# Developing Measurement Scales for Passwords Uniqueness and Secrecy

*Emergent Research Forum Paper (ERF)*

**Ali Vedadi**
Mississippi State University
Av540@msstate.edu

**Merrill Warkentin**
Mississippi State University
m.warkentin@msstate.edu

## Abstract

Password authentication represent a critical weakness in the security of online security. Research shows that individuals frequently engage in poor password hygiene behaviors, including choosing weak passwords, reusing passwords, writing down passwords, and facilitating password theft by not covering their keyboard or emailing passwords in clear text. Among these factors, password uniqueness and password secrecy are of great importance in maintaining password hygiene. Despite the importance of these factors, there are no rigorously-tested measurement scales in information security literature for these concepts; therefore, the goal of this research is to develop and empirically test two measurement scales for two key constructs in this context, namely password uniqueness and password secrecy.

***Keywords:*** *Password Hygiene, Uniqueness, Secrecy, Scale Development, LISREL*

## Introduction

In 21st century, password authentication represents a critical weakness in the security of online security. Research shows that individuals frequently engage in poor password hygiene behaviors, including choosing weak passwords, reusing passwords, writing down passwords, and facilitating password theft by not covering their keyboard or emailing passwords in clear text (Adams and Sasse 1999; Campbell et al. 2006, Guo 2013, Zhang et al. 2009). A recent Guardian report (Yadron 2016) indicates that about 272.3 million email usernames and passwords are stolen. Similar accounts about identity theft of online bank accounts have been reported, emphasizing the importance of password hygiene.

In 2004, Bill Gates declared "the password is dead." (Best, 2004) While the declaration may, at some time in the future, prove prophetic, passwords remain the de facto authentication mechanisms. Passwords are usually vulnerable to different threats. Passwords are easily shared with others, are easy to guess, are often written down for all to see, and are often forgotten by the user. The results of these problems include system insecurity, costly breaches, the inability to hold a specific user accountable for unauthorized actions on a system, lost time accessing a system, and increased cost for systems and help desk staff to reset passwords. However, despite the development various innovative methods for authentication, the knowledge-based authentication via password remains the primary method in use (Cheswick, 2013; De Angeli, Coventry, Johnson, and$ Renaud, 2005).

One of the most important issues in the password management area is the overemphasis on choosing strong passwords while underestimating other important factors. In many cases, Internet users, for the sake of convenience, may use a strong password for the most of their important online accounts. This could impose a significant threat to one's online security. For example, hackers can search for other accounts associated with that same person and if the password is reused, they can conveniently access all other accounts. Furthermore, even when an Internet user employs unique and strong passwords for each important online account, he or she may fail to maintain strict secrecy in entering passwords, and may even purposely share passwords with others. Furthermore, many users avoid using password managers for a variety of reasons such as not feeling comfortable with sharing such sensitive information with a third-party, the fear of master password getting compromised, etc. Despite the prevalence of these problems and thus, substantial need to investigate the password hygiene, there are no rigorously-tested measurement scales in information security literature for these concepts; therefore, the goal of this research is to develop and empirically test two measurement scales for two key constructs in this context, namely password uniqueness and password secrecy. The ultimate goal is to develop a full nomology of password hygiene constructs and their corresponding measurement scales as better measures of information security behaviors have been identified as a key research goal for information security scholars (Crossler, et al. 2013; Straub 2009; Warkentin, et al. 2012). The establishment of validated measurement scales for general use by scholars is a time-tested tradition, and is guided by key influential articles by established procedures (Churchill, 1979, MacKenzie, et al. 2011). These measurement scales can enable future studies in the information security context to quantify and scientifically measure these two constructs in various relevant studies. This paper is structured as follows. First, these two constructs are conceptualized. Then, measurement items are presented. Next, the process of data collection and analysis is described. Then, results are presented and discussed. Finally, future steps for enhancing this project are explained.

## Password Uniqueness

The proliferation of online accounts and smartphone apps has led to cognitive overload, such that the number of passwords users are required to remember is growing well beyond the cognitive ability of most users to remember. Adams and Sasse (1999) suggest that users are generally capable of remembering only four or five unique passwords, but the modern world require more than ten distinct passwords, on average (Zhange et al. 2009). This memory limitation is the primary reason for password reuse (Chiasson et al 2009; Duggan et al. 2012; Gaw and Felten 2006; Grawemeyer and Johnston 2011; Zhang et al. 2009). Password strength policies increase this difficulty by requiring users to select multiple composition attributes such as a mixture of uppercase and lowercase letters, numbers, symbols, and the avoidance of dictionary words. Although password composition policies may result in stronger individual passwords, the added difficulty encourages users to reuse passwords rather than create unique passwords for each service or web site. Reusing a password is tantamount to revealing passwords. Many high-profile breaches (Yahoo, Home Depot, Sony, LinkedIn) resulted in cross-site compromises because users maintain the same password for many different services. An individual practicing good password hygiene with respect to uniqueness will avoid reusing the same password across multiple services, and will periodically change their passwords. We define password uniqueness as: *Choosing a unique password for every important online account.*

## Password Secrecy

Passwords must be secret to be secure. To be an effective security tool, passwords must be strong and secret, and users must be able to remember them. However, to alleviate cognitive challenges associated with password strength, and the growing number of required password, secrecy is often a casualty to convenience (Zhang et al., 2009). Users often fail to create strong passwords and they also fail to keep them secret (Warkentin et al., 2004). Extant research has demonstrated that forcing users to conform to high levels of password complexity results in a greater tendency to store strong passwords on sticky notes attached to monitors or affixed to the bottom of keyboards (RAS, 2006a from Zhang 2009). Good password hygiene related to secrecy includes not sharing passwords with others and concealing the use of passwords. In this study, password secrecy is defined as: *The degree of vigilance and cautiousness about using passwords.*

## Methods

To develop the scales, we followed a process recommended by Devellis (1991) that starts with developing new and conceptually consistent definitions of the constructs (as are stated in the previous section). After writing several initial items for each construct, we formed an expert panel to gain the opinion of academic experts in instrumentation, as well as non-academics representative of the target population. After rigorous examination of items and dropping and rewriting some of the items, 5 items for password uniqueness and 6 items remained for the next step. Items are measured using a 7-point fully-anchored Likert scale. Target population can be any Internet user with at least one important online account (like an online bank account). The list of items appear in Table 1.

To test the psychometric properties of the survey, we distributed it to an IRB-approved subject pool in a large university in the southeastern United States and collected 190 complete and usable responses. The average age of participants was 23.5 years. 59% of the respondents were male and 41% were female. In the next sections, the results of 3 stages of validation are presented, including the preliminary investigation, confirmatory factor analysis, and assessing overall model fit. SPSS v23 was used for preliminary analysis including reliability analysis and exploratory factor analysis. LISREL 9.5 was also used for confirmation factor analysis and testing the overall model fit.

**Table 1. Measurement Items**

| Construct | Items |
|---|---|
| Uniqueness | 1. Each of my passwords is unique for each system I access.<br>2. I use a different password for more than one system.<br>3. I avoid using the same password for systems and websites I access.<br>4. I try to select new passwords for new websites I log into.<br>5. Even if I am not required to do so, I regularly select new passwords for key sites. |
| Secrecy | 1. I always monitor my surroundings when entering a password.<br>2. When entering my password, I make an effort to conceal it from others' view.<br>3. Prior to entering my password, I ensure no one is looking over my shoulder.<br>4. I never tell anyone my passwords.<br>5. No one knows my passwords but me.<br>6. I do not share my password with anyone. |

## Preliminary Investigation

The reliability analysis and the Cronbach's Alpha values for Uniqueness and Secrecy show adequate reliability for both scales (.88 and .91 accordingly) and dropping none of the items could increase the reliability; therefore, all items were used for exploratory factor analysis. In addition, by rotating the component matrix using Varimax, 70% of variance was explained by the model. Table 2 shows that items significantly (above .7) load on the expected factors with no cross-loading (the numbers below .4 are suppressed).

**Table 2. Rotated Component Matrix**

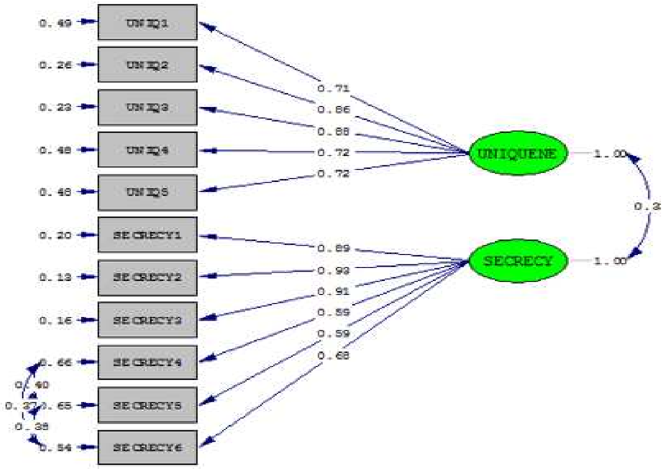
| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Component | 1 | 0.755 | 0.875 | 0.872 | 0.786 | 0.802 | | | | | | |
| | 2 | | | | | | 0.813 | 0.876 | 0.833 | 0.807 | 0.819 | 0.858 |
| Item | | UNQ1 | UNQ2 | UNQ3 | UNQ4 | UNQ5 | SEC1 | SEC2 | SEC3 | SEC4 | SEC5 | SEC6 |

### *Confirmatory Factor Analysis*

Table 3 shows that Squared Multiple Correlations (SMCs) for all items were above .5 (which is the acceptable minimum) except for Secrecy 4 and Secrecy 5. Furthermore, Modification Indices (MI) for LAMBDA-X were acceptable (no value above 10), but this was not the case with MI of Theta-Delta as they were 8 values above or well-above 10. As expected, goodness-of-fit statistics are not ideal and several statistics do not even meet the minimum cut-off point ($\chi$2 = 251, df=43, RMSEA=0.16, NNFI=.82, CFI=.86, CN=51, RMR=.08, GFI=.77). These statistics were improved after covarying the error terms (as can be seen below Figure 1). In addition, Secrecy 4, Secrecy 5 and Secrecy 6 had factor loadings below the minimum acceptable point (.7). Thus, the second iteration of analysis was necessary. After dropping Secrecy 4, Secrecy 5 and Secrecy 6, we ran the model again with the remaining items.

**Table 3. Squared Multiple Correlations**

| Item | Uniqueness 1 | Uniqueness 2 | Uniqueness 3 | Uniqueness 4 | Uniqueness 5 | Secrecy 1 | Secrecy 2 | Secrecy 3 | Secrecy 4 | Secrecy 5 | Secrecy 6 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SMC | 0.506 | 0.742 | 0.774 | 0.517 | 0.519 | 0.773 | 0.858 | 0.805 | 0.427 | 0.435 | 0.538 |



($\chi^2$ = 58, df= 40, NNFI= .98, CFI= .98, Critical N= 208, RMSEA= .04, RMR= .045, GFI= .94).

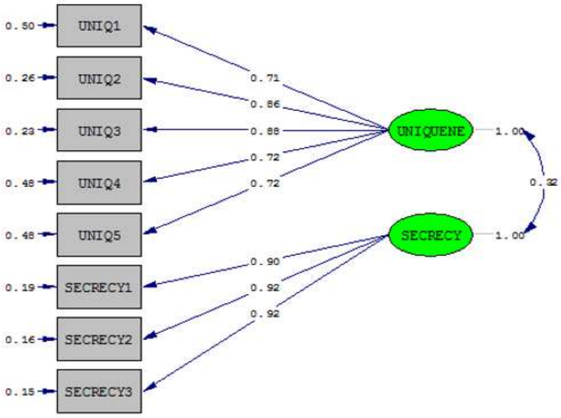
**Figure 1. Path Diagram**

### *Second Iteration of Analysis*

Table 4 shows that all SMCs are acceptably above .5. MI for both LAMBDA-X and Theta-Delta do not indicate any significant problem (no value above 10); therefore, there was no need to relax any constraint. In addition, goodness-of-fit statistics improved even further by dropping these items. Finally, factor loadings for all items are above .7 and for Secrecy items, they are all above .9 which is desirable.

**Table 4. Squared Multiple Correlations**

| Item | Uniqueness 1 | Uniqueness 2 | Uniqueness 3 | Uniqueness 4 | Uniqueness 5 | Secrecy 1 | Secrecy 2 | Secrecy 3 |
|---|---|---|---|---|---|---|---|---|
| SMC | 0.505 | 0.744 | 0.775 | 0.516 | 0.519 | 0.809 | 0.843 | 0.849 |



($\chi^2$ =27.62, df=19, NNFI=.98., CFI=.99., Critical N=248, RMSEA=.049, RMR=.03, GFI=.96).

**Figure 2. Revised Path Diagram**

## Conclusion and Future Steps

This ongoing research project, motivated by the significant dual problems of (1) password recall failures leading to password reuse and (2) failure to maintain password secrecy, has established two statistically reliable and valid scales for use in further studies into the nomological net of information security research into password selection and use phenomena. Researchers can utilize these validated scales in their present form or in forms modified or adapted to fit the context of related studies. Repeated direct or adapted use of these scales will further solidify their value in relevant research contexts.

Further efforts to expand this foundation include the identification of other key constructs within this nomology that require reliable and valid measures for scientific analysis and measurement. The constructs that we are currently working on include password generation methods, password maintenance methods and password recall. Research into the perceived psychological ownership of passwords or password management techniques may be valuable. Other psychological, attitudinal and dispositional constructs and factors that lack empirically tested measurement scales deserve more attention and doing so may shed light on this entire nomology; therefore, we are currently in the process of developing and refining items for the aforementioned constructs in order to present a comprehensive password hygiene measurement scale.

## REFERENCES

Adams, A., & Sasse, M. 1999. "Users Are Not the Enemy," *Communications of the ACM* (42:12), pp.41–46.

Best, J. 2004. "Gates: The password is dead", found at www.zdnet.com/article/gates-the-password-is-dead/

Campbell, J., Kleeman, D., & Ma, W. 2006. "Password Composition Policy: Does Enforcement Lead to Better Password Choices?" *Proceedings of the 17th Australasian Conference on Information Systems Password Composition Policy*, Adelaide, Australia.

Cheswick, W. 2013. "Rethinking Passwords". *Communications of the ACM*, (*56*:2), pp. 40-44.

Chiasson, S., Forget, A., Stobert, E., Van Orschot, P. C., & Biddle, R. 2009. "Multiple Password Interference in Text Passwords and Click-Based Graphical Passwords," *Proceedings of the 16th ACM conference on Computer and communications security*, CCS '09, Chicago, Illinois, pp. 500-511.

Churchill Jr, G. A. 1979. "A Paradigm for Developing Better Measures of Marketing Constructs". *Journal of Marketing Research*, pp. 64-73.

Crossler, R. E., Johnston, A. C., Lowry, P. B., Hu, Q., Warkentin, M., & Baskerville, R. 2013. "Future Directions for Behavioral Information Security Research," *Computers & Security* (32), pp. 90-101.

De Angeli, A., Coventry, L., Johnson, G., & Renaud, K. 2005. "Is a Picture Really Worth a Thousand Words? Exploring the Feasibility of Graphical Authentication Systems". *International Journal of Human-Computer Studies*, (*63*:1), pp. 128-152.

Devellis, R. 1991. "Scale Development: Theory and Applications". London: Sage.

Duggan, G. B., Johnson, H., & Grawemeyer, B. 2012. "Rational Security: Modelling everyday password use," *International Journal of Human-Computer Studies* (70), pp. 415–431.

Gaw, S. Felten, E. 2006. "Password Management Strategies for Online Accounts," *Proceedings of the Second Symposium on Usable Privacy and Security*. ACM Press, New York.

Grawemeyer, B., & Johnson, H. 2011. "Using and Managing Multiple Passwords: A Week to a View," *Interacting with Computers* (23), pp. 256-267.

Guo, K. H. 2013. "Security-Related Behavior in Using Information Systems in the Workplace: A Review and Synthesis," *Computers & Security* (32), pp. 242-251.

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. 2011. "Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques". *MIS Quarterly*, (*35*:2), pp. 293-334.

Straub, D.W. 2009. "Black Hat, White Hat Studies in Information Security. *Keynote Address at IFIP WG 8.11/11.13 Workshop on Information Security Research*. Cape Town, South Africa,.

Yadron, D. (2016). "Hacker collects 272m email addresses and passwords, some from Gmail," found at www.theguardian.com/technology/2016/may/04/gmail-yahoo-email-password-hack-hold-security

Warkentin, M., Straub, D., & Malimage, K. 2012. "Measuring Secure Behavior: A Research Commentary". *Proceedings of the Annual Symposium on Information Assurance*. Albany, New York, pp. 1–8.

Warkentin, Merrill, Kimberly Davis, & Ernst Bekkering. 2004. "Introducing the Check-Off Password System (COPS): An Advancement in User Authentication Methods and Information Security," *Journal of Organizational and End User Computing*, (16:3, pp. 41-58.

Zhang, J., Luo, X., Akkaladevi, S., and Ziegelmayer, J. 2009. "Improving Multiple Password Recall: An Empirical Study," *European Journal of Information Systems* (18:2) pp.165–176.