

References

- Agarwal, S., El-Gaaly, T., Farid, H., & Lim, S.-N. (2020). *Detecting Deep-Fake Videos from Appearance and Behavior*. <https://doi.org/10.48550/ARXIV.2004.14491>
- AlBadawy, E. A., Lyu, S., & Farid, H. (2019, June). Detecting AI-Synthesized Speech Using Bispectral Analysis. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.
- Ali, K., Li, C., Zain-ul-abdin, K., & Zaffar, M. A. (2022). Fake news on Facebook: Examining the impact of heuristic cues on perceived credibility and sharing intention. *Internet Research*, 32(1), 379–397. <https://doi.org/10.1108/INTR-10-2019-0442>
- Al-Khazraji, S. H., Saleh, H. H., Khalid, A. I., & Mishkhal, I. A. (2023). Impact of Deepfake Technology on Social Media: Detection, Misinformation and Societal Implications. *The Eurasia Proceedings of Science Technology Engineering and Mathematics*, 23, 429–441. <https://doi.org/10.55549/epstem.1371792>
- Altay, S., Lyons, B. A., & Modirrousta-Galian, A. (2025). Exposure to Higher Rates of False News Erodes Media Trust and Fuels Overconfidence. *Mass Communication and Society*, 28(2), 301–325. <https://doi.org/10.1080/15205436.2024.2382776>
- Aruni, S. F., & Hidayat, R. (2019). Psikologi Peluang Kewirausahaan: Proses Kognitif Pengusaha Startup Digital dalam Opportunity Recognition. *Jurnal Psikologi*, 46(1), 45. <https://doi.org/10.22146/jpsi.34608>
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human Memory: A Proposed System and its Control Processes. In *Psychology of Learning and Motivation* (Vol. 2, pp. 89–195). Elsevier. [https://doi.org/10.1016/S0079-7421\(08\)60422-3](https://doi.org/10.1016/S0079-7421(08)60422-3)
- Babiker, A., & Basta, T. (2024). *Deepfake Voice Implementation for Scams: Information Security Risks with Deepfake Technology*. KTH Royal Institute of Technology.
- Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417–423. [https://doi.org/10.1016/S1364-6613\(00\)01538-2](https://doi.org/10.1016/S1364-6613(00)01538-2)
- Baddeley, A. D., & Hitch, G. (1974). Working Memory. In *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). Elsevier. [https://doi.org/10.1016/S0079-7421\(08\)60452-1](https://doi.org/10.1016/S0079-7421(08)60452-1)
- Bago, B., Rand, D. G., & Pennycook, G. (2020). Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *Journal of Experimental Psychology: General*, 149(8), 1608–1613. <https://doi.org/10.1037/xge0000729>
- Bastick, Z. (2021). Would you notice if fake news changed your behavior? An experiment on the unconscious effects of disinformation. *Computers in Human Behavior*, 116, 106633. <https://doi.org/10.1016/j.chb.2020.106633>
- Batailler, C., Brannon, S. M., Teas, P. E., & Gawronski, B. (2022). A Signal Detection Approach to Understanding the Identification of Fake News. *Perspectives on Psychological Science*, 17(1), 78–98. <https://doi.org/10.1177/1745691620986135>
- Borji, A. (2023). Qualitative failures of image generation models and their application in detecting deepfakes. *Image and Vision Computing*, 137, 104771. <https://doi.org/10.1016/j.imavis.2023.104771>
- Bray, S. D., Johnson, S. D., & Kleinberg, B. (2022). *Testing Human Ability To Detect Deepfake Images of Human Faces* (Version 3). arXiv. <https://doi.org/10.48550/ARXIV.2212.05056>
- Brieger, A. (2021). *Taking Back Their Faces: The Damages of Non-Consensual Deepfake Pornography on Female Journalists*.

- Caramancion, K. M. (2020). Understanding the Impact of Contextual Clues in Misinformation Detection. *2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*, 1–6. <https://doi.org/10.1109/IEMTRONICS51293.2020.9216394>
- Chesney, B., & Citron, D. (2019). *Deep Fakes: A Looming Challenge for Privacy*. <https://doi.org/10.15779/Z38RV0D15J>
- Chesney, R., & Citron, D. (2019). Deepfakes and the New Disinformation War. *Foreign Affairs*, 98(1), 147–155. JSTOR.
- Cho, B., Woodward, L., & Afflerbach, P. (2020). *Qualitative Approaches to the Verbal Protocol Analysis of Strategic Processing*.
- Cofie, N., Braund, H., & Dalgarno, N. (2022). Eight ways to get a grip on intercoder reliability using qualitative-based measures. *Canadian Medical Education Journal*. <https://doi.org/10.36834/cmej.72504>
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46.
- Cowan, J. (2019). The potential of cognitive think-aloud protocols for educational action-research. *Active Learning in Higher Education*, 20(3), 219–232. <https://doi.org/10.1177/1469787417735614>
- Dane, E., Rockmann, K. W., & Pratt, M. G. (2012). When should I trust my gut? Linking domain expertise to intuitive decision-making effectiveness. *Organizational Behavior and Human Decision Processes*, 119(2), 187–194. <https://doi.org/10.1016/j.obhdp.2012.07.009>
- De Ruiter, A. (2021). The Distinct Wrong of Deepfakes. *Philosophy & Technology*, 34(4), 1311–1332. <https://doi.org/10.1007/s13347-021-00459-2>
- Delfino, R. (2019). Pornographic Deepfakes—Revenge Porn’s Next Tragic Act – The Case for Federal Criminalization. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3341593>
- Diel, A., Lalgı, T., Schröter, I. C., MacDorman, K. F., Teufel, M., & Bäuerle, A. (2024). Human performance in detecting deepfakes: A systematic review and meta-analysis of 56 papers. *Computers in Human Behavior Reports*, 16, 100538. <https://doi.org/10.1016/j.chbr.2024.100538>
- Djudin, T., & Hamdani, H. (2023). Revised-Verbal Protocols: A Valuable Technique for Verbalizing the Students’ Thoughts in Teaching-Learning Process. *International Journal of Research and Innovation in Social Science*, VII(IX), 602–612. <https://doi.org/10.47772/IJRISS.2023.70949>
- Ericsson, K. A. (2002). Towards a procedure for eliciting verbal expression of non-verbal experience without reactivity: Interpreting the verbal overshadowing effect within the theoretical framework for protocol analysis. *Applied Cognitive Psychology*, 16(8), 981–987. <https://doi.org/10.1002/acp.925>
- Ericsson, K. A. (2017). Protocol Analysis. In W. Bechtel & G. Graham (Eds.), *A Companion to Cognitive Science* (1st ed., pp. 425–432). Wiley. <https://doi.org/10.1002/9781405164535.ch33>
- Evans, J. St. B. T. (2008). Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. *Annual Review of Psychology*, 59(1), 255–278. <https://doi.org/10.1146/annurev.psych.59.103006.093629>
- Fletcher, J. (2018). Deepfakes, Artificial Intelligence, and Some Kind of Dystopia: The New Faces of Online Post-Fact Performance. *Theatre Journal*, 70(4), 455–471. <https://doi.org/10.1353/tj.2018.0097>

- French, A. M., Storey, V. C., & Wallace, L. (2025). The impact of cognitive biases on the believability of fake news. *European Journal of Information Systems*, 34(1), 72–93. <https://doi.org/10.1080/0960085X.2023.2272608>
- Green, A. (1998). *Verbal protocol analysis in language testing research: A handbook* (Vol. 5). Cambridge University Press.
- Groh, M., Epstein, Z., Firestone, C., & Picard, R. (2022). Deepfake detection by human crowds, machines, and machine-informed crowds. *Proceedings of the National Academy of Sciences*, 119(1), e2110013119. <https://doi.org/10.1073/pnas.2110013119>
- Guberman, S. (2017). Gestalt Theory Rearranged: Back to Wertheimer. *Frontiers in Psychology*, 8, 1782. <https://doi.org/10.3389/fpsyg.2017.01782>
- Halpern, D., Valenzuela, S., Katz, J., & Miranda, J. P. (2019). From Belief in Conspiracy Theories to Trust in Others: Which Factors Influence Exposure, Believing and Sharing Fake News. In G. Meiselwitz (Ed.), *Social Computing and Social Media. Design, Human Behavior and Analytics* (Vol. 11578, pp. 217–232). Springer International Publishing. https://doi.org/10.1007/978-3-030-21902-4_16
- Hörhan, M., & Eidenberger, H. (2021). Gestalt descriptions for deep image understanding. *Pattern Analysis and Applications*, 24(1), 89–107. <https://doi.org/10.1007/s10044-020-00904-6>
- Hughes, J., & Parkes, S. (2003). Trends in the use of verbal protocol analysis in software engineering research. *Behaviour & Information Technology*, 22(2), 127–140. <https://doi.org/10.1080/0144929031000081341>
- Huntsinger, J. R., & Ray, C. (2016). A flexible influence of affective feelings on creative and analytic performance. *Emotion*, 16(6), 826–837. <https://doi.org/10.1037/emo0000188>
- Kahneman, D. (2011). Thinking, fast and slow. *Farrar, Straus and Giroux*.
- Korteling, J. E. (Hans)., Van De Boer-Visschedijk, G. C., Blankendaal, R. A. M., Boonekamp, R. C., & Eikelboom, A. R. (2021). Human- versus Artificial Intelligence. *Frontiers in Artificial Intelligence*, 4, 622364. <https://doi.org/10.3389/frai.2021.622364>
- Li, Y., Yang, X., Sun, P., Qi, H., & Lyu, S. (2019). *Celeb-DF: A Large-scale Challenging Dataset for DeepFake Forensics* (Version 4). arXiv. <https://doi.org/10.48550/ARXIV.1909.12962>
- Mai, K. T., Bray, S., Davies, T., & Griffin, L. D. (2023). Warning: Humans cannot reliably detect speech deepfakes. *PLOS ONE*, 18(8), e0285333. <https://doi.org/10.1371/journal.pone.0285333>
- Malik, A., Kuribayashi, M., Abdullahi, S. M., & Khan, A. N. (2022). DeepFake Detection for Human Face Images and Videos: A Survey. *IEEE Access*, 10, 18757–18775. <https://doi.org/10.1109/ACCESS.2022.3151186>
- Mania, K. (2024). Legal protection of revenge and deepfake porn victims in the European Union: Findings from a comparative legal study. *Trauma, Violence, & Abuse*, 25(1), 117–129.
- Masip, J., Garrido, E., & Herrero, C. (2004). Defining deception. *Anales de Psicología*, 20, 147–171.
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), 276–282.
- Mori, M., MacDorman, K. F., & Kageki, N. (2012). The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine*, 19(2), 98–100. <https://doi.org/10.1109/MRA.2012.2192811>
- Muhly, F., Chizzonic, E., & Leo, P. (2025). AI-deepfake scams and the importance of a holistic communication security strategy. *International Cybersecurity Law Review*, 1–9.

- Müller, N. M., Pizzi, K., & Williams, J. (2022). Human Perception of Audio Deepfakes. *Proceedings of the 1st International Workshop on Deepfake Detection for Audio Multimedia*, 85–91. <https://doi.org/10.1145/3552466.3556531>
- Mustak, M., Salminen, J., Mäntymäki, M., Rahman, A., & Dwivedi, Y. K. (2023). Deepfakes: Deceptions, mitigations, and opportunities. *Journal of Business Research*, 154, 113368. <https://doi.org/10.1016/j.jbusres.2022.113368>
- Nas, E., & De Kleijn, R. (2024). Conspiracy thinking and social media use are associated with ability to detect deepfakes. *Telematics and Informatics*, 87, 102093. <https://doi.org/10.1016/j.tele.2023.102093>
- Noushad, B., Van Gerven, P. W. M., & De Bruin, A. B. H. (2024). Twelve tips for applying the think-aloud method to capture cognitive processes. *Medical Teacher*, 46(7), 892–897. <https://doi.org/10.1080/0142159X.2023.2289847>
- Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research. *Administration and Policy in Mental Health and Mental Health Services Research*, 42(5), 533–544. <https://doi.org/10.1007/s10488-013-0528-y>
- Pehlivanoglu, D., Lin, T., Deceus, F., Heemskerk, A., Ebner, N. C., & Cahill, B. S. (2021). The role of analytical reasoning and source credibility on the evaluation of real and fake full-length news articles. *Cognitive Research: Principles and Implications*, 6(1), 24. <https://doi.org/10.1186/s41235-021-00292-3>
- Pehlivanoglu, D., Zhu, M., Zhen, J., Gagnon-Roberge, A. A. C. M., Kern, R., Woodard, D., Cahill, B. S., & Ebner, N. C. (2024). *Is this real? Susceptibility to deepfakes in machines and humans*. <https://doi.org/10.31219/osf.io/etxzw>
- Pennycook, G., & Rand, D. G. (2021). The Psychology of Fake News. *Trends in Cognitive Sciences*, 25(5), 388–402. <https://doi.org/10.1016/j.tics.2021.02.007>
- Qureshi, J., & Khan, S. (2024). Deciphering Deception—The Impact of AI Deepfakes on Human Cognition and Emotion. *Journal of Advances in Artificial Intelligence*, 2(1). <https://doi.org/10.18178/JAAI.2024.2.1.101-107>
- Rana, M. S., Nobi, M. N., Murali, B., & Sung, A. H. (2022). Deepfake Detection: A Systematic Literature Review. *IEEE Access*, 10, 25494–25513. <https://doi.org/10.1109/ACCESS.2022.3154404>
- Schiff, K. J., Schiff, D. S., & Bueno, N. (2023). *The Liar's Dividend: The Impact of Deepfakes and Fake News on Trust in Political Discourse* (SocArXiv x43ph). Center for Open Science. <https://doi.org/10.31219/osf.io/x43ph>
- Schwarz, N. (2012). Feelings-as-Information Theory. In P. Van Lange, A. Kruglanski, & E. Higgins, *Handbook of Theories of Social Psychology: Volume 1* (pp. 289–308). SAGE Publications Ltd. <https://doi.org/10.4135/9781446249215.n15>
- Shin, S. Y., & Lee, J. (2022). The Effect of Deepfake Video on News Credibility and Corrective Influence of Cost-Based Knowledge about Deepfakes. *Digital Journalism*, 10(3), 412–432. <https://doi.org/10.1080/21670811.2022.2026797>
- Siemens, G., Marmolejo-Ramos, F., Gabriel, F., Medeiros, K., Marrone, R., Joksimovic, S., & De Laat, M. (2022). Human and artificial cognition. *Computers and Education: Artificial Intelligence*, 3, 100107. <https://doi.org/10.1016/j.caeai.2022.100107>
- Smith, H., & Mansted, K. (2020). *What's a deep fake?* (Weaponised Deep Fakes, pp. 05–10). Australian Strategic Policy Institute; JSTOR. <http://www.jstor.org/stable/resrep25129.6>
- Suri, H. (2011). Purposeful sampling in qualitative research synthesis. *Qualitative Research Journal*, 11(2), 63–75.

- Sutterlin, S., Lugo, R. G., Ask, T. F., Veng, K., Eck, J., Fritsch, J., Özmen, M.-T., Bärreiter, B., & Knox, B. J. (2022). The Role of IT Background for Metacognitive Accuracy, Confidence and Overestimation of Deep Fake Recognition Skills. In D. D. Schmorow & C. M. Fidopiastis (Eds.), *Augmented Cognition* (Vol. 13310, pp. 103–119). Springer International Publishing. https://doi.org/10.1007/978-3-031-05457-0_9
- Tarchi, P., Lanini, M. C., Frassinetti, L., & Lanatà, A. (2023). Real and Deepfake Face Recognition: An EEG Study on Cognitive and Emotive Implications. *Brain Sciences*, 13(9), 1233. <https://doi.org/10.3390/brainsci13091233>
- Taylor, E. (2018). PAUSED for thought? Using verbal protocol analysis to understand the situational and temporal cues in the decision-making of residential burglars. *Security Journal*, 31(1), 343–363. <https://doi.org/10.1057/s41284-017-0104-3>
- Trickett, S. B., & Trafton, J. G. (2009). *Chapter 18 A PRIMER ON VERBAL PROTOCOL ANALYSIS*. <https://api.semanticscholar.org/CorpusID:156051536>
- Unoki, M., Li, K., Chaiwongyen, A., Nguyen, Q.-H., & Zaman, K. (2024). Deepfake speech detection: Approaches from acoustic features related to auditory perception to deep neural networks. *IEICE Transactions on Information and Systems*, 2024MUI0001. <https://doi.org/10.1587/transinf.2024MUI0001>
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Social Media + Society*, 6(1), 205630512090340. <https://doi.org/10.1177/2056305120903408>
- Wang, M., & Deng, W. (2021). Deep face recognition: A survey. *Neurocomputing*, 429, 215–244. <https://doi.org/10.1016/j.neucom.2020.10.081>
- Wang, R., Juefei-Xu, F., Huang, Y., Guo, Q., Xie, X., Ma, L., & Liu, Y. (2020). *DeepSonar: Towards Effective and Robust Detection of AI-Synthesized Fake Voices* (Version 3). arXiv. <https://doi.org/10.48550/ARXIV.2005.13770>
- Wang, S., & Kim, S. (2022). Users' emotional and behavioral responses to deepfake videos of K-pop idols. *Computers in Human Behavior*, 134, 107305. <https://doi.org/10.1016/j.chb.2022.107305>
- Weikmann, T., Greber, H., & Nikolaou, A. (2024). After Deception: How Falling for a Deepfake Affects the Way We See, Hear, and Experience Media. *The International Journal of Press/Politics*, 19401612241233539. <https://doi.org/10.1177/19401612241233539>
- Westerlund, M. (2019). The Emergence of Deepfake Technology: A Review. *Technology Innovation Management Review*, 9(11), 39–52. <https://doi.org/10.22215/timreview/1282>
- Westerman, D., Spence, P. R., & Van Der Heide, B. (2014). Social Media as Information Source: Recency of Updates and Credibility of Information. *Journal of Computer-Mediated Communication*, 19(2), 171–183. <https://doi.org/10.1111/jcc4.12041>
- Wickens, C. D., & Carswell, C. M. (2021). INFORMATION PROCESSING. In G. Salvendy & W. Karwowski (Eds.), *HANDBOOK OF HUMAN FACTORS AND ERGONOMICS* (1st ed., pp. 114–158). Wiley. <https://doi.org/10.1002/9781119636113.ch5>
- Zhang, T. (2022). Deepfake generation and detection, a survey. *Multimedia Tools and Applications*, 81(5), 6259–6276. <https://doi.org/10.1007/s11042-021-11733-y>
- Zi, B., Chang, M., Chen, J., Ma, X., & Jiang, Y.-G. (2020). WildDeepfake: A Challenging Real-World Dataset for Deepfake Detection. *Proceedings of the 28th ACM International Conference on Multimedia*, 2382–2390. <https://doi.org/10.1145/3394171.3413769>