

Title: Effective Red Teaming: Psychological Safety's Role in the Production of Malevolent Creativity Idea Generation

Application for the 2024-2025 Graduate Research and Creative Activity (GRACA) Grant

MA Student: Emma Theobald

Faculty Advisor: Dr. Samuel Hunter

University of Nebraska at Omaha, Department of Psychology

Project Description

In 1984, the Irish Republican Army (IRA) attempted to assassinate the prime minister of the British government at the time, Margaret Thatcher. After the failed attempt, involving a bomb planted in the hotel Thatcher was staying at, the IRA stated “You have to be lucky all the time. We only have to be lucky once.” (Clutterbuck, 2008). Opposingly, counterterrorism officials, to protect against nonstate actors or terrorist organizations, must always be developing novel countermeasures to mitigate potential threats. It only takes one successful implementation of a plan for catastrophe to ensue, and nonstate actors and terrorists have been seen to adopt novel methods and generate creative solutions to increase chances of success. (Gill et al., 2013). For example, the 9/11 terrorist attack has been recognized as a “failure of the imagination” by the 9/11 Commission (National, 2011). One method to proactively mitigate emerging and potential threats is red teaming. Red teaming, put simply, is the act of thinking like the “enemy” or playing devil’s advocate. (Hoffman, 2017). Although most frequently utilized in cybersecurity, military, and counterterrorism industries, red teaming is impartial to one field or industry with applications existing in organizations such as brainstorming on an competitors marketing scheme to attenuate any loss of sales. Thus, the overarching goal or outcome of red teaming is to innovate “decision making by challenging assumptions and exploring new ideas.” (Mateski, 2013; Landry, 2017).

As nonstate actors and terrorists poise novel threats and adopt emerging technology, red teams must follow in suit to outmaneuver the enemy. This requires the generation of malevolent creative ideas or solutions. Malevolent creativity can be defined as “the generation of ideas that are novel and useful” (Gill et al., 2013; Gutworth et al., 2018) with the primary goal of causing physical or psychological harm (Hunter et al., 2022). Malevolent creativity is theorized to follow the same creative process (Mumford et al., 1991) as benevolent creativity. Red teaming poises as a tool for counterterrorism practitioners to reduce the likelihood of successful implementation of novel threats from nonstate actors by thinking like the enemy and generating malevolently novel solutions themselves.

Although red teaming has proven to be valuable in the past, there is a lack of evidence on how to maximize its potential. Particularly, psychological safety, “the shared belief amongst individuals as to whether it is safe to engage in interpersonal risk-taking in the workplace,” may be a variable to consider in red teaming exercises. Past literature has alluded to psychological safety having an impact on red teaming exercises, one military leader states that in a military context the best way to establish shared trust within a group is for the educator or leader to be willing to share “his or her own vulnerabilities and mistakes.” (Matherly, 2020). Based on the impression management theory (Leary & Kowalski, 1990), individuals may be able to tailor their behavior or communication within a group to fit the social norms of that group. This study is interested in whether individuals follow impression management theory when communicating malevolently creative ideas where the outcome is to cause harm to others as well as if empirical results with hold the same assumptions about psychological safety’s effect on red teaming exercises as current red teaming practitioners.

This study aims to first and foremost provide empirical evidence to red teaming practitioners to create the group climate and elicit the best possibility of valuable outcomes of malevolent creative solutions to benefit the end goal to protect against and mitigate novel threats. This study also poises to fill gaps in the literature including lack of experimental manipulation of red teaming exercises with only a few proposes of experimentally manipulated red teaming exercises (Behlendorf & Ackerman, 2023) existing at this time, as well as how psychological safety effects malevolent creativity processes. Therefore, the following hypotheses are proposed:

H1: In both the malevolent and benevolent creativity conditions, ideas generated will be significantly more creative than in the control condition.

H2: In both malevolent and benevolent conditions, psychological safety will significantly moderate the relationship between condition and creativity, such that higher psychological will elicit higher creativity and malevolent creativity.

H3: In the malevolent creativity condition, psychological safety will significantly moderate the relationships between condition and malevolent creativity, more so than in the relationship between the benevolent creativity conditions and creativity.

Method

The survey will take place on Qualtrics survey platform and participants will be recruited from the Prolific sampling platform. Participant inclusion criteria will be over 19 years old and English speaking. The participants will first be asked to complete demographic questions regarding age, ethnicity, gender, employment, and level of education.

This study will be using a 3×2 randomized design with six conditions, such that the three conditions are: creative group (benevolent creativity/malevolent creativity/control) and presence of psychological safety (no psychological safety/psychological safety). Participants in the malevolent and benevolent creativity conditions will read a prompt describing that they are an event security staff member at an outdoor concert taking part in an exercise to generate ideas. The control group will read a similar prompt but will ask to generate ideas on what activities the event security staff will take part in throughout the concert. The malevolent creativity condition will be asked to generate ideas on potential threats to the attendees at the outdoor concert, while the benevolent creativity condition will be asked to generate ideas on potential celebratory surprises that could occur for the attendees. Before reporting their ideas, the participants will be randomly put into either a psychological safety condition or no psychological safety condition based on Chatman et al., 1998's psychological safety manipulation prompt as seen in Deng et al., 2019. To ensure the psychological safety manipulation was successful, participants will complete Edmonson (1999) psychological safety questionnaire. Then, participants will report their ideas. Three independent raters will assess the malevolent creativity of participants' responses using indices of idea novelty, feasibility, utility, and harm (Amabile, 1982). All raters have been trained prior to coding on consensual assessment technique (CAT; Amabile, 1982) coding procedures for a minimum of 20 hours. For statistical analysis, a 3×2 ANOVA will be run in SPSS after proper preparation of designated variables such as data cleaning, conducting interrater correlation coefficients to ensure agreement among raters and recoding values of conditions to be in separate groups.

This methodology for participants to generate ideas based on a proposed prompt and experimental manipulation (Deng et al., 2019) has been developed in the creativity literature (Amabile, 1982) and has been previously used in the malevolent creativity literature (Nguyen et al., 2024). CAT coding has been regarded as "the gold standard" of analyzing creative ideas in an experimental study (Amabile, 1982), and leaders in the field of creativity literature have been consulted and have recommended this methodology for the proposed hypotheses.

The expected findings of this study are that psychological safety will moderate the relationship between the benevolent creativity condition and creative outcomes (i.e. novelty and feasibility), and in the malevolent creativity condition, psychological safety will moderate the relationship between malevolent creativity condition and creative outcomes (i.e. novelty, feasibility, utility, and harm) more so than in the benevolent creativity condition. As stated before, red teaming is vital to counteract, mitigate, and reduce novel threats brought on by nonstate actors and terrorists generating malevolently creative ideas. Understanding the best way to conduct red teaming exercises is vital to generating highly novel and feasible creative ideas. The generation of harmful ideas may pose a different challenge as those who perceive low psychological safety may not feel comfortable expressing their harmful ideas as producing harmful ideas is outside of group norms.

Project Timeline

Month	Planned Research Activities
November 2024	Submit study to IRB for approval. Begin developing study in Qualtrics survey platform
January 2025	Run internal pilot and pilot on Prolific sampling platform
May 2025	Collection of data.
June 2025	Analyze data.
July 2025	Write up results into publishable paper format.
August 2025	Finalize publishable paper and submit to best fitting Journal. Reformat for application to Society of Industrial and Organizational Psychology Annual Conference.

Student/Faculty Mentor Roles

Student Role: Emma Theobald, Department of Psychology, will focus on the following research activities: Designing and implementing surveys in Qualtrics platform, running data collection on Prolific sampling platform, analyzing data using SPSS and RStudio, writing results into publishable paper, and reformatting paper for conference submission.

Faculty Role: Dr. Samuel Hunter, Regents-Foundation Professor of Industrial and Organizational Psychology, will oversee data collection and analysis, review surveys and methodology, provide general supervision, provide feedback for paper submissions to both a journal and conference, and provide co-author support.

Previous Internal Funding

I have received no previous internal funding for this project.

Budget Justification

The budget request for this project includes a stipend of \$5,000 for participant expenses throughout data collection and compensation for hours spent on this project. I expect to work on this project throughout the Fall 2024, Spring 2025, and Summer 2025 semesters, and will be focusing many of my working hours to completing this project. As such, I will be unable to seek additional employment and will require an hourly wage for basic living expenses. This will allow me to complete this project in a timely manner and obtain the highest quality outcomes in both publication and conference submissions. The stipend is calculated based on \$12, the minimum wage in Nebraska, with 40 hours of work per week. A table of the complete budget is included below:

Budget Item	Budget Description	Justification for Expenses
November 2024	Hourly Wage Costs	Hourly Total: \$480 • 40 hours allocated • \$12 wage per hour
January 2025	Hourly Wage Costs Pilot Study	Hourly Total: \$480 • 40 hours allocated • \$12 wage per hour Pilot Study: \$150 • Pilot data for 10 participants (\$15 per participant) for fair compensation
May 2025	Hourly Wage Costs Data Collection	Hourly Total: \$480 • 40 hours allocated • \$12 wage per hour Data Collection: \$3,000 • Data for 200 participants (\$15 per participation) for fair compensation
August 2025	Hourly Wage Costs	Hourly Total: \$410 • 40 hours allocated • \$10.25 wage per hour (adjusted to remain under the maximum funding available)

References

- National Commission on Terrorist Attacks. (2011). The 9/11 Commission Report: Final Report of the National Commission on Terrorist Attacks Upon the United States (Authorized Edition). WW Norton & Company.
- Clutterbuck, L. (2008, November 30). Terrorists have to be lucky once; targets, every time | rand. <https://www.rand.org/pubs/commentary/2008/11/terrorists-have-to-be-lucky-once-targets-every-time.html>
- Nguyen, T. L., Walters, K. N., d'Amato, A. L., Miller, S. R., & Hunter, S. T. (2024). Target personification influences the positive emotional link between generating and implementing malevolently creative ideas. *Creativity Research Journal*, 36(1), 42-57.
- Hoffman, B. G. (2017). *Red Teaming: How your business can conquer the competition by challenging everything*. Crown business.
- Mateski, M. (2013). Red Teaming: A Balanced View. *The Red Team Journal*.
- Matherly, M. C. (2020). The Group Psychology of Red Teaming. *Journal of Military Learning*, 29-49.
- Mumford, M. D., Mobley, M. I., Reiter-Palmon, R., Uhlman, C. E., & Doares, L. M. (1991). Process analytic models of creative capacities. *Creativity research journal*, 4(2), 91-122.
- Landry, T. (2017). Embracing the devil: An analysis of the formal adoption of red teaming in the security planning for major events. *Homeland Security Affairs*.
- Gill, P., Horgan, J., Hunter, S. T., & D. Cushenbery, L. (2013). Malevolent creativity in terrorist organizations. *The Journal of Creative Behavior*, 47(2), 125-151.
- Gutworth, M. B., Cushenbery, L., & Hunter, S. T. (2018). Creativity for deliberate harm: Malevolent creativity and social information processing theory. *The Journal of Creative Behavior*, 52(4), 305-322.
- Hunter, S. T., Walters, K., Nguyen, T., Manning, C., & Miller, S. (2022). Malevolent creativity and malevolent innovation: A critical but tenuous linkage. *Creativity research journal*, 34(2), 123-144.

- Behlendorf, B., & Ackerman, G. (2023). DESSERT: A Novel Framework for Empirical Red Teaming at Scale. *Simulation & Gaming*, 54(1), 5-27.
- Chatman, J. A., Polzer, J. T., Barsade, S. G., & Neale, M. A. 1998. Being different yet feeling similar: The influence of demographic composition and organizational culture on work processes and outcomes. *Administrative Science Quarterly*, 43: 749-780.
- Deng, H., Leung, K., Lam, C. K., & Huang, X. (2019). Slacking off in comfort: a dual-pathway model for psychological safety climate. *Journal of management*, 45(3), 1114-1144.
- Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative science quarterly*, 44(2), 350-383.
- Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological bulletin*, 107(1), 34.
- Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of personality and social psychology*, 43(5), 997.
- Hoffman, B. (1996). Intelligence and terrorism: Emerging threats and new security challenges in the post-cold war era. *Intelligence and National Security*, 11(2), 207-223.



September 20th, 2024

Subject: Emma Theobald GRACA

Dear GRACA Selection Committee,

I am writing on behalf of Emma Theobald, a student applying for a GRACA award. I am currently Emma's PhD advisor. In the past year Emma has been my graduate student, she has worked closely with former and current senior students of mine, and they have spoken very highly of her. I highly recommend Emma for the position and feel she will do an outstanding job with this award.

Emma has an outstanding work ethic, taking on various projects and working effectively with fellow students and lab members. She volunteers to do work above and beyond what others are doing and I can consistently go to her with a challenging issue knowing she will step up and get the job done. I have found myself offering her the opportunity to work on more complex tasks and this is largely due to her ability to competently complete all tasks that come her way. Her effort and motivation are unmatched.

She is committed to conducting exceptional research and simply put, I need her expertise to get my own program of work running. I can rely on Emma as I feel confident in her ability to deliver outstanding research. Emma is a perfect choice for the GRACA, and I hope you are willing to support her. Emma will provide me with regular updates on her progress for the project to ensure that deadlines are being met, and I will provide mentor support when needed.

Best,

Samuel T. Hunter, PhD

Regents-Foundation Professor of Industrial and Organizational Psychology

Head of Strategic Operations

National Counterterrorism, Innovation, Technology, and Education (NCITE) center

University of Nebraska at Omaha

Director of NCITE's Innovation Incubator

Associate Editor, [*Creativity Research Journal*](#)

Associate Editor, [*Journal of Business and Psychology*](#)

samuelhunter@unomaha.edu