multilingual passphrase recall effectiveness was significantly rated above the mean, thus suggesting that participants found multilingual passphrases usable. However, multilingual

passphrase recall effectiveness was significantly rated as less usable compared to the mean rating of short password recall effectiveness. In addition, the mean ratings of multilingual passphrase recall effectiveness for females were significantly lower than those of males. This suggests that females faced more challenges in recalling multilingual passphrases than their counterparts. A study by Shay et al. (2016) found no significant difference between short password and passphrase recall difficulties, although the use of a blacklist by Shay et al. (2016) might have forced participants to generate completely new short passwords. As such, the challenges of recalling passphrases were on a par with those of recalling short passwords. On the other hand, passphrase generation requirements for this study might have forced participants to generate a new passphrase. Reduced password reuse in multilingual passphrase generation suggests that participants had to generate a new passphrase; hence the need for more time to learn and memorise the new passphrase.

Furthermore, participants displayed more typographical errors when logging in using multilingual passphrases. Typographical errors accounted for 26% of all unsuccessful log in attempts when using multilingual passphrases, compared to 10% when logging in using short passwords. Passphrases have many characters to be keyed in, something that increases the likelihood of typographical errors when compared to short passwords (Keith et al., 2009). This was exacerbated by the log in platform for this study which was case sensitive. Participants experienced failure to recall their multilingual passphrases in addition to typographical errors. On average, participants failed to recall their multilingual passphrases on two occasions compared to a short password recall average of 1.3 over a period of two weeks. Outcomes of multilingual passphrase recall failure were dominated by minor errors of failing to recall lower- or upper-case letters and forgetting to include a space between substrings. These results suggest that participants were still aware of their multilingual passphrases, but were failing to recall minor technical details about their structure. It could be argued that participants forgot the technical details of their passphrases as a result of not recalling the password rules that guided them during multilingual passphrase generation. These findings led to the following conclusion:

*Users are more likely to experience negative consequences in their attempt to accurately and completely recall multilingual passphrases than they will when recalling short passwords.*

### **7.3.2.2 Multilingual passphrase efficiency**

Data was gathered from the system logs to ascertain the usability of multilingual passphrases with regards to efficiency. This evaluation was done during multilingual passphrase generation and recall.

Data gathered on the number of multilingual passphrase generation attempts and the time taken to generate a multilingual passphrase was used to evaluate efficiency. It was assumed that more multilingual passphrase or short password generation attempts would suggest that the password policy was less usable (Shay et al., 2016). The results showed that it required 4.5 attempts, on average, to generate a multilingual passphrase whereas it had taken 2.1 attempts, on average, to generate a short password. These findings suggest that multilingual passphrases were found less usable with regard to efficiency when compared to short passwords. The average number of attempts that were required to generate a short password by users in this study was comparable to the average number of attempts reported in the literature: 1.2 and 2.4 attempts respectively (Melicher et al., 2016; Shay et al., 2016). However, participants in this study required far more attempts to generate a passphrase compared to the average of 2.1 and 1.92 that was reported in the literature (Melicher et al., 2016; Shay et al., 2016).

In addition, data on the average time taken to generate a multilingual passphrase and short password corroborated the average number of attempts required. Participants required, on average, 247 seconds to generate a multilingual passphrase and 82 seconds to generate a short password. A non-parametric test using a Wilcoxon signed ranks test showed that multilingual passphrase generation significantly required more time when compared to the time needed to generate short passwords. Short password reuse and a lack of exposure to passphrase policies could explain the huge difference between the two samples. Data was gathered on every

attempted short password and passphrase during generation. Analysis showed that participants went through different steps during the multilingual passphrase generation cycle. It was very rare for participants to adapt an initially thought passphrase as the final multilingual passphrase. For instance, participants often started by considering a short password which was deleted as they moved on to try other substrings for their multilingual passphrases. In addition, participants appeared to be testing the restrictiveness of the multilingual passphrase policy as they keyed in a single substring at first or it could be that users had forgotten to include a space between substrings of the multilingual passphrase. All these activities explain the huge difference in the average amount of time that was required to generate multilingual passphrases and short passwords with those reported in the literature. Furthermore, a bigger average multilingual passphrase generation time might boil down to differences in the demographics of participants and their exposure to ICTs. It is likely that participants in Melicher et al.'s (2016) and Shay et al.'s (2016) studies were from the First World, where the targeted participants might have had exposure to passphrase generation. For instance, all participants in Melicher et al. (2016) were from the USA and were engaged in using Amazon's Mechanical Turk crowdsourcing services, of which, Bonneau and Shutova (2012) noted that Amazon makes use of a two- word passphrase and a PIN as a form of authentication. Based on these findings it was concluded that:

*It requires more resources to efficiently generate multilingual passphrases than short passwords.*

Efficiency in multilingual passphrase recall was analysed. Data gathered by use of key logs on the time taken to type a multilingual passphrase and the number of multilingual passphrases recall attempts were used to evaluate efficiency. Table 14 shows that it took 133 seconds, on average, to completely and accurately key in a multilingual passphrase on the first login, three days after multilingual passphrase generation. At this stage, participants had required 30 seconds, on average, to key in their short passwords. The general trend was that the time required to key in a multilingual passphrase constantly went down until 49 seconds was reached on the last day of the experiment. Short password generation experienced a similar trend, as

participants had taken 21 seconds, on average, to key in their short passwords on the last day of the experiment. Possible explanations could be that users had rehearsed their multilingual passphrases over time to such an extent that they had become memorable, while the occurrence of typographical errors had also become less frequent. This is consistent with a proposition that constant rehearsal aids memorability (Bonneau & Schechter, 2014; Keith et al., 2007; Miller, 1956; Woods & Siponen, 2019). This was clearly demonstrated as participants required, on average, 2 attempts to recall a multilingual passphrase while 1.3 attempts were needed to recall short passwords. It was, therefore, concluded that:

*Multilingual passphrase users initially experience a high number of login failures due to typographical and memorability errors.*

#### **7.3.2.3 Multilingual passphrase user satisfaction**

Chapter 6 presented findings on user satisfaction during multilingual passphrase generation and recall. These findings are discussed below.

To evaluate multilingual passphrase generation user satisfaction, data was gathered on the perception of participants using questions shown in Table 8 in Chapter 6. This study findings from statistical analysis showed that multilingual passphrase generation was significantly usable according to user satisfaction ratings. However, multilingual passphrase generation was found significantly less usable when compared to short password generation. Challenges faced during users' attempts to accurately and completely generate multilingual passphrases might have negatively influenced their attitude towards the multilingual passphrase policy. The findings in Chapter 6 showed that password generation effectiveness has a positive relationship with user satisfaction. As such, it is possible that participants experienced more cognitive load during multilingual passphrase generation. The multilingual passphrase policy appears to have broken a common password generation tradition of password reuse, something that increased the mental effort needed to generate a multilingual passphrase. Shay et al. (2016) made a similar finding as they noted "that forcing users to break their

password-creation habits can increase difficulty” (p. 29). Based on these findings, the following two conclusions were made:

*Migrating from a short password policy to a multilingual passphrase policy will likely impact users’ attitude during passphrase generation.*

*Negative consequences experienced when attempting to accurately and completely generate a multilingual passphrase negatively affect a user’s attitude towards a multilingual passphrase policy.*

In addition, participants were asked to indicate their perception on multilingual passphrase recall user satisfaction. Questions in Table 9 of Chapter 6 were used to gather data for evaluating users’ satisfaction with multilingual passphrase recall. The results of the statistical analysis showed that both multilingual passphrase and short password’ user satisfaction recall were significantly rated usable. Although short password user satisfaction recall had a slightly higher mean rating than multilingual passphrase user satisfaction recall, the difference was not significant. In addition, there was no significant relationship between multilingual passphrase recall effectiveness and user satisfaction. A possible explanation could be that password recall does not have a major effect on users’ attitudes towards a password policy. Password recall failure is something that happens often in the literature (Choong et al., 2014). It could be that users have become accustomed to forgetting a new password to the extent that they consider this normal and acceptable. The literature suggests that users might not be overly worried about certain usability concerns because they have become accustomed to such concerns (Melicher et al., 2016; Stobert & Biddle, 2014). It is therefore concluded that:

*Migrating from a short password policy to a multilingual passphrase policy does not affect users’ attitudes towards password recall.*



#### 7.4 Multilingual passphrase length and security

Section 6.5.3, Table 21, shows that a passphrase of 14 characters long appeared to be more user-friendly to participants. Thus, multilingual passphrase usability ratings started to decline once the passphrase length exceeded 14 characters. However, a comparison of usability ratings between passwords with a maximum length of 14 characters and those with a length that ranged between 15 and 21 characters did not yield any significant difference (see Section 6.5.3). Furthermore, none of the 16-character long passphrases were guessed, as shown in Section 6.6.2. It is therefore concluded that an ideal multilingual passphrase length would be 16 characters. Shay et al. (2016) also recommend a two-word passphrase that is 16 characters long.

#### 7.5 Chapter summary

This chapter discussed research findings presented in Chapter 6 and made comparisons with findings in the literature. The chapter discussed password characteristics according to principles in socio-cultural theory. The analysis of the short password corpora showed that user-generated passwords were oriented to different languages within the researched context. This justified propositions by socio-cultural theory that appear to suggest that contextual symbolic tools (language) influence password characteristics. Just like western computer users, participants in this study adopted English language-oriented words and phrases as passwords. Furthermore, the use of African languages in password generation was mainly influenced by the adoption of personal information (for instance, names) as passwords. The use of symbols and digits in short passwords was done in predictable ways. Although multilingual passphrases reduced the practice of password reuse, the adaptation of names as passwords remained dominant. An analysis of password recall strategies showed that females were more likely to write down their passwords when compared to males.

An analysis of password strength showed that multilingual passphrases were stronger than short passwords. This finding was supported by previous findings in the literature. Passphrase strength in this study was attributed to the use of juxtaposed substrings, increasing passphrase length and using a dictionary check to enforce the use of multilingual passphrases. An analysis of usability showed that multilingual

passphrase generation and recall was usable. However, multilingual passphrase generation and recall were not as user-friendly as that of short passwords. It was reasoned that password reuse might have resulted in a biased opinion in favour of short passwords. As purported in theories of memorability, continued use of multilingual passphrases was shown to bridge the usability gap between multilingual passphrase and short password recall. The study findings indicated that effectiveness, efficiency and user satisfaction were important during multilingual passphrase generation, while only effectiveness and efficiency were found important to multilingual passphrase recall.

The next chapter uses the findings from this chapter to evaluate the proposed model of secure and usable multilingual passphrases.



University of Fort Hare  
*Together in Excellence*

## CHAPTER 8: RESEARCH CONTRIBUTIONS AND RECOMMENDATIONS

### 8.0 Introduction

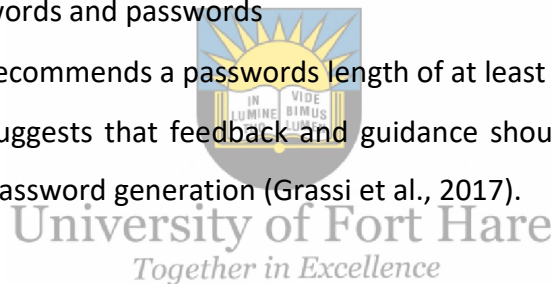
The idea of using passwords for authentication purposes can be traced to ancient Roman times (Adeka et al., 2013). Over the years, studies have found that text-based authentication mechanisms are associated with a number of security and usability limitations. This led to the proposition of different solutions aimed at enhancing the security and usability of passwords. This study developed a model for secure and usable multilingual passphrases, arguing that the problem of passwords could best be addressed by focusing on contextual factors in order to identify measures for attaining passphrase security and usability. In a way, this study moved away from generalising password security and usability challenges. Accordingly, this chapter presents the research contributions of this study, which followed design science research guidelines. Hence, this chapter reviews the way this study went through the critical steps of design science research, which can be broadly split into designing and evaluating (March & Smith, 1995, in Venable et al., 2012). The chapter starts with a synopsis of the study followed by a review of the theoretical foundation to reflect the use of kernel theories in the study as required by design science research. The chapter goes on to revisit the proposed model of secure and usable multilingual passphrases which symbolises the artefactual design of the study. An evaluation framework of the proposed model using primary data is included. The evaluation leads to the modification of the proposed model in line with primary data research findings. The chapter then makes recommendations for enhancing the security and usability of multilingual passphrases.

### 8.1 A synopsis of the study

The literature shows that password security and usability have been a cause for concern as far back as the 1970s, as users struggle to find a balance between secure and usable passwords. For instance, users base their passwords on a few selected semantics such as popular words, personal information (names, phone number, identification number, address and date of birth) and predictable keyboard patterns (AlSabah et al., 2018; Li et al., 2016; Renaud et al., 2019; Wang, Cheng, et al., 2015; Weir, 2010). Some

users base their passwords on unique cultural traits and native languages found within the context (AlSabah et al., 2018; Wang, Cheng, et al., 2015). One of the main highlights in the literature that addresses password security and usability is the NIST's Digital Identity Guideline of 2006 that motivated the use of principles in Shannon entropy. Millions of passwords that have been guessed using probabilistic guessing algorithms since 2009 exposed the limitations of using the principles in Shannon's entropy. It was found that users fulfilled password requirements that were guided by Shannon's entropy principles in predictable ways. This research progress saw the NIST proposing a different perspective on password security and usability. This was reflected by the NIST's publication of a SP 800-63B that

- discourages the use of keyboard patterns and personal information in passwords
- proposes the use of a blacklist to prohibit the adaptation of common words and passwords
- recommends a passwords length of at least eight characters and
- suggests that feedback and guidance should be given to users during password generation (Grassi et al., 2017).



Even though passphrases have been sounded for being secure and usable, there are highlights in the literature that shows users may not always be able to generate secure and usable passphrase (Bonneau & Shutova, 2012; Keith et al., 2007; Rao et al., 2013; Shay et al., 2016). The use of popular words in a language is one of the primary concerns raised in the literature when it comes to user generated passphrases (Shay et al., 2016). The literature recommends the use of a blacklist in order to restrict users from adapting popular words in a language as passphrases (Melicher et al., 2016; Shay et al., 2016). However, there are fears that a long blacklist may make passphrases less usable or popular words will change over time making a blacklist less effective or the fact that passwords differ with language may complicate the development of a complete blacklist (AlSabah et al., 2018; Blocki et al., 2013; Florêncio et al., 2014a).


This study motivated the use of multilingual passphrases in order to enhance security by increasing the passphrase search space. Design science research and mixed research methods were used. An experiment was used for data gathering and raw passwords were gathered by asking participants to generate a short password and a multilingual passphrase under pre-specified conditions. To demonstrate memorability, participants were asked to login to their profiles once every three days over a period of two weeks. Short passwords and multilingual passphrases were then sent for password guessing to test for strength using a PCFG algorithm. The password guessing results show that short passwords were weaker than multilingual passphrases. Fifty per cent of the short passwords were guessed compared to none of the multilingual passphrases. Multilingual passphrases in this study proved stronger against a PCFG when compared to those reported in the literature. For example, more than 20% of the 2word16 character passphrases were guessed in a study that was conducted by Shay et al. (2016). Using a blacklist that restricted the use of common words reduced the amount of guessed passphrases to below 20% even for passphrases that were generated using a mobile phone (Melicher et al., 2016; Shay et al., 2016). In terms of usability, short passwords appeared to be more user-friendly than multilingual passphrases. The following sections revisit the steps that were followed during the designing or building phase of the design science research that led to the proposed model.

#### **8.1.1 The theoretical foundation**

This study was grounded in socio-technical theory. Efforts were made to give a balanced view on both the social and technical subsystems in an attempt to address the password security and usability challenge. Chapter 3 used socio-cultural theory to explain the social subsystem of this study, while Chapter 4 explained the technical subsystem of this study. The following sections provide an overview of the social and technical subsystems for this study, leading to a review of the proposed model of secure and usable multilingual passphrases. It is important to give an overview of the activities that led to the proposed model as this creates a basis for understanding the evaluation framework for this study.

#### 8.1.1.1 The social subsystem

Chapter 3 of this study argued that Information Systems is a multidisciplinary subject domain where theories from other research disciplines can be used to address research problems. Accordingly, socio-cultural theory was adopted to explain the view of this study on the social subsystem. Socio-cultural theory argues that human mental development is guided by participation in a social environment. The theory proposes three principles for explaining psychological development within the context, namely, the genetic law of development, mediation and genetic domains. These principles have been widely used to explain high mental activities that are argued socially constructed rather than biologically constructed alone (Mercer & Howe, 2012; D. Shin, 2014). This study regards password generation as the use of a higher order mental activity whose occurrence can be explained by the principles in the socio-cultural theory. These principles of socio-cultural theory are reviewed below within the context of this study.



**The generic law of development.** The generic law of development argues that a computer user's settings as determined by "culture, language, history, peer groups and institutional structures at school or workplace play a critical role in shaping the initial human mental development" (Lantolf et al., 2015). A literature review in Chapter 3 demonstrated that the language landscape of Africa is characterised by a multilingual society. As such, computer users within this context are expected to speak and write at least two different languages. The use of multiple languages is expected to translate to passwords oriented in different languages that are found in the context. For example, passwords oriented in native Hebrew, Spanish, English, Chinese, Arabic languages etc were observed in the literature (AlSabah et al., 2018; Bonneau & Shutova, 2012; Li et al., 2016; Wang, Cheng, et al., 2015).

**Mediation.** Socio-cultural theory advances the notion that social interactions and cognitive activities are regulated and mediated by the use of different cultural artefacts (symbolic tools). These cultural artefacts differ from one context to the next, as indicated by the generic law of development. Research findings from the analysis of more than 100 million publicly leaked passwords support the argument advanced by the mediational principle that computer users from different contexts adopt different

symbolic tools. Chapter 3 observed that Chinese computer users orient their passwords to purely digit-based characters when compared to English users who base their passwords on concatenated English words and digits or words in the English dictionary. The use of Pinyin names in Chinese passwords, English names in the passwords of the English computer users and native Greek language-oriented passwords in the passwords of Greek computer users reflects the impact of the mediational principle. Furthermore, contextual cultural beliefs have been found to influence the choices of symbolic tools used in user-generated passwords. For example, Chinese computer users often include the numbers 6 and 8 as these are culturally seen as lucky numbers, while a 4 is believed to be an unlucky number and less frequently used (Yang et al., 2013). This study used the mediation principle to identify contextual symbolic tools of interest during password generation.

**The generic domains.** Socio-cultural theory proposes that higher order mental functionality is always in motion and goes through continuous change (Marginson & Dang, 2017). Similarly, the symbolic tools that are used to mediate interactions are also in constant flux as generations inherit and modify cultural artefacts. Within the context of passwords, such modifications can be necessitated by changes in password requirements. Jakobsson and Dhiman (2013) researched different password modification strategies in line with password requirements, noting that users make spelling mistakes, insertions, concatenate different character classes and replace different character classes. A longitudinal study by Von Zeszschwitz et al. (2013) also found that user-generated passwords often evolve over time as users reuse their passwords according to different password requirements. This study used the generic domain to model the practice of password reuse by adapting existing passwords to different password requirements.

#### **8.1.1.2 The technical subsystem**

Chapter 4 reviewed different authentication mechanisms based on what one knows, what one has and what one is. Passwords, a type of what one knows authentication mechanism, were found to be the most dominant form of authentication. Chapter 4 went on to identify online and offline password threats. The

popularity of offline password threats, explained in Section 4.2, has increased the importance of using stronger passwords. There is a general consensus in the literature that a secure password is one that is not easily guessed by a password guessing algorithm (Dell’Amico & Filippone, 2015). This view reveals a shift from the traditional theoretical view of defining a strong password as one that is made up of different character classes as perceived by Shannon’s entropy. A number of measures have since been suggested to promote the generation of strong and usable passwords. These include password composition policies, system assigned passwords, combined system and user-generated passwords, and PSMs. Associated limitations were discussed. Findings from Chapter 4 were used to inform the design of the proposed model in Chapter 5.

#### **8.1.1.3 A proposed model for secure and usable multilingual passphrases**

Chapter 5 of this study proposed a model that could be used to guide the generation of secure and usable multilingual passphrases. The proposed model is designed for environments where authentication and user identification are of high importance. Such an environment could correspond to the NIST’s Authentication Assurance Level Three, where security breaches of authentication credentials could expose the victim to high financial loss and inconvenience (Grassi et al., 2017). The model in Chapter 5 proposed that juxtaposing substrings from different languages, increasing passphrase length and using a dictionary check could be considered when promoting the generation of secure passphrases. This study adapted ideas of the NIST’s use of Shannon’s entropy and extended it to the generation of passphrases oriented to different languages. By so doing, the proposed model sought to tap into the social context (social subsystem) of this study, which is characterised by a multilingual user group, and encouraged the use of multilingual passphrases for enhancing passphrase security. In addition, the model proposed by this study adopted factors of usability that are in the ISO 9241-11 standard. These include effectiveness, efficiency and user satisfaction.



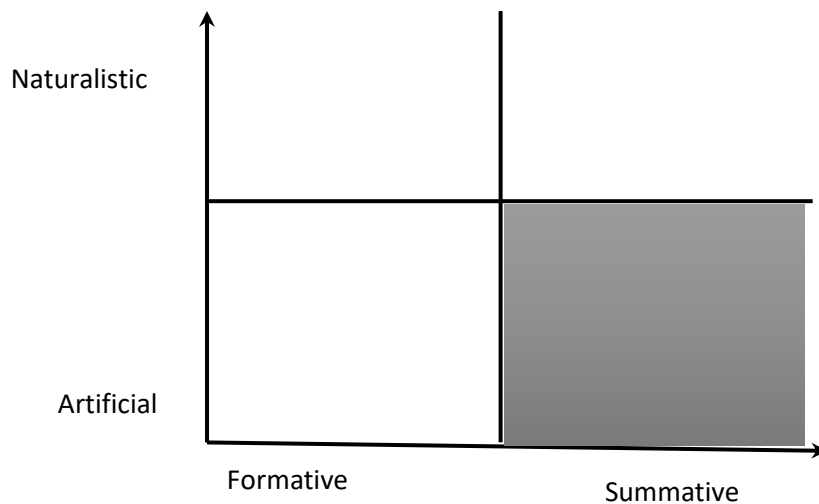
## 8.2 The evaluation framework

Evaluation is a crucial step in design science research as it validates the quality, utility and efficacy of the artefact and, to some extent, indicates why the proposed artefact is a better solution (Hevner et al., 2004; Pries-Heje, Baskerville, & Venable, 2008). Venable et al. (2012) also state that evaluation is the central and most important activity of design science research. Without evaluation, design science research outputs are unsubstantiated and their use cannot be justified. Furthermore, the evaluation process is the one that qualifies design science as a scientific research. March and Smith (1995, in Pries-Heje et al., 2008) define evaluation as “the development of criteria and the assessment of the artefact’s performance in comparison to the criteria” (p. 258). This section presents the development of evaluation criteria that were used in this study. The focus of the evaluation is on quality, efficacy and utility. This chapter goes on to use the evaluation criteria to assess the proposed model.

While this study followed the design science research process proposed by Peffers et al.(2008), Venable et al. (2012) noted that Peffers et al. (2007) did include a guideline to be followed when choosing different evaluation methods. There is a general consensus in the literature that available design science research process models and guidelines do not provide adequate guidance for choosing evaluation methodologies and strategies (Prat et al., 2015; Pries-Heje et al., 2008; Venable et al., 2012). This study adapted a comprehensive and widely validated evaluation framework that was proposed by Venable et al. (2012). Venable, Pries-Heje and Baskerville (2016) named the evaluation framework FEDS (Framework for Evaluation in Design Science Research).



University of Fort Hare  
*Together in Excellence*



**Figure 35. FEDS dimensions**

The FEDS is composed of two-dimensional characteristics for evaluating design science research as shown in Figure 35. On one end of the dimension is a functional purpose of conducting an evaluation (formative or summative evaluation), while the other end of the dimension shows different paradigms with which the evaluation could be conducted (artificial or naturalistic). According to Venable (2006, in Baskerville et al., 2015), artificial evaluation relates to a non-realistic way of evaluating a solution such as an experiment, while naturalistic evaluation involves the evaluation of a technological solution in its real environment using the actual users. The use of an experiment in this study suggests that the study assumed an artificial evaluation.

The evaluation process constitutes different strategies for conducting the evaluation. The choice of a strategy depends on contextual factors such as costs, time and the potential risks. In addition, Venable et al. (2012) propose the following four steps for conducting the FEDS process:

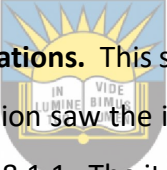
1. Define the purpose or goal of evaluation.
2. Identify the evaluation strategy or strategies.
3. Define the attributes to be evaluated.
4. Design individual evaluations.

### 8.2.1 The evaluation framework for this study

This study assumed the four steps proposed by Venable et al. (2012) in order to conduct the FEDS process as follows:

- i. **Define the purpose or goal of evaluation.** Venable et al. (2012) suggest at least four goals of evaluating a technological solution. These include rigour, uncertainty and risk reduction, ethics and efficiency. This study focused on attaining rigour and ethics. For the purposes of rigour, the study ensured that the proposed model encouraged the generation of secure and usable passphrases when compared to the short password policy. The aim was to establish the utility, efficacy and quality (or lack thereof) of the proposed model such that areas of improvements could be enlisted. It should be noted that this study gathered data using two separate password policies that paved the way for the comparison of the utility, efficacy and quality of the policies, based on their outcomes. For ethical reasons, the study ensured that participants were not disadvantaged following their participation in the study.
- ii. **Identify the evaluation strategy or strategies.** Venable et al. (2016) identified different strategies for evaluating a technological solution. Early formative evaluations were conducted in this study to ensure rigour. Given that this is a scientific study, users had to be engaged at some point. As such, this study assumed the technical risk and efficacy strategy where an experiment was used as the testbed. This strategy was adopted to ensure rigour, while reducing the costs associated with evaluating the proposed model (artefact) using real users in a real setting.
- iii. **Define the attributes to be evaluated.** This study evaluated the security and usability constructs in the proposed model. The choices of evaluations were guided by the literature in line with a recommendation by Venable et al. (2012). Password security was measured in terms of the number of guessing attempts required to guess user-generated passwords in this study. A comparison, based on the guessed passwords, was done to establish a more secure password policy

of the two policies that were considered in this study. Password characteristics were also evaluated to establish the influence of the password policy and the contextual factors. In regard to usability, t-tests were used to evaluate the perception of users on password policy effectiveness and user satisfaction. Potential correlations between the constructs were also evaluated. Data gathered using the system logs was used to evaluate the efficiency of the researched password policies. Given that two policies were considered in this study, any significant difference between the mean ratings of the samples from different password policies was interpreted as a reflection of the importance of the factor under evaluation. It was also important to evaluate the methodology used in the study. Accordingly, Cronbach's alpha coefficient and factor analysis were used to evaluate the reliability and validity of the data gathered for this study, according to explanations in Section 2.6.

- 
- iv. **Design individual evaluations.** This study was conducted in two iterations. The first iteration of evaluation saw the initial design of the proposed model which was explored in Section 8.1.1. The iteration was formative and artificial as it did not include any users in evaluating the artefact. The initial design of the artefact ensured that propositions from the literature were made. The second iteration saw a summative and artificial evaluation of the proposed model and induced rigour by engaging users through an experiment. Findings from this iteration were used to modify the initially proposed model. The next section presents the second iteration of the artefact's evaluation.

### **8.3 Evaluating a model for secure and usable multilingual passphrases**

This section reports on the evaluation of the proposed model. The first part of the evaluation focuses on multilingual passphrase security followed by multilingual passphrase usability.

### 8.3.1 Evaluating multilingual passphrase security

The proposed model identified juxtaposing substrings, passphrase length and dictionary check as factors for enhancing passphrase strength. These factors are evaluated below to establish the utility of the proposed model.

In particular with regard to **juxtaposing substrings**, the model of secure and usable multilingual passphrases in Chapter 5 made the following proposition:

***Proposition P1: Multilingual passphrases generated by juxtaposing substrings from different languages are more secure.***

This proposition was found to be true; as a result, the proposition was retained. It was noted that all multilingual passphrases that were based on juxtaposed substrings resisted password guessing. The justification for accepting proposition P1 is explained below:



Chapter 6 showed that African language-oriented short passwords were slightly more resistant to guessing when compared to English language-oriented short passwords. As such, combining substrings from African and Indo-European oriented languages is expected to have enhanced the security of passphrases. Moreover, the inclusion of popular substrings in the multilingual passphrases was not enough to compromise the overall security of resulting passphrases. It is therefore argued that juxtaposing substrings from different languages compensated for any possible security weaknesses that resulted from the use of common substrings. However, the use of personal information was found to be dominant in the multilingual passphrase corpora, something that might compromise security in a targeted attack. The literature showed that a trawling attack might struggle to guess passphrases based on personal information, whereas targeted attacks are more effective in guessing passwords based on personal information (Li et al., 2016; Veras et al., 2014; Wang et al., 2016). It is, therefore, recommended that the use of personal information in multilingual passphrases should be restricted. This suggestion calls for the modification of Proposition P 1 on juxtaposing substrings as follows:

**Proposition P1:** *Restricting the use of personal information in user generated passphrases by juxtaposing substrings from different languages can enhance security.*

In relation to **passphrase length**, the model in Chapter 5 proposed that:

**Proposition P2:** *Passphrase length and underlying passphrase structures can enhance passphrase security.*

This proposition was found to be true; hence, it was retained in the original model. The reasons for returning this proposition are as follows:

*The results of this study discussed in Chapter 6 showed that there is a constant drop in short password guessing effectiveness as password length increases from eight to fifteen characters. None of the short passwords that were at least sixteen characters long was guessed. In addition, none of the multilingual passphrases that were at least sixteen characters long was guessed. Hence, it can be concluded that a passphrase length of sixteen characters in multilingual passphrases yields better security. Hence, proposition P2 was adjusted as follows:*

**Proposition P2:** *Increasing the passphrase length to at least 16 characters in a multilingual passphrase that is based on at least two substrings enhances passphrase security.*

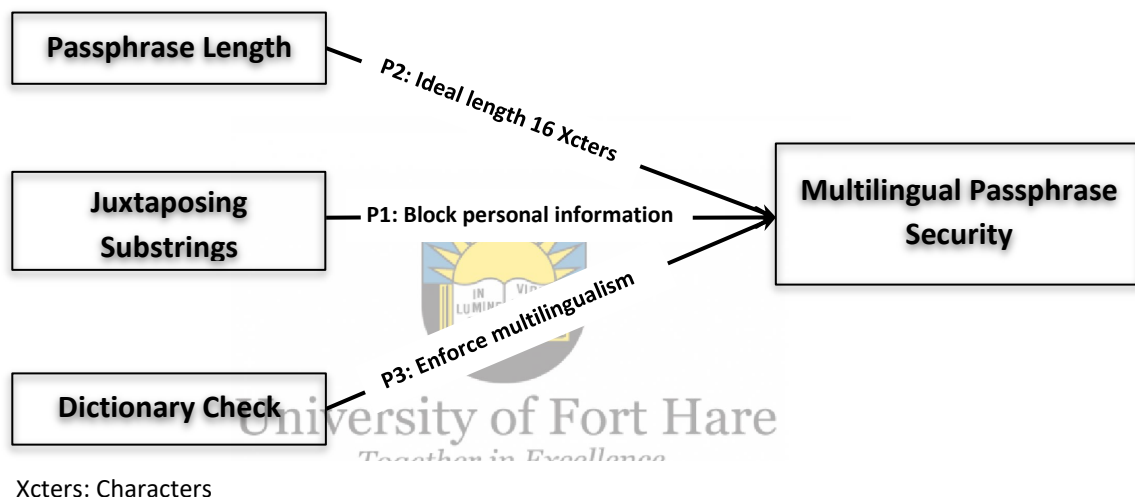
Chapter 5 of this study used a **dictionary check** to enforce the generation of passphrases based on multilingualism. As such, it was proposed that:

**Proposition P3:** *The use of dictionary checks can motivate users to base their passphrases on multiple languages.*

The study found this proposition to be true. A dictionary check in conjunction with pre-specified password rules forced participants to generate passphrases based on multilingualism, which contributed to secure passphrases. As stated earlier, none of

the resultant multilingual passphrases was guessed by the password guessing algorithm used in this study.

In light of the above findings on propositions for multilingual passphrase strength, this chapter proposed a modified model termed the *Multilingual passphrase security model*, shown in Figure 36. The model shows that passphrase length, juxtaposing substrings from different languages and a dictionary check contribute to passphrase security. The model also indicates that it is important that the juxtaposed substrings are not based on personal information such as user names.



**Figure 36. Multilingual Passphrase Security Model**

### 8.3.2 Multilingual passphrase usability

The factors of multilingual passphrase usability that were investigated included effectiveness, efficiency and user satisfaction. These factors were derived from the ISO 9241-11 standard, as explained in Chapter 5. System logs and a questionnaire were used to gather data for evaluating factors of usability. This section evaluates the factors of multilingual passphrase usability focusing on effectiveness, efficiency and user satisfaction. The evaluation is limited to two phases of a password lifecycle, that is, passphrase generation and recall.

### 8.3.2.1 Multilingual passphrase generation usability

This section evaluates the usability of multilingual passphrase generation in terms of effectiveness, efficiency and user satisfaction.

In particular with regard to **passphrase effectiveness**, the model of secure and usable multilingual passphrases in Chapter 5 proposed that:

*Proposition P4: The ability to effectively generate, memorise and type in a multilingual passphrase without experiencing negative consequences enhances passphrase usability.*

This proposition is evaluated in this section, focusing on passphrase generation effectiveness. A reliability analysis using a Cronbach's alpha of 0.916\*\* in Table 19, Chapter 6, shows a high reliability coefficient for the instrument that was used to assess multilingual passphrase effectiveness. In addition, Chapters 6 and 7 showed that **multilingual passphrase generation effectiveness** was significantly rated above the mean. However, the mean ratings of multilingual passphrase generation effectiveness were found to be significantly lower than those of short passwords. This finding suggests that even though participants considered multilingual passphrase generation usable, they found the process much more demanding when compared to generating short passwords. Cognitive load theory suggests that, in light of the limited capacity of the short-term memory, multilingual passphrases may have required more effort to generate than short passwords. This is a plausible explanation given the magnitude of password reuse that was observed during the generation of short passwords. Nevertheless, proposition P4 was modified to reflect the study finding as follows:

***Proposition P4(i):** The ability to effectively generate a multilingual passphrase without experiencing negative consequences positively influences passphrase usability.*

In addition, Chapter 5 made the following proposition on **passphrase efficiency**:

*Proposition P5: Efficacy in multilingual passphrase generation, recall and typing in leads to passphrase usability.*



Proposition P5 was evaluated using data in Chapters 6 and 7. Accordingly, it was shown that **multilingual passphrase generation efficiency** required more resources when compared to short password generation. For example, participants required more time and attempts to generate multilingual passphrases when compared to short passwords. However, as is the case with multilingual passphrase generation effectiveness, caution should be exercised with regard to this finding given the magnitude of password reuse during short password generation. Nevertheless, proposition P5 was modified to reflect the research findings as follows:

***Proposition P5(i): Efficacy during multilingual passphrase generation positively influences passphrase usability.***

Furthermore, Chapter 5 made the following proposition in relation to **multilingual passphrase user satisfaction**:

***Proposition P6: User satisfaction with a multilingual passphrase policy leads to passphrase usability.***



University of Fort Hare  
Together in Excellence

A reliability analysis using a Cronbach's alpha of 0.921\*\* in Table 19, Chapter 6, shows a high reliability coefficient for the instrument that was used to gather data on multilingual passphrase user satisfaction. Chapters 6 and 7 indicated that **multilingual passphrase generation user satisfaction** was significantly rated above the mean. However, the mean ratings for multilingual passphrase generation user satisfaction were found to be significantly lower than those of short passwords. This finding suggests that even though participants considered multilingual passphrase generation usable, they were demoralised by the complications that were associated with multilingual passphrase generation. Accordingly, proposition P6 was adjusted as follows:

***Proposition P6: User satisfaction with a passphrase policy during multilingual passphrase generation leads to passphrase usability.***

In addition, an interesting observation emerged from the data following a correlation analysis. A correlation analysis revealed a significant positive ( $p = 0.0001$ ) relationship between multilingual passphrase generation effectiveness and user satisfaction. This finding suggests that experiencing little to no negative consequences from completely and accurately generating a multilingual passphrase leads to user satisfaction and vice-versa. Accordingly, a proposition P7 was formulated to reflect this finding:

***Proposition P7: Effective multilingual passphrase generation positively influences user satisfaction with the passphrase policy.***

In light of multilingual passphrase generation effectiveness and efficiency challenges, it is recommended that the design of a multilingual passphrase generation prompt should include an option for users to display the passphrase as it is being keyed in. In addition, providing real-time feedback on multilingual passphrase length during generation could inform users whether they are meeting the length requirements. It is also argued that these measures for improving multilingual passphrase generation effectiveness and efficiency will reflect positively on user satisfaction.

#### **8.3.2.2 Multilingual passphrase recall usability**

This section evaluates the usability of multilingual passphrase recall (memorability) in terms of effectiveness, efficiency and user satisfaction.

The previous section indicated that Chapter 5 made the following proposition in relation to **multilingual passphrase effectiveness**:

***Proposition P4: The ability to effectively generate, memorise and type in a multilingual passphrase without experiencing negative consequences enhances passphrase usability.***

A reliability analysis using a Cronbach's alpha of 0.805\*\* in Table 19, Chapter 6, shows a high reliability coefficient for the instrument that was used on multilingual passphrase recall effectiveness. Chapters 6 and 7 showed that **multilingual passphrase**

**recall effectiveness** was significantly rated above the mean. However, the mean ratings of multilingual passphrase recall effectiveness were found to be significantly lower than those of short passwords. This finding suggests that, even though participants consider multilingual passphrase recall usable (effective), they found the process much more demanding when compared to recalling short passwords. Participants struggled to recall upper-case or lower-case letters and forgot to include a space between substrings in a multilingual passphrase. Furthermore, participants struggled to recall the sequence of substrings in a multilingual passphrase or completely forgot other substrings in a multilingual passphrase. Proposition P4 was therefore modified to reflect these findings as follows:

***Proposition P4(ii): Effectively recalling a multilingual passphrase without experiencing negative consequences leads to passphrase usability.***

As indicated earlier, Chapter 5 made the following proposition on **multilingual passphrase efficiency**:



***Proposition P5: Efficacy in multilingual passphrase generation, recall and typing in leads to passphrase usability.***

Chapters 6 and 7 showed that **multilingual passphrase recall efficiency** required more resources at first when compared to short password recall. For example, participants required more time to recall multilingual passphrases and made more login attempts when compared to short passwords. In addition to memorability challenges, participants experienced typographical errors when logging in using multilingual passphrases. However, the average amount of time needed to recall a multilingual passphrase became shorter over time, something that could be attributed to users who were learning and becoming accustomed to their new multilingual passphrases. In light of these findings, proposition P5 was modified as follows:

***Proposition P5(ii): Repeated use of a multilingual passphrase over time positively influences the usability of passphrases.***

Lastly, the previous section indicated that Chapter 5 made the following proposition in relation to **multilingual passphrase user satisfaction**:

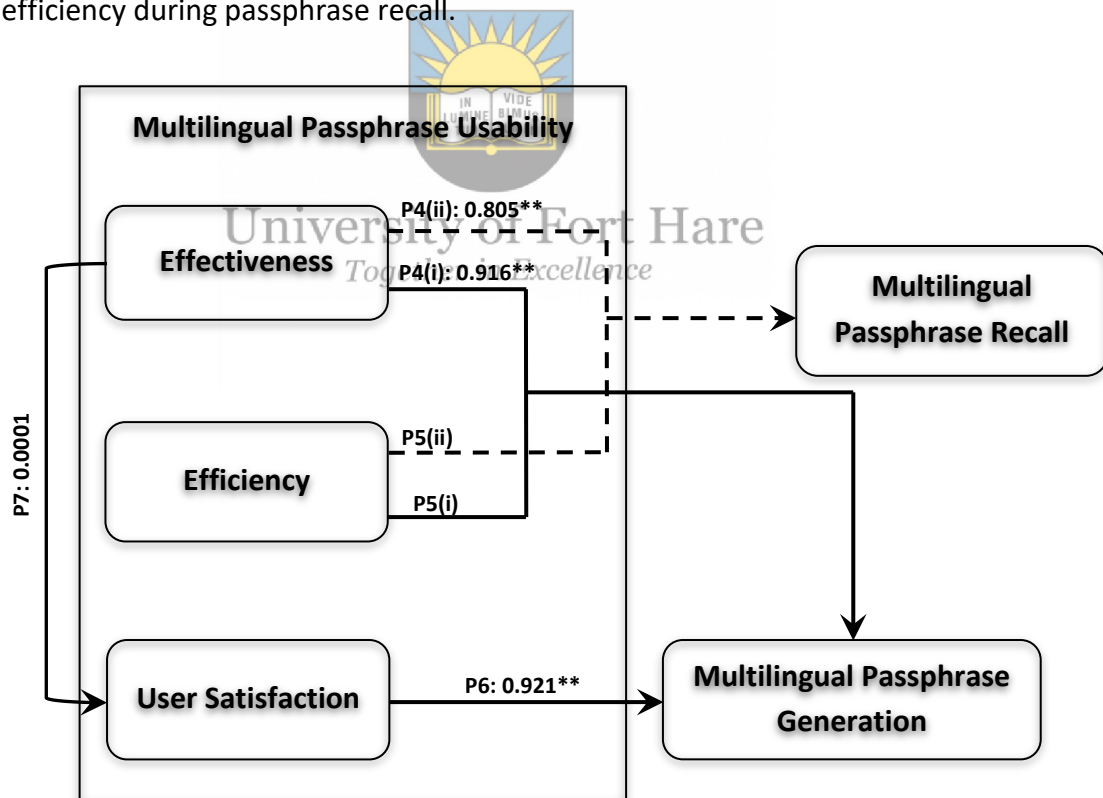
***Proposition P6:*** *User satisfaction with a multilingual passphrase policy leads to passphrase usability.*

Chapters 6 and 7 showed that **multilingual passphrase recall user satisfaction** was significantly rated above the mean. Furthermore, it was noted that the mean ratings of multilingual passphrase recall user satisfaction were not significantly lower than those for short passwords. A subsequent correlation analysis between multilingual passphrase recall effectiveness and user satisfaction did not show any significant relationship. This finding suggests that even though users found multilingual passphrases difficult to recall when compared to short passwords, they were not overly concerned by the challenges faced. This was justified by the adoption of similar password recall strategies for both the seemingly more user-friendly short password policy, according to user satisfaction ratings, and the less user-friendly multilingual passphrase policy. Based on these arguments, it was concluded that migrating from a short password policy to a multilingual passphrase policy will not affect users' perceptions of multilingual passphrase recall user satisfaction. Hence, this factor (user satisfaction) was deemed not to be important to multilingual passphrase recall.

In light of these findings on multilingual passphrase recall effectiveness, efficiency and user satisfaction, recommendations were made to improve the usability of multilingual passphrases. For instance, making provision for users to display a multilingual passphrase as it is being keyed into the login prompt could help address typographical errors and other memorability challenges. Ignoring the space between multilingual passphrase substrings and removing case sensitivity while logging in could also help alleviate some of the multilingual passphrase recall challenges. In addition, displaying multilingual passphrase generation requirements while logging in could be considered as an alternative solution for helping users recall technical details such as passphrase length, the requirement of a space between substrings and, in a way,

reminding users that the passphrase is supposed to have upper- and lower-case letters. Lastly, a warning could be displayed when logging in with a multilingual passphrase that warns users to be wary of transposing substrings in their passphrases.


Based on the findings on multilingual passphrase generation and recall, Figure 37 shows a revised model for usable multilingual passphrases, indicating that the usability factors proposed in the ISO 9241-11 are important during passphrase generation. Organisations implementing a multilingual passphrase should realise that the factors of effectiveness, efficiency and user satisfaction are important during passphrase generation. However, migrating from a short password to a multilingual passphrase does not significantly change users' perceptions of user satisfaction passphrase recall. As such, this study concludes that organisations wishing to adopt a multilingual passphrase policy should adopt measures that promote effectiveness and efficiency during passphrase recall.



**Figure 37. A model for Usable Multilingual Passphrases**

#### 8.4 Chapter summary

This chapter presented the study contributions, giving a synopsis of the study and revisiting its theoretical foundation. The chapter went on to give an overview of the proposed model of secure and usable multilingual passphrases. Revisiting the formulation of the proposed model was critical for the model's evaluation. The chapter then outlined a framework that was used in this study to evaluate the model. Artefact evaluation is a critical step when conducting a design science research. Hence, the proposed model in Chapter 5 was evaluated with a focus was on the constructs of security and usability, and their relationships. Some of the propositions in Chapter 5 were found to be not true. Hence, the model in Chapter 5 had to be modified to reflect the research findings. The final revised model was presented in two parts to clearly reflect on the modifications that had been effected as a result of the findings from the data analysis and interpretation. These parts are *A Multilingual Passphrase Security Model* and *A Model for Usable Multilingual Passphrases*.



The Multilingual Passphrase Security Model showed that passphrase length, juxtaposing substrings and using a dictionary check can enhance the security of multilingual passphrases. Furthermore, it was noted that restricting the use of personal information in juxtaposed substrings of a multilingual passphrase is critical. A Model for Usable Multilingual Passphrases showed that factors of usability differ with respect to the activities of passphrase generation and recall. The results of this study showed that effectiveness, efficiency and user satisfaction are critical during multilingual passphrase generation. While effective passphrase generation was found to be positively influencing user satisfaction, only the usability factors of effectiveness and efficiency were found critical during multilingual passphrase recall, with user satisfaction being found not to be relevant. In light of these findings, recommendations were made that could be considered in order to enhance the usability of multilingual passphrases.

The next chapter concludes the study and make suggestions for future research.

## CHAPTER 9: CONCLUSION

### 9.0 Introduction


This study developed a model of secure and usable multilingual passphrases for a multilingual user group. The literature suggests that the research on password security and usability dates back to the late 1970s following a publication by Morris and Thompson. Technological advancements have seen different authentication mechanisms being proposed with the aim of addressing the security and usability predicaments. These include biometrics and token-based authentications. Nevertheless, text-based authentication mechanisms in the form of passwords remain a dominant form of authentication. There is general consensus that passwords (text-based) will remain the primary form of authentication into the foreseeable future (AlSabah et al., 2018; Dell’Amico et al., 2010; Li et al., 2014; Mazurek et al., 2013; Wang, Cheng, et al., 2015; Woods & Siponen, 2019).

Various approaches to text-based password authentication have been advanced with the aim of enhancing security and usability (ISACA, 2015; AlFayyadh et al., 2012). This study motivated the use of multilingual passphrases. This chapter concludes the research by giving an overview of the study, followed by revisiting the research sub-questions and reviewing the study contributions. The limitations of this study are explained and areas for possible future research are identified. Finally, the concluding remarks signify the end of this chapter.

### 9.1 An overview of this study

This study was grounded in socio-technical theory. Chapter 3 and 4 looked at the socio- and technical subsystems following the primary principle of socio-technical theory of giving equal attention to the social and technical aspects of the phenomena under study. The literature review provided a plethora of evidence that supported the thought that certain password characteristics are a reflection of a user’s language. For example, the Markov chain password guessing algorithm exploits character distribution patterns in a user’s language to guess user-generated passwords (Narayanan & Shmatikov, 2005). Wang, Cheng, et al. (2015) observed differences in password

character distribution as a result of a users' language. They concluded that "passwords from different languages are intrinsically different from each other in letter distributions, and that passwords are close to their native language" (Wang, Cheng, et al., 2015, p. 5). Similarly, Bonneau and Xu (2012) found that local user languages influence the characteristics of user-generated passwords. In addition, the culture of a user has been found traceable in user-generated passwords (AlSabah et al., 2018). As such, Chapter 3 of this study sought to understand the influence of contextual factors in one's language development. Hence, socio-cultural theory was used to explain how language development is a result of contextual factors. The context of this study was found to be characterised by a multilingual user group. Text messages were used to reflect the magnitude of multilingualism within the research context. The study argued that exploiting multilingualism could promote the generation of secure and usable passphrases.




Chapter 4 explored the technical subsystem of the study and identified offline and online password threats as the security threat model for this study. It was found that probability and heuristic-based password guessing algorithms could be used to exploit user-generated passwords in an offline password attack. As such, a password that could resist password guessing was considered strong and secure in this study. A literature review of password guidelines, best practices and policies was carried out. Chapter 4 also established that while there are some interesting propositions on password guidelines, best practices and policies, attaining a secure and usable password remains a challenge. User behaviours that sought to circumvent the spirit of generating secure passwords were observed. Similarly, it was found that the private sector is reluctant to adopt password policies and guidelines that are commensurate with the levels of risks associated with authentication security breaches. There are suggestions that the available password policies are not usable; hence, their implementation might frustrate clients.

Chapter 5 used the findings in Chapters 3 and 4 to propose a model of secure and usable multilingual passphrases, which was arrived at through an abduction reasoning process. The proposed model exploited the characteristics of the research



context and motivated the use of multilingual passphrases. Passphrase security was considered to be a factor of juxtaposing substrings from different languages, increasing passphrase length and using a dictionary check to enforce the use of substrings from different languages in a passphrase. The literature motivates the use of passphrases without paying attention to the language orientation in resulting passphrases (Bonneau & Shutova, 2012; Komanduri, 2016; Melicher et al., 2016; Rao et al., 2013; Shay et al., 2016). This is a huge concern given the effects native languages and culture have had on short password structure and security something that could be extended to passphrases (AlSabah et al., 2018; Wang, Cheng, et al., 2015). There are already concerns that users often base their passphrases on popular words in a language (Bonneau & Shutova, 2012; Shay et al., 2016). As such, this study proposed the use of multilingual passphrases with the hope that, an increase in passphrase search space will enhance security. Effectiveness, efficiency and user satisfaction were considered measures of usability as defined in the ISO 9241-11 standard of usability.



Chapter 6 presented the findings gathered on the proposed model, using an experiment survey as explained in Chapter 2. The experiment saw users generating a short password and a multilingual passphrase with the aim of evaluating a better approach to text-based password authentications. Participants were asked to give their opinion on the usability of password generation under the different policies. Chapter 7 went on to discuss the findings of this study, while Chapter 8 used the findings from Chapters 6 and 7 to evaluate the proposed model. Finally, Chapter 8 presented a revised model. The model in Chapter 8 addresses the main research question of this study:

*How can local languages be exploited in order to improve the security and usability of passphrases?*

Design science research guidelines were followed in this study, which were explained in Chapter 2.

## 9.2 Research questions and findings

For a research study to be complete, it is important that all the research sub-questions are addressed. Chapter 1 outlined the main research question from which four sub-questions were derived. Chapter 2 of this study indicated that the secondary and primary data were gathered in order to address the research sub-questions and meet the research objectives. Table 22 below summarises the activities of data gathering that were done to address the research sub-questions of this study.

**Table 22. Data collection techniques and research sub-questions of this study**

Research sub-questions	Literature review	Experiment	survey
What are the different password policies in use?	X		
What are the language characteristics that could be considered to enhance the security of user-generated passphrases?	X	X	X
What are the factors affecting the usability of passphrases?	X	X	X
What are the password characteristics of a multilingual user group?	X	X	

The research activities that were carried out to address each research question are explained below:

- *What are the different password policies in use?*

Chapter 4, Sections 4.4.1 and 4.4.2 explored the different password guidelines, best practices and policies in use. The observed password guideline frameworks included the NIST's SP800-63B; the eID Interoperability for Pan European Electronic Government Services for the European Union; the Framework for Authentication and Non-Repudiation in Electronic Communication for the Norwegian public sector, and the National e-Authentication Framework for the Australian government (AlFayyadh et al., 2012; Grassi et al., 2017). The NIST authentication guidelines (NIST SP800-63 and recently the SP800-63B) are the most influential guidelines to date for designing of password policies (Wheeler, 2016). A combination of propositions in these password

guidelines include an outline of measures for motivating users to generate secure and usable passwords, a recommendation for storing passwords in encrypted format and enforcement of regular password changing.

A further literature review indicated that some of the principles in the password guidelines and best practices may not assure the security of user-generated passwords from online and offline password attacks. For instance, the security contributions made by storing passwords in encrypted format could be jeopardised by service providers' tendency to use weak hashing algorithms or users' behaviour in terms of generating simple and easy-to-guess passwords (Bauman et al., 2015; Florêncio et al., 2014a). In addition, regular password changing was seen to be easily compromised by password reuse and users found the approach frustrating (Chiasson & van Oorschot, 2015; Choong et al., 2014; Rinn et al., 2015; Zhang et al., 2010). These findings suggested a need for password policies that encourage users to generate strong and usable passwords.



Section 4.4.2.1 went on to explore the different password policies. The policies found in the literature were the password composition policy, the system and user-generated password policy, the system assigned password policy and PSMs (Houshmand & Aggarwal, 2012; Wang & Wang, 2015; Weir et al., 2010). The main security limitations were found to be the use of personal information and fulfilling password composition policies in predictable ways (Wang et al., 2015; Weir et al., 2010). In this regard, the literature recommended the use of a blacklist or passphrases to enhance the security of passwords generated under a password composition policy (Bonneau & Shutova, 2012; Braunstein, 2015; Grassi et al., 2017; Shay, Cranor, et al., 2016). In addition, frustration during password generation, password reuse, failure to recall passwords and authentication challenges (typographical errors) were found to be some of the usability limitations of the password composition policy (Choong et al., 2014; Melicher et al., 2016). The literature recommended the use of feedback and guidance during password generation (Furnell et al., 2018; Grassi et al., 2017). The use of passphrases was also recommended as a measure that could promote the usability of a password composition policy (Shay et al., 2016). System and user-generated

password policies are yet to be validated while system assigned password policy has been found difficult to recall. Furthermore, Mwangabi et al. (2014) found that PSMs merely nudge users into generating a strong password if a user perceives there is a need to do so. In addition, implementing PSM restrictively might frustrate users during password generation (Ur et al., 2012).

Given these findings on password guidelines, best practices and policies, this study sought to explore the security and usability of multilingual passphrases through this sub-question:

- *What are the language characteristics that could be considered to enhance the security of user-generated passphrases?*

Section 3.4.2 in Chapter 3 explored the practice of code-switching in the text messages of African computer users. It was observed that users within the research context of this study had the potential to generate a phrase constituting substrings from different languages. This study went on to extend ideas in the NIST use of Shannon's entropy to promote passwords based on multiple character classes/sets, thus motivating the use of multilingual passphrases. Section 4.2 in this study identified different password threats. Subsequently, offline password threats that could make unlimited numbers of password guessing attempts were considered a password threat model for this study. As such, to be considered secure, multilingual passphrases were expected to resist offline password guessing. A proposed model in Chapter 5 conceptualised the use of multilingual passphrases.

User-generated multilingual passphrases that were gathered using an experiment were exposed to a password guessing algorithm (PCFG) in an offline password attack. The findings from password guessing showed that the use of multilingual phrases enhanced password security, as none of the passphrases were guessed. Most of the participants who generated multilingual passphrases also generated short passwords. The short passwords were also exposed to an offline password attack. Consequently, just over 50% of the short passwords were guessed in comparison to none of the

multilingual passphrases. Similarly, just over 50% of the short passwords in Shay et al. (2016) were guessed by the PCFG used in this study. However, unlike findings in the literature that less than 20% of passphrases were guessed using PCFG (Melicher et al., 2016; Shay et al., 2016), none of the multilingual passphrases in this study were guessed. Based on these findings, it was concluded that passphrases based on multilingualism helped users generate secure passphrases.


- *What are the factors affecting the usability of passphrases?*

Sections 4.4.2 and 4.4.3 explored different usability challenges faced by users during password generation and logging in. Section 5.3.2.1 went on to consolidate all the usability challenges that were observed in the literature. These challenges were then aligned to factors of usability in the ISO 9241-11 standard. This standard proposes three usability factors, namely, effectiveness, efficiency and user satisfaction. Accordingly, accurately generating a passphrase, accurately typing in a passphrase, accurately recalling passphrases, meeting passphrase requirements, the use of a passphrase reminder, storage of passphrases on other media, use of semantic information and passphrase reuse were considered measures of effectiveness. Time taken to type in a passphrase, passphrase creation attempts, passphrase recall attempts, time taken to generate a passphrase and the number of passphrase re-entry attempts were considered measures of efficiency. User satisfaction was measured by participants' attitude towards the passphrase policy.

Data collection and analysis showed that effectiveness, user satisfaction and efficiency usability were important during passphrase generation. It was further found that passphrase generation effectiveness positively influenced user satisfaction. In addition, effectiveness and efficiency were found to be significant for passphrase recall, while user satisfaction was found not to be significant in this regard. Females struggled more than their male counterparts to effectively recall their passphrases and short passwords.

- *What are the password characteristics of a multilingual user group?*

Sections 4.2.2.2, 4.4.2 and 4.4.3 identified different characteristics of user-generated passwords. For example, Weir et al. (2009) in Section 4.2.2.2 noted that user-generated passwords could be split into password structures such as digits (D), symbols (S) and alphabetic letters (L). Other researchers describe password characters according to character composition such as LUDS where, L: lower-case alphabetic letter, U: upper-case alphabetic character, D: digits and S: symbols (Wheeler, 2016). Furthermore, the use of semantic information and keyboard patterns were found to be other common password characteristics in the literature (Li et al., 2016; Wang et al., 2016). Jakobsson and Dhiman (2013) suggest another approach to classifying password characteristics, observing that user-generated passwords are a result of concatenation, replacement, spelling mistakes and insertion. In addition, Bonneau (2012) identified what he refers to as global passwords to reflect common passwords.



A content analysis of 224 short passwords and 176 passphrases revealed some of the password characteristics that were observed in the literature. Twenty-three per cent and 21% of the short passwords reflected LDS and LSD password structures respectively, according to a definition by Weir et al. (2009), while 79% of the short passwords were generated following Jakobsson and Dhiman's (2013) concatenation. Close to 50% of the participants indicated that they had adapted a name to generate a short password and 3% of the short password corpora resembled global passwords. The short password corpora in this study reflected the principles of socio-cultural theory with user-generated passwords being found to be oriented to languages in the context.

In addition, the multilingual passphrase corpora showed that the majority (45%) of passphrases were based on lower-case alphabetic letters (L), 15% had passphrases that were based on lower- and upper-case letters (LU) and 14% were based on LUDS. These passphrases were mainly based on two substrings (78%), while 4% of the passphrase corpora had substrings that could be traced to global passwords.

### 9.3 Research contributions

This study makes contributions to the field of text-based authentication mechanisms by proposing a model of secure and usable multilingual passphrase generation for a multilingual user group. Given that this study followed design science research guidelines, it is worthwhile to identify the contribution it makes in line with the purported design science research contributions. With respect to a design science knowledge contribution framework proposed by Gregor and Hevner (2013), this study contributed an “improvement” (p. 245) or an improved solution to an existing problem. Thus, password security and usability challenges have been known since the 1970s and this study proposed an improved model of using multilingual passphrases when compared to the currently popular eight-character password requirement and monolingual passphrases mainly generated with a use of a restrictive blacklist. Figure 4 in Chapter 2 shows the contribution of this study according to Gregor and Hevner (2013).



Furthermore, studies on passphrases in the literature focus on the generation of long passwords with substrings separated by a space (Melicher et al., 2016; Rao et al., 2013; Shay et al., 2016). No study, to the best of our knowledge, has explored the use of multilingual passphrases. The solution proposed by this study improves passphrase security, although passphrases were found not to be as usable as short passwords. Unlike studies in the literature, for example Shay et al. (2016), this study did not make use of a blacklist of short passwords; this is something that may have exaggerated the usability of the short password policy in this study. In addition, a high rate of password reuse during short password generation might have influenced the outcome. Cognitive load theory suggests that reusing the same information reduces the amount of effort needed during password generation while aiding memorability (Woods & Siponen, 2019). This is very important given the limited capacity of the short-term memory, which is where password generation, retaining and recall occur (Atkinson & Shiffrin, 1968; Miller, 1956; Woods & Siponen, 2019). Nevertheless, it was encouraging to observe that participants were better able to login to their profiles over time, something that emphasised continued improvement in passphrase usability as participants proceeded with the experiment.

In addition to developing a new solution, the following knowledge contributions were made:

- The study showed that females are more likely to write down or store passwords in some media in case they forget. This is even more pronounced when females are exposed to a multilingual passphrase generation policy.
- The study used the same participants during short password and multilingual passphrase generation. Accordingly, practices of password reuse when participants migrated from a short password to a multilingual passphrase policy were observed. For example, where a user reused a password in a two-word multilingual passphrase there was a greater chance that the first word or substring on the far left of the multilingual passphrase would be the user's short password and only the second substring in the multilingual passphrase would be a newly generated substring. The second strategy applied to reuse passwords during multilingual passphrase generation was the adaptation of a short password that involved the removal of digits and symbols on the far right of the short password. Again, the adapted short password would constitute the first substring on the far left of the multilingual passphrase. Only once in the multilingual passphrase corpora was it observed that a participant based the passphrase on a short password that was repeatedly used as the first and second substring of the multilingual passphrase. However, the multilingual passphrase policy led to a reduction in password reuse when compared to short passwords.
- Another knowledge contribution was the establishment of adapting names during multilingual passphrase generation. It was observed that the chances of adapting a name as a multilingual passphrase remained the same following a migration from a short password to a multilingual passphrase policy.
- In addition, it was found that the attitude of a user towards a password policy is not likely to be affected by challenges faced during password recall. However,



challenges associated with password generation would certainly affect a user's attitude towards the password policy.

- This study also made important contributions on multilingual passphrase recall challenges. It was found that common challenges faced by users included a failure to remember upper-and lower-case letters, forgetting to include a space between substrings as well as the transposing of substrings in the original multilingual passphrase. Addressing these three challenges would improve multilingual passphrase usability by close to 60%.
- Finally, this study took a different perspective on passwords and the way they are used by adopting socio-cultural theory. Hence, the study demonstrated the influence of contextual factors, with a primary focus on language, on password generation and use.

#### 9.4 Limitations of the study

This study made important contributions. However, the following limitations were noted:

- Firstly, the experiment and data collection were limited to university students. It might be possible that the students, on average, may have produced better results given that they are at a university.
- Secondly, the experiment in the study could not use mild deception in such a way that participants would have been unaware that they were participating in a password-related study. Rather, all the engaged participants were aware that they were taking part in an experiment. The magnitude of impact that this might have had on ecological validity is not clear. However, Fahl, Harbach, Acar, and Smith (2013, in Shay et al., 2016) found that experimental passwords revealed similar characteristics to real passwords.
- A third important limitation of this study was that the sample that took part in the experiment was relatively small when compared to samples reported in the literature. However, it has to be noted that this study used the same group of



University of Fort Hare  
*Together in Excellence*

participants for both multilingual passphrase and short password generation and use.

- The fourth limitation is that, the study did not gather basic long passphrases to establish if it was not length alone that promoted the strength of passphrases in this study. Future studies can explore the security and usability contributions of African monolingual passphrases when compared to multilingual passphrases. However, it has to be highlighted that the findings in the present study that the majority of users adopted short passwords that were oriented towards the English language suggest that, African computer users are likely to orient their passphrases in Indo-European languages that are the first written and official languages. Hence, it is possible that resultant passphrases of African user would most likely assume characteristics that have already been reported in the literature of adapting common words in a language something that compromise security (Bonneau & Shutova, 2012; Rao et al., 2013; Shay et al., 2016). In addition, multilingual passphrases reported in this study appear to be more secure when compared to those in the literature (Melicher et al., 2016; Shay et al., 2016).
- Lastly, the study used a trawling attack to test password security. While this is a dominant practice described in the literature (Melicher et al., 2016; Shay et al., 2016, 2014; Weir et al., 2009), studies that used targeted password attacks have been shown to yield better results (Houshmand et al., 2015; Li et al., 2016; Rao et al., 2012; Veras et al., 2014; Wang et al., 2016).

### 9.5 Direction for future research

This study proposed a model of secure and usable passphrases. This section suggests future research areas for the model:

- Firstly, it is important to apply the model in a real environment using deception in such a way that participants are unaware that they are participating in an experiment. Although Shay et al. (2016, in Maoneke et al., 2018) argue that password experiments could be “designed in such a way that participants can simulate password generation and treat the process in the same manner they would when generating real passwords” (p. 37), it would be worthwhile to

establish how the model proposed by this study would perform in a real environment.

- Secondly, the proposed model could be implemented with a PSM to establish its influence on passphrase strength and usability. The PSMs are widely regarded as recommendation policies used to nudge users into generating strong and usable passwords (von Zezschwitz et al., 2013). As such, it would be interesting to establish how users would behave if they were to be asked to generate a multilingual passphrase using a PSM. It should be noted that the multilingual passphrase policy used in this study saw users being required to meet the multilingual passphrase policy requirements instead of being simply “asked” to do something, which is common with PSMs.
- Lastly, increasing the sample size may help the generalisability of the propositions of the study. Research on multilingual passphrases is scarce and further validation with a bigger sample would further consolidate findings from this study. While some studies have based their conclusions on fewer than 100 participants, for example Von Zezschwitz et al. (2013), other studies have based their conclusions on bigger samples (Kelley et al., 2012; Melicher et al., 2016; Shay et al., 2016, 2014).

## 9.6 Conclusion

This study proposed a model of secure and usable multilingual passphrases for a multilingual user group. The study argues that understanding the language terrain in the context of the users can play a pivotal role in enhancing password strength and usability. This is supported by the literature that shows that a poor focus on the use of linguistic properties in password generation could potentially compromise security (AlSabah et al., 2018; Rao, Jha, & Kini, 2013; Wang, Cheng, et al., 2015). For instance, the dominant use of popular words in a language has been found common in studies that explore security contributions of passphrases (Bonneau & Shutova, 2012; Shay et al., 2016; Veras et al., 2014). Accordingly, probability-based guessing algorithms take advantage of this skewed distribution in users’ selection of words during passphrase generation. While the use of a blacklist could prohibit the adaptation of common words

in a language, there are concerns that common words in a language differ from one context to the other (Blocki et al., 2013; Florêncio et al., 2014a). Besides, the list of common words or passwords may change overtime. This is critical to systems administrators who do not have access to the password database in order to influence the generation of evenly distributed short passwords and passphrases using a blacklist. This study promotes the generation and use of a multilingual passphrase. The study argues that using multilingual passphrases can increase the passphrase search space thereby enhancing security. Participants in this study showed that it may be possible for a multilingual user to generate strong multilingual passphrases when compared to short passwords based on LUDS. The study showed that, while the dominance of English language within the context may influence password choices, adopting multilingual passphrases can give users an opportunity to generate secure and usable passphrases. However, a closer look at the password corpora showed that users are likely to adapt personal identification information as passwords even if a multilingual passphrase policy is adopted. From a usability point of view, it was encouraging to note that users were able to learn, memorise and recall their multilingual passphrases overtime. Accordingly, this chapter presented key contributions of this study including contributions to the design science methodology of proposing an improved artefact, theoretical contributions made by extending socio-cultural theory to passwords, as well as knowledge contributions. This was followed by an overview of the study limitations. In conclusion, Section 9.5 presented a direction for future research, which could help validate the proposed model and enhance its generalisability.

## REFERENCES

- Adeka, M., Shepherd, S., & Abd-alhameed, R. (2013). Resolving the Password Security Purgatory in the Contexts of Technology, Security and Human Factors. In *International Conference on Computer Applications Technology* (Vol. 2013). Sousse, Tunisia: IEEE.
- Ågerfalk, P. (2010). Getting pragmatic. *European Journal of Information Systems*, 19(3), 251–256.
- Al-Ameen, M. N., Fatema, K., Wright, M., & Scielzo, S. (2015). The Impact of Cues and User Interaction on the Memorability of System-Assigned Recognition-Based Graphical Passwords. In *2015 Symposium on Usable Privacy and Security (SOUPS) 2015*, (pp. 185–196). Ottawa: USENIX Association.
- Al-Ameen, M. N., Wright, M., & Scielzo, S. (2015). Towards Making Random Passwords Memorable: Leveraging Users' Cognitive Ability Through Multiple Cues. In *Enhanced Security with Passwords & CAPTCHAs* (pp. 2315–2324). Seoul: ACM.
- AlFayyadh, B., Thorsheim, P., Jøssang, A., & Klevjer, H. (2012). Improving Usability of Password Management with Standardized Password Policies. In *7th Conference on Network and Information Systems Security*. Cabourg.
- Alomari, R., & Thorpe, J. (2019). On Password Behaviours and Attitudes in Different Populations. *Journal of Information Security and Applications*, 45(2019), 79–89.
- AlSabah, M., Oligeri, G., & Riley, R. (2018). Your Culture is in Your Password: An Analysis of a Demographically-Diverse Password Dataset. *Computers & Security*, 77(2018), 427–441.
- Andersson, D., & Saedén, D. (2013). *Authentication with Passwords and Passphrases-Implications on Usability and Security*. Retrieved from <https://www.rlvision.com/blog/wp-content/uploads/2013/12/Authentication-with-Passwords-Passphrases-Implications-on-Usability-and-Security.pdf>
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human Memory: A proposed System and Its Control Processes. In K. W. Spence & J. T. Spence (Eds.), *The Psychology of Learning and Motivation. Advances in Research and Theory* (Vol. 2, pp. 89–191). New York: Academic Press.
- Babb, J., Keith, M., & Steinbart, P. (2016). Can Relaxing Security Policy Restrictiveness Improve User Behavior? A Field Study of Authentication Credential Usage. In *49th*

- Hawaii International Conference on System Sciences* (pp. 4803–4812).  
Washington DC: IEEE Computer Society.
- Bailey, D. V., Dürmuth, M., & Paar, C. (2014). Statistics on Password Re-use and Adaptive Strength for Financial Accounts. In M. Abdalla & R. De Prisco (Eds.), *Proceedings of the International Conference on Security and Cryptography for Networks* (pp. 218–235). Amalfi: Springer.
- Bang, Y., Lee, D., Bae, Y., & Ahn, J. (2012). Improving Information Security Management: An Analysis of ID–Password Usage and a New Login Vulnerability Measure. *International Journal of Information Management*, 32(5), 409–418.
- Baskerville, R. L., Kaul, M., & Storey, V. C. (2015). Genres of Inquiry in Design-science Research: Justification and Evaluation of Knowledge Production. *MIS Quarterly*, 39(3), 541–564.
- Bauman, E., Lu, Y., & Lin, Z. (2015). Half a Century of Practice: Who is Still Storing Plaintext Passwords? In *Information Security Practice and Experience*. (pp. 253–267). Beijing: Springer.
- Bell, E., Bryman, A., & Harley, B. (2015). *Business Research Methods. Fifth Edition*. Oxford: Oxford University Press.
- Bevan, N., Carter, J., Earthy, J., Geis, T., & Harker, S. (2016). New ISO Standards for Usability, Usability Reports and Usability Measures. In M. Kurosu (Ed.), *International Conference on Human-Computer Interaction* (pp. 268–278). Cham: Springer.
- Bevan, N., Carter, J., & Harker, S. (2015). ISO 9241-11 Revised: What Have We Learnt About Usability Since 1998? In *International Conference on Human-Computer Interaction* (pp. 143–151). Cham: Springer.
- Blanchard, N., Malaingre, C., & Selker, T. (2018). Improving Security and Usability of Passphrases With Guided Word Choice. In *Proceedings of the 34th Annual Computer Security Applications Conference* (pp. 723–732). San Juan: ACM New York.
- Blocki, J., Komanduri, S., Procaccia, A. D., & Sheffet, O. R. (2013). Optimizing Password Composition Policies. In *Proceedings of the 14th ACM Conference on Electronic Commerce* (pp. 105–122). Philadelphia: ACM.
- Bonneau, J. (2012). The Science of Guessing: Analyzing an Anonymized Corpus of 70

- Million Passwords. In *2012 IEEE Symposium on Security and Privacy* (pp. 538–552). California: IEEE Computer Society.
- Bonneau, J., & Schechter, S. (2014). Towards Reliable Storage of 56-bit Secrets in Human Memory. In *Proceedings of the 23rd USENIX Security Symposium* (pp. 607–623). San Diego: USENIX Association.
- Bonneau, J., & Shutova, E. (2012). Linguistic Properties of Multi-word Passphrases. In J. Blyth, S. Dietrich, & J. L. Camp (Eds.), *International Conference on Financial Cryptography and Data Security* (pp. 1–12). Berlin: Springer.
- Bonneau, J., & Xu, R. (2012). Of contrase nas , תואמסיס , and 密码 Character encoding issues for web passwords. *Citeseer*, 1–8.
- Braunstein, P. (2015). *Making Secure Easy-to-Remember Passwords*. Retrieved from <http://www.cs.tufts.edu/comp/116/archive/fall2015/pbraunstein.pdf>
- Campbell, J., Ma, W., & Kleeman, D. (2011). Impact of Restrictive Composition Policy on User Password Choices. *Behaviour & Information Technology*, 30(3), 379–388.
- Carrier, L. M., & Benitez, S. Y. (2010). The Effect of Bilingualism on Communication Efficiency in Text Messages (SMS). *Multilingua-Journal of Cross-Cultural and Interlanguage Communication*, 29(2), 167–183.
- Carstens, D. S., Malone, L. C., & McCauley-Bell, P. (2006). Applying Chunking Theory in Organizational Password Guidelines. *Journal of Information, Information Technology, and Organizations*, 1, 97–113.
- Castelluccia, C., Dürmuth, M., & Perito, D. (2012). Adaptive Password-Strength Meters From Markov Models. In *Network and Distributed System Security Symposium 2012* (pp. 1–14). San Diego: Internet Society.
- Chatterjee, R., Bonneau, J., Juels, A., & Ristenpart, T. (2015). Cracking-Resistant Password Vaults using Natural Language Encoders. In *IEEE Symposium on Security and Privacy* (pp. 481–498). San Jose: IEEE Computer Society.
- Chiasson, S., & van Oorschot, P. C. (2015). Quantifying the Security Advantage of Password Expiration Policies. In *Designs, Codes and Cryptography* (pp. 401–408). Springer.
- Chiluwa, I. (2008). Assessing the Nigerianness of SMS Text-Messages in English. *English Today*, 24(1), 39–44.



- Choong, Y.-Y., Theofanos, M., & Lui, H.-K. (2014). *United States Federal Employees' Password Management Behaviors – a Department of Commerce Case Study*. Retrieved from [https://ws680.nist.gov/publication/get\\_pdf.cfm?pub\\_id=914843](https://ws680.nist.gov/publication/get_pdf.cfm?pub_id=914843)
- Ciampa, M. (2013). A Comparison of Password Feedback Mechanisms and Their Impact on Password Entropy. *Information Management & Computer Security*, 21(5), 344–359.
- Collis, J., & Hussey, R. (2013). *Business Research: A Practical Guide for Undergraduate and Postgraduate Students*. Nature.
- Cowan, N. (2000). The Magical Number 4 in Short-term Memory: A Reconsideration of Mental Storage Capacity. *Behavioral and Brain Sciences*, 24(4), 87–185.
- Cowan, N. (2014). Working Memory Underpins Cognitive Development, Learning, and Education. *Educational Psychology Review*, 26(2), 197–223.
- Das, A., Bonneau, J., Caesar, M., Borisov, N., & Wang, X. (2014). The Tangled Web of Password Reuse. In *The Network and Distributed System Security* (pp. 23–26). San Diego: Internet Society.
- Davis, M. C., Challenger, R., Jayewardene, D. N. W., & Clegg, C. W. (2014). Advancing Socio-Technical Systems Thinking: A Call For Bravery. *Applied Ergonomics*, 45(2014), 171–180.
- de Carnavalet, X. de C., & Mannan, M. (2014). From Very Weak to Very Strong: Analyzing Password-Strength Meters. In *In Network and Distributed System Security Symposium* (Vol. 14, pp. 23–26). San Diego: Internet Society.
- Dell'Amico, A., Michiardi, P., & Roudier, Y. (2010). Password Strength: An Empirical Analysis. In *Proceedings of the 2010 IEEE INFOCOM* (pp. 1–9). San Diego: IEEE.
- Dell'Amico, M., & Filippone, M. (2015). Monte Carlo Strength Evaluation: Fast and Reliable Password Checking. In *22nd ACM Conference on Computer and Communications Security* (pp. 158–169). Denver: ACM.
- Deumert, A., & Lexander, K. V. (2013). Texting Africa: Writing as Performance. *Journal of Sociolinguistics*, 17(4), 522–546.
- Deumert, A., & Masinyana, S. O. (2008). Mobile Language Choices — The Use of English and isiXhosa in Text Messages (SMS) Evidence from a Bilingual South African Sample. *English World-Wide*, 29(2), 117–147.
- Doherty, N. F. (2014). The Role of Socio-Technical Principles in Leveraging Meaningful



- Benefits From IT Investments. *Applied Ergonomics*, 45(2), 181–187.
- Duermuth, M., Angelstorf, F., Castelluccia, C., Perito, D., Duermuth, M., Angelstorf, F., ... Chaa-, A. (2015). OMEN: Faster Password Guessing Using an Ordered Markov Enumerator. In *International Symposium on Engineering Secure Software and Systems*. Milan: Springer.
- Duggan, G. B., Johnson, H., & Grawemeyer, B. (2012). Rational Security: Modelling Everyday Password Use. *Journal of Human Computer Studies*, 70(6), 415–431.
- Durkin, M., Mulholland, G., & McCartan, A. (2015). A Socio-Technical Perspective on Social Media Adoption: A Case From Retail Banking. *International Journal of Bank Marketing*, 33(7), 944–962.
- Dürmuth, M., Chaabane, A., Perito, D., & Castelluccia, C. (2013). When Privacy Meets Security: Leveraging Personal Information for Password Cracking (pp. 1–19). arXiv preprint arXiv.
- Dyers, C., & Davids, G. (2015). Post-modern “languagers”: the Effects of Texting by University Students on Three South African Languages. *Southern African Linguistics and Applied Language Studies*, 33(1), 21–30.
- Egelman, S., Sotirakopoulos, A., Muslukhov, I., Beznosov, K., & Herley, C. (2013). Does My Password Go up to Eleven? The Impact of Password Meters on Password Selection. In *Conference on Human Factors in Computing Systems* (pp. 2379–2388). Paris: ACM New York.
- Eisenhardt, K. M., & Graebner, M. E. (2007). Theory Building From Cases: Opportunities and Challenges. *Academy of Management*, 50(1), 25–32.
- España, L. (2016). Effects of Password Type and Memory Techniques on User Password Memory. *PSI CHI Journal of Psychological Research*, 21(4), 269–276.
- Everitt, K. M., Bragin, T., Fogarty, J., & Kohno, T. (2009). A Comprehensive Study of Frequency, Interference, and Training of Multiple Graphical Passwords. In *SIGCHI Conference on Human Factors in Computing Systems* (pp. 889–898). Boston: ACM.
- Florêncio, D., & Herley, C. (2010). Where Do Security Policies Come From? In *Symposium on Usable Privacy and Security (SOUPS)*. Redmond: ACM.
- Florêncio, D., Herley, C., & van Oorschot, P. C. (2014a). An Administrator’s Guide to Internet Password Research. In *The Proceedings of the 28th Large Installation*

- System Administration Conference (LISA14)* (pp. 34–52). Seattle, WA: USENIX Association.
- Florêncio, D., Herley, C., & van Oorschot, P. C. (2014b). Password Portfolios and the Finite-Effort User: Sustainably Managing Large Numbers of Accounts. In *23rd USENIX Security Symposium* (pp. 575–590). San Diego: USENIX.
- Furnell, S. (2016). The usability of security – revisited. *Computer Fraud & Security Bulletin*, (9), 5–11.
- Furnell, S., Khern-am-nuai, W., & Esmael, R. (2018). Enhancing Security Behaviour by Supporting the User. *Computers & Security*, 75(2018), 1–9.
- Goes, P. B. (2014). Editor’s Comments. Design Science Research in Top Information Systems Journals. *MIS Quarterly*, 38(1).
- Grassi, P. A., Garcia, M. E., & Fenton, J. L. Digital Identity Guidelines. NIST Special Publication 800-63-3 (2017). United States of America.
- Gregor, S., & Hevner, A. R. (2013). Positioning and Presenting Design Science Research for Maximum Impact. *MIS Quarterly*, 37(2), 337–355.
- Gruszka, A., & Orzechowski, J. (2016). Meta-Analysis of the Research Impact of Baddeley’s Multicomponent Working Memory Model and Cowan’s Embedded-processes Model of Working Memory: A Bibliometric Mapping Approach. *Polish Psychological Bulletin*, 47(1), 1–11.
- Guo, Y., Zhang, Z., & Guo, Y. (2019). Optiwords: A New Password Policy for Creating Memorable and Strong Passwords. *Computers & Security*, 85, 423–435.
- Harris, S., & Maymí, F. (2019). *All in one CISSP Exam Guide* (8th ed.). New York: McGraw Hill Education.
- Hayashi, E., & Hong, J. I. (2011). A Diary Study of Password Usage in Daily Life. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2627–2630). Vancouver: ACM New York.
- Helkala, K., & Bakås, T. H. (2013). National Password Security Survey: Results. In S. M. Furnell, N. L. Clarke, & V. Katos (Eds.), *Proceedings of the European Information Security Multi-Conference (EISMC)* (pp. 23–33). Lisbon: University of Plymouth Press.
- Henson, R. K., & Roberts, J. K. (2006). Use of Exploratory Factor Analysis in Published Research. Common Errors and Some Comment on Improved Practice.

- Educational and Psychological Measurement*, 66(3), 393–416.
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science In Information Systems Research. *MIS Quarterly*, 28(1), 75–105.
- Houshmand, S., & Aggarwal, S. (2012). Building Better Passwords Using Probabilistic Techniques. In *28th Annual Computer Security Applications Conference* (pp. 109–118). Orlando, Florida, USA: ACM New York.
- Houshmand, S., Aggarwal, S., & Flood, R. (2015). Next Gen PCFG Password Cracking. *IEEE Transactions on Information Forensics and Security*, 10(8), 1776–1791.
- Iacono, J. C., Brown, A., & Holtham, C. (2011). The Use of the Case Study Method in Theory Testing: The Example of Steel Emarketplaces. *Electronic Journal of Business Research Methods*, 9(1), 57–65.
- Iivari, J. (2007). A Paradigmatic Analysis of Information Systems As a Design Science. *Scandinavian Journal of Information Systems*, 19(2), 39–64.
- Iivari, J., & Venable, J. R. (2009). Action Research and Design Science Research - Seemingly Similar But Decisively Dissimilar. In *European Conference on Information Systems (ECIS)* (pp. 1–13). Verona.
- Inglesant, P. G., & Sasse, M. A. (2010). The True Cost of Unusable Password Policies: Password Use in the Wild. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI '10* (pp. 1–10). Atlanta, Georgia, USA: ACM.
- ISACA. (2015). *CISA - Review Manual 26th Edition*. Rolling Meadows: ISACA.
- Jakobsson, M., & Dhiman, M. (2013). The Benefits of Understanding Passwords. In *Mobile Authentication Problems and Solutions* (Vol. 2013, pp. 5–24). New York: Springer.
- Ji, S., Yang, S., Hu, X., Han, W., Li, Z., & Beyah, R. (2017). Zero-Sum Password Cracking Game: A Large-Scale Empirical Study on the Crackability, Correlation, and Security of Passwords. *IEEE Transactions on Dependable and Secure Computing*, 14(5), 550–564.
- Ji, S., Yang, S., Wang, T., Liu, C., & Lee, W. (2015). PARS: A Uniform and Open-source Password Analysis and Research System. In *Proceedings of the 31st Annual Computer Security Applications Conference* (pp. 321–330). Los Angeles: ACM.
- Juang, K., & Greenstein, J. (2018). Integrating Visual Mnemonics and Input Feedback

- With Passphrases to Improve the Usability and Security of Digital Authentication. *Human Factors*, 60(5), 658–668.
- Kang, P. (2015). The Effects of Different Alphabets on Free Text Keystroke Authentication: A Case Study on the Korean – English Users. *The Journal of Systems & Software*, 102(2015), 1–11.
- Keith, M., Shao, B., & Steinbart, P. (2009). A Behavioral Analysis of Passphrase Design and Effectiveness. *Journal of the Association for Information Systems*, 10(2), 63–89.
- Keith, M., Shao, B., & Steinbart, P. J. (2007). The usability of passphrases for authentication: An empirical field study. *International Journal of Human-Computer Studies*, 65(2007), 17–28.
- Kelley, P. G., Komanduri, S., Mazurek, M. L., Shay, R., Vidas, T., Bauer, L., ... Julio, L. (2012). Guess Again (and Again and Again): Measuring Password Strength by Simulating Password-Cracking Algorithms. In *IEEE Symposium on Security and Privacy* (pp. 523–537). San Francisco: IEEE.
- Keong, Y. C., Gill, S. K., Noorezam, M., & Abdulrazaq, A. (2012). Gender Differences and Culture in English Short Message Service Language among Malay University Students. *The Southeast Asian Journal of English Language Studies*, 18(2), 67–74.
- Khan, S. N. (2014). Qualitative Research Method: Grounded Theory. *International Journal of Business and Management*, 9(11), 224–233.
- Komanduri, S. (2016). *Modeling the Adversary to Evaluate Password Strength With Limited Samples*. Carnegie Mellon University.
- Komanduri, S., Shay, R., Cranor, L. F., Herley, C., & Schechter, S. (2014). Telepathwords: Preventing Weak Passwords by Reading Users' Minds. In *Proceedings of the 23rd USENIX Security Symposium* (pp. 591–606). San Diego: USENIX Association.
- Komanduri, S., Shay, R., Kelley, P. G., Mazurek, M. L., Bauer, L., Christin, N., ... Egelman, S. (2011). Of Passwords and People: Measuring the Effect of Password-Composition Policies. In *SIGCHI Conference on Human Factors in Computing Systems* (pp. 2595–2604). Vancouver: ACM New York.
- Kovács, G., & Spens, K. M. (2005). Abductive Reasoning in Logistics Research. *International Journal of Physical Distribution & Logistics Management*, 35(2),

132–144.

- Lantolf, J., Thorne, S. L., & Poehner, M. (2015). Sociocultural Theory and Second Language Development. In B. van Patten & J. Williams (Eds.), *Theories in Second Language Acquisition* (pp. 207–226). New York: Routledge.
- Lantolf, J. P. (2000). Introducing sociocultural theory. *Sociocultural Theory and Second Language Learning*, 1, 1–26.
- Lantolf, J. P., Thorne, S. L., & Poehner, M. E. (2015). Sociocultural Theory and Second Language Development. In B. van Patten & J. Williams (Eds.), *Theories in Second Language Acquisition* (Eds, pp. 207–226). New York: Routledge.
- Lee, J. S., Pries-Heje, J., & Baskerville, R. (2011). Theorizing in Design Science Research. In H. Jain, A. P. Sinha, & P. Vitharana (Eds.), *Proceedings of the International Conference on Design Science Research in Information Systems* (pp. 1–16). Berlin: Springer.
- Lexander, K. V. (2011). Texting and African Language Literacy. *New Media & Society*, 13(3), 427–443.
- Li, Y., Wang, H., & Sun, K. (2016). A Study of Personal Information in Human-chosen Passwords and Its Security Implications. In *The 35th Annual IEEE International Conference on Computer Communications* (pp. 1–9). San Francisco: IEEE.
- Li, Z., Han, W., & Xu, W. (2014). A Large-Scale Empirical Analysis of Chinese Web Passwords. In *Proceedings of the 23rd USENIX Security Symposium* (pp. 559–574). San Diego: USENIX Association.
- Lin, T. T. C., Paragas, F., Goh, D., & Bautista, J. R. (2016). Developing Location-Based Mobile Advertising in Singapore: A Socio-Technical Perspective. *Technological Forecasting & Social Change*, 103(2016), 334–349.
- Lund, B. A. M. (2001). Measuring Usability with the USE Questionnaire. *Usability Interface*, 8(2), 3–6.
- Ma, J., Yang, W., Luo, M., & Li, N. (2014). A Study of Probabilistic Password Models. In *Proceedings - IEEE Symposium on Security and Privacy* (pp. 689–704). San Jose: IEEE Computer Society.
- Malone, D., & Maher, K. (2012). Investigating the Distribution of Password Choices. In *Proceedings of the 21st International Conference on World Wide Web* (pp. 301–310). Lyon: ACM New York.

- Maoneke, P. B., Flowerday, S., & Isabirye, N. (2018). The Influence of Native Language on Password Composition and Security: A Socioculture Theoretical View. In L. J. Janczewski & M. Kutyłowski (Eds.), *33rd IFIP TC 11 International Conference, SEC 2018* (pp. 33–46). Poznan: Springer Nature Switzerland AG 2018.
- Maoneke, P. B., & Flowerday, S. (2018). Password Policies Adopted by South African Organizations: Influential Factors and Weaknesses. In *17th International Conference, ISSA 2018 Pretoria, South Africa, August 15–16, 2018 Revised Selected Papers* (pp. 30–43). Pretoria: Springer.
- March, S. T., & Smith, G. F. (1995). Design and Natural Science Research on Information Technology. *Decision Support Systems*, 15(1995), 251–266.
- Marginson, S., & Dang, T. K. A. (2017). Vygotsky's Sociocultural Theory in the Context of Globalization. *Asia Pacific Journal of Education*, 37(1), 116–129.
- Mazurek, M. L., Komanduri, S., Vidas, T., Bauer, L., Christin, N., Cranor, L. F., ... Ur, B. (2013). Measuring Password Guessability for an Entire University Categories and Subject Descriptors. In *2013 ACM SIGSAC conference on Computer & Communications Security* (pp. 173–186). Berlin: ACM New York.
- Melicher, W., Kurilova, D., Segreti, S. M., Kalvani, P., Shay, R., Ur, B., ... Mazurek, M. L. (2016). Usability and Security of Text Passwords on Mobile Devices. In *Proceedings of the 2016 Conference on Human Factors in Computing Systems* (pp. 527–539). San Jose, CA, USA: ACM New York.
- Mercer, N., & Howe, C. (2012). Explaining the Dialogic Processes of Teaching and Learning: The Value and Potential of Sociocultural Theory. *Learning, Culture and Social Interaction*, 1(1), 12–21.
- Mertler, C., & Vannatta, R. (2004). *Advanced and Multivariate Statistical Methods: Practical Application and Interpretation*. (V. Mertler, Ed.) (3rd ed.). New York: Taylor & Francis.
- Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information. *The Psychological Review*, 63(2), 81–97.
- Mkansi, M., & Acheampong, E. A. (2012). Research Philosophy Debates and Classifications: Students' Dilemma. *The Electronic Journal of Business Research Methods Volume*, 10(2), 132–140.
- Montalvão, J., Freire, E. O., Bezerra, M. A., & Garcia, R. (2015). Contributions to



- Empirical Analysis of Keystroke Dynamics in Passwords. *Pattern Recognition Letters*, 52(2015), 80–86.
- Morel, E., Bucher, C., Doehler, S. P., & Siebenhaar, B. (2012). SMS Communication as Plurilingual Communication Hybrid Language use as a Challenge for Classical Code-Switching Categories. *Lingvisticae Investigationes*, 35(2), 260–288.
- Mwagwabi, F., McGill, T., & Dixon, M. (2014). Improving Compliance with Password Guidelines: How User Perceptions of Passwords and Security Threats Affect Compliance with Guidelines. In *The 47th Hawaii International Conference on Systems Security* (pp. 3188–3197). Washington DC: IEEE Computer Society.
- Narayanan, A., & Shmatikov, V. (2005). Fast Dictionary Attacks on Passwords Using Time-Space Tradeoff. In *12th ACM conference on Computer and communications security* (pp. 364–372). Virginia: ACM.
- Ndlovu, A. (2016). Code-switching in whatsapp chat messages in Botswana. *NAWA Journal of Language & Communication*, 10(1), 132–140.
- Neuman, W. L. (2014). *Social Research Methods: Qualitative and Quantitative Approaches* (7th ed.). Harlow: Pearson Education Limited.
- Niederman, F., & March, S. T. (2012). Design Science and the Accumulation of Knowledge in the Information Systems Discipline. *ACM Transactions on Management Information Systems*, 3(1), 1–15.
- Niehaves, B. (2007). On Episemological Diversity in Design Science: New Vistas for a Design-Oriented IS Research? In *International Conference on Information Systems* (pp. 1–14).
- Paas, F., & Ayres, P. (2014). Cognitive Load Theory: A Broader View on the Role of Memory in Learning and Education. *Educational Psychology Review*, 26(2), 191–195.
- Peppers, Ken, Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2008). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, 24(3), 45–77.
- Peters, A. N., Winschiers-Theophilus, H., & Mennecke, B. E. (2015). Cultural Influences on Facebook Practices: A Comparative Study of College Students in Namibia and the United States. *Computers in Human Behavior*, 49(2015), 259–271.
- Pilar, D. R., Jaeger, A., Gomes, C. F. A., & Stein, L. M. (2012). Passwords Usage and

- Human Memory Limitations: A Survey Across Age and Educational Background. *PLoS ONE*, 7(12), 1–7.
- Prat, N., Comyn-Wattiau, I., & Akoka, J. (2015). A Taxonomy of Evaluation Methods for Information Systems Artifacts. *Journal of Management Information Systems*, 32(3), 229–267.
- Pries-Heje, J., Baskerville, R., & Venable, J. R. (2008). Strategies for Design Science Research Evaluation. In *Proceedings of the 2008 European Conference on Information Systems* (pp. 1–13).
- Rao, A., Jha, B., & Kini, G. (2013). Effect of Grammar on Security of Long Passwords. In *Proceedings of the 3rd ACM Conference on Data and Application Security and Privacy* (pp. 317–324). San Antonio: ACM New York.
- Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor Analysis and Scale Revision. *Psychological Assessment*, 12(3), 287–297.
- Renaud, K., Otondo, R., & Warkentin, M. (2019). “This is the way ‘I’ create my passwords” ... does the endowment effect deter people from changing the way they create their passwords? *Computers & Security*, 82(2019), 241–260.
- Rinn, C., Summers, K., Rhodes, E., Virothaisakun, J., & Chisnell, D. (2015). Password Creation Strategies Across High- and Low-literacy Web Users. In *Proceedings of the 78th ASIS&T Annual Meeting: Information Science With Impact: Research in and for the Community* (pp. 52:1--52:9). St. Louis: American Society for Information Science Silver Springs.
- Roberts, P., Priest, H., & Traynor, M. (2006). Reliability and Validity in Research. *Nursing Standards*, 20(44), 41–45.
- Saunders, M., Lewis, P., & Thornhill, A. (2009). *Research Methods for Business Students, 5th Ed. Research methods for business students*.
- Saunders, M., Lewis, P., & Thornhill, A. (2016). *Research methods for business students (7th ed)*. New York: Pearson.
- Sawyer, S., & Jarrahi, M. H. (2013). Socio-Technical Approaches to the Study of Information Systems. In A. Tucker & H. Topi (Eds.), *CRC Handbook of Computing* (pp. 1–39). Syracuse: Chapman and Hall.
- Schweppe, J., & Rummer, R. (2014). Attention, Working Memory, and Long-Term Memory in Multimedia Learning: An Integrated Perspective Based on Process



- Models of Working Memory. *Educational Psychology Review*, 26(2014), 285–306.
- Scott, S., & Palincsar, A. (2013). Sociocultural Theory. education.com. Retrieved from <http://www.education.com/reference/article/sociocultural-theory/> 1
- Shay, R., Kelley, P. G., Komanduri, S., Mazurek, M. L., Ur, B., Vidas, T., ... Cranor, L. F. (2012). Correct Horse Battery Staple: Exploring the Usability of System-assigned Passphrases. In *Symposium on Usable Privacy and Security (SOUPS) 2012* (pp. 1–20). Washington DC: ACM New York.
- Shay, R., Komanduri, S., Durity, A. L., Huh, P. (Seyoung), Mazurek, M. L., Segreti, S. M., ... Cranor, L. F. (2016). Designing Password Policies for Strength and Usability. *ACM Transactions on Information and System Security*, 18(4), 1–34.
- Shay, R., Komanduri, S., Durity, A. L., Huh, P. S., Mazurek, M. L., Segreti, S. M., ... Cranor, L. F. (2014). Can Long Passwords Be Secure and Usable? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2927–2936). Toronto, ON, Canada: ACM New York.
- Shay, R., Ur, B., Bauer, L., Christin, N., Cranor, L. F., Forget, A., ... Segreti, S. M. (2015). A Spoonful of Sugar? The Impact of Guidance and Feedback on Password-Creation Behavior. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 2903–2912). Seoul: ACM New York.
- Shin, D.-H., & Song, H.-R. (2012). The Switchover to Digital Broadcasting in Korea. *Technological Forecasting & Social Change*, 79(8), 1447–1461.
- Shin, D. (2014). A Socio-Technical Framework for Internet-of-Things Design: A Human-Centered Design for the Internet of Things. *Telematics and Informatics*, 31(4), 519–531.
- Stobert, E., & Biddle, R. (2014). The Password Life Cycle: User Behaviour in Managing Passwords. In *Proceedings of the Tenth Symposium On Usable Privacy and Security (SOUPS)* (pp. 243–255). Menlo Park: USENIX.
- Tagg, C., Baron, A., & Rayson, P. (2012). “I Didn’t Spel that Wrong did i. Oops”: Analysis and Normalisation of SMS Spelling Variation. *Lingvisticie Investigaciones*, 35(2), 367–388.
- Tam, L., Glassman, M., & Vandenwauver, M. (2010). The Psychology of Password Management: a Tradeoff Between Security and Convenience. *Behaviour and Information Technology*, 29(3), 233–244.

- Taneski, V., Heričko, M., & Brumen, B. (2014). Password security – no change in 35 years? In *37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* (pp. 1360–1365). Opatija: IEEE.
- Thurlow, C., & Brown, A. (2003). Generation Txt? The sociolinguistics of young people's text-messaging. *Discourse Analysis Online*, 1(1), 1–27.
- Ur, B., Bees, J., Segreti, S. M., Bauer, L., Christin, N., & Cranor, L. F. (2016). Do Users' Perceptions of Password Security Match Reality? In *Proceedings of the Conference on Human Factors in Computing Systems* (pp. 3748–3760). San Jose: ACM New York.
- Ur, B., Kelley, P. G., Komanduri, S., Lee, J., Maass, M., Mazurek, M. L., ... Cranor, L. F. (2012). How Does Your Password Measure Up? The Effect of Strength Meters on Password Creation. In *21st USENIX Security Symposium*. Bellevue, Washington, USA: USENIX.
- Ur, B., Noma, F., Bees, J., Segreti, S. M., Shay, R., Bauer, L., ... Cranor, L. F. (2015). "I Added '!' at the End to Make It Secure": Observing Password Creation in the Lab. In *Symposium on Usable Privacy and Security (SOUPS) 2015* (pp. 123–140). Ottawa: USENIX Association.
- Ur, B., Segreti, S., Bauer, L., Christin, N., Cranor, L. F., Komanduri, S., ... Shay, R. (2015). Measuring Real-World Accuracies and Biases in Modeling Password Guessability. In *24th USENIX Security Symposium (USENIX Security 15)* (pp. 463–481). Washington DC: USENIX.
- Vaishnavi, V. K., & Kuechler, W. jr. (2015). *Design Science Research Methods and Patterns. Innovating Information and Communication Technology* (2nd ed.). Florida: CRC Press.
- Vaithyasubramanian, S., & Christy, A. (2015). A Scheme to Create Secured Random Password Using Markov Chain. In *Artificial Intelligence and Evolutionary Algorithms in Engineering Systems* (pp. 809–814). New Delhi: Springer.
- Vance, A., Eargle, D., Ouimet, K., & Straub, D. (2013). Enhancing Password Security through Interactive Fear Appeals: A Web-based Field Experiment. In *46th Hawaii International Conference on System Sciences Enhancing* (pp. 2988–2997). Wailea: IEEE.

- Venable, J., Pries-Heje, J., & Baskerville, R. (2012). A Comprehensive Framework for Evaluation in Design Science Research. In K Peffers, M. Rothenberger, & B. Kuechler (Eds.), *Design Science Research in Information Systems. Advances in Theory and Practice* (Vol. 7286, pp. 423–438). Berlin: Springer.
- Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: A Framework for Evaluation in Design Science Research. *European Journal of Information Systems*, 25(1), 77–89.
- Venkatesh, V., Brown, S. A., & Bala, H. (2013). Bridging The Qualitative-Quantitative Divide: Guidelines For Conducting Mixed Methods Research In Information Systems. *MIS Quarterly*, 37(1), 21–54.
- Veras, R., Collins, C., & Thorpe, J. (2014). On the Semantic Patterns of Passwords and their Security Impact. In *Network and Distributed System Security (NDSS) Symposium* (pp. 1–16). San Diego: Internet Society.
- von Zezschwitz, E., De Luca, A., & Hussmann, H. (2013). Survival of the Shortest: A Retrospective Analysis of Influencing Factors on Password Composition. In *14th IFIP TC 13 International Conference on Human-Computer Interaction* (pp. 460–467). Cape Town: Springer, Berlin, Heidelberg.
- Voyiatzis, A. G., Fidas, C. A., Serpanos, D. N., & Avouris, N. M. (2011). An Empirical Study on the Web Password Strength in Greece. In *2011 Panhellenic Conference on Informatics* (pp. 212–216). Kastoria: IEEE Computer Society.
- Wang, D., Cheng, H., Gu, Q., & Wang, P. (2015). Understanding Passwords of Chinese Users: Characteristics, Security and Implications. In *Proceedings of the ChinaCrypt 2015* (pp. 1–14).
- Wang, D., Jian, G., Cheng, H., Gu, Q., Zhu, C., & Wang, P. (2015). Zipf's Law in Passwords. *ACM Transactions on Information and System Security*, 1(1), 1–19.
- Wang, D., & Wang, P. (2015). The Emperor's New Password Creation Policies: An Evaluation of Leading Web Services and the Effect of Role in Resisting Against Online Guessing. In G. Pernul, P. Y. A. Ryan, & W. Edgar (Eds.), *20th European Symposium on Research in Computer Security (ESORICS)* (pp. 456–477). Vienna: Springer.
- Wang, D., Zhang, Z., Wang, P., Yan, J., & Huang, X. (2016). Targeted Online Password Guessing: An Underestimated Threat. In *Conference on Computer and Communications Security* (pp. 1242–1254). Vienna: ACM New York.

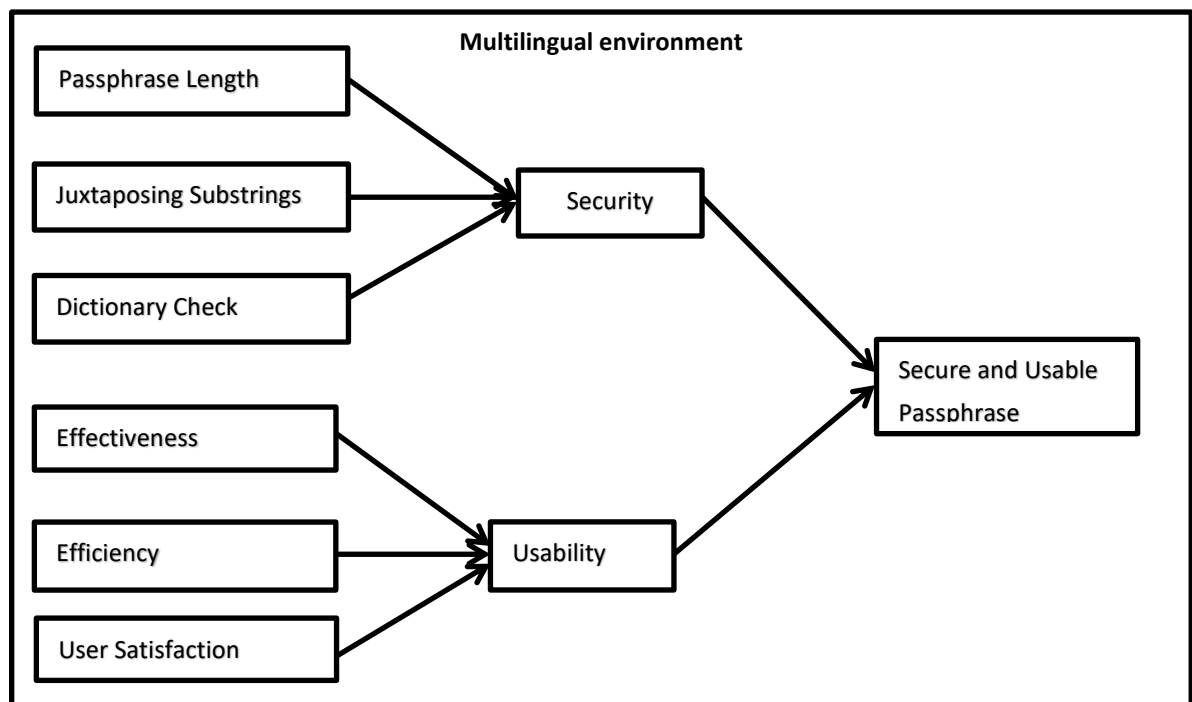
- Wassenaar, D. R. (2006). Ethical issues in social science research. In M. T. Blanche, K. Durrheim, & D. Painter (Eds.), *Research in practice: Applied methods for the social sciences* (2nd Editio, pp. 94–100). Cape Town: UCT Press.
- Weber, J. E., Cloud, S. T., Guster, D., Cloud, S. T., Safonov, P., & Cloud, S. T. (2008). A Developmental Perspective On Weak Passwords and Password Security. *Journal of Information Technology Management*, XIX(3), 1–8.
- Weir, C. M. (2010). *Using Probabilistic Techniques to Aid in Password Cracking Attacks*. The Florida State University.
- Weir, M., Aggarwal, S., Collins, M., & Stern, H. (2010). Testing Metrics for Password Creation Policies by Attacking Large Sets of Revealed Passwords. In *17th ACM conference on Computer and Communications Security* (pp. 162–175). Chicago: ACM New York.
- Weir, M., Aggarwal, S., de Medeiros, B., & Glodek, B. (2009). Password Cracking Using Probabilistic Context-Free Grammars. In *Proceedings of the 30th IEEE Symposium on Security and Privacy* (pp. 391–405). Oakland, California: IEEE.
- Wheeler, D. L. (2016). zxcvbn: Low-Budget Password Strength Estimation. In *Proceedings of the 25th USENIX Security Symposium* (pp. 156–173). Austin: USENIX.
- White, A. M., Shaw, K., Monroe, F., & Moreton, E. (2014). Isn't that Fantabulous: Security, Linguistic and Usability Challenges of Pronounceable Tokens. In *Proceedings of the 2014 workshop on New Security Paradigms Workshop* (pp. 25–38). Laurel Point Victoria: ACM.
- Williams, B., Brown, T., & Onsmann, A. (2012). Exploratory Factor Analysis: A Five-step Guide for Novices. *Australasian Journal of Paramedicine*, 8(3), 1–10.
- Winschiers-Theophilus, H., & Bidwell, N. J. (2013). Toward an Afro-Centric Indigenous HCI Paradigm. *International Journal on Human-Computer Interaction*, 29(2013), 243–255.
- Woods, N. (2017). Frequently Using Passwords Increases Their Memorability – A False Assumption or Reality? In *Twenty-third Americas Conference on Information Systems* (pp. 1–5). Boston: AIS Electronic Library (AISeL).
- Woods, N., & Siponen, M. (2018). Too many passwords? How understanding our memory can increase password memorability. *International Journal of Human -*

- Computer Studies*, 111(2017), 36–48.
- Woods, N., & Siponen, M. (2019). Improving Password Memorability, While not Inconveniencing the User. *International Journal of Human Computer Studies*, 128(2019), 61–71.
- Wu, P. P., Fookes, C., Pitchforth, J., & Mengersen, K. (2015). A Framework for Model Integration and Holistic Modelling of Socio-Technical Systems. *Decision Support Systems*, 71(2015), 14–27.
- Yang, C., Hung, J., & Lin, Z. (2013). An Analysis View on Password Patterns of Chinese Internet Users. *Nankai Business Review International*, 4(1), 66–77.
- Yang, Y., Lindqvist, J., & Oulasvirta, A. (2014). Text Entry Method Affects Password Security. In *Proceedings of Learning from Authoritative Security Experiment Result* (pp. 11–20). Arlington, Virginia: USENIX Association.
- Yilmaz, K. (2013). Comparison of Quantitative and Qualitative Research Traditions: Epistemological, Theoretical, and Methodological Differences. *European Journal of Education*, 48(2), 311–325.
- Yin, R. K. (2003). *Case Study Research design and methods*. (3rd ed.). Thousand Oaks (CA): Sage.
- Zhang, J., Luo, X., Akkaladevi, S., & Ziegelmayer, J. (2009). Improving Multiple-Password Recall: An Empirical Study. *European Journal of Information Systems*, 18(2), 165–176.
- Zhang, Y., Monroe, F., & Reiter, M. K. (2010). The Security of Modern Password Expiration: An Algorithmic Framework and Empirical Analysis. In *Proceedings of the 17th ACM conference on Computer and communications security* (pp. 176–186). Chicago: ACM.
- Zuengler, J., & Miller, E. R. (2006). Cognitive and Sociocultural Perspectives: Two Parallel SLA Worlds? *TESOL Quarterly*, 40(1), 35–58.

## APPENDIX A: THE EXPERIMENT DESIGN AND OVERVIEW

### 1.0 Experiment overview

The experiment for this study was aimed at gathering data for evaluating the utility, efficacy and quality of the model presented in Figure 1. The conducting of the experiment resulted in data being gathered for evaluating the security and usability contributions of user-generated passphrases.



**Figure 1. A model for generating usable and secure passphrases**

Data for each measurement construct was gathered using different data collection techniques that included quantitative and qualitative data. Throughout the experiment, either one or both of these techniques were used to gather data for assessing usability and security. In particular to passphrase security, measures had to be put in place to ensure that participants created passphrases based on juxtaposed substrings.

In addition, word and password dictionaries were used in the background to check whether a passphrase entered by a user during passphrase generation was oriented to a single Indo-European language. Accordingly, during passphrase generation the

dictionaries from the OpenWall were used for auto word checking in the background of the web-application. The use of opensource dictionaries such as the OpenWall for words and passwords is a common practice in the literature (Ma et al., 2014; Mazurek et al., 2013; Shay et al., 2015). Given that the research context for this study contained two Indo-European languages, English and Afrikaans, dictionaries with words from these languages were integrated in the web application such that pure English or Afrikaans passphrases were prohibited. The following Uniform Resource Locators on the OpenWall were used to access the English, Afrikaans and wordlist dictionaries that were integrated with the web-application for this study:

<http://download.openwall.net/pub/wordlists/languages/English/>

<http://download.openwall.net/pub/wordlists/languages/Afrikaans/>

<http://download.openwall.net/pub/wordlists/passwords/>

## 1.2 PASSWORD POLICIES

The experiment in this study was done in two parts with the same participants generating passwords following two different password generation policies in succession. Each part involved activities of password generation, logging in to test recall and the completion of two questionnaires for gathering the data used to assess usability. Passwords generated under different password policies were stored separately, together with data from the online survey, to compare contributions to usability and security. The functionality of the web-based application platform was designed in such a way that participants had to meet password and passphrase requirements for each policy (Keith et al., 2009). The password generation policies for the two parts are as follows:

**Part One: Comprehensive 8 (Comp8).** The Comp8 password policy is one of the most popular policies and was designed following a guideline by NIST (Melicher et al., 2016; Shay et al., 2016). For this study, participants were required to generate a short password that

- must be at least 8 characters long, without any spaces

- must contain at least one capital letter, one lowercase letter, one number and one special character (symbols like &, \$, @, #, ! and \*)
- must not contain your username or personal details.

**Part Two: 2word16.** Participants were required to create passphrases with the following characteristics:

- Have at least two words.
- The words making a password should be from at least two different languages, for example an English word and the other can be an Afrikaans word.
- The words making the passphrase should be separated by at least one space or non-letter sequence.
- The password should have at least sixteen characters



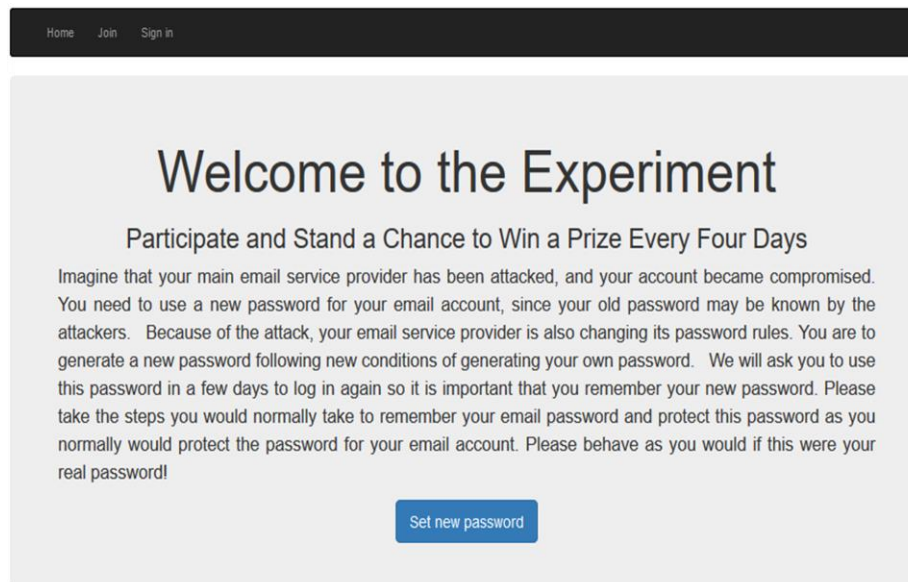
University of Fort Hare  
*Together in Excellence*



## ACTIVITIES OF THE EXPERIMENT (PASSPHRASE GENERATION AND RECALL)

### DAY ONE OF THE EXPERIMENT

**Step 1: Scenario (Welcome page).** Participants were presented with a hypothetical scenario to motivate the generation of realistic passwords and passphrases as recommended in the literature (Komanduri et al., 2011; Melicher et al., 2016; Shay et al., 2016). After reading the scenario, participant then had to click the “Set new password” button.



Home Join Sign in

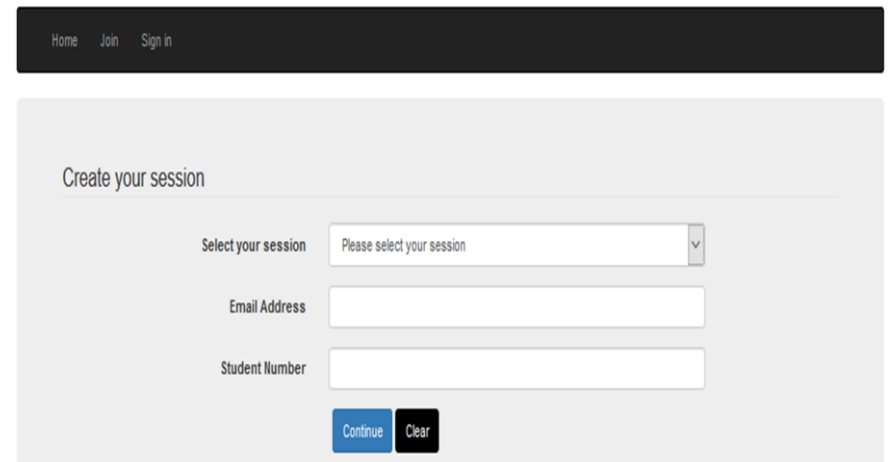
# Welcome to the Experiment

## Participate and Stand a Chance to Win a Prize Every Four Days

Imagine that your main email service provider has been attacked, and your account became compromised. You need to use a new password for your email account, since your old password may be known by the attackers. Because of the attack, your email service provider is also changing its password rules. You are to generate a new password following new conditions of generating your own password. We will ask you to use this password in a few days to log in again so it is important that you remember your new password. Please take the steps you would normally take to remember your email password and protect this password as you normally would protect the password for your email account. Please behave as you would if this were your real password!

Set new password

**Step 2: Selecting a password generation session.** Participants were required to register for the experiment by selecting a session (short password or long password), and providing an email address and a student number. Participants started with short password generation. Hence, a session for short password generation was selected for the first two weeks.



Home Join Sign in

### Create your session

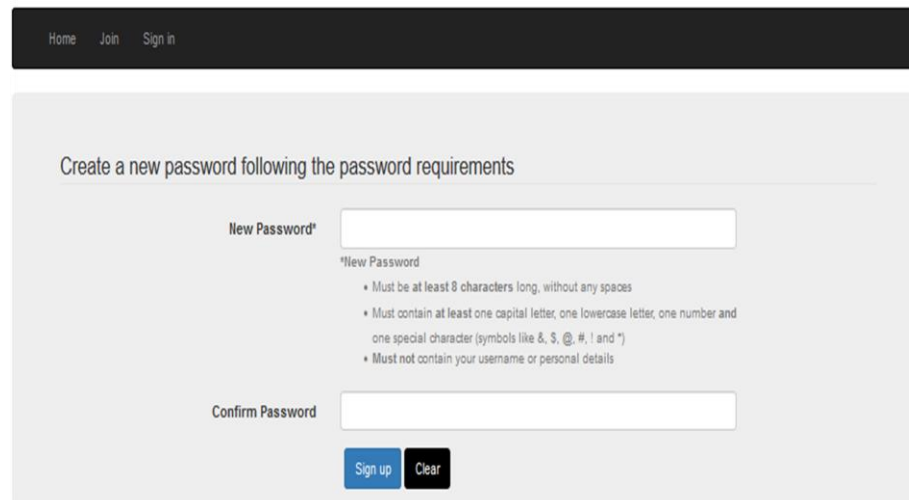
Select your session

Email Address

Student Number

Continue Clear

**Step 3: Password generation and confirmation.** Participants were requested to generate and confirm their passwords. Password generation was guided by pre-specified conditions as shown below.



Home Join Sign in

Create a new password following the password requirements

New Password\*

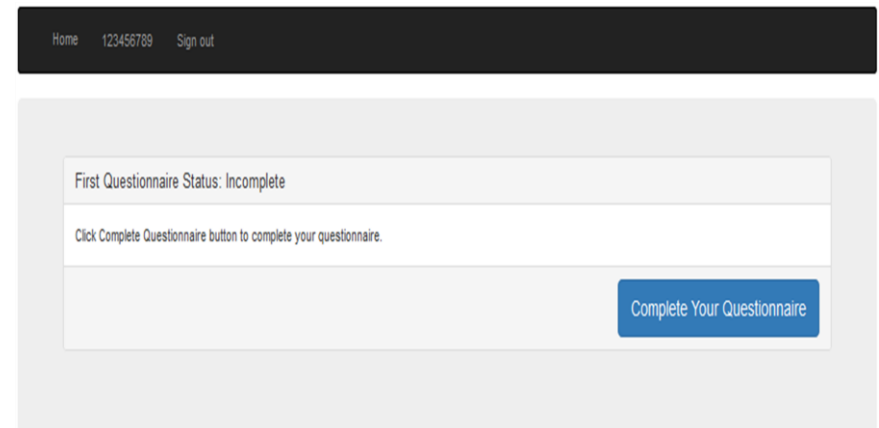
\*New Password

- Must be at least 8 characters long, without any spaces
- Must contain at least one capital letter, one lowercase letter, one number and one special character (symbols like &, \$, @, #, ! and \*)
- Must not contain your username or personal details

Confirm Password

[Sign up](#) [Clear](#)

**Step 4: Survey 1.** Participants were asked to complete an online survey. During this step, participants provided their demographics and indicated their sentiments towards password generation. Appendix B, Survey 1 shows the questionnaire that was completed by participants. This step was also used as a distractor task prior to asking participants to recall their passwords.



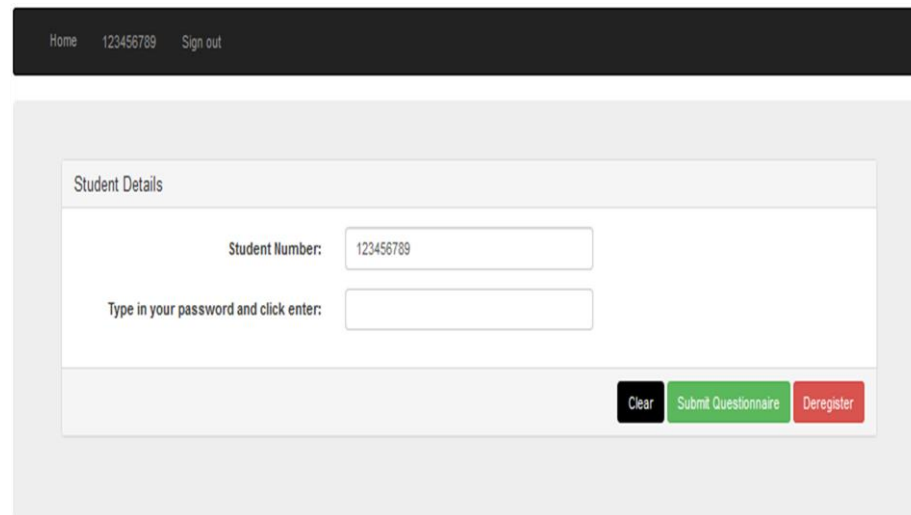
Home 123456789 Sign out

First Questionnaire Status: Incomplete

Click Complete Questionnaire button to complete your questionnaire.

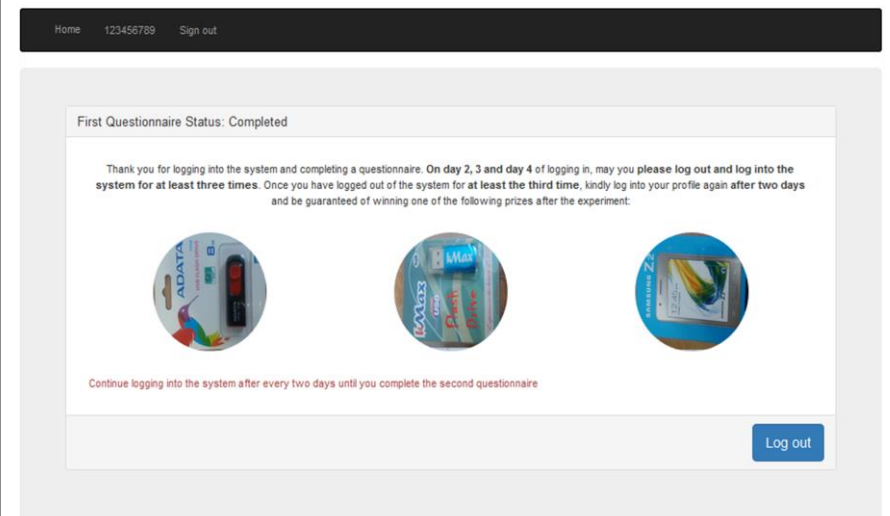
[Complete Your Questionnaire](#)

**Step 5: Password recall 1.** Participants were asked to enter their passwords to establish whether they could still recall their newly generated passwords. Participants who failed to type in their passwords in five attempts were assumed to have forgotten their passwords. A window would then appear on the screen giving the user permission to recover the password by sending it to the participant's email address.



The screenshot shows a web interface with a dark header bar containing 'Home', '123456789', and 'Sign out'. Below the header is a light gray box titled 'Student Details'. Inside this box, there is a form with two input fields. The first field is labeled 'Student Number:' and contains the text '123456789'. The second field is labeled 'Type in your password and click enter:' and is empty. At the bottom right of the form, there are three buttons: 'Clear' (black), 'Submit Questionnaire' (green), and 'Deregister' (red).

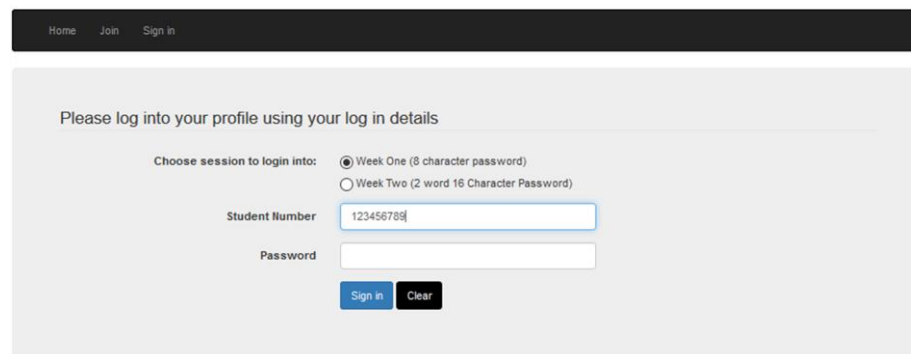
**Step 6: The end of day one activities.** Participants were presented with details about the prizes available to be won and a reminder to login after two days whereupon a second survey was completed.



The screenshot shows a web interface with a dark header bar containing 'Home', '123456789', and 'Sign out'. Below the header is a light gray box titled 'First Questionnaire Status: Completed'. Inside this box, there is a message: 'Thank you for logging into the system and completing a questionnaire. On day 2, 3 and day 4 of logging in, may you please log out and log into the system for at least three times. Once you have logged out of the system for at least the third time, kindly log into your profile again after two days and be guaranteed of winning one of the following prizes after the experiment:'. Below the message are three circular images of prizes: a USB drive, a box of tissues, and a box of tissues. At the bottom right of the box, there is a 'Log out' button.

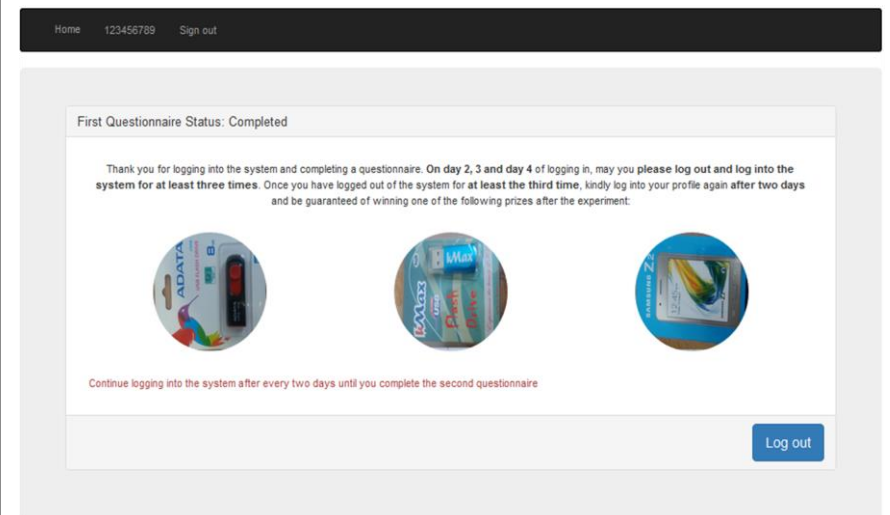
## DAY TWO OF THE EXPERIMENT.

**Step 7: Password recall 2.** Participants were asked to login to their accounts. This step helped to gather data on password recall and typing patterns. Participants had a maximum of five attempts to type in their passwords and those who could not recall them had the option to request that their passwords be sent to their email addresses.



The screenshot shows a login interface with a dark header containing 'Home', 'Join', and 'Sign in' links. Below the header, a light gray box contains the text 'Please log into your profile using your log in details'. Inside this box, there are two radio buttons for 'Choose session to login into:'. The first option is 'Week One (8 character password)' which is selected. The second option is 'Week Two (2 word 16 Character Password)'. Below these, there is a text input field for 'Student Number' containing the value '123456789'. Below that is a text input field for 'Password'. At the bottom of the form are two buttons: 'Sign in' and 'Clear'.

**Step 8: Reminder.** After logging in to confirm that the participant still recalled the password, an announcement was displayed on the participant's screen inviting him/her to login to the profile again after two days.




The screenshot shows a reminder page with a dark header containing 'Home', '123456789', and 'Sign out' links. Below the header, a light gray box contains the text 'First Questionnaire Status: Completed'. Below this, there is a paragraph of text: 'Thank you for logging into the system and completing a questionnaire. On day 2, 3 and day 4 of logging in, may you please log out and log into the system for at least three times. Once you have logged out of the system for at least the third time, kindly log into your profile again after two days and be guaranteed of winning one of the following prizes after the experiment:'. Below the text are three circular images of prizes: a USB drive, a box of Max, and a box of Max. Below the images is a red text line: 'Continue logging into the system after every two days until you complete the second questionnaire'. At the bottom right of the box is a blue button labeled 'Log out'.

Steps 7 and 8 above were repeated until the participant had completed a second questionnaire (survey 2). Survey 2 was set to automatically appear ten days after the day of password generation.

## PASSPHRASE GENERATION

After completing the short password experiment. The participants were invited to generate and recall a passphrase. The steps 1 to 8 were repeated and only the activities in Steps 2, 3 and 7 had to be changed. Thus, participants had to select a session that allowed them to generate a passphrase during Step 2. This allowed Step 3 to display a window for passphrase generation as shown below. Steps 4, 5 and 6 remained unchanged except that participants were using passphrases for logging in and out. Similarly, Step 7 saw participants selecting the option marked: week 2 (2-word 16 character password) in order to activate logging in using a passphrase. The word 'password' was used as shown below to avoid confusion as it was assumed that participants were more familiar with the word password than passphrase.



[Home](#) [Join](#) [Sign in](#)

### Create a new password following the password requirements

New Password\*

\*New Password

- Have at least two words.
- The words making a password should be from at least two different languages for example an English word and the other can be an Afrikaans word
- The words making a password should be separated by at least one space or non-letter sequence.
- The password should have at least sixteen characters

Confirm Password

Sign up

Clear

## APPENDIX B: THE QUESTIONNAIRE

### Survey 1: Demographics and Password Creation Process Usability

#### 1.0 DEMOGRAPHIC DATA

1. What is your gender? Male ☐ Female ☐

2. What is your age group?

18–25	<input type="checkbox"/>
26–35	<input type="checkbox"/>
36 and above	<input type="checkbox"/>

3. On a scale of 1 (novice) to 5 (expert), how would you rate your level of experience using computers?

1	2	3	4	5
---	---	---	---	---

4. A) State your mother tongue or native language: \_\_\_\_\_

B) State your second preferred language: \_\_\_\_\_

5. Indicate your ethnic group e.g. Xhosa: \_\_\_\_\_

## 1.2 DATA FOR ASSESSING PASSWORD CREATION

### 1.2.1 Password creation strategy

For each of the following statements, rate the extent you agree or disagree with the statement on a scale of 1 to 5:	<div>Strongly Agree ←————→ Strongly Disagree</div>				
	1	2	3	4	5
1. I created my password using non-standard spelling.					
2. The password I created is based on words written using different languages.					
3. I created my password using spelling abbreviations (slang or colloquial terms).					
4. I created my password based on another password I already know.					
5. I created my password based on an address that I know of.					
6. I created my password based on more than one word in English.					
7. I created my password based on the name of someone or something I know.					
8. I created my password based on words in my mother tongue.					
9. I created my password based on a phone number.					
10. The password I created is based on one or more words in a language other than English.					
11. The password I created is based on a date of birth.					

### 1.3 DATA FOR EVALUATING PASSWORD CREATION USABILITY

#### 1.3.1 Data for evaluating password creation effectiveness

For each of the following statements, rate the extent you agree or disagree with the statement on a scale of 1 to 5:		Strongly Agree					Strongly Disagree
		1	2	3	4	5	
1	Creating a password for this study was easy.						
2	Creating a password for this study was simple.						
3	Password creation requirements for this study are user-friendly.						
4	It required few steps for me to create a password for this study compared to when I am using other policies.						
5	Password creation requirements for this study were flexible.						
6	Creating a password for this study was effortless.						
7	I could quickly and easily recover from the mistakes I made during password creation.						
8	I can successfully create a password using requirements specified for this study every time.						

  
**University of Fort Hare**  
*Together in Excellence*

#### 1.3.2 Data for evaluating password creation user satisfaction.

For each of the following statements, rate the extent you agree or disagree with the statement on a scale of 1 to 5:		Strongly Agree					Strongly Disagree
		1	2	3	4	5	
9	I was satisfied with the password creation process for this study.						
10	I would recommend the password creation process for this study to a friend.						
11	Creating a password for this study was fun.						
12	Password creation process for this study works the ways I want it to work.						
13	Creating a password for this study was wonderful.						
14	I would prefer to use this study's way of creating a password on Facebook and my email.						
15	It was pleasant to create passwords for this study.						



## SURVEY 2: DATA FOR EVALUATING PASSWORD RECALL/MEMORABILITY USABILITY

### 2.1 Data for evaluating password memorability user satisfaction

For each of the following statements, rate the extent you agree or disagree with the statement on a scale of 1 to 5:		Strongly Agree					Strongly Disagree
		1	2	3	4	5	
1	I would recommend the password creation process for this study to a friend because it helps one create a password that is easy to remember.						
2	I was satisfied with recalling passwords for this study.						
3	The process of remembering the password I created for this study occurred the way I wanted it to occur.						
4	I would not need to write down or store my passwords as much if I always used a password format like the one I used for this study.						
5	I could remember more passwords at once if I always used a password format like the one I used for this study.						
6	The password format I used for this study helped me create a password that was easy to remember.						

### 2.2 Data for evaluating password memorability effectiveness

For each of the following statements, rate the extent you agree or disagree with the statement on a scale of 1 to 5:		Strongly Agree					Strongly Disagree
		1	2	3	4	5	
1	Recalling the password for this study was easy.						
2	Remembering a password for this study was simple.						
3	It required a few steps for me to remember the password I was using for this study compared to when I use other policies.						
4	Remembering a password for this study was effortless.						
5	I could quickly and easily recover from the mistakes I made during logging in.						
6	I could successfully remember the password I was using for this study every time.						

### 2.3 Data for evaluating password memorability strategy

For each of the following statements, rate the extent you agree or disagree with the statement on a scale of 1 to 5:		Strongly Agree					Strongly Disagree
		1	2	3	4	5	
1	I managed to memorise and remember the password I was using for this study.						
2	I had to write down my password on a piece of paper in case I failed to remember my password.						
3	I had to save my password for this study on my mobile phone in case I forgot it.						
4	I saved the password I used for this study on the internet browser because I could not memorise it.						
5	I wrote and saved my password somewhere on the computer because I could not remember it.						
6	I had to share my password for this study with a colleague in case I forgot it.						



University of Fort Hare  
Together in Excellence

## APPENDIX C: THE INFORMED CONSENT FORM

### INFORMED CONSENT

I hereby agree to participate in research regarding passphrase authentication. I understand that I am participating freely and without being forced in any way to do so. I also understand that I can stop participating in the survey at any point should I not want to continue and that this decision will not in any way affect me negatively.

I understand that this is a research project whose purpose is not necessarily to benefit me personally.

I have received the telephone number of a person to contact should I need to speak about any issues which may arise in this interview.

I understand that this consent form will not be linked to the questionnaire, and that my answers will remain confidential.

I understand that if at all possible, feedback will be given to my community on the results of the completed research.

.....  
**Signature of participant**

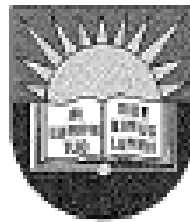
**Date:**.....

I hereby agree to the tape recording of my participation in the study



.....  
**Signature of participant** *Together in Excellence* **Date:**.....

## APPENDIX D: ETHICAL CLEARANCE CERTIFICATE



**University of Fort Hare**  
*Together in Excellence*

### **ETHICAL CLEARANCE CERTIFICATE** **REC-270710-028-RA Level 01**

Certificate Reference Number: FLO0815MAC01

Project title: **A model for using cultural knowledge in creating secure and usable passphrases.**

Nature of Project: PhD in Information Systems

Principal Researcher: Pardon Blessings Maoneke

Supervisor: Prof S Flowerday

Co-supervisor: Dr N Isabirye

On behalf of the University of Fort Hare's Research Ethics Committee (UREC) I hereby give ethical approval in respect of the undertakings contained in the above-mentioned project and research instrument(s). Should any other instruments be used, these require separate authorization. The Researcher may therefore commence with the research as from the date of this certificate, using the reference number indicated above.

Please note that the UREC must be informed immediately of

- Any material change in the conditions or undertakings mentioned in the document
- Any material breaches of ethical undertakings or events that impact upon the ethical conduct of the research

The Principal Researcher must report to the UREC in the prescribed format, where applicable, annually, and at the end of the project, in respect of ethical compliance.

**Special conditions:** Research that includes children as per the official regulations of the act must take the following into account:

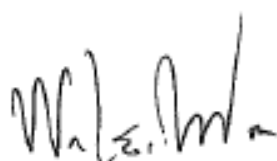
Note: The UREC is aware of the provisions of s71 of the National Health Act 61 of 2003 and that matters pertaining to obtaining the Minister's consent are under discussion and remain unresolved. Nonetheless, as was decided at a meeting between the National Health Research Ethics Committee and stakeholders on 6 June 2013, university ethics committees may continue to grant ethical clearance for research involving children without the Minister's consent, provided that the prescripts of the previous rules have been met. This certificate is granted in terms of this agreement.

The UREC retains the right to

- Withdraw or amend this Ethical Clearance Certificate if
  - Any unethical principal or practices are revealed or suspected
  - Relevant information has been withheld or misrepresented
  - Regulatory changes of whatsoever nature so require
  - The conditions contained in the Certificate have not been adhered to
- Request access to any information or data at any time during the course or after completion of the project.
- In addition to the need to comply with the highest level of ethical conduct principle investigators must report back annually as an evaluation and monitoring mechanism on the progress being made by the research. Such a report must be sent to the Dean of Research's office

The Ethics Committee wished you well in your research.

Yours sincerely



**Professor Wilson Akpan**  
**Acting Dean of Research**

02 March 2017