2013

Genetic Basis of Behavioral Security

Rachel Chung
Chatham University, rchung@chatham.edu

Dennis F. Galletta
University of Pittsburgh - Main Campus, galletta@katz.pitt.edu

Genetic Basis of Behavioral Security

Ting-ting (Rachel) Chung
Chatham University
rchung@chatham.edu

Dennis Galletta
University of Pittsburgh
galletta@katz.pitt.edu

ABSTRACT

Behavioral genetics offers numerous opportunities to bridge gaps in biological research of IS and to shed light on the nature versus nurture debate. This study seeks to explain persistent weaknesses in behavioral security from a genetic perspective. A synthesis of current literatures on cognitive neuroscience, decision making, and fraud victimization suggests a genetic basis for user susceptibility to security risks such as phishing scams. Using the classic twin design, this study reports estimated heritability of behavioral security to be at least 29.15% by comparing concordance between 144 pairs of monozygotic (MZ) twins and that between 52 pairs of same-sex dyzygotic (DZ) twins. Zygosity of the twin pairs serves as the primary independent variable in the behavioral genetics analysis, while performance on a behavioral security test serves as the dependent measure. Implications of the study results are discussed with respect to IS research as well as managerial practices.

Keywords

Behavioral genetics, Phishing, Security, Human factors, Human-computer interaction.

INTRODUCTION

Vulnerabilities in cybersecurity are present in both personal realms and business settings. For example, employees at the grocery chain SuperValu received a fraudulent message notifying them about new (bogus) account numbers for wire transfers to two vendors, Frito-Lay and American Greetings. The employees complied and sent $10 million to the fraudulent accounts[1]. Such cybersecurity issues, having a behavioral source, are of obvious concern to many businesses.

Researchers have concentrated on two major aspects of security. The first area of study involves the technology, which includes tools for encryption, identity assurance, and threat discovery and elimination. However, as the SuperValu case illustrates, an important second area is behavioral security, which examines two groups of people involved in the threats: the perpetrators and the victims. Studies of perpetrators attempt to predict the likelihood of unleashing a threat, identifying their identity, and dissuading attacks by threatening likely and severe punishment. Studies of victims, however, investigate what kinds of attacks are most persuasive (e.g., Dhamija et al. 2006), who might fall for these attacks (e.g., Jagatic et al., 2007; Galletta, et al., 2011), and how people might be better armed against them (e.g., (Kirlappos and Sasse, 2012).

Methods to arm people against attacks often include education or training. However, the effectiveness of these methods has been rather disappointing so far. For example, Weirich (2006) found no significant differences between two groups in their responses to a survey questionnaire five months after a warning flyer manipulation was administered, thus concluding that fear appeals are ineffective for behavioral security in an organizational setting. Herley (2009) demonstrated that the estimated costs of learning to create strong passwords and to understand URLs outweigh the actual benefits. User education might lack potency because of user misconceptions about security. Those misconceptions might reflect deep-rooted decision-making heuristics and biases, such as trust halo effect or anchoring effect, rather than simple lack of knowledge (Kirlappos and Sasse, 2012).

Individual differences in security behavior are striking. In many fields, especially psychology and medicine, when individual differences become problematic, research studies often attempt to assess the extent to which nature (biology: i.e., genetics) is more important than nurture (environment: i.e., training or education) to understand and predict those differences. In most cases, both nature and nurture play important roles in shaping a complex behavior. However, in the case of behavioral security, the extent of nature's influence, compared to that of nurture, is largely unknown in the literature.

The study reported here was designed on the premise that a better understanding of the roles of nature versus nurture in shaping behavioral security would help improve our efforts. Novel studies in areas such as NeuroIS (e.g., Dimoka et al., 2011), to physiological studies (e.g., Haney et al., 2007), and to the four drives of human nature (Abraham et al., 2009), the hunt for biological artifacts underlying cognitive processes is evidenced by a growing theoretical and empirical literature (Kock, 2004, Kock, 2009, Pirolli and Card, 1999). However, these approaches do not tell the entire story; other approaches can both broaden and deepen our understanding of the problem.

[1] For details, please visit http://tinyurl.com/lds35bj

Taking the novel physiological evidence and evolutionary psychology approaches can provide estimates for the population, rather than explanations for individual differences. Although understanding the population in general is important, this strategy provides little insight into those aforementioned factors that lead to differences in individual behaviors in the modern business environment. To estimate the potential impact of interventions such as training or fear appeals, it is important to understand the causes of those individual differences in the extent to which individuals have adopted desirable measures, have avoided risks, and have sensed or failed to sense an imminent risk.

Our specific research questions therefore are: To what extent is security behavior malleable by environmental forces (e.g., interventions), versus genetic makeup? Why are some individuals easier prey for cybercrime than others? How much can training be expected to help?

THE CASE OF BEHAVIORAL SECURITY

The twin design, a classic behavioral genetics research method, can be applied to understand the role of nature versus nurture in the IS user's propensity to engage in secure behavior. Recent studies on security behavior, such as diligence in backing up data files (Boss and Galletta, 2008) and resistance to impulsive clicking on phishing messages (e.g., Galletta et al., 2011) reveal that, even after years of publicizing the need for making backups and avoiding clicking on phishing messages, users often fail to heed those warnings. Many users fail to make backups even after indicating an intention to do so. Many users also continue to click through to websites that are likely to be hostile, even when they have generic subject headers and generic text, and even if they are from strangers and have numeric URL links to content. Furthermore, some users believe outrageous claims of winnings, commissions for transferring large sums of money, or inheritances from unknown decedents. There might be a genetic explanation that can be found using a twin study.

Common across twin studies in business research is the finding that genetic determinants explain a significant portion of individual differences in prudence (Simonson and Sela, 2011), risk taking (Cesarini et al., 2010) and cooperative behavior (Cesarini et al., 2008). Moreover, molecular genetics research has identified allelic correlates of personality traits such as novelty seeking (Cloninger et al., 1996). Because these traits may underlie many of the behavioral issues in IS security, we anticipate genes to play a considerable role in explaining why some individuals are more prone to experience security breaches than others.

Preliminary findings from the limited literature on phishing suggest the presence of personality correlates of security behavior. In an experimental study, Galletta et. al. (2011) found that specific personality traits, such as financial risk propensity, distrusting stance, and the ability to focus, predicted the user's propensity to click on a phishing message. Heritability of financial risk propensity, and other financial decision making behaviors, has been estimated to be around .25 (Cesarini et al., 2010). In other words, 25% of individual differences can be attributed to genetic variations. The rest can be attributed to nurture, or variations in the environment where these individuals have developed. Distrusting stance is likely to be a subset of the agreeableness dimension, whereas the ability to focus can be a manifestation of the conscientiousness dimension of the Big-Five Personality theory. As indicated earlier, heritability estimates for personality dimensions are reported to be around 50% in the literature with great consistency (Plomin and Caspi, 1999). These findings give strong reason to suspect that individual differences in IS users' propensity to engage in secure behavior can be at least partially due to genetic differences.

Studies of fraud victimization, information security, and behavioral genetics allude to potential genetic bases for behavioral security for two theoretical reasons: Risk taking and decision biases. Training has also been studied.

Risk taking studies have found genetic influences on prudence (Simonson and Sela, 2011), risk-taking (Cesarini et al., 2010), novelty-seeking (Cloninger et al., 1996), financial risk propensity (Cesarini et al., 2010), and financial gambling (Dreber et al., 2011).

Decision biases also play a role. Framing, in particular, is one of the decision making biases that correlate with cognitive ability (Stanovich and West, 2008), a highly heritable characteristic (Plomin and Spinath, 2004). The framing effect was recently found to have a genetic basis (Roiser et al., 2009), traced to the Serotonin transporter-linked polymorphic region (5-HTTLPR). Specifically, this genetic variation affects decision making biases driven by contextual cues and uncertainty (Roiser et al., 2009). The amygdala plays a critical role in regulating dopamine levels, which influence decision making and choices (Rogers, 2011). Therefore, these empirical studies suggest that there might be a genetic basis for decision biases that make individuals susceptible to phishing schemes.

Finally, studies have evaluated the effectiveness of training programs to improve the rate of secure behavior. Fear appeals (Boss and Galletta, 2008), anti-phishing educational websites (Sheng et al., 2010), PhishGuru, a specialized training program (Kumaraguru et al., 2009), and Solid training (Kirlappos and Sasse, 2012) have all shown low to moderate impact on secure behavior improvement.

These findings suggest that both nature and nurture could play significant roles in shaping security behavior. By employing the classic twin design of behavioral genetics research, this study will uncover the relative contribution of genetic makeup versus environmental factors to individual differences in security behavior.

RESEARCH METHODOLOGY

The empirical data collection efforts involved participant enrollment, questionnaire data collection, and a phishing IQ test, described below, at the annual Twins Days festival in Twinsburg, Ohio, the largest annual gathering for twins and other multiples in the world. This annual event routinely attracts more than 2,000 pairs of twins who have opportunities to participate in research studies and social events. Behavioral genetic researchers have frequently collected research data at the festival (e.g., Ashenfelter and Krueger, 1994, Settle et al., 2009, Wise et al., 2007).

Participants

Four hundred and five individuals participated in this research study. Zygosity data were missing from five pairs of twins, and data from the twin sibling were missing from three participants. The final dataset included a total of 196 pairs of twins, which consisted of 144 pairs of MZ twins, and 52 pairs of same-sex DZ twins, resulting in a sample size of 392 participants. The participants' ages ranged from 18 to 80, (μ=33.5, median=27). Fifty-three (13.52%) participants reported having been fraud victims.

One method of ensuring that MZ and DZ twins were drawn from comparable populations is to examine the distributions of relevant demographic measures between the two groups. If there are significant differences between the MZ and DZ twins, we would include the relevant variable in subsequent SEM analysis as a covariate. Summary statistics in Table 1 illustrate that the MZ and DZ twin participants are comparable in terms of age and gender. They are also comparable in terms of Facebook account ownership, and whether they have been fraud victims. Inferential statistics suggest no significant differences along any of these dimensions between the two twin groups. Thus, while the sample size is relatively small, especially for DZ twins, there appears to be no systematic bias in the sample that would compromise the resulting estimates of the SEM analysis.

Measures

Zygosity (i.e., whether the twins are MZ or DZ) as a dichotomy factor is the primary independent variable and is measured with an answer to the question "Is your twin brother/sister an identical twin? That is, are you monozygotic twins?" (Ashenfelter and Krueger 1994). Self-report measures of zygosity have been shown to correlate nearly perfectly with genetic verification (Wise et al., 2007). Therefore, this self-report item should be sufficient for measuring zygosity for the purpose of this study. Demographic information, such as gender, age, education, occupation, and zip-code, were also collected as control variables.

The primary dependent variable for this study is the participant's performance on eight questions in the behavioral security test discussed below.

Procedure

Participants answered a questionnaire regarding their fraud victimization experience, Internet technology usage, and personality assessments, as well as the eight-question behavioral security test. Participants were shown e-mail messages or webpages and asked to determine if each of them was legitimate or not. The participants responded to all stimuli on site. After both twins completed the study, a $10 participation fee was paid to each twin.

RESULTS

Performance on the behavioral security test is summarized in Table 1. The highest possible score on the test is 8. MZ and DZ twins performed at comparable levels, with no statistically significant difference between the types.

Table 1. Total Scores (Out of 8) on the Behavioral Security Test

	MZ twins (N = 284)	DZ twins (N = 108)	*Inferential Statistics*
Total Score (range: 0 - 8)	4.72 (1.20)	4.65 (1.10)	$t(360) = .524$ $p = .60$

To evaluate the heritability of performance on the behavioral security test, two behavioral genetic analyses were performed. First, consistent with the classic behavioral genetics literature, a simple comparison of intra-class correlation was performed as a first test of the rate of twin concordance in behavior (Alford, Funk, and Hibbing 2005; Settle et al 2009). Our analysis reveals that the intra-class correlation is different for MZ twins (.51) and DZ twins (.32) (see Table 3). As heritability can be estimated as twice the difference between the MD and DZ correlations, h^2 for the behavioral security test performance was estimated to be .38. In other words, genetic variations account for about 38 percent of the variance in behavioral security test results (see Table 2).

Table 2. Genetic and Environmental Influences on the Behavioral Security Test Total Scores, based on Intra-Class Correlations

	Intra-class Correlation		Heritability (h^2 or a^2)	Shared Environment (c^2)	Non-shared Environment (e^2)
Phenotype	MZ	DZ	2*(MZ-DZ)	(2*DZ)-MZ	1-MZ
Total Test Score	.51	.32	.38	.26	.46

The second analysis involved more sophisticated tools that have recently examined the same parameters using covariance-based univariate SEM (e.g., Nicolaou, 2008a; 2008b; Settle, 2009). The maximum likelihood method, as operationalized in the OpenMx library (Boker et al., 2011) of the R statistical package, estimated the relative contributions of genetics (A), shared environment (C), and unshared environment (E) to twin resemblance in behavioral security test performance (Neale and Cardon, 1992, Falconer and MacKay, 1996). This technique allows us to generate parameter estimates for the magnitude of the ACE components separately, along with the size of the errors of these estimates, while at the same time testing and comparing the fit of various models (Neale and Cardon, 1992).

The first step is to test the full ACE model. The genetic effect component A measured in the model serves as a

primary indicator of the genetic basis of behavioral security. Component A would be one if differences in behavioral security measures are completely determined by genetic variation, and zero if the measures are driven entirely by environmental factors. The fit of the overall ACE model is assessed based on Log Likelihood, and the Akaiki information criterion (AIC) (Akaike, 1987).

Table 3 summarizes results for the full ACE model and two submodels: CE, and AE. The -2 Log Likelihood value and the Akaike information criterion (AIC) (Akaike, 1987) revealed that the best fitting model for explaining individual differences in behavioral security test performance included A, C and E. The two submodels (CE and AE), both lacking the genetic influence component A, are significantly different from the full ACE model. Results indicate that 29.15% of the variance in the dependent measure is explained by genetic differences. 34.96% of the variance can be attributed to shared environments, and 35.89% can be attributed to non-shared environments.

Table 3. ACE Model Comparisons: Behavioral Security Test Performance

Model	A	C	E	-2 Log-Likelihood	AIC	*df*	*p*-value
ACE	29.15%	34.96%	35.89%	-1365.188	-2153.188	394	--
CE	--	99.08%	1%	-358.794	-1148.794	395	< 0.001
AE	100%	--	0%	717.92	-72.08	395	< 0.001

To summarize, both the classic intra-class correlation approach and more recent, more sophisticated ACE modeling via SEM reveal that variations in performance on the behavioral security test are partially explained by variations in genetic makeup. The first approach provides a larger estimate of heritability (38.4%) while the second approach provides a smaller, yet still substantial estimate (29.15%). Both represent the proportion of behavior that is not expected to be changed due to training.

DISCUSSION

Results of this study have several implications for research. First of all, findings about the genetic influences may suggest stability of relevant traits over time. Genetic manifestation is durable, and so if the genetic makeup influences at least some of the variance in security behavior, the unsecure user behavior we observe may be more durable than previously anticipated.

Theoretical understanding of the mechanisms underlying the genetic influences could help us unpack the specific sources of such influences. Future studies should also explore multivariate models using other measures of behavioral security to triangulate the findings.

Future studies will also be able to adopt the twin study technique to distinguish biological versus environmental antecedents to a host of IS usage outcomes. Examples are attitudes such as negativism and impatience, intentions such as adoption or system usage, and behaviors such as error rates, learning speed, cooperation, systems misuse (e.g., D'Arcy et al., 2009), and reactions to managerial interventions such as fear appeals (Boss and Galletta, 2008). These are but a few of the potential breakthroughs that a twin study can afford. Future studies can also be extended to family studies, revealing a more detailed environmental component.

One implication of our findings is that previous emphasis on security education or training may need to be reconsidered. Notions that behavioral security weaknesses can be resolved through education or training alone may be problematic with taking into consideration the biological basis of unsecure behavior. In spite of the growing and frequent warnings, many users just do not seem to internalize training tips. By knowing the extent to which differences in these secure behaviors are determined by genetic makeup versus environmental forces, this research can help managers specify areas where managerial intervention (i.e., one form of environmental influence) may be the most (as well as least) fruitful.

While practitioners often provide warnings and educational experiences to aid users in working more safely with information technologies, our study would provide perhaps more realistic expectations for such programs, and even serve to persuade management to deploy funds towards alternative solutions. For instance, if workers insist on sharing passwords, and do not heed warnings about that practice, management might redeploy training funds towards inexpensive fingerprint recognition devices. Also, if workers fail to make backups, software that makes automatic network backups might be purchased instead.

Findings from twin studies will also benefit educators or parents, who strive to understand how to foster secure online behavior and reduce the rate of cyberbullying or other forms of cybercrime. Twin and adoption studies have found very little influence that the shared family environment has on shaping personality or intelligence (Plomin and Caspi, 1999, Plomin and Spinath, 2004). These findings sent shock waves through the research community on parenting, which always theorized a strong relationship between a positive family environment with positive outcomes in personality or intelligence. Moreover, mental disorders such as ADHD or autism were blamed on poor parenting until behavioral genetics research provided evidence of strong heritability. The finding that these traits are highly heritable and are resistant to family influence not only challenged parenting theories, but also transformed parenting practices. Behavioral genetic findings of IS security behavior may prove to challenge fundamental assumptions in a wide array of IS practices.

Findings from the study reported here must be considered in light of several methodological limitations. First, while the sample size for MZ twins is reasonable, the DZ group is considerably smaller. Unequal sample sizes are rather common in published twin studies, especially those with data collected from the Twins Days Festival. This is probably because the festival events tend to attract MZ rather than DZ twins. There is no reason to suspect that

the unequal samples sizes have compromised our analysis in any way, but our future work will involve recruiting more DX twins into the study at the 2013 Festival to make sure the results are robust. Moreover, current findings are based on a single measure of behavioral security. Estimates of multivariate ACE models would not only strengthen our understanding, they will also help us uncover mediating or moderating processes that shed light on biological mechanisms underlying genetic influences.

CONCLUSION

Behavioral genetics, the study of the genetics of behavior, offer many opportunities to bridge these gaps in biological research of IS. By employing the classic twin design, commonly used in behavioral genetics research, this study is among the first to unpack the genetic versus environmental determinants of individual differences in information systems, using the context of security behavior. The study of genetic versus environmental influences have the potential promise to push the boundaries and challenge the basic assumptions of many IS theories beyond the security literature. A fuller understanding of genetic influences will only improve theoretical explanations for IS user behavior.

ACKNOWLEDGMENTS

We thank Institute for Fraud Prevention, Carlow University's Geibel Institute for Social Justice and Responsibility, and the Dean's Office at Katz Graduate School of Business of the University of Pittsburgh for financial support for this research. We would also like to thank Jennifer Bourne, David Carroll, Manika Kumari, Michelle Lu, Chenjui Su, Amy Timo, and Jocelyn Inlay for their assistance with data collection and management.

REFERENCES

1. Abraham, C., Junglas, I., Watson, R. T. and Boudreau, M.-C. (2009) In *International Conference on Information Systems*Phoenix, Arizona.
2. Akaike, H. (1987) *Psychometrika*, **52,** 317-332.
3. Ashenfelter, O. and Krueger, A. (1994) *American Economic Review*, **84,** 1157-1173.
4. Boker, S., Neale, M. C. C. L., Maes, H., Wilde, M., Spiegel, M., Brick, T., Spies, J., Estabrook, R., Kenny, S., Bates, T., Mehta, P. and Fox, J. (2011) *Psychometrika*, **76,** 306-317.
5. Boss, S. and Galletta, D. F. (2008) In *Ninth International Research Symposium on Accounting Information Systems*Paris, France.
6. Cesarini, D., Dawes, C. T., Fowler, J. H., Johannesson, M., Lichtenstein, P. and Wallace, B. (2008) *Proceedings - National Academy of Sciences USA*, **105,** 3721-4003.
7. Cesarini, D., Johannesson, M., Lichtenstein, P., Sandewall, Ã. r. and Wallace, B. (2010) *The Journal of Finance*, **65,** 1725-1754.
8. Cloninger, C. R., Adolfsson, R. and Svrakic, N. M. (1996) *Nature Genetics*, **12,** 3-4.
9. D'Arcy, J. A., Hovav, A. and Galletta, D. F. (2009) *Information Systems Research*, **20,** 79-98.
10. Dimoka, A., Pavlou, P. A. and Davis, F. (2011) *Information Systems Research*, **22,** 687-702.
11. Dreber, A., Rand, D. G., Wernerfelt, N., Garcia, J. R., Vilar, M. G., Lum, K. and Zeckhauser, R. (2011) *Journal of Risk and Uncertainty*, **43,** 19-38.
12. Falconer, D. and MacKay, T. (1996) *Introduction to Quantitative Genetics*, Longmans Green, Essex UK.
13. Galletta, D. F., Moody, G., Dunn, B. K. and Walker, J. (2011) In *International Conference on Information Systems*Shanghai, China.
14. Haney, M., Pike, J., Galletta, D., Polak, P. and Chung, T. R. (2007) *ICIS 2007 Proceedings*.
15. Herley (2009) In *the 2009 workshop on New security paradigms workshop* Oxford, United Kingdom.
16. Kirlappos, I. and Sasse, M. A. (2012) *IEEE Security & Privacy*, **10,** 24-32.
17. Kock, N. (2004) *Organization Science*, **15,** 327-348.
18. Kock, N. (2009) *Management Information Systems Quarterly*, **33,** 395-418.
19. Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L. F. and Hong, J. (2009) *ACM Transactions on Internet Technology*, **10,** 1-31.
20. Neale, M. C. C. L. and Cardon, L. R. (1992) *Methodology for Genetic Studies of Twins and Families*, Kluwer Academic Publishers.
21. Pirolli, P. and Card, S. (1999) *Psychological Review*, **106,** 643-675.
22. Plomin, R. and Caspi, A. (1999) In *Handbook of personality: Theory and research*(Eds, Pervin, L. A. and John, O. P.) Guilford Press, New York, NY, pp. 251-276.
23. Plomin, R. and Spinath, F. M. (2004) *Journal of Personality and Social Psychology*, **86,** 112-129.
24. Rogers, R. (2011) *Nerupsychopharmacology*, **36,** 114-132.
25. Roiser, J. P., de Martino, B., Tan, G. C. Y., Kumaran, D., Seymour, B., Wood, N. W. and Dolan, R. J. (2009) *Journal of Neoroscience*, **29,** 5985-5991.
26. Settle, J. E., Dawes, C. T. and Fowler, J. H. (2009) *Political Research Quarterly* **62,** 601-613
27. Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L. F. and J., D. (2010) In *the 28th international conference on Human factors in computing systems* Atlanta, GA.
28. Simonson, I. and Sela, A. (2011) *Journal of Consumer Research*, **37,** 951-966.
29. Stanovich, K. E. and West, R. F. (2008) *Journal of Personality and Social Psychology*, **94,** 672-695.
30. Weirich, D. (2006) In *Department of Computer Science*, Vol. Doctor of Philosophy University College London, London, UK.
31. Wise, P. M., Hansen, J. L., Reed, D. R. and Breslin, P. A. S. (2007) *Chemical senses*, **32,** 749-754.