HUSO 2025

The Eleventh International Conference on Human and Social Analytics

ISBN: 978-1-68558-238-8

March 9th –13th, 2025

Lisbon, Portugal

HUSO 2025 Editors

Dennis J. Folds, Lowell Scientific Enterprises (LSE), USA

HUSO 2025

Forward

The Eleventh International Conference on Human and Social Analytics (HUSO 2025), held between March 9th, 2025, and March 13th, 2025, in Lisbon, Portugal, continued a series of international events bridging the concepts and the communities dealing with emotion-driven systems, sentiment analysis, personalized analytics, social human analytics, and social computing.

The recent development of social networks, numerous ad hoc interest-based virtual communities, and citizen-driven institutional initiatives, raise a series of new challenges in considering human behavior, both in personal and collective contexts.

There is a great possibility to capture particular and general public opinions, allowing individual or collective behavioral predictions. This also raises many challenges, on capturing, interpreting, and representing such behavioral aspects. While scientific communities now face new paradigms, such as designing emotion-driven systems, dynamicity of social networks, and integrating personalized data with public knowledge bases, the business world looks for marketing and financial prediction.

We take here the opportunity to warmly thank all the members of the HUSO 2025 technical program committee, as well as all the reviewers. The creation of such a high-quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and effort to contribute to HUSO 2025. We truly believe that, thanks to all these efforts, the final conference program consisted of top-quality contributions. We also thank the members of the HUSO 2025 organizing committee for their help in handling the logistics of this event.

We hope that HUSO 2025 was a successful international forum for the exchange of ideas and results between academia and industry for the promotion of progress in the field of human and social analytics.

HUSO 2025 Chairs

HUSO 2025 Steering Committee
Christian Bourret, University of Paris East - Marne la Vallée (UPEM), France
Nitin Agarwal, University of Arkansas at Little Rock, USA
Dennis J. Folds, Lowell Scientific Enterprises (LSE), USA

HUSO 2025 Publicity Chairs
Francisco Javier Díaz Blasco, Universitat Politècnica de València, Spain
Ali Ahmad, Universitat Politècnica de València, Spain
José Miguel Jiménez, Universitat Politècnica de València, Spain
Sandra Viciano Tudela, Universitat Politècnica de València, Spain

HUSO 2025
Committee

HUSO 2025 Steering Committee

Christian Bourret, University of Paris East - Marne la Vallée (UPEM), France
Nitin Agarwal, University of Arkansas at Little Rock, USA
Dennis J. Folds, Lowell Scientific Enterprises (LSE), USA

HUSO 2025 Publicity Chairs

Francisco Javier Díaz Blasco, Universitat Politècnica de València, Spain
Ali Ahmad, Universitat Politècnica de València, Spain
José Miguel Jiménez, Universitat Politècnica de València, Spain
Sandra Viciano Tudela, Universitat Politècnica de València, Spain

HUSO 2025 Technical Program Committee

Paul Abbiati, Founding Fellow of the EUROPEAN LAW INSTITUTE, Austria
Nitin Agarwal, University of Arkansas at Little Rock, USA
Harry Agius, Brunel University London, UK
Jessica Barfield, University of Tennessee - Knoxville, USA
Balbir Barn, Middlesex University, London, UK
Christian Bourret, Université Gustave Eiffel, France
Anita Chandra, IIT Patna, India
Dickson Chiu, The University of Hong Kong, Hong Kong
Claudia Da Re, Gustave Eiffel University - Marne-la-Vallée Campus, France
Chen Ding, Ryerson University, Canada
Birgitta Dresp-Langley, Centre National de la Recherche Scientifique, France
Thierry Edoh, RFW-Universität Bonn, Germany
Fernanda M. Eliott, Noyce Science Center - Grinnell College, USA
Diana Florea, Lucian Blaga University of Sibiu, Romania
Silvia Florea, Lucian Blaga University of Sibiu, Romania
Dennis J. Folds, Lowell Scientific Enterprises (LSE), USA
Matteo Francia, University of Bologna, Italy
Enrico Gallinucci, University of Bologna, Italy
Luca Giraldi, EMOJ srl, Ancona, Italy
Sergiy Gnatyuk, National Aviation University, Kyiv, Ukraine
Damian Gordon, Technological University Dublin, Ireland
Denis Gracanin, Virginia Tech, USA
Shatha Ali A. Hakami, University of Birmingham, UK
Hafida Hammadi, Gustave Eiffel University, France
Tzung-Pei Hong, National University of Kaohsiung, Taiwan
Rodrigo Huerta-Quintanilla, Cinvestav, Mexico
Domenico Talia, Università della Calabria, Rende, Italy
Emilio Insfran, Universitat Politecnica de Valencia, Spain
Yasushi Kambayashi, Sanyo-Onoda City University, Japan

Table of Contents

Countering the Ripple Effects: Strategies for Decoding and Disrupting Emotional Triggers in Online Rumor Trust

Yi-Chen Lee
Department of Information and Management
National Dong Hwa University
Hualien, Taiwan, ROC
email: ycleeim@gms.ndhu.edu.tw

***Abstract*—This study investigates how emotional factors influence the message believability of online fake news, focusing on the individual and interaction effects of emotional arousal, valence, and social contagion. Grounded in the "Feelings-as-Information" theory, the research explores how emotions serve as heuristic cues that shape cognitive evaluations and processes. The study examines how varying emotional conditions affect message believability through an experimental design simulating fake message dissemination. The findings aim to provide insights into the role of emotions in driving misinformation and offer strategies to mitigate its spread. Additionally, the outcomes will inform media literacy efforts and guide social media platform policies to better address the challenges posed by emotionally driven misinformation.**

Keywords: arousal cues; emotional valence; social contagion; message believability; fake news dissemination.

I. INTRODUCTION

Online fake news has become a significant concern, worsened by the rise of social media platforms [1]. False messages on these platforms disrupt public understanding and influence societal decisions. Accurate information is essential for informed behavior and positive social outcomes, but the fast spread of false information makes this difficult. Studies show misleading content spreads faster and broader than verified facts [2]. This is troubling, as unverified claims and rumors hinder effective online communication. The absence of strict fact-checking and editorial oversight further fuels the spread of misinformation [3]. Despite ongoing efforts, online false messages remain a challenge to clear and accurate communication.

Emotion significantly influences how information is evaluated and shared online. Social media often reflects emotional reactions to life events [4]. Messages that trigger strong emotions like fear, disgust, or surprise are more likely to be shared, as these emotions capture attention and promote distribution [5]. This is especially true for fake news or rumors, often using exaggerated language and vivid imagery to manipulate emotions [6]. During the COVID-19 pandemic, emotionally charged messages significantly increased the spread of misinformation, confusing the public [7]. This raises an important question: How do emotional expressions affect the message's believability and dissemination? This points to the need for further research on how emotional factors shape the evaluations on social platforms.

Emotional expressions are essential in information evaluation, but research findings are mixed. Some studies show positive effects, while others suggest the opposite. For instance, Yin, Bond, and Zhang (2017) [8] found that online reviews with low arousal levels were perceived as more helpful. In contrast, Ye and Motoki (2024) [9] observed that high-arousal messages are more effective in the USA, while low-arousal ones are preferred in Japan, particularly in discussions about healthy food. These differences highlight the complexity of emotional dynamics in evaluating information. However, these studies focus primarily on factual content, leaving a gap in understanding how emotional expressions affect the believability of online fake news and sharing behavior.

This study explores how emotional arousal cues, valence, and social contagion in online fake news impact perceived message believability. Analyzing these emotional dimensions together seeks to uncover their individual and combined effects. The research is based on the Feelings-as-Information Theory, which suggests that emotions act as mental shortcuts influencing cognitive evaluations. This framework is well-suited for investigating how arousal cues shape the believability of online fake news, especially when moderated by the emotional valence and social contagion within the messages.

The rest of this paper is organized as follows. Section II describes the literature review and hypothesis development. Section III addresses the planned research method. Section IV goes into the expected contributions of this study.

II. LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

A. Feelings-as-Information Theory

The Feelings-as-Information Theory, introduced by Schwarz and Clore (1983) [10] and later expanded by Schwarz (2012) [11], posits that emotions act as informational cues that shape judgments and decisions. Emotions function as heuristics, enabling individuals to interpret complex information efficiently by offering quick, often subconscious, feedback on the relevance and significance of stimuli. Research supports this framework, demonstrating that people rely on emotions to assess events

 ISBN: 978-1-68558-238-8

[10]. Forgas (1995) [12] found that emotions influence tasks like memory recall and risk assessment, while Watson and Spence (2007) [13] explored how emotions affect consumer behavior. Anninou (2018) [14] further explained how different emotional appraisals impact decision-making. This theory is crucial for understanding how emotional expressions influence the perceived believability of online rumors, particularly on social platforms. Since users frequently engage with emotional content, the theory offers insight into how arousal, valence, and social contagion shape message believability and behavior intentions. It also provides a foundation for designing social media algorithms to enhance credible information dissemination and mitigate rumor spread.

B. Hypothesis development

Emotional arousal plays a crucial role in how people evaluate and engage with online information. Research shows that emotionally charged content is better remembered [15], more trusted [16], and spreads faster on social media [5]. The Feelings-as-Information Theory [11] suggests that heightened emotional engagement increases the significance and believability of messages. In contrast, the absence of arousal cues lowers emotional intensity, reducing cognitive engagement and the likelihood of accepting and sharing the message. This indicates that strategies aimed at reducing emotional arousal in online content could lower the perceived believability of fake news. Thus, hypothesis 1 is stated as follows:

Hypothesis 1: The absence of emotional arousal cues in online fake messages will lower their perceived message believability compared to messages containing such cues.

Emotional valence refers to the positive or negative dimension of the emotion conveyed in messages. Positive emotions generally enhance trust and acceptance of the message, while negative emotions can induce skepticism and critical evaluation [17, 18]. Additionally, Lerner and Keltner (2000) [19] demonstrated that positive emotions increase heuristic processing and trust, making individuals more receptive to the message content, while negative emotions foster more systematic and critical processing. When people encounter positive valence messages, they tend to lower their cognitive defenses, which can create a favorable cognitive bias, leading them to view the message as more credible.

Based on these findings, this study proposes that positive emotions are associated with increased trust and acceptance, thereby enhancing the perceived message believability of online rumors. Conversely, negative emotions can induce skepticism and critical evaluation, reducing the perceived message believability of these rumors. When individuals experience negative emotions, they are more likely to engage in systematic and analytical thinking, which helps them scrutinize the content more closely and identify potential inconsistencies or falsehoods. This critical evaluation process is essential for recognizing online rumors, as it promotes a more cautious and discerning approach to information processing. Therefore, we hypothesize that online rumors presented with negative emotional valence will perceived as less believable compared to those offered with positive emotional valence. Thus, Hypothesis 2 is stated as follows:

Hypothesis 2: Online fake messages with negative emotional valence will be perceived as less believable than those with positive emotional valence.

Although emotional arousal cues influence information evaluation, emotional valence may shape its effect. Research shows that arousal combined with positive valence enhances satisfaction and approach behaviors, while arousal with negative valence promotes skepticism and avoidance [20, 21]. Positive experiences with high arousal are more memorable and believable, enhancing message impact [22]. Thus, emotional valence may moderate the effect of arousal cues on the believability of online rumors. Arousal with positive valence boosts favorable evaluations, while arousal with negative valence fosters analytical thinking, helping individuals detect misinformation.

Consequently, this heightened scrutiny can reduce the likelihood of misinformation being accepted as accurate. Hence, the impact of emotional arousal cues on the perceived message believability is moderated by emotional valence. We proposed the hypothesis 3 as follows:

Hypothesis 3: Arousal cues have a weaker effect on message believability under the negative valence conditions than under the positive valence conditions.

Emotional contagion refers to the spread of emotions from one person to another, amplifying social validation and making messages seem more credible through shared emotions [23]. It occurs in both face-to-face interactions and online communication, such as social media [4]. Emotional contagion theory explains how emotions transfer among individuals, influencing group behaviors and beliefs, especially in social media settings. Social cognitive theory further suggests that observing emotional expressions can shape one’s emotions and behaviors through observational learning [24]. Research shows that emotional contagion enhances conformity, promoting message acceptance and credibility [25]. Shared emotions foster perceived consensus and social validation, reinforcing belief in the message’s validity [24]. Individuals who are more susceptible to emotional contagion tend to mimic others’ expressions, strengthening emotional responses and credibility perceptions. However, minimizing emotional contagion can help curb the spread of misinformation, as increased social attention makes fake news harder to control. Thus, Hypothesis 4 is stated as follows:

Hypothesis 4: Fake news containing emotional contagion cues will be perceived as more believable compared to fake news without such cues.

Emotional arousal shapes how people evaluate information, and emotional contagion can amplify this effect. Emotional contagion reinforces heightened emotional states caused by arousal, increasing the perceived credibility of information within a community [5]. Research shows that collective emotions on social media influence public opinion and behavior, enhancing the believability of emotionally charged messages [26]. When content aligns with the audience's feelings, it is more likely to be shared and trusted [27]. Hatfield, Cacioppo, and Rapson (1993) [23] found that emotional contagion amplifies individual emotions, strengthening collective responses and boosting message credibility. This feedback loop intensifies arousal, making messages more compelling and memorable. However, reducing arousal and contagion cues can help curb the spread of misinformation by limiting emotional resonance. Online rumors and fake news gain intensity through social attention and the effects of emotional contagion. This heightened emotional arousal exacerbates their dramatization, making the spread of such misinformation increasingly challenging to control. Therefore, this study suggests that the impact of emotional arousal cues on perceived message believability is less when emotional contagion cues are absent. Thus, Hypothesis 5 is stated as follows:

Hypothesis 5: The impact of arousal cues on perceived message believability is reduced in the absence of emotional contagion cues compared to when such cues are present.

Drawing from the Feelings-as-Information Theory and previous literature, the degree to which individuals trust the information plays a critical role in shaping their actions, such as whether they share it on social platforms. When a message is recognized as unreliable or dubious, individuals are less motivated to share it to avoid spreading misinformation or protect their reputation within their online communities. On the other hand, when a message is perceived as more believable, users feel validated and socially responsible, which can encourage sharing and disseminating what they believe is accurate information. This highlights the importance of message credibility in driving engagement and contributes to understanding how misinformation spreads online. Thus, Hypothesis 6 is stated as follows:

Hypothesis 6: The lower the perceived message believability, the lower the behavioral intention to share the message on social media.

III. RESEARCH METHOD

A. Task 1: Preparation and Design of Experimental Materials

- Collection of Online Fake News:

The objective is to gather a comprehensive list of significant domestic and international online rumors from 2024. The selection criteria for these rumors will include factors such as the number of shares, mentions, and public impacts. This will involve monitoring popular social media platforms, news websites, and fact-checking organizations. Additionally, a survey will be conducted with approximately 3-40 online readers to assess their awareness and familiarity with these rumors. Based on the survey results, the most recognized and impactful rumors will be selected for the experimental design.

- Selection of Online Platforms:

The objective is to identify the most popular online platforms for spreading and encountering online rumors. A preliminary survey will be conducted to determine current platform preferences among a representative sample of internet users (e.g., Facebook, Twitter, Instagram, Reddit, Line, WeChat, etc.). The platforms with the highest usage and relevance will then be selected to host the experiment design.

- Pre-testing of Multiple Independent Variables:

The objective is to ensure a clear definition and effective manipulation of independent variables for the experiment. Pre-tests will be conducted to distinguish between the presence and absence of arousal cues, the presentation of positive versus negative emotional valence, and with versus without social contagion cues. Feedback from these pre-tests will refine the experimental materials and ensure that participants can effectively recognize and respond to the different conditions.

- Collection and Development of Measurement Items:

The objective is to create reliable and valid measurement items for assessing the impact of emotional dimensions on message believability. Existing measurement items from relevant literature will be adapted, focusing particularly on emotional responses and perceived believability. All items will be rated on a 7-point Likert scale. A pre-test involving experts in human-computer interactions and electronic commerce will be conducted to validate these measurement items.

The measurements of perceived message believability are adopted from Appelman and Sundar (2016) [28]. Participants were asked: "To what extent do you agree that the content you just read is accurate?", "To what extent do you agree that the content you just read is authentic?", "To what extent do you believe the content you just read is believable?". In order to measure the participants' behavior intention, we will adopt a 3-item scale derived from Lee and Ma (2012) [29]. Participants indicated whether they agree "I intend to share the information on social media," "I expect to share the information on social media," and "I plan to share the information. The questionnaires will be revised based on expert feedback and translation checks to enhance reliability and validity.

B. Task 2: Examine the Causal Effects of Message Believability

- Experimental Setup:

The second task is to examine the causal effects of message believability. We will conduct a between-subject design with a 2 (arousal cues: present vs. absent) x 2 (emotional valence: positive vs. negative) x 2 (emotional contagion cues: with vs. without) factorial design. This setup will create eight different experimental conditions to investigate the main and interaction effects of these emotional dimensions on the perceived believability of online rumors.

- Participants:

We will start by posting announcements and advertisements in university group chats and forums to recruit participants. The subjects should be representative of e-commerce users because the largest population of online users is 20 to 40 years old. Interested participants will register for the experiment by signing up through a designated online registration form. Before the experiment, participants will be asked to confirm their attendance and receive reminder emails about their scheduled session.

We will recruit 320-350 subjects to participate in the experiment. To ensure that the sample is representative of the broader online population, we will strive for demographic diversity in terms of age, gender, education level, and internet usage habits. Participants will be randomly assigned to one of the eight experimental conditions.

- The Procedure of Online Experiment

Pre-experimental Briefing: Participants will receive an email or notification with detailed instructions about the study. The briefing will include an introduction to the study's purpose and the task they will be performing. Participants will be informed that they will be evaluating different online messages to understand their perceived message believability. They will also be notified that their participation is voluntary and that their responses will be kept confidential.

Experiment: Participants will be assigned to one of the eight experimental conditions in a randomized manner. They will be instructed to carefully browse the online messages presented to them, which have been tailored according to their assigned condition (varying by emotional arousal cues, emotional valence, and social contagion cues). After browsing, participants will complete an online questionnaire designed to measure their emotional responses and perceived believability of the messages and behavior intention.

Debriefing: Upon completing the questionnaire, participants will receive a debriefing message explaining the true purpose of the study and emphasizing the importance of critically evaluating online information. They will be instructed not to discuss the experiment with others to maintain the integrity of the study. Participants will be thanked for their participation and provided with a digital participation gift or the chance to enter a raffle for a prize.

- Data Analysis:

Before analyzing the data, we will first check the quality of the questionnaire responses. This involves identifying and excluding invalid responses, such as those with excessively short or long completion times, incomplete answers, or answers that show a lack of attention or bias. After ensuring the data quality by filtering out these invalid responses, we will proceed with the statistical analysis. ANOVA will be used to examine manipulation checks, as well as the main effects and interaction effects of arousal cues, emotional valence, and emotional contagion cues on perceived believability. Additionally, moderation analyses will investigate the role of emotional valence and contagion cues in moderating the relationship between arousal cues and perceived believability. To assess the reliability of the dependent variable, message believability, we will calculate Cronbach's alpha to ensure the internal consistency of the measurement items. By conducting these analyses, we aim to understand how different emotional factors influence the believability of online messages, providing insights that can inform strategies to counteract the spread of misinformation.

IV. CONCLUSION

This study aims to contribute to the understanding of how emotional factors influence the believability and spread of online rumors. By examining emotional arousal, valence, and contagion simultaneously, it provides insights into their individual and combined effects. The findings will enhance knowledge of how emotions shape cognitive evaluations and drive misinformation. Additionally, the study offers practical implications for designing strategies to mitigate the spread of fake news by identifying key emotional triggers that affect believability. This research will inform future efforts in media literacy and platform policies to better address the challenges posed by emotionally driven misinformation.

ACKNOWLEDGMENT

This work was supported by the National Science and Technology Council of the Republic of China under Contract NSTC 113-2410-H-259-062-MY2.

REFERENCES

[1] B. Zhong, "Going beyond fact-checking to fight health misinformation: A multi-level analysis of the Twitter response to health news stories," *International Journal of Information Management,* vol. 70, pp. 102626, 2023.

[2] E. Aïmeur, S. Amri, and G. Brassard, "Fake news, disinformation and misinformation in social media: a review," *Social Network Analysis and Mining,* vol. 13, no. 1, pp. 30, 2023.

[3] H. Allcott and M. Gentzkow, "Social Media and Fake News in the 2016 Election," *Journal of Economic Perspectives,* vol. 31, no. 2, pp. 211-236, 2017.

[4] R. Lin, and S. Utz, "The emotional responses of browsing Facebook: Happiness, envy, and the role of tie strength," *Computers in human behavior,* vol. 52, pp. 29-38, 2015.

[5] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," *science,* vol. 359, no. 6380, pp. 1146-1151, 2018.

[6] Y. Chen, N. J. Conroy, and V. L. Rubin, "Misleading online content: recognizing clickbait as" false news"," in Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection, 2015, pp. pages 15-19.

[7] M. Charquero-Ballester, J. G. Walter, I. A. Nissen, and A. Bechmann, "Different types of COVID-19 misinformation have different emotional valence on Twitter," *Big Data & Society,* vol. 8, no. 2, pp. 20539517211041279, 2021.

[8] D. Yin, S. D. Bond, and H. Zhang, "Keep your cool or let it out: Nonlinear effects of expressed arousal on perceptions of consumer reviews," *Journal of Marketing Research,* vol. 54, no. 3, pp. 447-463, 2017.

[9] Y. Ye, and K. Motoki, “Effects of emotions on the helpfulness of online reviews of healthy and unhealthy food in Japan and the United States,” *Journal of Consumer Behaviour*, 2024.
[10] N. Schwarz, and G. L. Clore, “Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states,” *Journal of personality and social psychology,* vol. 45, no. 3, pp. 513, 1983.
[11] N. Schwarz, “Feelings-as-information theory,” *Handbook of theories of social psychology,* vol. 1, pp. 289-308, 2012.
[12] J. P. Forgas, “Mood and judgment: the affect infusion model (AIM),” *Psychological bulletin,* vol. 117, no. 1, pp. 39, 1995.
[13] L. Watson, and M. T. Spence, “Causes and consequences of emotions on consumer behaviour: A review and integrative cognitive appraisal theory,” *European Journal of Marketing,* vol. 41, no. 5/6, pp. 487-511, 2007.
[14] I. Anninou, “The application of appraisal theories of emotions in the case of consumer confusion: theoretical and practical implications,” *Management Decision,* vol. 56, no. 8, pp. 1696-1715, 2018.
[15] A. Lang, K. Dhillon, and Q. Dong, “The effects of emotional arousal and valence on television viewers’ cognitive capacity and memory,” *Journal of Broadcasting & Electronic Media,* vol. 39, no. 3, pp. 313-327, 1995.
[16] K. Kircanski, N. Notthoff, M. DeLiema, G. R. Samanez-Larkin, D. Shadel, G. Mottola, L. L. Carstensen, and I. H. Gotlib, “Emotional arousal may increase susceptibility to fraud in older and younger adults,” *Psychology and aging,* vol. 33, no. 2, pp. 325, 2018.
[17] J. A. G. Folse, M. Porter III, M. B. Godbole, and K. E. Reynolds, “The effects of negatively valenced emotional expressions in online reviews on the reviewer, the review, and the product,” *Psychology & Marketing,* vol. 33, no. 9, pp. 747-760, 2016.
[18] J. Kim, and P. Gupta, “Emotional expressions in online user reviews: How they influence consumers' product evaluations,” *Journal of Business Research,* vol. 65, no. 7, pp. 985-992, 2012.
[19] J. S. Lerner, and D. Keltner, “Beyond valence: Toward a model of emotion-specific influences on judgement and choice,” *Cognition & emotion,* vol. 14, no. 4, pp. 473-493, 2000.
[20] A. W. Brooks, and M. E. Schweitzer, “Can Nervous Nelly negotiate? How anxiety causes negotiators to make low first offers, exit early, and earn less profit,” *Organizational Behavior and Human Decision Processes,* vol. 115, no. 1, pp. 43-54, 2011.
[21] E. Harmon-Jones, T. F. Price, and P. A. Gable, “The influence of affective states on cognitive broadening/narrowing: Considering the importance of motivational intensity,” *Social and Personality Psychology Compass,* vol. 6, no. 4, pp. 314-327, 2012.
[22] L. A. Cavanaugh, J. R. Bettman, and M. F. Luce, “Feeling love and doing more for distant others: Specific positive emotions differentially affect prosocial consumption,” *Journal of Marketing Research,* vol. 52, no. 5, pp. 657-673, 2015.
[23] E. Hatfield, J. T. Cacioppo, and R. L. Rapson, “Emotional contagion,” *Current directions in psychological science,* vol. 2, no. 3, pp. 96-100, 1993.
[24] A. Bandura, “Social foundations of thought and action,” *Englewood Cliffs, NJ,* vol. 1986, no. 23-28, pp. 2, 1986.
[25] K. Peters, Y. Kashima, and A. Clark, “Talking about others: Emotionality and the dissemination of social information,” *European Journal of Social Psychology,* vol. 39, no. 2, pp. 207-222, 2009.
[26] A. Goldenberg, D. Garcia, E. Halperin, and J. J. Gross, “Collective emotions,” *Current directions in psychological science,* vol. 29, no. 2, pp. 154-160, 2020.
[27] R. E. Guadagno, D. M. Rempala, S. Murphy, and B. M. Okdie, “What makes a video go viral? An analysis of emotional contagion and Internet memes,” *Computers in human behavior,* vol. 29, no. 6, pp. 2312-2319, 2013.
[28] A. Appelman, and S. S. Sundar, “Measuring message credibility: Construction and validation of an exclusive scale,” *Journalism & Mass Communication Quarterly,* vol. 93, no. 1, pp. 59-79, 2016.
[29] C. S. Lee, and L. Ma, “News sharing in social media: The effect of gratifications and prior experience,” *Computers in human behavior,* vol. 28, no. 2, pp. 331-339, 2012.

 ISBN: 978-1-68558-238-8

Resilience and Node Impact Assessment in YouTube Commenter Networks Leveraging Focal Structure Analysis

Md Monoarul Islam Bhuiyan*, Shadi Shajari*, Nitin Agarwal*†
*COSMOS Research Center, University of Arkansas at Little Rock, Arkansas, USA
†International Computer Science Institute, University of California, Berkeley, California, USA
e-mail: {mbhuiyan, sshajari, nxagarwal}@ualr.edu

***Abstract*—Communication networks play a pivotal role in shaping information dissemination across social media platforms. Identifying influential groups or key players within these networks is essential for understanding how information flows and spreads. YouTube, as the leading video-sharing platform, offers a vast and dynamic environment for such studies. Our extended research centers on Focal Structure Analysis (FSA), aiming to identify core commenter groups within 35 YouTube channels discussing the Indo-Pacific region. By analyzing a dataset containing 308,890 videos, 726,078 commenters, and 1,536,284 comments, we apply two distinct FSA methods, namely FSA 1.0 and FSA 2.0, to detect influential network structures. We further evaluate the impact of these structures using network resilience metrics, including flow robustness and the giant component ratio. Our findings indicate that removing key focal structures results in a more fragmented and sparse network, significantly impairing information flow. This suggests that these core commenter groups act as critical bridges, facilitating communication and enhancing the cohesion of the network. By extending our prior work, this study offers deeper insights into the mechanisms of information spread on YouTube, providing a more comprehensive understanding of the platform's commenter dynamics.**

Keywords-Focal Structure Analysis; Social Network Analysis; YouTube; Network Resiliency.

I. INTRODUCTION

With the rise of social media platforms and their sophisticated recommendation algorithms, several aspects including content creation and sharing, news consumption, community engagement, societal influence, narrative propagation [1] and many other activities have gained wide popularity. This rapid adoption has become possible due to massive user engagement over content, driven by semiotics [2]. Every day, enormous amounts of information are generated through these platforms. While this rapid growth plays a pivotal role in the data sources for researchers, it is also crucial to find the best actionable knowledge from these data sources. Additionally, extracting actionable insights has widely been researched through the topology of complex social networks. As of 2024, YouTube is the second-most popular social media platform, the number one video-sharing platform globally, and available in over 100 countries and 80 languages its prominence in its user base has become streamlined due to its users massive engagement (views, comments, likes, shares, subscriptions, etc.) over the actual content [3]. Among these, YouTube's comment section provides a platform for constructive discourse, enabling viewers to share insights and directly connect with content creators. Despite this, the public discussion space can often lead to negativity and unproductive comments, which in turn can impair the user experience.

This paper conducts a comparative analysis of two distinct versions of Focal Structure Analysis namely FSA 1.0 [4] and FSA 2.0 [5], which is a social network analysis methodology designed to identify core sets of commenter groups within the co-commenter network of YouTube channels [6]. Initially, this study compares the outcomes of these two FSA approaches to evaluate their effectiveness in extracting relevant focal structures. Following this comparison, it addresses two research questions:

- **RQ1:** How do focal structures within a complex social network impact its resilience, as measured by network resilience metrics?
- **RQ2:** How much does each node in a particular focal structure contribute to the overall robustness?

By exploring the significance of these core groups and their impact on network resilience, this study aims to provide insights into the structural dynamics and robustness of social networks.

The rest of the paper is organized as follows. Section II reviews existing studies on identifying focal structures, detecting authoritative and community approaches, and measuring network resiliency metrics. Section III outlines the methods used for collecting data in this study. Section IV describes the experimental methodologies applied, while Section V presents the findings of our research. Finally, Section VI summarizes the study and suggests directions for future research.

II. RELATED WORK

This section is divided into two parts. The first part discusses the relevant literature related to identifying important nodes in the social network, and the second part covers the metrics available for measuring network resiliency.

A. Identifying Important Structures

Identifying key individuals who are best connected or most influential in a social network is crucial for extracting actionable knowledge. Consequently, various methods have been proposed to identify these key nodes. While Hyperlink-Induced Topic Search (HITS) determines hubs and authorities [7], PageRank assigns a numerical weight for each node in the network [8]. Both of these approaches can be used to identify influential nodes. On the other hand, identifying the communities [9] and clusters from a social network perspective has also been widely studied. Generally, in a community, similar nodes are more clustered together than nodes that do not share commonalities. Previous researchers have also worked on a more sophisticated approach where their focus shifted from identifying the influential nodes or communities to detecting smaller key sets of players who maximized the information diffusion. The authors in [4] devised a methodology where they identified focal patterns leveraging the Louvain method that gave them more relevant information about the network than obtained from the influential nodes [9]. When applying this method to large biological networks, they found more prominent, smaller, and relevant structures in protein-protein interaction networks [10]. An online analysis and visualization-based tool has also been built for the ease of analyzing these small and pertinent focused structures [11]. Since this method could not extract structures with lower connection density, researchers extended their approach by combining highly connected candidate focal structures based on similarity values. This allowed the identification of both cliquish and small sparse, yet connected, structures [12]. An advanced version of this approach was proposed by [13], where the authors combined user-level centrality and group-level modularity methods to create a bi-level maximization network model that overcame the shortcomings of the previously described focal structures analysis methods.

B. Network Resiliency Metrics

Network resilience, like influential node and community identification, is crucial in Social Network Analysis (SNA), denoting a network's ability to withstand disruptions while maintaining core functions. The study by Bertoni et al. [14] employs social network analysis to identify key contributors to resilience in an intensive care unit, integrating SNA-derived indicators with non-network attributes, whereas another research comprehensively reviews resilience functions and regime shifts in complex systems across various domains through empirical observations, experimental studies, and theoretical analysis [15]. Several metrics have also been developed to quantify network resilience in the face of disruptions, such as flow robustness [16], and giant component ratio [17].

However, a key gap exists in current research. While these metrics effectively measure network resilience, they have not been extensively applied to the context of social networks like YouTube. Our work aims to bridge this gap by incorporating network resilience approaches into the analysis of social networks, offering a more comprehensive understanding of their ability to adapt and function under various stresses.

III. Data Collection

The data for this study was collected using a specialized tool designed to collect information from YouTube through its API [18]. The collection process involved retrieving videos, comments, and channel data based on specific keywords. These keywords were selected through a thoughtful process that involved reviewing commonly used terms and phrases relevant to discussions in the Indo-Pacific region. While no formal methodology was employed, the selection was guided by careful consideration of the linguistic and cultural context to ensure the keywords captured a broad range of relevant topics. Examples of keywords used include "Komunis Cina | China pengaruh Indonesia", "Muhammadiyah Cina | China | Tiongkok | Tionghoa", "Kejam Uighur | Uyghur", and "Muslim Brother | Indonesia Uighur | Uyghur". The final dataset comprised 35 YouTube channels, 308,890 videos, 726,078 commenters, and 1,536,284 comments.

IV. Methodology

This section outlines the methods used in our study. First, it details the creation of a co-commenter network and introduces focal structure analysis alongside the problem statement. Finally, it concludes by explaining various network resiliency metrics.

A. Co-commenter Network Creation

The analysis started with creating co-commenter networks for each YouTube channel. These networks connect commenters who have commented on the same video across one or more channels, as described in [19]. The edges between commenters are weighted based on the number of shared videos they have commented on. Only commenters who have engaged with at least 5 videos are included in the network, as this threshold ensures the analysis focuses on active and consistent users, minimizing noise from sporadic commenters, as shown in Figure 1.

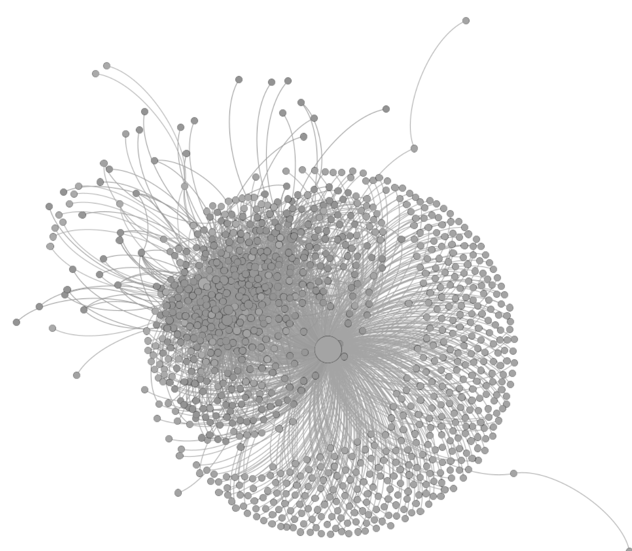


Figure 1. A YouTube co-commenter network where nodes represent commenters, and edges indicate shared commenting on the same video for the channel with ID `"UCfWNZIJkm268rLO_yeRlcww"`.

B. Focal Structure Analysis

Focal Structure Analysis (FSA) is a social network analysis method that aims to find key sets of individuals rather than a set of key individuals within a social network. FSA aims to extract minimal influential groups in a network, thereby enhancing the knowledge discovery process. The earliest version of FSA (i.e., FSA 1.0) utilizes global and local interconnectedness-based algorithms to identify focal patterns [4]. After partitioning the network into focal structures, FSA 1.0 stitches interconnected structures using Jaccard's Coefficient [12]. FSA 1.0 groups nodes with similar clustering coefficients into focal structures, collectively identifying core-focused groups distinct from traditional community detection methods [9]. However, the current version of FSA 2.0 employs a bi-level maximization network model to identify authoritative individuals and cohesive communities within the network [13]. This analysis identifies key sets of influential commenters by leveraging degree centrality and clustering coefficient methods at the commenter level, and spectral modularity at the group level. Removing these focused core groups from the network may disrupt information flow or break down important connections, potentially compromising the overall effectiveness of its structure. Figures 2 and 3 illustrate the key focal structures identified through FSA 1.0 and FSA 2.0, respectively.

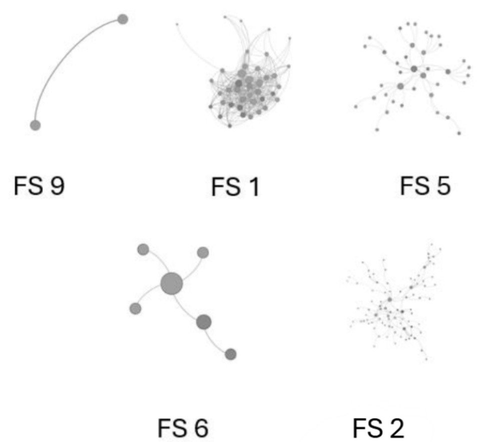


Figure 2. Several prominent Focal Structures (FS) detected using Focal Structure Analysis 1.0 from the channel with ID "UCfWNZIJkm268rLO_yeRlcww".

C. Problem Statement

A Focal Structure (FS) is a key set of individuals who may be responsible for organizing information diffusion. A focal structure contains a set of vertices (at least two) and edge(s). These individuals from the focused core groups may not be the most influential on their own but by interacting together form a compelling power. Consider a social network $G = (V, E)$, where V is the set of vertices and E is the set of edges,

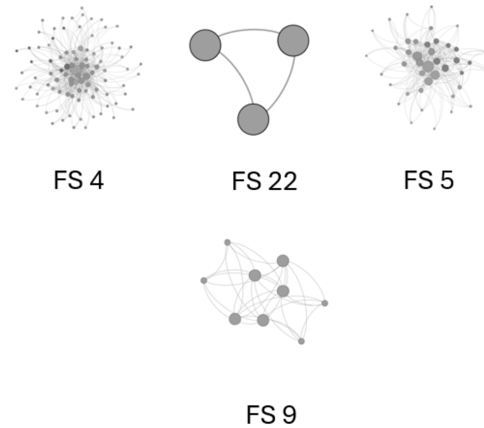


Figure 3. Notable Focal Structures (FS) uncovered through Focal Structure Analysis 2.0 from the channel with ID "UCfWNZIJkm268rLO_yeRlcww".

where a focal structure can formally be defined as follows: Focal structures in G are defined by $F = \{G'\}$, where $G' = (V', E')$ and $V' \subseteq V$ and $E' \subseteq E$. For all i and j, $i \neq j$, $G_i \in F$ and $G_j \in F$, such that no two focal structures can subsume each other, or $G_i \not\subset G_j$ and $G_j \not\subset G_i$.

D. Network Resiliency Metrics

This section describes metrics used to quantify network resilience.

*1) **Flow Robustness**:* Flow robustness serves as an imperative graph metric, quantifying the resilience of a network by evaluating the proportion of reliable flows against the total flow count [16]. A flow is called reliable if it maintains at least one uninterrupted path despite potential link or node failures. It offers insight into the network's capacity to sustain communication between nodes following the removal of nodes. Flow robustness values range between 0 and 1, with 1 denoting seamless communication across all nodes and 0 indicating a lack of inter-nodal communication, indicative of a network devoid of connections. The flow robustness (FR) of a graph $G(V, E)$ is computed using:

$$FR(G) = \frac{\sum_{i=1}^{n} |C_i|(|C_i| - 1)}{|n|(|n| - 1)}, \qquad 0 \leq FR \leq 1 \quad (1)$$

*2) **Giant Component Ratio**:* The Giant Component Ratio (GCR) is a key metric in network resilience analysis, measuring the ratio of nodes within the Largest Connected Component (LCC) to the total number of nodes in the network. It is computed using:

$$GCR(G) = \frac{N_{LCC}}{N}, \qquad 0 < GCR < 1 \qquad (2)$$

where the N_{LCC} represents the number of nodes in the largest connected component and N denotes the total number of nodes in the network. This metric also serves as a critical indicator of a network's ability to maintain structural cohesion and connectivity upon the removal of focused core groups.

Copyright (c) IARIA, 2025. ISBN: 978-1-68558-238-8

*3) **Isolated Nodes and Cluster Analysis**:* The impact of the commenter's removal from the communication network will also be evaluated through two metrics. While isolated node count measures network fragmentation that may hinder information flow, cluster analysis is performed to identify potential community fracturing and its impact on network cohesion and dynamics.

E. Spearman's rank correlation coefficient

We use Spearman's rank correlation coefficient [20], implemented to evaluate the monotonic relationship between variables without assuming linearity. The method ranks the data, assigns tied values their average rank, and computes the correlation coefficient ρ as:

$$\rho = 1 - \frac{6\sum_{i=1}^{n}(R(X_i) - R(Y_i))^2}{n(n^2-1)},$$

where $R(X_i)$ and $R(Y_i)$ are the ranks of observations in X and Y, d_i is the rank difference, and n is the number of pairs. In our analysis, the correlation between flow robustness and the giant component ratio was found to be 0.92, indicating a strong positive monotonic relationship. Given this high correlation, we selected flow robustness as the primary metric to assess the impact of nodes in each focal structure while also focusing on how the removal of a specific focal structure affects network resiliency.

V. RESULTS

This section evaluates the impact and resiliency of focal structures identified by FSA 1.0 and 2.0. We assess the influence of these structures on key metrics like flow robustness, giant component ratio and compare their performance against standard methods, such as PageRank and Louvain community detection. While both FSA 1.0 and FSA 2.0 produce key focal structures, the impact of the focal structures identified by FSA 2.0 is more prominent. Additionally, FSA 2.0 generates a greater number of such impactful structures compared to FSA 1.0.

A. Node Impact Assessment

Our study assesses the impact of each focal structure through the nodes associated with it. At first, we employed the provisional removal of each focal structure from the network and observed changes in Flow Robustness (FR) and the Giant Component Ratio (GCR). Given the strong correlation (0.92) between GCR and FR, we chose to focus on the flow robustness metric to simplify the analysis. After that, we calculated the impact score by dividing the complement of FR by the number of nodes in each focal structure. This approach allowed us to rank focal structures based on the impact of nodes within each focal structure.

Our findings reveal a noteworthy outcome where the focal structures identified by FSA 2.0 demonstrated a higher overall impact than those identified by FSA 1.0, indicating the more significant influence of nodes within these structures. Overall, these differences highlight the varying capabilities of FSA 1.0 and FSA 2.0 in revealing critical focal structures, with FSA 2.0 offering a more extensive and impactful identification of key groups within the network, as demonstrated in Figures 4 and 5.

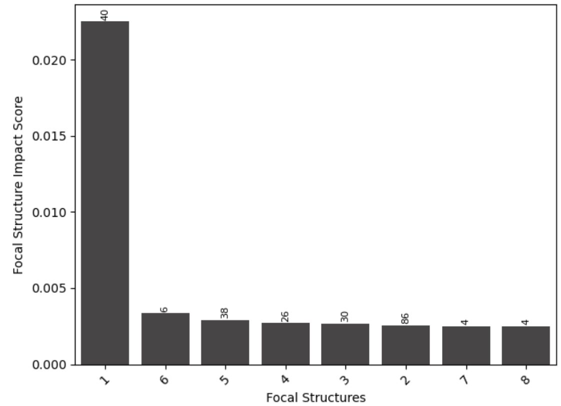


Figure 4. The calculated impact scores for focal structures identified by FSA 1.0 show the relative influence of each structure in maintaining network robustness.

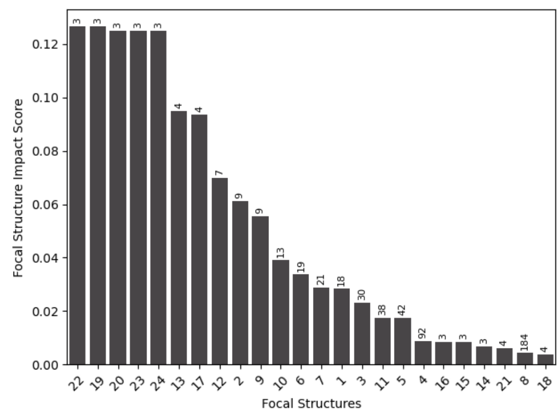


Figure 5. The impact scores of focal structures identified by FSA 2.0 demonstrate the significant role of individual nodes in affecting network resilience.

B. Network Resiliency Assessment

In this study, we also assessed the impact of focal structures identified by FSA 2.0 on network resiliency using flow robustness and Giant Component Ratio (GCR). For comparison, we evaluated the resiliency of structures detected by PageRank, the Louvain community detection algorithm, and FSA 2.0. FSA 2.0 identified 24 focal structures for our YouTube co-commenter network, compared to 7 detected by the Louvain community detection algorithm, with the top 24 influential nodes from PageRank also included for visualization purposes. Our analysis revealed that while larger community-based structures, such as those identified by the Louvain algorithm, contained more nodes, they did not exhibit the same impact on the network as the focal structures identified by FSA 2.0. The focal structures from FSA 2.0 consistently outperformed the community structures regarding flow robustness and GCR. Figure 6 reveals that focal structures from FSA 2.0 consistently result in a greater reduction in flow robustness, highlighting their critical role in maintaining information flow, whereas community and PageRank nodes exhibit comparatively lower impact. Furthermore, Figure 7 shows a pronounced decrease in GCR upon the removal of FSA 2.0 focal structures, underscoring their significant influence in sustaining the largest connected component, while community and PageRank nodes exhibit less disruptive effects. This finding underscores the unique contribution of focal structures built on individual and group-based node features. In addition to that, focal structures play a crucial role in bridging communities and maintaining overall network connectivity regardless of their size.

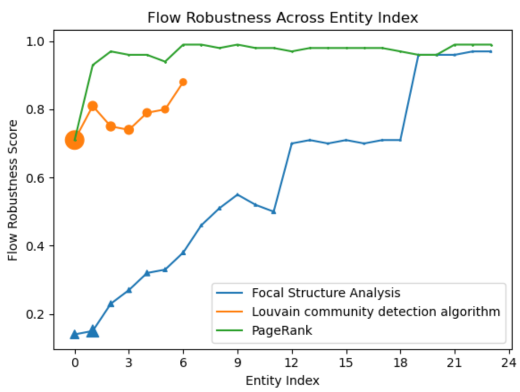


Figure 6. Comparison of the impact on network flow robustness when removing key structures identified by FSA 2.0, Louvain community detection, and PageRank.

The network's modularity increased when we removed the impactful focal structures that FSA 2.0 had found. This, in turn, indicates that these structures are essential to maintaining the information flow across communities. On the other hand, the removal of the larger community structures did not have the same impact. Their function as crucial gatekeepers in the spread of information is further highlighted by the network fragmentation brought about by the removal of smaller, well-positioned comments from FSA 2.0 networks.

This comparison demonstrates how much better FSA 2.0 is at locating critical structures that have a big impact on network resilience. The focal structures identified by FSA 2.0 continuously shown noticeable influence on the network, showing their crucial role in preserving information flow and network cohesion, even if community-based structures had a larger number of nodes.

Lastly, when provisionally removing focal structures identified by FSA 2.0, it caused considerable fragmentation within the network, leading to the isolation of nodes from the overall network and the formation of numerous clusters. For instance, the removal of one focal structure resulted in 611 clusters and 605 isolated nodes, while even a focal structure containing only 3 nodes was able to isolate 431 nodes, as shown in Table 1. These findings underscore the imperative influence of focal

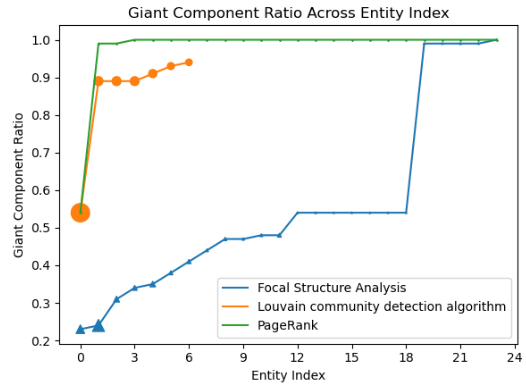


Figure 7. Evaluation of network resilience through changes in the Giant Component Ratio (GCR) following the removal of nodes detected by FSA 2.0, Louvain community detection, and PageRank.

TABLE I. IMPACT OF REMOVING PROMINENT FOCAL STRUCTURES IDENTIFIED BY FSA 2.0 ON NETWORK FRAGMENTATION.

Focal Structure	Nodes	No. of Clusters	Isolated Nodes
5	42	611	605
9	9	486	483
22	3	434	431

structures in the network, revealing how their removal can disproportionately disrupt connectivity and lead to significant fragmentation, even when the focal structure itself is relatively small. As a result, focal structures play a pivotal role in preserving network connectivity and highlight their significant impact on maintaining overall network cohesion.

VI. Conclusion and Future Work

In this study, we first obtained the focal structures of YouTube co-commenter's network by leveraging two distinct versions of focal structures analysis, FSA 1.0 and FSA 2.0. Furthermore, through various network resiliency metrics, we delved deeper to assess how these focal structures were crucial to the overall success of information dissemination for the defined networks. By examining the flow robustness and giant component ratio, we demonstrated that the focal structures detected by FSA 2.0 exhibit a significantly higher impact on the network compared to those identified by both FSA 1.0 and traditional community detection algorithms. Despite the larger size of community-based structures, they failed to match the influence of smaller, strategically positioned focal structures identified by FSA 2.0. Our evaluation not only justified that removing some of the focal structures from the network made it more sparse, fragmented, and less cohesive but also that the information flow of the co-commenter network was disrupted heavily, which means that these focal structures were acting as a bridge between other commenters of the communication network.

In future, to advance our understanding of commenter networks and focal commenter groups, our research should

 ISBN: 978-1-68558-238-8

utilize the contextual representation of commenter networks by incorporating content, engagement, and other attributes. Utilizing contextualized focal structure analysis could thus help enhance the comprehensive discovery and interpretability of focal commenter structures.

ACKNOWLEDGEMENTS

This research is funded in part by the U.S. National Science Foundation (OIA-1946391, OIA-1920920), U.S. Office of the Under Secretary of Defense for Research and Engineering (FA9550-22-1-0332), U.S. Army Research Office (W911NF-23-1-0011, W911NF-24-1-0078), U.S. Office of Naval Research (N00014-21-1-2121, N00014-21-1-2765, N00014-22-1-2318), U.S. Air Force Research Laboratory, U.S. Defense Advanced Research Projects Agency, Arkansas Research Alliance, the Jerry L. Maulden/Entergy Endowment at the University of Arkansas at Little Rock, and the Australian Department of Defense Strategic Policy Grants Program. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organizations. The researchers gratefully acknowledge the support.

REFERENCES

[1] M. I. Gurung, M. M. I. Bhuiyan, A. Al-Taweel, and N. Agarwal, "Decoding youtube's recommendation system: A comparative study of metadata and gpt-4 extracted narratives," in *Companion Proceedings of the ACM on Web Conference 2024*, 2024, pp. 1468–1472.

[2] M. I. Gurung, N. Agarwal, and M. M. I. Bhuiyan, "How does semiotics influence social media engagement in information campaigns?," 2025.

[3] S. Shajari, M. Alassad, and N. Agarwal, "Characterizing suspicious commenter behaviors," in *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining*, 2023, pp. 631–635. DOI: 10.1145/3625007.3627309.

[4] F. Sen, R. T. Wigand, N. Agarwal, D. Mahata, and H. Bisgin, "Identifying focal patterns in social networks," in *2012 Fourth International Conference on Computational Aspects of Social Networks (CASoN)*, IEEE, 2012, pp. 105–108.

[5] M. Alassad, M. N. Hussain, and N. Agarwal, "Finding fake news key spreaders in complex social networks by using bi-level decomposition optimization method," in *Modeling and Simulation of Social-Behavioral Phenomena in Creative Societies: First International EURO Mini Conference, MSBC 2019, Vilnius, Lithuania, September 18–20, 2019, Proceedings 1*, Springer, 2019, pp. 41–54.

[6] S. Shajari, R. Amure, and N. Agarwal, "Analyzing anomalous engagement and commenter behavior on youtube," *AMCIS 2024 Proceedings. 6*, 2024.

[7] J. M. Kleinberg, "Authoritative sources in a hyperlinked environment," *Journal of the ACM (JACM)*, vol. 46, no. 5, pp. 604–632, 1999.

[8] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," *Computer networks and ISDN systems*, vol. 30, no. 1-7, pp. 107–117, 1998.

[9] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of statistical mechanics: theory and experiment*, vol. 2008, no. 10, P10008, 2008.

[10] F. Şen, R. T. Wigand, N. Agarwal, M. Mete, and R. Kasprzyk, "Focal structure analysis in large biological networks," in *IPCBEE, ser. 1, vol 70, IACSIT Press.*, 2014.

[11] F. Sen, N. Nagisetty, T. Viangteeravat, and N. Agarwal, "An online platform for focal structures analysis-analyzing smaller and more pertinent groups using a web tool," in *2015 AAAI Spring Symposium Series*, 2015.

[12] F. Şen, R. Wigand, N. Agarwal, S. Tokdemir, and R. Kasprzyk, "Focal structures analysis: Identifying influential sets of individuals in a social network," *Social Network Analysis and Mining*, vol. 6, pp. 1–22, 2016.

[13] M. Alassad, M. N. Hussain, and N. Agarwal, "Comprehensive decomposition optimization method for locating key sets of commenters spreading conspiracy theory in complex social networks," *Central European Journal of Operations Research*, vol. 30, no. 1, pp. 367–394, 2022.

[14] V. B. Bertoni, T. A. Saurin, and F. S. Fogliatto, "How to identify key players that contribute to resilient performance: A social network analysis perspective," *Safety science*, vol. 148, p. 105 648, 2022.

[15] X. Liu *et al.*, "Network resilience," *Physics Reports*, vol. 971, pp. 1–108, 2022.

[16] M. J. Alenazi and J. P. Sterbenz, "Comprehensive comparison and accuracy of graph metrics in predicting network resilience," in *2015 11th international conference on the design of reliable communication networks (DRCN)*, IEEE, 2015, pp. 157–164.

[17] B. Bollobás and B. Bollobás, *Random graphs*. Springer, 1998.

[18] O. Adeliyi, I. Solaiman, S. Shajari, U. Onyepunuka, and N. Agarwal, "Detecting and characterizing inorganic user engagement on youtube," in *Workshop Proceedings of the 18th International AAAI Conference on Web and Social Media*, 2024. DOI: 10.36190/2024.01.

[19] S. Shajari, N. Agarwal, and M. Alassad, "Commenter behavior characterization on youtube channels," in *Proceedings of the eKNOW International Conference on Information, Process, and Knowledge Management, Venice Italy*, 2023, pp. 59–64.

[20] C. Spearman, "The proof and measurement of association between two things," *The American journal of psychology*, vol. 100, no. 3/4, pp. 441–471, 1987.

 ISBN: 978-1-68558-238-8

An Investigation of Inconsistent Expectations of Horse Racing Experts

Yasuhiko Watanabe, Hideaki Nakanishi, Yoshihiro Okada
Ryukoku University
Seta, Otsu, Shiga, Japan
email: watanabe@rins.ryukoku.ac.jp, t170517@mail.ryukoku.ac.jp, okada@rins.ryukoku.ac.jp

***Abstract*—In recent years, statistical researches often showed even experts can make mistakes although they have a wealth of knowledge and experience. In this study, we focus on horse racing experts, such as racing horse owners and trainers, and investigate whether they have inconsistent expectations on their professional issue. Using sire line, distance of races, and order of finish as clues, we analyze the 36922 horses registered with Japan Racing Association (JRA) from 2010 to 2017 statistically. The results of the statistical analysis showed that horse racing experts had inconsistent expectations on the problem of which race distance they thought were favorable for horses of a certain sire line. We think this is because experts' unconscious minds affected their expectations. Even for experts, it is difficult to consciously notice what they have unconsciously felt.**

Keywords*—*decision making; expert; Thoroughbred horse; sire line; race distance.

I. Introduction

Unlike most of us, experts have a wealth of knowledge and experience. However, even experts can sometimes make mistakes. For example, in the past, baseball coaches often taught players to aim for grounders rather than fly balls. However, in recent years, statistical researches brought a new batting approach that batters should aim for big fly balls rather than grounders. The new approach, known as the "fly-ball revolution", has surprised many baseball coaches and players around the world. The reason they were surprised is because they had a firm expectation on this issue and it was incorrect. The point is that they had one expectation on one issue. A question now arises whether experts have inconsistent expectations on one issue. In this study, we focus on horse racing experts, such as racing horse owners and trainers. In order to win horse races and get the prize money, they want to find races where their horses are more likely to win.

In order to analyze horse racing experts' inconsistent expectations, we focus on sire line, distance of races, and order of finish. A sire line is a term that refers to the paternal lineage or ancestry of a horse, especially a racehorse. Many people, especially horse racing experts, often say that a sire line can indicate the potential abilities or characteristics of a horse, such as which distance races they are good at.

The rest of this paper is organized as follows: In Section II, we survey the related works. In Section III, we survey information about racehorses and show how to collect it. In Section IV, we show how to analyze racehorse information statistically and discuss whether horse racing experts have inconsistent expectations on their professional issue. Finally, in Section V, we present our conclusions.

II. Related work

Thoroughbred horses originated from a small number of Arab, Barb, and Turk stallions and native British mares approximately 300 years ago [1]–[3]. Since then, they have been selectively bred to improve speed and stamina, and are consequently superior competitive racehorses. Wade et al. reported a high-quality draft sequence of the genome of the horse and suggested that the horse was domesticated from a relatively large number of females, but few males [4]. McGivney et al. reported that centuries of selection for favourable athletic traits among Thoroughbreds acts on genes with functions in behavior, musculoskeletal conformation, and metabolism [5]. Recently, some genomic regions were identified as a candidate region influencing racing performance in racehorses [6]. Many researchers applied statistical models to evaluate various parameters on racing performance in Thoroughbred horses [7]. Martin, Strand and Kearney reported that the most influential parameter was distance raced [8]. Cheetham et al. investigated whether both race earnings and number of race starts were associated with horse signalment (age, sex, and breed), gait, and race surface [9]. Wells, Randle and Williams investigated how temporal, behavior, and loading related factors associated with the period before the start of the race influences racehorse performance [10]. Statistical researches are conducted not only in horse racing but also in other sports, such as baseball. In recent years, statistical researches brought a new batting approach that batters should aim for big fly balls rather than grounders [11]. Kato and Yanai reported that Shohei Otani, the Japanese superstar slugger in Major League Baseball (MLB), always aims for hitting fly balls [12]. This new batting approach, the so-called "fly-ball revolution", shows that even experts may make mistakes. It is important to discuss how and why experts made mistakes. Yerkes and Dodson studied the relationship between arousal and performance and showed that a little stress can help we perform a task, however, too much stress degrades our performance [13]. However, experts have a wealth of knowledge and experience, and usually have staff to share their stresses and consider issues with them. Aircraft pilots are under a great deal of mental stress when they are flying their planes. Shappell and Wiegmann focused on preventing errors in aviation, including decision errors, and

 ISBN: 978-1-68558-238-8

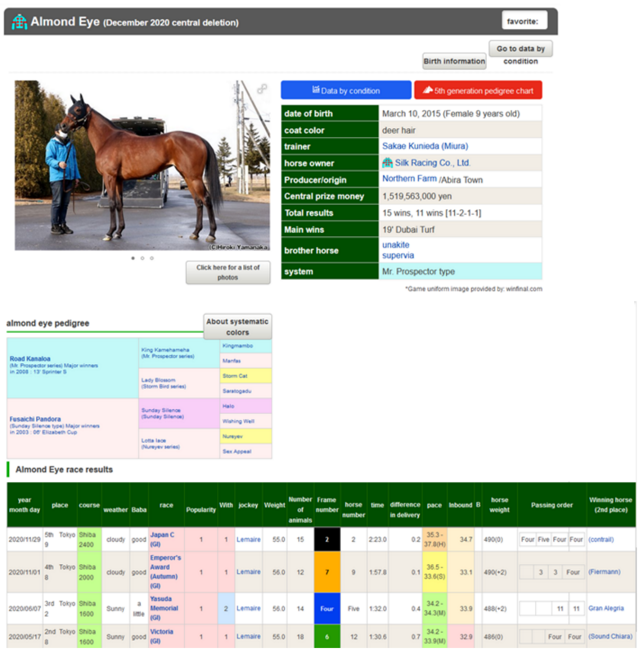


Figure 1. An example of horse information provided by Keiba Lab.

propose a framework for analyzing and classifying human errors [14]. Kang and Yoon studied the types of errors that both younger and older adults make when learning how to use new technologies [15]. They found that older adults used different strategies than younger adults. However, they did not report how experts made mistakes. Bechara et al. studied unconscious mental processing and reported unconscious minds picked up danger first [16]. However, they did not study whether unconscious minds affect experts' expectations.

III. A COLLECTION OF RACEHORSE INFORMATION

Keiba Lab [17] is one of the most popular horse racing information sites in Japan. This site records various information about all racehorses registered with Japan Racing Association (JRA) and registered users can freely access it. Figure 1 shows an example of horse information provided by Keiba Lab. As shown in Figure 1, the horse information provided by Keiba Lab consists of personal information and race results. Personal data consists of name, date of birth, age, sex, coat color, breeder, birth place, owner, trainer, ancestors up to three generations ago, sire line, career statistics, career prize money, and so on. Race results consist of venue, event date, distance, weather, racetrack, surface, race name, favourite, order of finish, jockey, weight, horse number, frame number, time, and so on. In order to discuss whether horse racing experts have inconsistent expectations on their professional issue. we collected information about 36922 horses registered with JRA from 2010 to 2017 from Keiba Lab. Table I shows the number of horses registered with JRA from 2010 to 2017.

On Keiba Lab, various sire lines are used to classify horses. We surveyed how racehorse sire lines diverged and grouped

TABLE I. THE NUMBER OF HORSES REGISTERED WITH JRA FROM 2010 TO 2017.

year	number of registered horses
2010	4470
2011	4524
2012	4505
2013	4595
2014	4672
2015	4663
2016	4738
2017	4755
Total	36922

TABLE II. THE NUMBER OF HORSES CLASSIFIED INTO THE THREE MAIN SIRE LINE TYPES.

sire line	number of horses
Native Dancer Line	8799
Nearctic Line	6383
Royal Charger Line	18104
others	3636
Total	36922

them into Native Dancer Line, Nearctic Line, Royal Charger Line, and others. For example, Figure 1 shows that the sire line of *Almond Eye* was Mr. Prospector Line. It branched out from Native Dancer Line. As a result, in this study, we determined that the sire line of *Almond Eye* was Native Dancer Line. Then, we classified 36922 horses registered with JRA from 2010 to 2017 into these four types. Table II shows the number of horses classified into these four sire line types. As shown in Table II, 90 percent of the 36922 horses were classified into the three main sire lines: Native Dancer Line, Nearctic Line, and Royal Charger Line.

36922 horses had competed in races of various distances. We grouped the race distances into five types: 1000 – 1399m, 1400 – 1799m, 1800 – 2199m, 2200 – 2799m, and more than 2800m. Then, we investigated which distance races and how many times the 36922 horses had competed in. For example, *Almond Eye* had competed in one 1000–1399m race, six 1400–1799m races, four 1800–2199m races, and four 2200 – 2799m races. Table III shows the number of times the 36922 horses of four sire lines had competed in races of various distances.

Horse owners get prize money when their horses place in the top five in races held by JRA. As a result, we investigated which distance races and how many times the 36922 horses of four sire lines had placed in the top and the top five in races held by JRA. Tables IV and V show the number of times the 36922 horses of four sire lines had placed in the top and the top five in the races of various distances, respectively.

IV. ANALYSIS OF INCONSISTENT EXPECTATIONS OF HORSE RACING EXPERTS

Horse racing experts have the problem of which distance races are favorable or unfavorable for racehorses of a certain sire line. Also, they have expectations on this problem. In this section, we investigate whether horse racing experts have inconsistent expectations on this problem.

TABLE III. THE NUMBER OF TIMES THE 36922 HORSES OF FOUR SIRE LINES HAD COMPETED IN RACES OF VARIOUS DISTANCES.

sire line	race distance 1000-1399m	1400-1799m	1800-2199m	2200-2799m	2800m-	Total
Native Dancer	27008	31619	28568	4173	2511	93879
Nearctic	18710	22444	20072	2838	1647	65711
Royal Charger	42525	67514	71758	13181	5848	200826
others	9879	12058	10780	1817	876	35410
Total	98122	133635	131178	22009	10882	395826

TABLE IV. THE NUMBER OF TIMES THE 36922 HORSES OF FOUR SIRE LINES HAD FINISHED IN FIRST PLACE IN THE RACES OF VARIOUS DISTANCES.

sire line	race distance 1000-1399m	1400-1799m	1800-2199m	2200-2799m	2800m-	Total
Native Dancer	1947	2261	2121	341	188	6858
Nearctic	1347	1511	1399	206	143	4606
Royal Charger	2580	4767	5496	1078	495	14416
others	677	855	671	105	52	2360
Total	6551	9394	9687	1730	878	28240

TABLE V. THE NUMBER OF TIMES THE 36922 HORSES OF FOUR SIRE LINES HAD FINISHED IN TOP FIVE PLACE IN THE RACES OF VARIOUS DISTANCES.

sire line	race distance 1000-1399m	1400-1799m	1800-2199m	2200-2799m	2800m-	Total
Native Dancer	9345	10912	10552	1748	1120	33677
Nearctic	6462	7700	7112	1070	728	23072
Royal Charger	13893	23937	26949	5369	2713	72861
others	3203	4054	3564	655	317	11793
Total	32903	46603	48177	8842	4878	141403

A. Basic idea

It is widely recognized that inherited variation in physical and physiological characteristics is responsible for variation in individual aptitude for race distance. Many horse racing experts would agree that if the best race distance of ancestors is known, the offspring's best race distance is most likely to take after them. As a result, we focus on three factors of racehorses:

- sire line,
- race distance, and
- order of finish.

In this section, we first investigate whether horse racing experts entered their horses of certain sire lines into races of certain distances too many times or too few times. The result of this investigation shows which distance races the experts thought were favorable or unfavorable for racehorses of a certain sire line. Then, we investigate whether horses of certain sire lines won or lost races of certain distances too many times. The result of this investigation shows which distance races were fovorable or unfavorable for racehorses of a certain sire line. Next, we investigate whether horse racing experts entered their horses into races of a certain distance too many times. The result of this investigation shows experts' judgements of horses' performance. Finally, we compare the results of statistical analyses on experts' race selection, the race results, and experts' judgements of horses' performance, and detect inconsistent expectations of horse racing experts.

B. Detection of race distance and sire line combinations that horse racing experts selected too many times or too few times

In order to detect cases where horse racing experts entered their horses of certain sire lines into races of certain distances too many times or too few times, we conduct the statistical analysis by using Hypothesis *ES*.

Hypothesis *ES* If experts did not enter too many times or too few times their racehorses of certain sire lines into races of certain distances, we would expect that experts entered their horses of sire line s_i into races of distance d_j at most $N_{ES}(s_i, d_j)$ times

$$N_{ES}(s_i, d_j) = P_{ES}(d_j) \times \sum_j N_{entry}(s_i, d_j) \tag{1}$$

where d_j is the type of race distance. We classified race distances into five types:

d_1	1000 – 1399m
d_2	1400 – 1799m
d_3	1800 – 2199m
d_4	2200 – 2799m
d_5	2800m –

$N_{entry}(s_i, d_j)$ is the number of times horses of sire line s_i were entered into races of distance d_j, as a result, $\sum_j N_{entry}(s_i, d_j)$ is the total number of times horses of sire line s_i were entered into races. $P_{ES}(d_j)$ is the probability that an expert enters his/her horse into a race of distance d_j. $P_{ES}(d_j)$ is

$$P_{ES}(d_j) = \frac{\sum_i N_{entry}(s_i, d_j)}{\sum_i \sum_j N_{entry}(s_i, d_j)} \tag{2}$$

where $\sum_i N_{entry}(s_i, d_j)$ is the total number of times horses were entered into races of distance d_j and $\sum_i \sum_j N_{entry}(s_i, d_j)$ is the total number of times horses were entered into races.

If this hypothesis is rejected by an two-sided binomial test [18], we determine that experts entered their horses of sire lines s_i into races of distance d_j too many times or too few times.

C. Detection of race distance and sire line combinations that gave good or poor results for racehorse experts too many times

In order to detect cases where horses of certain sire lines won or lost races of certain distances too many times, we conduct the statistical analysis by using Hypothesis *RR*.

Hypothesis *RR* If horses of certain sire lines did not perform well too many times or too few times in races of certain distances, we would expect that horses of sire line s_i

Copyright (c) IARIA, 2025. ISBN: 978-1-68558-238-8

TABLE VI. The p-values of experts' race selections for horses of Native Dancer Line.

sire line	race distance				
	1000–1399m	1400–1799m	1800–2199m	2200–2799m	2800m–
Native Dancer	1.0000	0.3024	0.0000	0.0000	0.0825

TABLE VII. The p-values of race results of horses of Native Dancer Line.

result	race distance				
	1000–1399m	1400–1799m	1800–2199m	2200–2799m	2800m–
first place	0.9997	0.8036	0.6069	0.7820	0.1506
top five place	0.9999	0.0890	0.7712	0.9542	0.1472

finished within $rank$-th place in races of distance d_j at most $N_{RR}(s_i, d_j, rank)$ times

$$N_{RR}(s_i, d_j, rank) = P_{RR}(d_j, rank) \times N_{entry}(s_i, d_j) \quad (3)$$

where d_j is the type of race distance. We classified race distances into five types in the same way that we did in Hypothesis *ES*. $N_{entry}(s_i, d_j)$ is the number of times horses of sire line s_i were entered into races of distance d_j. $P_{RR}(d_j, rank)$ is the probability that a horse finished within $rank$-th place in a race of distance d_j. $P_{RR}(d_j, rank)$ is

$$P_{RR}(d_j, rank) = \frac{\sum_i N_{result}(s_i, d_j, rank)}{\sum_i N_{entry}(s_i, d_j)} \quad (4)$$

where $N_{result}(s_i, d_j, rank)$ is the number of times horses of sire line s_i finished within $rank$-th place in races of distance d_j. As a result, $\sum_i N_{result}(s_i, d_j, rank)$ is the total number of times horses finished within $rank$-th place in races of distance d_j. Furthermore, $\sum_i N_{entry}(s_i, d_j)$ is the total number of times horses were entered into races of distance d_j.

If this hypothesis is rejected by an two-sided binomial test, we determine that horses of sire line s_i finished too many times or too few times within $rank$-th place in races of distance d_j.

D. Detection of horses that horse racing experts judged to have performed well

If a horse perform well in a race of a certain distance, experts will try to enter the horse into another race of a similar distance. As a result, if horses are judged to have performed well in races of a certain distance, experts may enter them into races of a similar distance repeatedly. In order to detect cases where horse racing experts entered their horses into races of certain distances too many times or too few times, we conduct the statistical analysis by using Hypothesis *EJ*.

Hypothesis *EJ* If an expert did not enter too many times or too few times his/her racehorse of a certain sire line into races of a certain distance, we would expect that the expert entered horse h_k into races of distance d_j at most $M_{EJ}(h_k, d_j)$ times

$$M_{EJ}(h_k, d_j) = P_{EJ}(s_i, d_j) \times M_{entry}(h_k, d_j) \quad (5)$$

where s_i is the sire line of horse h_k and d_j is the type of race distance. We classified race distances into five types in the same way that we did in Hypothesis *ES*. $M_{entry}(h_k, d_j)$ is the number of times horse h_k were entered into races of distance d_j. $P_{EJ}(s_i, d_j)$ is the probability that an expert enters a horse of sire line s_i into a race of distance d_j. $P_{EJ}(s_i, d_j)$ is

$$P_{EJ}(s_i, d_j) = \frac{N_{entry}(s_i, d_j)}{\sum_j N_{entry}(s_i, d_j)} \quad (6)$$

where $N_{entry}(s_i, d_j)$ is the number of times horses of sire line s_i were entered into races of distance d_j. As a result, $\sum_i N_{entry}(s_i, d_j)$ is the total number of times horses were entered into races of distance d_j.

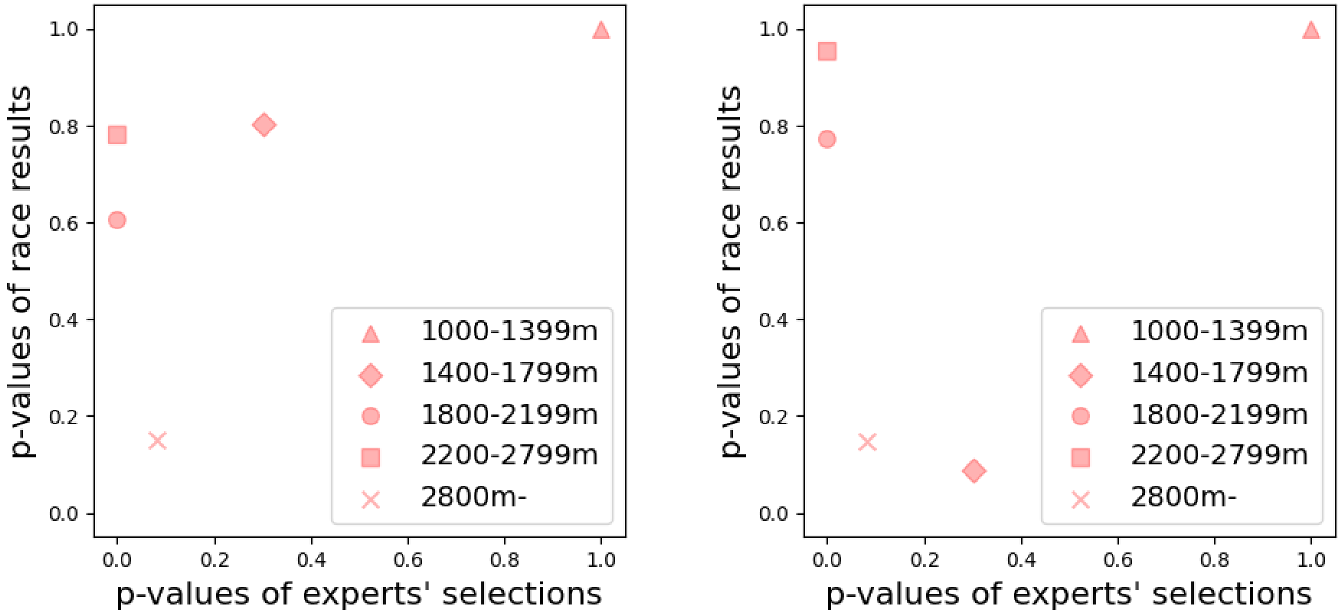
If this hypothesis is rejected by an two-sided binomial test [18], we determine that an expert entered his/her horse h_k of sire lines s_i into races of distance d_j too many times or too few times.

E. Results of the investigation

In order to investigate whether horse racing experts have inconsistent expectations, we apply Hypothesis *ES*, *RR*, and *EJ* tests on the 8799 horses of Native Dancer Line registered with JRA from 2010 to 2017, as shown in Table I. The significance levels for Hypothesis *ES*, *RR*, and *EJ* were 0.05. First, we calculated the p-values of experts' race selections, the race results, and experts' judgements of horses' performance by applying Hypothesis *ES*, *RR*, and *EJ*, respectively. Table VI shows the p-values of experts' race selections for horses of Native Dancer Line. Table VII show the p-values of race results (first place and top five place) of horses of Native Dancer Line. Figure 2 shows the p-values of experts' race selections vs the race results (first place and top five place) for horses of Native Dancer Line. Table VIII shows the number of the 8799 horses of Native Dancer Line competed in races of various distances and the number of times the horses had competed in the races and finished in first place and top five place. Table IX shows the number of horses of Native Dancer Line determined by Hypothesis *EJ* to have repeatedly competed in races of various distances and the number of times the horses had competed in the races and finished in first place and top five place.

First, we consider the results obtained by applying Hypothesis *ES*. Table VI shows

- the p-value of race distance type d_1 (1000 – 1399m) was more than 0.975. As a result, experts entered horses of Native Dancer Line into 1000 – 1399m races too many times. In other words, many experts strongly thought horses of Native Dancer Line were favorable to win in 1000 - 1399m races.
- the p-values of race distance type d_3 (1800 – 2199m) and d_4 (2200 – 2399m) were less than 0.025. In addition, the p-value of race distance type d_5 (2800m –) was low, 0.0825. As a result, many experts strongly thought horses of Native Dancer Line were unfavorable to win in races over 1800m.



(a) first place

(b) top five place

Figure 2. The p-values of experts' race selections vs race results (Native Dancer Line).

TABLE VIII. The number of horses (Native Dancer Line) competed in races of various distances and the number of times the horses had competed in the races and finished in first place and top five place.

	race distance				
	1000-1399m	1400-1799m	1800-2199m	2200-2799m	2800m-
horses (Native Dancer Line)	5045	7135	5599	1269	574
races competed in	27008	31619	28568	4173	2511
times in first place	1947	2261	2121	341	188
times in top five place	9345	10912	10552	1748	1120

TABLE IX. The number of horses (Native Dancer Line) determined by Hypothesis *EJ* to have repeatedly competed in races of various distances and the number of times the horses had competed in the races and finished in first place and top five place.

	race distance				
	1000-1399m	1400-1799m	1800-2199m	2200-2799m	2800m-
horses competed repeatedly	2320	1575	1940	628	376
races competed in	20005	13491	18207	3308	2248
times in first place	1593	1285	1643	308	188
times in top five place	7552	5615	7725	1502	1051

Next, we consider the results obtained by applying Hypothesis *EJ*. Figure 3 shows

- the number of horses repeatedly competed in the races of distance d_4 (2200 – 2799m) and d_5 (2800m –) was half or more than the number of all horses competed in the respective races. It is probable that many experts carefully considered which horses were favorable to win in races over 2200m. In other words, many experts thought horses of Native Dancer Line were unfavorable to win in races over 2200m.
- the number of horses repeatedly competed in the races of distance d_1 (1000 – 1399m), d_2 (1400 – 1799m), and d_3 (1800 – 2199m) was less than half the number of all horses competed in the respective races. It is probable that, compared to races over 2200m, many experts did not carefully considered which horses were favorable to win in races under 2200m. In other words, many experts thought that horses of Native Dancer Line were favorable to win in races of distances under 2200m compared to races over 2200m.

We focused on experts' expectations for races of distance d_3 (1800 – 2199m). This is because their expectations were inconsistent for races of this distance, as shown below.

- Table VI, the results obtained by applying Hypothesis *ES*, shows that many experts thought horses of Native Dancer Line were unfavorable to win in races of distance d_3 (1800 – 2199m).
- Figure 3, the results obtained by applying Hypothesis *EJ*, shows that many experts thought horses of Native Dancer Line were favorable to win in races of distance d_3 (1800 – 2199m).

We thought the reason for this inconsistent expectations is that many experts unconsciously knew horses of Native Dancer Line were unfavorable to win in races of this distance. Actually, Table VII, the results obtained by applying Hypothesis *RR*, shows that horses of Native Dancer Line were unfavorable to win in races of distance d_3 (1800 – 2199m) compared to races of distance d_1 (1000 – 1399m). Many experts may have

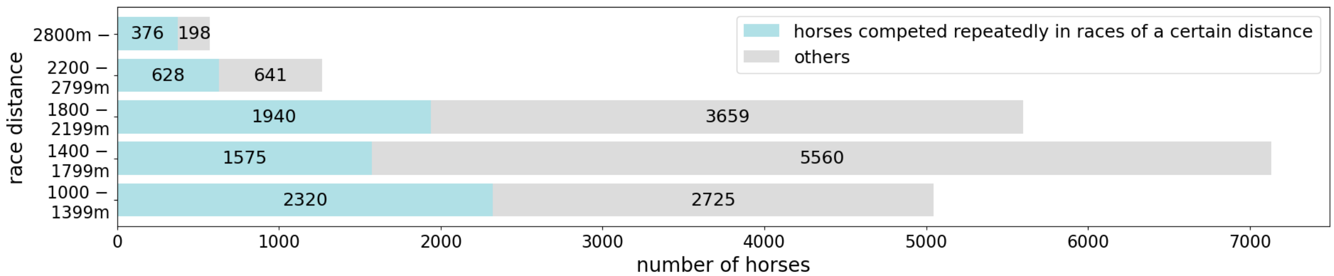


Figure 3. The number of horses (Native Dancer Line) competed repeatedly and others in races of various distances.

unconsciously avoided selecting races of distance d_3 (1800 – 2199m). Also, their conscious minds may not have known that horses of Native Dancer Line were unfavorable to win in races of this distance. As a result, many experts may not have carefully considered which horses were favorable to win in races of this distance.

V. Conclusion

Although experts have a wealth of knowledge and experience, they sometimes make mistakes. However, not enough research has been done on how and why experts made mistakes. We thought that one of the reasons why they made mistakes is that they have inconsistent expectations. As a result, in this paper, we investigated whether experts have inconsistent expectations on their professional issue. We analyzed sire lines, race distances, and race results of the 36922 horses statistically and showed that horse racing experts had inconsistent expectations on the problem of which race distance they thought were favorable for horses of a certain sire line. We think this is because experts' unconscious minds affected their expectations. Even for experts, it is difficult to consciously notice what they have unconsciously felt. To generalize this finding, we intend to analyze race performance data in other sire lines and compare the results with those obtained in this study.

References

[1] M. A. Bower *et al.*, "The cosmopolitan maternal heritage of the thoroughbred racehorse breed shows a significant contribution from british and irish native mares," *Biology letters*, vol. 7, no. 2, pp. 316–320, 2011. [Online]. Available: https://doi.org/10.1098/rsbl.2010.0800 [accessed: 2025-01-23]

[2] E. Cunningham, J. J. Dooley, R. Splan, and D. Bradley, "Microsatellite diversity, pedigree relatedness and the contributions of founder lineages to thoroughbred horses," *Animal genetics*, vol. 32, no. 6, pp. 360–364, 2001. [Online]. Available: https://doi.org/10.1046/j.1365-2052.2001.00785.x [accessed: 2025-01-23]

[3] E. Hill *et al.*, "History and integrity of thoroughbred dam lines revealed in equine mtdna variation," *Animal genetics*, vol. 33, no. 4, pp. 287–294, 2002. [Online]. Available: https://doi.org/10.1046/j.1365-2052.2002.00870.x [accessed: 2025-01-23]

[4] C. Wade *et al.*, "Genome sequence, comparative analysis, and population genetics of the domestic horse," *Science*, vol. 326, no. 5954, pp. 865–867, 2009. [Online]. Available: https://doi.org/10.1126/science.1178158 [accessed: 2025-01-23]

[5] B. A. McGivney *et al.*, "Genomic inbreeding trends, influential sire lines and selection in the global thoroughbred horse population," *Scientific Reports*, vol. 10, no. 1, p. 466, 2020. [Online]. Available: https://doi.org/10.1038/s41598-019-57389-5 [accessed: 2025-01-23]

[6] E. W. Hill, B. A. McGivney, J. Gu, R. Whiston, and D. E. MacHugh, "A genome-wide snp-association study confirms a sequence variant (g. 66493737c>t) in the equine myostatin (mstn) gene as the most powerful predictor of optimum racing distance for thoroughbred racehorses," *BMC genomics*, vol. 11, no. 1, pp. 1–10, 2010. [Online]. Available: https://doi.org/10.1186/1471-2164-11-552 [accessed: 2025-01-23]

[7] C. Wylie and J. Newton, "A systematic literature search to identify performance measure outcomes used in clinical studies of racehorses," *Equine veterinary journal*, vol. 50, no. 3, pp. 304–311, 2018. [Online]. Available: https://doi.org/10.1111/evj.12757 [accessed: 2025-01-23]

[8] G. Martin, E. Strand, and M. Kearney, "Use of statistical models to evaluate racing performance in thoroughbreds," *Journal of the American Veterinary Medical Association*, vol. 209, pp. 1900–1906, 1996. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/8944806/ [accessed: 2025-01-23]

[9] J. Cheetham, A. Riordan, H. Mohammed, C. McIlwraith, and L. Fortier, "Relationships between race earnings and horse age, sex, gait, track surface and number of race starts for thoroughbred and standardbred racehorses in north america," *Equine Veterinary Journal*, vol. 42, no. 4, pp. 346–350, 2010. [Online]. Available: https://doi.org/10.1111/j.2042-3306.2010.00032.x [accessed: 2025-01-23]

[10] I. Wells, H. Randle, and J. M. Williams, "Does the start of flat races influence racehorse race performance?" *Applied Animal Behaviour Science*, vol. 253, p. 105682, 2022. [Online]. Available: https://doi.org/10.1016/j.applanim.2022.105682 [accessed: 2025-01-23]

[11] T. Sawchik, *Has the Fly-Ball Revolution Begun?*, FanGraphs Baseball, 2017. [Online]. Available: https://blogs.fangraphs.com/has-the-fly-ball-revolution-begun/ [accessed: 2025-01-23]

[12] M. Kato and T. Yanai, "Launch fly balls for better batting statistics: Applicability of "fly-ball revolution" to Japan's professional baseball league," *International Journal of Performance Analysis in Sport*, vol. 22, no. 3, pp. 437–453, 2022. [Online]. Available: https://doi.org/10.1080/24748668.2022.2075302 [accessed: 2025-01-23]

[13] R. M. Yerkes and J. D. Dodson, "The relation of strength of stimulus to rapidity of habit-formation," *Journal of Comparative Neurology and Psychology*, vol. 18, no. 5, pp. 459–482, 1908. [Online]. Available: https://doi.org/10.1002/cne.920180503 [accessed: 2025-01-23]

[14] S. A. Shappell and D. A. Wiegmann, *The human factors analysis and classification system–HFACS*, U.S. Department of Transportation Federal Aviation Administration, 2000. [Online]. Available: https://www.skybrary.aero/sites/default/files/bookshelf/1481.pdf [accessed: 2025-01-23]

[15] N. E. Kang and W. C. Yoon, "Age- and experience-related user behavior differences in the use of complicated electronic devices," *International journal of human-computer studies*, vol. 66, no. 6, pp. 425–437, 2008. [Online]. Available: https://doi.org/10.1016/j.ijhcs.2007.12.003 [accessed: 2025-01-23]

[16] A. Bechara, H. Damasio, D. Tranel, and A. R. Damasio, "Deciding advantageously before knowing the advantageous strategy," *Science*, vol. 275, no. 5304, pp. 1293–1295, 1997. [Online]. Available: https://doi.org/10.1126/science.275.5304.1293 [accessed: 2025-01-23]

[17] *Keiba Lab*, Keiba Lab. [Online]. Available: https://www.keibalab.jp/ [accessed: 2025-01-23]

[18] M. Hollander and D. A. Wolfe, *Nonparametric Statistical Methods, 2nd Edition*. Wiley-Interscience, Jan. 1999.