

Differential Analysis: Crew Resource Management and Profiles on the Balanced Inventory of Desirable Responding

Charalambos C. Cleanthous, Ryan Sain, Tabitha Black, Stephen Vera, Suzanne Milton

Abstract—A concern when administering questionnaires is whether the participant is providing information that is accurate. The results may be invalid because the person is trying to present oneself in an unrealistic positive manner referred to as ‘faking good’, or in an unrealistic negative manner known as ‘faking bad’. The Balanced Inventory of Desirable Responding (BIDR) was used to assess commercial pilots’ responses on the two subscales of the BIDR: impression management (IM) and self-deceptive enhancement (SDE) that result in high or low scores. Thus, the BIDR produces four valid profiles: IM low and SDE low, IM high and SDE low, IM low and SDE high, and IM high and SDE high. The various profiles were used to compare the respondents’ answers to crew resource management (CRM) items developed from the USA Federal Aviation Administration’s (FAA) guidelines for CRM composition and training. Of particular interest were the results on the IM subscale. The comparisons between those scoring high (lying or faking) versus those low on the IM suggest that there were significant differences regarding their views of the various dimensions of CRM. One of the more disconcerting conclusions is that the high IM scores suggest that the pilots were trying to impress rather than honestly answer the questions regarding their CRM training and practice.

Keywords—USA commercial pilots, crew resource management, faking, social desirability.

I. INTRODUCTION

THE concern with social desirability, that is, responding in a manner that makes one’s behavior appear socially desirable, is long standing due to the concerns with the implications of the attempt at lying [1]-[3]. There is also a dearth of studies examining social desirability responding and pilots. The problem is so important that a number of techniques have been designed to counteract the phenomenon [1], [4]. At least eight methods have been devised to help manage socially desirable responding. Edwards [5] recommended the forced-choice method. The participant is forced to choose between two items that concern different topics but present socially desirable responding equally. A second method involves presenting items that are supposedly neutral in their presentation [5]. Both the forced choice and neutral items have been criticized on the ground that

participants may not necessarily find such items to be socially desirable. A third method suggests that the threatening items used to assess social desirability may be counteracted by the use of the randomized response technique introduced by [6], [7]. Research suggests that there is less underreporting on very threatening items when the randomized response technique is used [8]. However, the reduction of social desirability itself is not particularly convincing, thus the technique is of limited usability [9]. A fourth method is to have participants self-administer the social desirability questionnaire. The data for this method are not very convincing in their ability to control social desirability responding [8]. The bogus pipeline is a fifth method that has been devised to help control socially desirable responding [10], [11]. The method attempts to convince the participant that the electrodes that are attached to one’s body can detect whether one is telling the truth or not [12]. The method appears relatively efficacious [13], [14], [11], but it leaves open the question as to which form of social desirability it is controlling [15]. A sixth method involves the use of interviewers to administer the questionnaire, the idea is that by establishing rapport with the participant, the social desirable responding will be controlled. The results of such interviews, however, indicate that when the participant and the interviewer are similar in terms of social distance, socially desirable responding is not reduced [16], [17]. A seventh method to emerge is the use of proxy participants, that is, a person familiar with the target participant answers questions about the target participant. The research suggests that when it concerns observable behavior, the proxy produces satisfactory results; however, this is not the case with attitudes [18]. The overall conclusion is that self-administered questionnaires reduce socially desirable responding the best, while telephone interviews maximize such responding [19], [1]. With the advent of computers, computer assessment has entered into the picture. Two computer assisted assessment methods have been developed. The first method allows the interviewer to type responses, this is known as computer-assisted personal interviewing (CAPI). The second involves the participant typing in one’s responses; this is referred to as computer-assisted self-interviewing (CASI). Initially, it was assumed that these methods, and in particular CASI, would provide the participants with a greater sense of anonymity. In turn, one would expect greater forthrightness than with the aforementioned methods [20]. Although the data do provide some cautions, overall the computer administered questionnaires are superior in lowering the rates of SDR [4].

Charalambos C. Cleanthous is with Eastern Washington University, Cheney, WA 99004 USA (phone: 509-359-2816; fax: 509-359-4366; e-mail: ccleanthous@ewu.edu).

Ryan Sain is with PsychCore, Spokane, WA 99005 USA (e-mail: ryansain@gmail.com).

Tabitha Black, Stephen Vera, and Suzanne Milton are with the Eastern Washington University, Cheney, WA 99004 USA (e-mail: trahder@eagles.ewu.edu, stephenvera@ewu.edu, smilton@ewu.edu).

Crew resource management (CRM) has undergone a number of evolutions since its inception as cockpit resource management [21], [22]. Today, the concept is used in a more inclusive manner that better captures the evolution of the interactive nature of human factors. Lauber [23] defined CRM as “using all available resources, information, equipment, and people to achieve safe and efficient flight operations” (p. 20). CRM [32] has been a growing concern, especially as so many accidents, in both commercial and general aviation, have been due to “pilot error” rather than “equipment error.” The modern foundation of CRM has begun at the National Aeronautics and Space Administration [24]. Although there have been evolving definitions and concerns [21], the primary goal has been to optimize the human-machine interface and related interpersonal activities [25]. Interventions using CRM as the focus have had relatively clear outcomes, CRM works. However, it is not clear that CRM works in all domains [26]. In part, this may be due to the lack of standardization in content, design, delivery, and evaluation [27]. Over the history of CRM, there has always been a small subset of commercial pilots who have rejected the concept and strategies taught [28], or have clashed with cultural practices [21], [3]. Moreover, employment applicants may distort their responses; the direction such distortion takes place tends to match the norms of the desired role rather than necessarily socially desirable [29].

The human factor, with all its strengths and limitations, requires greater examination especially when socially desirable responding can complicate the results. The concern is about developing interventions, whether program-wide or for the individual, that do not reflect the actual situation due to the social desirability bias. The present study sought to examine whether pilots who endorsed socially desirable items had significantly different profiles than those who did not, and where, if anywhere, they differed on endorsing or not various CRM criteria as promoted by the USA Federal Aviation Administration [25].

II. METHODS

A. Participants

The participants were drawn from two online sites following permission from the respective administrators. The two sites were the Flights Above the Pacific Northwest Facebook page and the Airline Pilot Central Forums. The participation requirements stipulated that the participant had to have a USA certificate as a commercial pilot, and had to have been active as a commercial pilot within the past 10 years. The reason for the latter requirement was that it increased the probability that the pilot would have had CRM training history per the FAA regulations [25]. Given these criteria, 58 males and five females were retained from a total of 66; three did not identify their gender. The male-female respondent distribution matched the ratio of the USA commercial pilot rate [30].

B. Measures

Three measures were used. The first was about the

participants’ demographics that included: gender, post-secondary education, and total flight time. The second, the *CRM Training Survey* (CRMTS) training survey was developed (Black) from the FAA CRM criteria. The resulting questionnaire had seven subscales: The *pilot-in-command* (PIC) appraised the pilot’s understanding of the authority of a pilot in the command role. The items reflect the pilot’s understanding on degree of authority, aid to this authority, and protocols of standard operating procedures regarding PIC authority. The second subscale was concerned with *communication* (COM). The items refer to interactions between the pilot and all other human or equipment interface for the purposes of the flight mission. The communication factors included coordination of activities, decisions, conflict resolution, clarification of responsibility, and pre-post briefings. Third, the *management of a flight team* (MGT) items dealt with flight management attitude and ability questions. Part of the focus was on the pilot’s use of leadership skills as encouraged by FAA CRM training (Federal Aviation Administration, 2004). *Workload and time management* (TM) was the fourth subscale and was concerned with the pilot’s grasp of task overload (“staying ahead of the airplane”), and an understanding of situational awareness. The fifth and sixth subscales dealt with *stress* (STR) and *fatigue* (FTG). The items pertain to pilot compliance and degree of acceptance of FAA rulings regarding pilot stress and fatigue. The areas assessed included recognition of behavioral markers, prevention, and recovery actions. Finally, *aeronautical decision making* (ADM) involved items that tapped into a pilot’s understanding and ability to effectively evaluate information, even if contradictory or incomplete, while compartmentalizing biases and other cognitive factors.

The *Balanced Inventory of Desirable Responding-6* (BIDR) [2] was used to assess socially desirable responding. There were two reasons for selecting the BIDR. First, it views SDR as multifaceted thus making it current with views regarding SDR [31], and second, it is a self-administered questionnaire that the literature has determined the better method for assessing SDR [4], [1]. In addition, the questionnaires were administered via computer; regarding SDR in particular, the computer method is superior in lowering SDR than the paper-and-pencil method [4]. The BIDR-6 produces two subscales: *Self-deception enhancement* (SDE) and *impression management* (IM). The former is concerned with the presentation of oneself as a person with superior psychological adjustment and personal superiority; these people tend to exaggerate positive attributes and attempt to hide negative attributes about themselves. Those who endorse items that result in IM attempt to portray oneself as a person who is very socially responsible; the items on this subscale describe observable behaviors.

III. RESULTS

A. Participants

The demographic information indicated that the gender of the present self-selected sample matched the rates of the USA

commercial pilot population. However, statistically it is impossible to draw any definitive conclusions regarding the females as their number (five) was very small. Any discussion about the females will simply be about suggestive trends.

The data indicated that on the BIDR-6, 17 of the 63 participants scored in socially desirable ways: 13 scored high on the SDE subscale (faking good), and four on the IM subscale (faking bad). Thus, 26.2% of the sample selected items that would exaggerate their social status. As the literature indicates, such scores may jeopardize the results. We tested the assumption of normalcy using the Kolmogorov-Smirnov test. Results indicated that the distribution was not significant, subsequently a non-parametric analysis was used to evaluate differences between the groups. Two analyses were done using the Mann-Whitney U Test for independent samples. Comparing the two socially desirable scoring groups against the typical responders on the CRMTS, indicated that there were significant differences on the communication ($p < 0.027$) and fatigue subscales ($p < 0.000$), and on the overall CRM scale ($p < 0.02$). However, when the socially desirable responders were combined into one group, the results indicated there were now five significant differences when compared to the non-socially desirable responders. The differences included the above three subscales with the significance levels changing. The results indicated that the statistically significant differences existed on the communication subscale ($p < 0.027$ vs. $p < 0.009$ (combined)), fatigue subscale ($p < 0.000$ vs. $p < 0.002$ (combined)), managing a flight team ($p < 0.002$), aeronautical decision making ($p < 0.006$), and on the overall CRM scale ($p < 0.02$ vs. $p < 0.003$ (combined)). Thus, on all the subscales but fatigue, the differences were greater when the two groups were combined (See Table I).

IV. CONCLUSION

The results speak to the necessity of using a scale that assesses socially desirable responding. Although a variety of procedures have been developed to counteract socially desirable responding, all leave much room for improvement. In the present case, rather than excluding the socially desirable responders, their results were made the comparison group allowing us to detect important differences when compared to the typical responders.

A limitation of the present study is the small number of female respondents. Although the rate in the sample matches that of the USA female pilot population, the small number makes it impossible to do any meaningful statistical analyses. However, the small number that did respond in a socially desirable manner does raise important concerns as those of the male pilot who responds in a socially desirable manner. Greater efforts do need to be made for greater female representation in future research as it would allow for scientifically sound comparisons on both attitudes and performance.

TABLE I
CRM ANALYSES BY BIDR RESPONDENT TYPE

CRM Category	Mean (SD) Valid	Mean (SD) Socially Desirable	Mann-Whitney p
Communication	38% (.137) correct	27% (.130) correct	.009
Managing a Flight Team	37% (.141) correct	28% (.119) correct	.029
Fatigue	93% (.098) correct	74% (.272) correct	.002
Aeronautical Decision Making	77% (.156) correct	59% (.251) correct	.006
Overall CRM	49% (.139) correct	38% (.139) correct	.003

The literature also indicates that socially desirable responding occurs under conditions of the occupational context. This, too, is an important phenomenon and ties in with the concern over whether interventions may be developed based on inadequate data. One way of controlling for the higher rates of socially desirable responding is to use one of the methods for controlling such behavior, with the computerized version currently receiving the greater confirmation for reducing socially desirable responding.

By examining socially desirable responding as a multifaceted concept with the ability to determine direction of impression management, investigative research may be made in reference to the motives of the participant. Such research may provide insight to prevention and identification of the bias. Examining such differences would allow for the development of interventions to examine in-cockpit performance that would enhance CRM.

REFERENCES

- [1] Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology*, 15(3), 263-280.
- [2] Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). New York: Academic Press.
- [3] Paulhus, D. L. (2002). Socially desirable responding: the evolution of a construct. In H. I. Braun, N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 46-69). Mahwah NJ: Lawrence Erlbaum.
- [4] National Aeronautics and Space Administration (2007). *National Aviation Operations Monitoring Service—Reference Report: Concepts, methods, and development roadmap*. Retrieved on February 17, 2017 from https://www.nasa.gov/pdf/207238main_NAOMS%20Reference%20Report_508.pdf
- [5] Edwards, A. L. (1970). *The measurement of personality traits by scales and inventories*. Holt, Reinhard and Winston. New York.
- [6] Himmelfarb, S. and Lickteig, C. (1982). Social desirability and the randomized response technique. *Journal of Personality and Social Psychology*, 43, 710-717.
- [7] Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60, 63-69.
- [8] Bradburn, N. M. & Sudman, S. (1979). *Improving interview methods and questionnaire design*. San Francisco, CA: Josey-Bass.
- [9] Miller, J. P. (1981). Complexities of the randomized response solution. *American Sociological Review*, 46, 928-930.
- [10] Jones, E. E. & Sigall, H. (1971). The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin*, 76, 349-364.
- [11] Roese, N. J., & Jamieson, D. W. (1993). Twenty years of bogus pipeline research: A critical review and meta-analysis. *Psychological Bulletin*, 114, 363-375.
- [12] Ries, M., Kalle, R. J., & Tedeschi, J. T. (1981). The bogus pipeline and attitude moderation following induced compliance: Misattribution of dissonance arousal or impression management inhibition? *Journal of Social Psychology*, 115, 247-258.
- [13] Aguinis, H., Pierce, C. A., & Quigley, B. M. (1995). Enhancing the

- validity of self-reported alcohol and marijuana consumption using a bogus pipeline procedure: A meta-analytic review. *Basic and Applied Social Psychology*, 16, 515-527.
- [14] Arkin, R. M. and Lake, E. A. (1983). Plumbing the depths of the bogus pipeline: A reprise. *Journal of Research in Personality*, 17, 81-88.
- [15] Paulhus, D. L. (1982). Individual differences, impression management, and cognitive dissonance: Their concurrent operation in forced compliance. *Journal of Personality and Social Psychology*, 43, 838--852.
- [16] Weiss, C. H. (1969). Interviewer biasing effects: Toward a reconciliation of findings. *Public Opinion Quarterly*, 33, 127-129.
- [17] Williams, J. A. (1969). Interviewer biasing effects: Toward a reconciliation of findings. *Public Opinion Quarterly*, 33, 125-127.
- [18] Sudman, S. & Bradburn, N. M. (1974). *Response effects in surveys*. Chicago, IL: Aldine Press.
- [19] Calsyn, R. J.; Roades, L. A.; and Calsyn, D. S. (1992). Acquiescence in needs assessment studies of the elderly. *The Gerontologist*, 32, 246-252.
- [20] Nicholls, W. L., II, Baker, R. P., & Martin, J. (1997). The effect of new data collection technologies on survey data quality. In L. E. Lyberg, P. P. Biemer, M. Collins, E. D. de Leeuw, C. Dippo, N. Schwarz, and D. Trewin (Eds.), *Survey measurement and process quality* (pp. 221-248). New York: John Wiley & Sons.
- [21] Helmreich, R.L., Merritt, A.C., & Wilhelm, J.A. (1999). The evolution of crew resource management training in commercial aviation. *International Journal of Aviation Psychology*, 9(1), 19-32.
- [22] Helmreich, R. L. & Foushee, H. C. (2010). Why CRM? Empirical and theoretical bases of human factors training. In B. G. Kanki (Ed.), *Crew resource management* (2nd ed., pp. 3-57). San Diego, CA: Elsevier Science.
- [23] Lauber, J. K. (1984). Resource management in the cockpit. *Air Line Pilot*, 53, 20-23.
- [24] Cooper, G. E., White, M. D., & Lauber, J. K. (1980). *Resource management on the flight deck: Proceedings of a NASA/industry workshop*. (NASA CP-2120). Moffett Field, CA: NASA Ames Research Center.
- [25] Federal Aviation Administration. (2004). Advisory Circular System 120-51E (2004) – Crew resource management training. Retrieved on August 4, 2016 from http://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentid/22879.
- [26] Salas, E., Wilson, K. A., Burke, C. S., Wightman, D. C. (2006). Does crew resource management training work? An update, an extension, and some critical needs. *Human Factors*, 48(2), 392-412.
- [27] Salas, E., Prince, C., Bowers, C. A., Stout, R. J., Oser, R. L., & Cannon-Bowers, J. A. (1999). A methodology for enhancing crew resource management training. *Human Factors*, 41(1), 161-172. doi: 10.1518/001872099779577255.
- [28] Helmreich, R. L., & Wilhelm, J. A. (1991). Outcomes of crew resource management training. *International Journal of Aviation Psychology*, 1(4), 287-300.
- [29] Birkeland, S. A., Manson, T. M., Kisamore, J. L., Brannick, M. T., & Smith, M. A. (2006). A meta-analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment*, 14, 317-335.
- [30] Bureau of Labor Statistics (2016). *Labor Force Statistics from the Current Population Survey*. Transportation and material moving—aircraft pilot and flight engineers. Retrieved on February 26, 2017 from <https://www.bls.gov/cps/cpsaat11>
- [31] Calsyn, R. J., Roades, L. A., & Calsyn, D. S. (1992). Acquiescence in needs assessment studies of the elderly. *The Gerontologist*, 32, 246-252.
- [32] Hart, C. M., Ritchie, T. D., Hepper, E. G., & Gebauer, J. E. (2015). The Balanced Inventory of Desirable Responding Short Form (BIDR-16). *SAGE Open*, 5(4), 1-9. Retrieved on November 4, 2106, from sco.sagepub.com. DOI: 10.1177/2158244015621113.
- [33] Maurino, D. E. & Murray, P. S. (2010). Crew resource management. In J. A. Wise, V. D. Hopkin, and D. J. Garland (Eds.), *Handbook of aviation human factors* (2nd ed., pp. 10-1-20). Boca Raton, FL: CRC Press, Taylor and Francis Group.