Proceedings of the 5th International Conference on Interdisciplinary Humanities and Communication Studies
DOI: 10.54254/2753-7064/2026.37216

The Gaze Asymmetry Framework: Asymmetry of Social Defence Mechanisms under Human and Algorithmic Gazes

Enning Shen

School of Art & Design, Wuhan Institute of Technology, Wuhan, China
10091401@wit.com.cn

Abstract. In contemporary public spaces, individuals automatically activate social defences when facing the human gaze, yet they relax when confronted with ubiquitous cameras. The difference between humans and machines in triggering social defence has not been explained in a systematic way; existing theories can explain traditional discipline and defence, but they do not clearly explain how algorithmic surveillance in digital contexts, which has no physical body, may cause the structural weakening of individuals' social perception. This paper uses a conceptual approach based on cognitive psychology and micro-sociology, and it proposes the Gaze Asymmetry Framework to fill this gap; this framework explains the phenomenon through a simple three-part logic: the human gaze triggers defences because it contains natural cues of subjectivity; the algorithmic gaze removes subjectivity in a systematic way, and this disrupts the "imagined audience", so traditional defence mechanisms are turned off; long-term exposure to this environment without a subject may lead to a cognitive spillover effect, which can possibly affect individuals' broader social perception abilities. This study offers a theoretical tool for understanding the basic mechanisms behind modern surveillance, and it helps explain how digital contexts change the boundaries of human interaction.

Keywords: algorithmic gaze, surveillance studies, social defence mechanism, cognitive spillover, dramaturgical theory

1. Introduction

People in modern public spaces often show contradictory behaviours; when they notice the gaze of other people, for example catching a stranger staring at them on a crowded subway car or being watched by passersby while sitting in a cafe, they often feel social unease, and this unease leads them to unconsciously change their posture, pull in their personal space, or turn on social defence mechanisms through "civil inattention" [1]; but when people face common machine sensors like cameras, they often do not activate the front-stage impression management process described in dramaturgical theory [2], because there are no social or environmental signals to set it off; this difference creates a key behavioural contradiction in modern public spaces; even though algorithmic surveillance is everywhere, existing research gives little understanding of why machine watching does not cause the same social and emotional reactions as the human gaze; the main goal of this paper is to address this theoretical gap.

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-7064/2026.37216

No one has combined these two into a single framework that can systematically explain the mental process behind the different activation of social defence and its long-term effects; theories of traditional human gaze and algorithmic surveillance provide useful ideas about the social defences and control methods people use when observed [2-4], but they cannot fully explain the current digital environment; in this new situation, algorithmic monitoring does not focus on a real person, it collects information that is then taken out of its original social setting [5]; these invisible algorithms lack signs of personal subjectivity or moral judgement, so people being watched lose the reason to turn on usual social defences; as a result, current literature has two clear research gaps, one is the lack of a single theoretical framework to explain the mental mechanisms behind this asymmetric phenomenon, and the other is about the structural weakening of social perception when people are exposed long-term to an algorithmic environment without social evaluation, which remains unanswered.

To address these gaps, this paper focuses on two main research questions:

RQ1: Why does this gaze asymmetry occur, and what cognitive mechanisms explain how human gaze triggers distinct social anxiety and defensive behaviours, while algorithmic surveillance fails to trigger the same social defence response?

RQ2: How might long-term exposure to algorithmic environments influence individuals' social perception and interpersonal interactions?

To deal with this gap, the paper looks at the theoretical mechanisms behind the difference between human and algorithmic gazes when they trigger social and emotional responses; the study argues that algorithms naturally lack the ability to signal social connection or moral agency, so people do not mentally recognise them as valid judges, and as a result, traditional social defence mechanisms such as civil inattention [1] lose the conditions that set them off; accordingly, the lack of defences shown by people under the algorithmic gaze reflects a weakened perception of being watched, which is caused by the technological neutralisation of social pressure. Starting from this point, the study takes a conceptual and theory-building approach and develops the Gaze Asymmetry Framework to compare and analyse this psychological change; in particular, the research discusses in a systematic way how long-term exposure to algorithmic environments may harm people's ability to perceive social cues in public space. The paper offers three main theoretical contributions: one contribution is that it sets up a single conceptual framework, the Gaze Asymmetry Framework, which centres on how subjectivity is attributed; another contribution is that it clearly explains the mental processes behind why social defences fail when people face algorithmic surveillance that has no body; a further contribution is that it gives a theoretical explanation for the long-term structural effects that algorithmic environments have on human social thinking. This study does not test the proposed relationship empirically, but instead develops a conceptual framework for future empirical research.

2. Literature review

Existing research on the gaze has mostly centred on explaining the internal ways by which it sets off social defences; more precisely, studies establish social gaze as a basic communicative cue in human interaction [6], and these communicative cues, which rely on other people's sight, facial expressions, and body direction, give individuals the cognitive basis to view gaze as something socially meaningful, so they act as the necessary trigger for defensive reactions.

Building on this basis, the human gaze sets off an individual's psychological defences mainly because it carries clear signs of subjectivity and puts the observed person under constant moral and social judgement; rather than seeing these mechanisms separately, existing literature treats the

DOI: 10.54254/2753-7064/2026.37216

impact of the human gaze through three connected thematic layers, namely social evaluation, social discipline, and social interaction. One thematic layer is to view gaze as a form of social evaluation, and Sartre [7] points out the ontological weakness of being reduced to an object of judgement under the gaze of the Other; this evaluative pressure makes individuals actively monitor and adjust their behaviour, and as shown by Snyder's self-monitoring theory and Argyle and Dean's [8] Intimacy Equilibrium Model, people keep regulating non-verbal cues such as eye contact frequency and physical distance to keep psychological safety and manage how they are evaluated. Moving from the interpersonal level to the structural level, the gaze works as a mechanism of discipline; Foucault [3] explains how the widespread gaze centralises power relations and creates a feeling of oppression that encourages internalised self-discipline among the observed; these evaluative and disciplinary pressures together form a practical framework for everyday social navigation. Because the human gaze places individuals under both interpersonal judgement and structural discipline, it acts as a functional trigger that forces them to actively manage their social boundaries; this need to manage the gaze connects directly to the third thematic layer, which treats gaze as interaction, where individuals use specific front-stage social defences and impression management strategies, as described in Goffman's [2] dramaturgical theory.

Further elaboration is provided by the defensive logic behind the performance perspective [2] and civil inattention [1]; when people sense a gaze that carries moral judgement, they subconsciously start a front-stage act, managing how they appear or pulling back their physical space as a safety strategy to protect the boundaries of interaction in public places; this felt judgement may then lead them to behave more carefully in social situations; in addition, Snyder's [9] self-monitoring theory supports the idea that people with high self-monitoring traits are more clearly sensitive to this kind of evaluative gaze.

The 'imagined audience' idea put forward by Marwick and Boyd [10] is closely connected to impression management; from this point of view, people do not just respond to someone being physically there, but they actively build that presence in their mind. When individuals notice social cues, they judge the moral position and ability to evaluate of possible observers, and they place these qualities onto the audience they imagine; this imagined audience sets off social defences because people take in outside social rules and possible judgements. The mental expectation of being judged by this constructed audience creates a felt risk of social punishment or shame; this expected judgement, not just the actual presence of other people, works as the mental signal that pushes individuals to start front-stage behaviours. By actively projecting in this way, individuals establish subjectivity, and this becomes a necessary condition for social defence mechanisms to turn on; when the above theories are combined, it can be worked out that the subjectivity and moral evaluation capacity inside the human gaze are the main conditions that set off traditional defence mechanisms. But the spread of digital watching technologies changes this existing pattern in a basic way; when the watching source moves from a human placed in a social setting to a machine with no body, the social cues needed to build an evaluative imagined audience are completely removed. This missing part naturally leads to an important question: how do people mentally deal with and react to an algorithmic gaze that has no subjectivity or moral agency?

Traditional social defence mechanisms are not activated automatically; instead, they are highly conditional, and as Nass and Moon's [4] key premise says, unconscious human social rules are activated mainly because of the cues of subjectivity that the other person gives, so when this core condition is broken, an individual's defence mechanisms stop working as a result. Consequently, when the gaze moves to algorithms and machines in today's digital world, the algorithmic gaze takes away signals of subjectivity at the physical and systemic levels, which makes traditional defence

DOI: 10.54254/2753-7064/2026.37216

mechanisms completely ineffective; it is worth noting that Nass and Moon [4] mainly deal with interactive, human-like interfaces, whereas this framework looks at non-interactive, background monitoring technologies with no human-like cues, and the human gaze sets off social defence because it has cues that let people infer the presence of a subject who can judge them socially.

At the physical level, as Koskela [11] notes, modern cameras create a "gaze without eyes"; this bodiless gaze removes the eye contact that is part of normal human interaction, and it directly produces a surface-level lack of subjectivity signals. The Cuelessness Model by Kemp and Rutter [12] gives a useful theoretical way to look further at how this shortage of emotional signals works, even though it was first made to study mediated communication; its basic idea can be reasonably applied to today's digital surveillance, and it shows that cameras also lack important emotional and expressive hints, which leaves people feeling quite disconnected when they face these machines and makes them see this kind of watching as fully "de-emotionalized." Still, existing studies point out this physical removal of emotion clearly, but they mostly stay at the level of describing the experience; a key gap remains in explaining how this physical lack disturbs the first mental triggers of social defence.

On the systemic level, the traditional panoptic theory [3] suggests that the gaze creates oppression and encourages self-discipline; however, Haggerty and Ericson [13] point out that today's algorithmic gaze no longer aims at a person with a body, but instead turns the individual into a "data double" [13] from which social context has been removed, and this reduction of the embodied person into data pieces takes away the algorithmic gaze's capacity for moral judgement. The discussion by Lyon [5] of "invisible surveillance algorithms", together with Burrell's [14] analysis of the "structural opacity" of algorithms, shows how algorithms systematically hide the oppressive nature of the traditional panopticon; these studies effectively break down the structural opacity and datafication of algorithmic surveillance, but their attention stays on the technological device itself, so a critical theoretical gap remains. The existing literature seldom discusses how these systemic changes, along with the resulting lack of moral judgement, disturb the way a person mentally recognises the algorithm as a proper judge; as a result, the deeper reasons behind the failure of impression management and social defence are mostly left unexplained.

To summarise, previous scholars have produced a large body of work on human social defence and on criticisms of algorithmic surveillance, but two main research gaps still exist.

Existing studies have seldom compared the human gaze and the algorithmic gaze within one single framework; this makes it hard to analyse how asymmetric mechanisms set off social and emotional responses. They have also not looked at the structural effects of being exposed to such environments over a long period; in psychology, the habituation effect [15] says that prolonged exposure to a specific stimulus raises the perceptual threshold and lowers sensitivity. In the context of digital surveillance being everywhere, this suggests that the constant presence of cameras may work as a steady, habituating stimulus; however, most current literature still remains at the level of describing how algorithms take away subjectivity in practice. It is still not addressed whether chronic exposure to a bodiless algorithmic gaze, which has no social evaluation qualities, can make individuals' social perception in public spaces become perceptually dulled and structurally changed; existing research either explains the social results of the human gaze or the structural features of algorithmic surveillance, but it has paid little attention to the mental process that connects the two. To fill this gap, this paper connects the type of observer to the activation or failure of social defence, and it puts forward the Gaze Asymmetry Framework; this framework discusses why psychological defence mechanisms do not work under the algorithmic gaze and what this failure means for human social perception.

DOI: 10.54254/2753-7064/2026.37216

3. Theoretical framework

This framework is a theoretical analytical tool based on the meeting point of cognitive psychology and micro-sociology, and its main job is to compare and analyse the structural differences, or asymmetry, between the human gaze and the algorithmic gaze in turning on individuals' social defence mechanisms. To give a clear overall picture, the framework is made up of three connected conceptual layers: one layer contrasts the conditions that set things off, where cues of subjectivity are present or absent; another layer looks at the mental processes that follow, where the 'imagined audience' is either successfully formed or fails to appear; and a further layer considers the behavioural and long-term results, where social defences are turned on, or defence fails and perceptual dulling takes place. The main theoretical contribution of the paper is the idea that perceived subjectivity and evaluative capacity are a key condition for starting social defence mechanisms, and here subjectivity attribution means the mental process by which a person believes that an observer has social agency and the ability to judge. The three-layer structure of the proposed framework is shown in Figure 1.

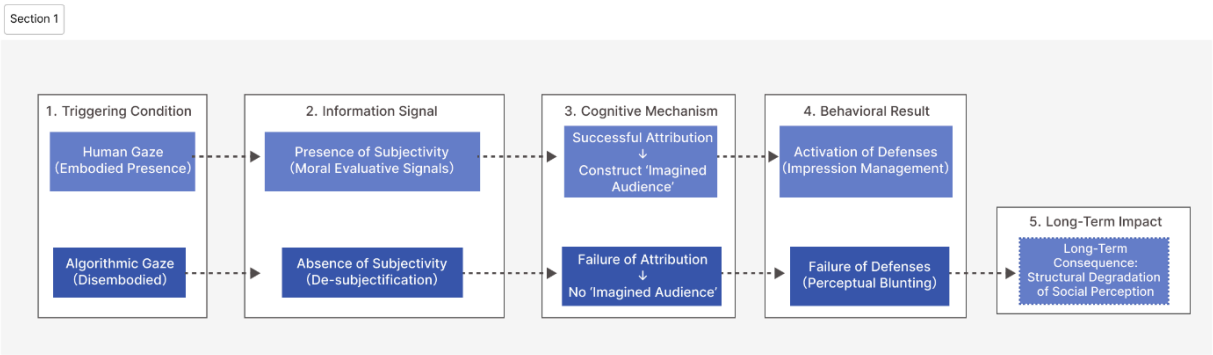

Figure 1. The three-layer structure of the Gaze Asymmetry Framework

The logical starting point of this framework lies in the defence-triggering mechanisms of traditional embodied interactions; in this context, subjectivity attribution refers to the mental process by which individuals pick up physical communicative signals, such as eye contact and facial expressions, and combine them into one overall signal showing that the observer has social ability and moral judgement; following Goffman's [2] dramaturgical theory, when people perceive this moral checking, social anxiety is generated, and this anxiety pushes them to manage their front-stage impressions so as to actively regulate social boundaries; therefore, recognising the moral-evaluative capacity of the watching subject is the necessary condition for activating traditional defence mechanisms; this condition is closely tied to the mental creation of an "imagined audience" [10]; as a result, an important theoretical point appears: if an individual meets a gaze where these subjectivity cues are missing, the ability to project and form this imagined audience breaks down; this failure of construction basically stops the mental basis needed to trigger any social defence.

Within digital surveillance, the second layer of this framework looks at the asymmetric ways in which the algorithmic gaze operates; when a machine is doing the watching, the cues of subjectivity that people need to turn on their defences are taken away in two related ways; one way is that physical emotion is removed, because cameras produce a "gaze without eyes" that takes away eye contact and emotional expressions; another way is that the system removes subjectivity, since algorithms turn the real person into a bodiless "data double" [13] and hide the direct pressure of

DOI: 10.54254/2753-7064/2026.37216

panoptic watching; these two kinds of loss together lead to a deep failure in how people attribute subjectivity; since the algorithm has neither emotional feeling nor a systemic moral background, the person cannot mentally see the machine as a real social being; this early failure of attribution then leads straight to the next mental stage, where the building of the "imagined audience" [10] fails; because the algorithm naturally has no social qualities or moral ability, people have no human-like features to place onto the watcher; as a result, the person cannot build an imagined audience that feels morally watchful in the mind; this missing part is thought to stop the defensive reflex, since without an imagined audience the person expects no social check, moral judgement, or risk from others; the mental pressure and possible fear of social punishment that would push a person into a front-stage act disappear; in the end, the person loses the inner reason to actively manage social boundaries, and they settle into an unguarded, relaxed condition when dealing with algorithmic watching; this mental breakdown shows up as a complete failure of defence.

Based on the differences in how the two types of gaze trigger defences, the last part of this framework (Layer 3) works out the wider structural effects of being exposed for a long time to the bodiless algorithmic gaze; when a person is constantly placed in digital surveillance environments, such as networks of visual monitoring cameras, the mental impact goes beyond a simple description of experience and moves through a particular mental chain; the early failure of defence under algorithmic watching may lead to repeated non-use of traditional impression management and defensive practices; drawing on the "Disuse" mechanism [16], this long-term lack of practice weakens the person's sensitivity to cues of subjectivity; importantly, this repeated non-use does have effects; over time, the absence of mental construction and defensive practice turns into a settled state of being unguarded. To explain how this habit carries back into human society, the framework brings in the behavioural mindset theory by Wyer & Xu [17]; this theory shows that mental or behavioural steps often used in one setting can join together into a general behavioural mindset; driven by this mindset, a transfer across situations takes place; the habitual mental process of weakened surveillance perception, first created in human-machine interaction, does not stay within that boundary; rather, it is used automatically when the person meets later, quite different situations, namely daily face-to-face contact in physical public spaces; this cross-context carry-over effect points towards a possible structural reshaping of human thinking, which may lead to an overall weakening of social perception; because people habitually operate with unclear defensive boundaries, their social trust network slowly erodes, and this then shows up as a reduced ability to read emotions and a loss of proper social distance in everyday interpersonal contact.

4. Discussion

To sum up, the Gaze Asymmetry Framework is proposed in this study, and it aims to explain the differences in thinking and behaviour that appear when people are watched by humans instead of by algorithms; the importance of this framework in theory comes from the fact that most earlier research on surveillance has concentrated on the controlling effects of physical devices or the large-scale process of turning people into data, but this research has not answered a key question clearly, and that question is how watching that has no body can switch off a person's mental defences; when "subjectivity attribution" is placed at the centre, the framework brings together small-scale social contact and digital watching, and it shows that a breakdown in the mental construction of an imagined audience is the basic mental mechanism that leads to defence failure, which gives a new analytical tool that previous surveillance studies have not paid attention to.

An interactive installation called 'Gaze Reverse' [18] can work as a practical extension and a way to explain this framework through real use; this installation is not just a design exercise, but it works

DOI: 10.54254/2753-7064/2026.37216

like a thinking space where the main parts of the framework can be changed and looked at; when people first walk into the space and come across hidden cameras, they do not feel psychologically guarded, and this matches exactly with the second layer of the framework, because algorithmic watching has no signs of subjectivity, so the 'imagined audience' is not triggered and defence mechanisms may not work properly; conversely, when the installation clearly shows images of crowds looking back at the participants, it adds back 'subjectivity signals' on purpose, and this immediately causes mental discomfort and defensive reactions; this quick change in behaviour shows in a practical way the main idea of the first layer of the framework, which is that seeing subjectivity and moral judgement ability in the watcher is the necessary condition for turning on traditional social defence mechanisms; this installation, as shown in Figure 2, is offered as an example of how the framework works, not as proof that its proposed mechanisms are true.

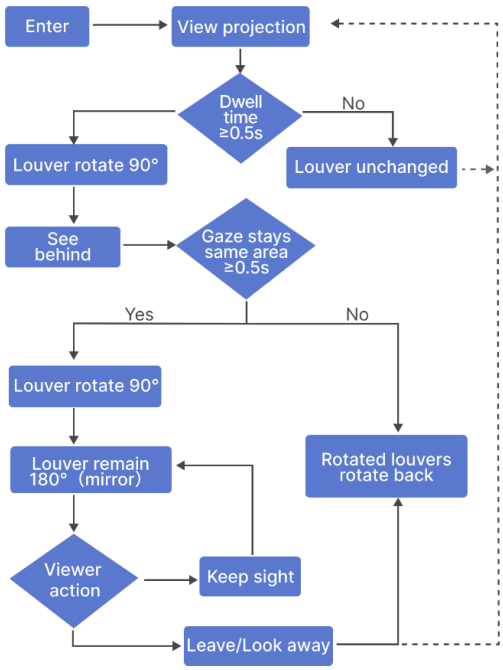

Figure 2. "Gaze Reverse" interactive installation demonstrating the framework's first and second layers

Although this framework mostly works at a theoretical level, it still gives a useful conceptual way to think about the wider structural results of modern surveillance; the theoretical reasoning suggests that spending a long time in algorithmic settings with no clear watching subject could lead to mental dulling and a weakening of social perception; to keep the study academically careful, though, clear theoretical limits need to be set; the present framework focuses on visual monitoring that is everywhere in physical public spaces, such as CCTV and security cameras; it suggests that these effects are stronger when the watching feels new or personally aimed at someone, and they get weaker as the watching becomes a normal part of the background, which matches the getting-used-to process described in Layer 3; the framework does not fully cover the algorithmic watching that is built into bigger digital systems, for example platform recommendation systems run by AI or short-video algorithms; the cognitive spillover effect discussed in this paper gives a possible way to look

DOI: 10.54254/2753-7064/2026.37216

at these things, but the strongly tailored workings of platform algorithms need their own separate theoretical explanation.

This study recognises some limitations despite its conceptual contributions; one limitation is that the third layer of the framework, which concerns the structural weakening of social perception, comes mostly from theoretical reasoning rather than systematic empirical research, so future work should combine quantitative tracking with qualitative methods; given the sociological nature of this question, future studies need to include strong qualitative approaches such as in-depth interviews and online fieldwork, because these can capture the subtle shifts in people's lived experiences, trust networks, and social boundaries within digital surveillance settings; another limitation is that the micro-sociological and psychological theories used in this paper are based on particular cultural contexts, and under different cultural patterns, public tolerance of privacy and social norms varies a great deal, which means the cross-cultural usefulness of these defence mechanisms still needs further study.

5. Conclusion

In conclusion, this study puts forward the Gaze Asymmetry Framework to explain how the change from being watched by a person with a body to being watched by algorithmic surveillance without a body may affect people's social perception and defence mechanisms; compared with older theories of constant watching, such as Foucault's, this research has an important advantage, because it does not only look at large-scale social control or the abstract idea of the "data double", but instead places "subjectivity attribution" at the centre of the analysis; this perspective gives a unified mental explanation for why defence fails, and it shows that when signs of subjectivity are absent, the construction of an "imagined audience" is conceptually stopped; going beyond immediate behavioural reactions, the framework turns these processes into a model of spillover across contexts, which suggests that the unguarded mindset formed during human-machine contact may slowly reduce people's sensitivity to the real human gaze, and in this way the social and mental boundaries and the basic structures of human interaction are reshaped.

Even though this framework has some ability to explain things in theory, its use is limited in certain ways; it is mainly made for visual watching tools such as security cameras in physical public spaces, and putting it directly onto hidden back-end AI systems or platform recommendation algorithms would need careful adjustment; to go beyond reasoning alone, future research has to test these mechanisms with real data through long-term tracking and experiments; future empirical work is expected to improve, widen, or even question this framework as it is used across different technological systems and different cultural settings, which keeps an open and changing discussion in surveillance studies.

References

[1] Goffman, E. (1963). *Behavior in public places*. Free Press.
[2] Goffman, E. (1959). *The presentation of self in everyday life*. Doubleday Anchor Books.
[3] Foucault, M. (1975). *Surveiller et punir: Naissance de la prison*. Gallimard.
[4] Nass, C., & Moon, Y. (2002). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153
[5] Lyon, D. (2014). Surveillance, Snowden, and big data: Capacities, consequences, critique. *Big Data & Society, 1*(2).
[6] Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews, 24*(6), 581–604. https://doi.org/10.1016/s0149-7634(00)00025-7
[7] Sartre, J.-P. (1943). *L'être et le néant: Essai d'ontologie phénoménologique*. Gallimard.

DOI: 10.54254/2753-7064/2026.37216

[8] Argyle, M., & Dean, J. (1965). Eye-contact, distance and affiliation. *Sociometry, 28*(3), 289. https://doi.org/10.2307/2786027
[9] Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology, 30*(4), 526–537. https://doi.org/10.1037/h0037039
[10] Marwick, A. E., & Boyd, D. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society, 13*(1), 114–133. https://doi.org/10.1177/1461444810365313
[11] Koskela, H. (2000). 'The gaze without eyes': Video-surveillance and the changing nature of urban space. *Progress in Human Geography, 24*(2), 243–265.
[12] Kemp, N. J., & Rutter, D. R. (1982). Cuelessness and the content and style of conversation. *British Journal of Social Psychology, 21*(1), 43–49.
[13] Haggerty, K. D., & Ericson, R. V. (2000). The surveillant assemblage. *British Journal of Sociology, 51*(4), 605–622. https://doi.org/10.1080/00071310020015280
[14] Burrell, J. (2016). How the machine "thinks": Understanding opacity in machine learning algorithms. *Big Data & Society, 3*(1), 2053951715622512.
[15] James, W. (1890). *The principles of psychology (Vol. 1)*. Henry Holt.
[16] Thorndike, E. L. (1913). The psychology of learning. In *Educational psychology: Vol. 2*. Teachers College.
[17] Wyer, R. S., & Xu, A. J. (2010). The role of behavioral mind-sets in goal-directed activity: Conceptual underpinnings and empirical evidence. *Journal of Consumer Psychology, 20*(2), 107–125. https://doi.org/10.1016/j.jcps.2010.01.003
[18] Shen, E. N., Zhao, X. Y., & Tian, H. Y. (2025). Gaze Reverse [Interactive installation]. Human-Machine Co-Shaping, Sphinx International Art Education, Shanghai, China.