Uniglobal of Journal Social Sciences and Humanities
Journal Homepage: www.ujssh.com

Explanation Is Not Redress: A Critical Integrative Review of Fairness in Algorithmic Platform Work

Gao, Yang[1*] & Mohd Shukri, Amer Fareed[1]

[1]SEGi University, Kota Damansara, 47810 Petaling Jaya, Selangor, Malaysia

*Corresponding author email: 1620014682@qq.com

To link to this article: https://doi.org/10.53797/ujssh.v5i2.22.2026

Abstract: Algorithmic platform work raises two distinct governance questions: whether workers can understand a decision and whether they can challenge and correct it. This critical integrative review examines how explanations, human review and contestability relate to procedural justice. A purposive, source-checked synthesis brings together 34 substantive publications spanning organisational justice, platform labour and human–computer interaction, supported by two methodological sources. Evidence was differentiated by design, decision setting, outcome and source-access depth rather than pooled as a single intervention effect. The synthesis identifies three recurring problems. Informational and procedural fairness are treated as interchangeable, and preferences for review are mistaken for effective redress. Findings from simulated decisions are also transferred to consequential platform work without testing the relevant conditions. Experimental evidence supports context-specific benefits of explanations, but not a universal transparency effect or an established explanation–appeal synergy. Qualitative and organisational research places these responses within continuing relationships of control, dependence and collective voice. The article develops a three-capability framework separating intelligibility, contestability and effective redress, together with an evidence-to-outcome matrix for evaluating each capability. Malaysian platform work, now governed by the Gig Workers Act 2025 and its grievance, conciliation and tribunal routes, provides an application context, not a population for which treatment effects are claimed. The central implication is that favourable fairness ratings should complement, rather than replace, evidence of accessible challenge, authorised review and justified correction.

Keywords: Algorithmic management, procedural justice, explainability, contestability, platform work

1. Introduction

A platform can explain an adverse decision without allowing a worker to challenge it. It can also display a review option without giving the reviewer sufficient authority to change an error. These possibilities make explanation, appeal and redress analytically distinct. For a worker who loses access to a task or receives an unfavourable evaluation, understanding what happened is not the same achievement as obtaining a justified reconsideration. The central question of this review is therefore not whether transparency is desirable in the abstract, but what different governance features enable and what evidence establishes their value.

Algorithmic management is embedded in the organisation of work, not confined to an isolated prediction. Lee, Kusbit, Metsky, and Dabbish (2015) describe drivers' responses to automated assignment and evaluation, including social sensemaking through online forums. Rosenblat and Stark (2016) connect information asymmetries to the indirect control of drivers. At a broader level, Kellogg, Valentine, and Christin (2020) integrate direction, evaluation and discipline within an account of algorithmic control and resistance. Political theory sharpens the stakes. Muldoon and Raekstad (2023) argue that algorithmic systems in ride-hailing and food delivery can sustain relationships of domination, partly because consequential decisions such as account deactivation may be imposed with no meaningful route to contest them. Together, these perspectives direct attention to the relationship in which a decision is communicated, rather than treating the notice as a neutral container for information.

The relevant literatures are already substantial. Starke, Baleis, Keller, and Marcinkowski (2022) systematically reviewed 58 empirical studies of algorithmic fairness perceptions and documented heterogeneity in concepts, measurements and settings. Kadolkar, Kepes, and Subramony (2025) integrate research on algorithmic management in the gig economy, while Keegan and Meijerink (2025) connect it to organisational psychology, organisational behaviour and human resource management. This article does not claim to introduce either algorithmic fairness or contestability. Its

**Corresponding author: 1620014682@qq.com*
https://ujssh.com/ All right reserved.

narrower contribution is to align the evidence supporting worker-facing governance features with the particular outcomes those features are supposed to deliver. The question has also become a regulatory one in the review's application context. Malaysia's Gig Workers Act 2025 (Act 872) came into force on 31 March 2026. It introduced an internal grievance mechanism, conciliation and a Gig Workers Tribunal for disputes that include account suspension (Bernama, 2026; Ius Laboris, 2026). Formal routes to challenge now exist on paper; whether they deliver intelligible reasons and justified correction is precisely the evidential question this review addresses.

That alignment matters because apparently similar findings can answer different questions. A recruitment vignette may establish a change in perceived informational fairness; a repeated-choice experiment may identify preferences among appeal procedures; interviews may reveal why people need support to contest a decision. None of these observations alone establishes that an operational system identifies errors or delivers an appropriate remedy. Conversely, a review procedure need not maximise favourable ratings to be valuable: a justified refusal may remain unpopular. The synthesis therefore treats perceived fairness as an important outcome in its own right, but not as a substitute for institutional performance.

Three questions organise the review. First, how do explanations, human review and contestability differ conceptually and in their measured effects? Second, what does the available evidence support about their independent and combined contributions to fairness? Third, what changes when these findings are applied to recurrent, economically consequential platform work, including the Malaysian context? The resulting framework distinguishes intelligibility, contestability and effective redress. It also separates functional complementarity between governance capabilities from a statistically demonstrated positive interaction between experimental treatments. This distinction prevents an attractive institutional argument from being reported as an established behavioural effect.

The article proceeds from an explicit review method to a synthesis of experimental, worker-centred and conceptual evidence. It then develops an evidence-to-outcome framework, practical evaluation criteria and a focused research agenda. The unit of contribution is the critical integration of published evidence; no new worker sample or experiment is presented.

2. Review Methodology

2.1 Review purpose and scope

A critical integrative design was selected because the question crosses empirical and conceptual traditions whose outcomes are not directly interchangeable. Snyder (2019) distinguishes review designs according to their purpose, and Torraco (2016) emphasises the reasoning through which critique becomes synthesis. Accordingly, this review sought a defensible conceptual integration rather than an exhaustive inventory or a pooled average effect. It is not presented as a systematic review, meta-analysis or preregistered evidence synthesis.

The final reference base contains 36 scholarly publications: 34 substantive works and two review-methodology papers. Two further non-scholarly sources describe Malaysian legislation and are used for context only. The substantive selection combines organisational-justice foundations, studies of algorithmic platform work, experiments on decision explanations and review, and conceptual analyses of contestability. Most focal digital-work and decision-design publications date from 2015–2026, and 24 of the 36 scholarly publications (67%) appeared in 2022 or later; Colquitt's (2001) construct-validation paper supplies an earlier theoretical anchor. The counts describe the selected reference base, not the number of records retrieved, screened or excluded in a systematic search.

2.2 Retrieval, selection and source access

Targeted web retrieval and source checking were completed on 8 October 2026. References in the supplied research protocol served as discovery seeds, not as verified evidence. Additional retrieval combined algorithmic management, fairness, transparency, explanations, contestability, human review, appeal and platform-worker terms; a separate search addressed Malaysian food-delivery contestation. Exact-title and DOI look-ups resolved primary publications and available author or institutional copies. A supplementary benchmark search on 22 November 2026 targeted journals of six major publishers (Elsevier, Wiley, Springer Nature, Sage, Emerald and Taylor & Francis) to identify recent field and longitudinal studies that the original retrieval had missed. Ten publications were added at this stage, and every bibliographic record was re-resolved against Crossref. Representative searches and the extraction fields appear in Appendix A. No subscription-database search, comprehensive citation export or bibliometric coverage is claimed.

Publications were retained purposively when they clarified a central construct, directly examined a relevant decision feature, supplied worker-context evidence, or exposed a consequential inference boundary. General technical explainability papers without a decision-subject or governance connection were outside the focal synthesis. Non-scholarly advocacy and news were not used as treatment-effect evidence. Selection continued until the focal distinctions could be compared across experimental, organisational and contestability research; this was an analytical stopping decision, not a claim that the literature had been exhausted.

Source access was recorded rather than flattened into a claim of uniform full-text review. Relevant original sections were inspected for 19 publications, including the two methodological sources. Seventeen publications, including the ten

added in the supplementary search, were used at abstract level for background concepts or contextual findings, not for detailed assessments of experimental effects. Detailed comparisons of the principal decision-feature experiments rely on original methods and results, not search snippets. Abstract-level sources support the stated design, context or broad finding only. The accompanying source registry identifies the access route, locator, permitted claim and limitation for every reference.

2.3 Extraction, comparison and critical synthesis

Extraction recorded the decision setting, participant or documentary population, design, manipulated or observed feature, fairness dimension, principal finding and inference boundary. The central comparison asked whether a study measured understanding, perceived treatment, a preference for review, actual contestation, or realised correction. Experiments were assessed for construct alignment and the scope of randomisation; survey findings were treated as associations; qualitative and conceptual work was used to interpret context and specify governance requirements. The review did not assign numerical quality scores or combine dissimilar study designs into one evidential ranking.

An AI-assisted analytical pass compared the extracted claims and was followed by a source-and-manuscript consistency check. There was no independent second-reviewer screening or inter-rater reliability calculation. Contradictory and null results were retained, especially where they challenged a universal transparency benefit or presumed explanation–appeal synergy. Existing reviews supplied conceptual orientation, but their constituent studies were not counted as additional independently inspected evidence. The three-capability framework was developed by comparing what each type of evidence could establish with what the proposed governance outcome would require. It is an interpretive synthesis, not an empirically validated causal model.

3. Results of the Evidence Synthesis

3.1 Fairness is not a single interchangeable outcome

Organisational justice provides an essential starting distinction. Colquitt (2001) differentiates procedural, distributive, interpersonal and informational justice. For the present question, these concern how a decision is made, how benefits and burdens are allocated, how a person is treated, and the adequacy of information supplied. A specific explanation may address information while leaving the allocation rule unchanged. A review route may increase the opportunity to challenge the procedure without making the original reason clearer. Treating either change as an improvement in a single undifferentiated fairness construct obscures what has actually changed.

The distinction extends to the outcome against which a governance feature is evaluated. Understanding, perceived fairness, trust, acceptance and successful correction are not equivalent endpoints. Table 1 translates this observation into an evidence map. Its categories are analytical distinctions for this review, not a new psychometric scale. In particular, effective redress means a justified corrective response where the original decision, supporting data or procedure warrants correction; it does not mean that every appeal should be upheld.

Table 1. Concepts, appropriate evidence and invalid substitutions.

Concept	Evidence that addresses the concept	What should not stand in for it
Intelligibility	The affected worker can identify the relevant reason, data and decision criterion.	Disclosure volume or a claim that the algorithm is transparent.
Perceived procedural fairness	A distinct measure of the worker's judgement of the decision process.	Informational satisfaction or liking the allocation outcome.
Contestability	A usable route for questioning information, application or the underlying rule.	A human-supervision label or a generic contact button.
Effective redress	Competent review with authority and a justified response to a demonstrated error or defect.	Stated appeal availability, preference for a reviewer or an unqualified reversal rate.
Worker consequences	Observed work access, income consequences and continued participation over a defined period.	Hypothetical task acceptance or a single continuation-intention item.

Source: author synthesis. These distinctions organise evidence; they do not constitute a validated measurement instrument.

3.2 Explanations have conditional, dimension-specific benefits

The experimental evidence does not support the proposition that any additional explanation improves every fairness dimension. Binns et al. (2018) report three studies in which explanation format and presentation context shape justice judgements. Ochmann, Michels, Tiefenbeck, Maier, and Laumer (2024), using 801 participants in simulated recruiting, find informational-fairness benefits of transparency without corresponding procedural or distributive benefits. The latter result is particularly important for work research: a better account of why a decision occurred is not automatically perceived as a better procedure for reaching it.

Mirbabaie, Langer, Rieskamp, and Hofeditz (2026) examine four transparency conditions with 234 analysed participants on a simulated study-allocation platform. Distributive transparency improved distributive-fairness ratings, but the study did not establish an overall fairness advantage across transparency conditions. An exploratory pairwise informational-fairness difference for the combined condition should not be confused with a supported omnibus informational effect. This was a Prolific-based prototype experiment, not a field intervention among delivery workers. Its contribution is evidence against a simple monotonic transparency assumption, not proof that explanations are generally ineffective.

A contrasting result comes from Henning and Langenbach (2026). In public-administration vignettes with 4,250 participants, individualised reasons improved perceived fairness and reduced the human–AI fairness gap. Their principal presentation combines outcome and procedural ratings in an index that was not preregistered, although separate measures were also examined. This stronger positive evidence argues against replacing an unqualified transparency optimism with an equally unqualified scepticism. It still concerns reactions to described decisions, not actual administrative correction or platform-worker outcomes.

Field and longitudinal studies of workers add an important qualification. In a survey of truck drivers, Bujold, Parent-Rocheleau and Gaudet (2022) found that transparency of algorithmic surveillance was associated with procedural justice, whereas transparency of algorithmic performance management was associated with distributive justice. This dimension-specific pattern parallels the experimental results. Using a longitudinal design with gig workers, Hu, Zeng, Wang and Teng (2024) reported an inverted-U relationship. Greater transparency at first accompanied higher fairness perceptions and lower resistance, but beyond a point it was associated with more resistance. Wesche et al. (2024) likewise found that explanations improved trust and acceptance in a hypothetical selection study but not in a second study with real incentives. Grimmelikhuijsen (2023) showed that the effect of explainability on the perceived trustworthiness of automated decisions differed between low- and high-discretion administrative scenarios. These studies do not test redress. They do, however, make an unqualified “more transparency is better” assumption difficult to sustain even outside the laboratory.

Taken together, these studies suggest that the decision criterion explained, the degree of individualisation, the comparison condition and the outcome measure all matter. They do not establish a common effect size that can be transported to platform dispatch. The relevant design question is therefore precise: what does a particular explanation enable a particular decision subject to understand or do? This formulation retains the potential value of explanation while preventing a change in informational satisfaction from being used as evidence for a remedy that was never tested.

3.3 Contestability is distinct from human oversight

Yurrita et al. (2023) offer a direct experimental separation of explanations, human oversight and contestability in loan decisions, with 267 participants. Explanations and contestability relate to different fairness dimensions, while the reported procedural-fairness analysis does not support an explanation-by-contestability interaction. Human oversight is not equivalent to an actionable challenge route. The study therefore supports distinguishing the features, but it does not justify reporting their combination as an established synergistic treatment. A feature can have a useful separate contribution without multiplying the effect of another feature.

Lyons, Wijenayake, Miller, and Velloso (2022) examine preferences among review procedures. Their 100 participants made 1,200 repeated choices; the choices must not be treated as 1,200 independent people. Human review, shorter delays and the opportunity to provide new information were preferred. Such findings identify desirable procedural attributes, but preference may partly reflect expectations about obtaining a favourable outcome. They do not demonstrate that human reviewers are more accurate, independent or able to deliver an appropriate remedy in practice.

Documentary and qualitative work broadens this picture. Lyons, Velloso, and Miller (2021) analyse submissions to Australia’s proposed AI Ethics Framework and find that contestability is recognised as a safeguard but understood in different ways. Yurrita et al. (2025), drawing on 21 interviews about an anticipated rental-detection system, identify cooperative sensemaking, social support in contestation and distributed responsibility before deployment. The latter participants were not platform workers, and the study did not evaluate successful appeals against a deployed system. Nevertheless, it identifies forms of assistance that an isolated interface button cannot itself supply.

The combined implication is that “human in the loop” is an underspecified governance description. A supervisor who checks aggregate performance, a reviewer who hears an individual challenge and an authority who can change a defective rule occupy different roles. The first does not logically entail the second, and the second does not guarantee the third.

Table 2 summarises the central empirical comparisons without collapsing them into a single effect or counting repeated responses as additional participants.

Table 2. Core empirical comparisons and their evidential boundaries.

Study and setting	Design or sample	Relevant result	Boundary
Yurrita et al. (2023): loans	Factorial vignette; 267 participants	Separates explanation, oversight and contestability; no supported explanation–contestability interaction.	Perception, not operational redress.
Lyons et al. (2022): loan review	Conjoint choices; 100 people, 1,200 choices	Human review, less delay and new information were preferred.	Repeated preferences, not review accuracy.
Ochmann et al. (2024): recruiting	Simulation; 801 participants	Transparency benefited informational fairness, not all fairness dimensions.	Not real hiring or dispatch.
Mirbabaie et al. (2026): study allocation	Prototype; 234 analysed participants	Dimension-specific effects; no supported overall fairness benefit.	Prolific setting; exploratory contrasts need qualification.
Henning and Langenbach (2026): public administration	Vignettes; 4,250 participants	Individualised reasons improved perceived fairness.	Combined index not preregistered; no realised redress.
Jabagi et al. (2025): Uber drivers	Survey; 435 drivers	Fairness perceptions were associated with worker-attitude outcomes.	Associations; abstract-level use here.
Cameron (2024): ride-hailing	Longitudinal qualitative study	Worker choice coexists with continued managerial control.	Contextual explanation, not a treatment effect.
Yurrita et al. (2025): rental detection	Interviews; 21 participants	Contestability involved cooperation, support and shared responsibility.	Anticipated system; no observed appeal outcomes.

Source: the cited original publications. Sample numbers belong to those studies, not to new data collected for this review.

3.4 Platform work changes the relevant relationship and time horizon

Worker-centred research explains why a one-off decision experiment cannot settle the institutional question. Möhlmann, Zalmanson, Henfridsson, and Gregory (2021) locate tensions in the combination of market matching and organisational control. Duggan, Sherman, Carbery, and McDonnell (2020) situate app-work within employment relations and distinguish it from other platform arrangements. A worker's encounter with an allocation notice can therefore be part of an ongoing managerial relationship, rather than an isolated consumer choice between interfaces.

Cameron's (2024) longitudinal qualitative study shows how constant but confined choices can sustain workers' participation within algorithmic control. Wood, Graham, Lehdonvirta, and Hjorth (2019), studying remote platform work, describe flexibility alongside constraints and adverse job-quality features. Rani and Furrer (2021) analyse work allocation, remuneration and working-time issues among 675 workers across five online platforms in 27 developing countries. These settings are not interchangeable with Malaysian ride-hailing. Their common relevance is to show why continued participation, flexibility or the availability of some choices cannot alone certify fair governance.

Jabagi, Croteau, Audebrand, and Marsan (2025) connect fairness perceptions of algorithmic human-resource decisions to perceived support and job satisfaction in a survey of 435 Uber drivers. This provides a worker-specific bridge to organisational justice. It does not identify whether explanations or appeal mechanisms caused those outcomes. The worker-specific association is relevant to the review, but it cannot fill the institutional gap between perceived treatment and an observed remedy. Three-wave data from 435 European crowdworkers show a related pathway. Perceived algorithmic surveillance undermined trust and fairness and increased privacy concerns, and these in turn informed intentions to comply with, alter or resist surveillance (van Zoonen, von Bonsdorff, & van der Heijden, 2026). Zhou, Lei, Cooke, Huang and Zhang (2025) similarly show that the same algorithmic practices can be attributed by platform workers either to the platform's commitment or to its control, with different consequences. Where formal channels are weak, contestation is therefore likely to move into workarounds and collective "algoactivism". Dasgupta, Carbery, McDonnell and Jooss (2026) examine this from the platform's side as a conflict-management problem, rather than as worker deviance alone.

The synthesis consequently distinguishes an immediate notice response from a cumulative governance experience. A clear explanation today may be evaluated differently after repeated unsuccessful reviews. An appeal may be formally

available but practically costly because it competes with working time. These are plausible transportability conditions derived from the relationship between the literatures, not measured effects established by this review. They imply that evidence should follow workers across decision, challenge, review and subsequent work, rather than ending at the moment a notice is rated.

3.5 Malaysian evidence concerns collective contestation, not treatment effects

Uchiyama, Furuoka, and Omar (2025) compare Malaysian food-delivery blackouts in 2022 and 2024, examining the differing continuity and robustness of collective contestation. The publisher abstract supports the relevance of organised worker responses, but not a causal estimate of explanations, appeal access or successful correction. This evidence should neither be ignored because it lacks an interface experiment nor be recast as evidence for the effects of one. It locates a collective dimension that a purely individual fairness questionnaire would miss.

For Malaysian application, the appropriate inference is a set of questions rather than a transplanted coefficient. Does the worker understand the reason in the language used? Can a challenge be made without jeopardising access to work? Can a representative help identify a recurring problem? Does review address a single case, inaccurate data or a contested allocation rule? None of these conditions is established merely by naming Malaysia as the setting. The available sources motivate locally grounded validation while leaving the relative effectiveness of explanation and appeal among Malaysian platform workers unresolved. The Gig Workers Act 2025 changes the practical baseline without answering that question. Reported provisions include written notice of account suspension with an investigation period of no more than 14 days, and partial compensation where a deactivated worker is cleared. Disputes pass through a 30-day internal grievance stage, conciliation through the Industrial Relations Department and a Gig Workers Tribunal whose hearings are public (Bernama, 2026; Ius Laboris, 2026). Commentary on the Act also describes transparency and human-review rights for automated monitoring and decision-making systems (Ius Laboris, 2026). These are contestability and redress provisions in the sense used here. Their existence is not evidence that they are accessible, that reviewers have authority over the disputed data or rule, or that corrections follow. The Act therefore creates an unusually clear opportunity to observe the transition from formal channel to realised remedy.

4. Integrative Framework and Theoretical Discussion

4.1 Three separately auditable governance capabilities

The synthesis yields a three-capability framework: intelligibility, contestability and effective redress. Its starting point is consistent with Alfrink, Keller, Kortuem, and Doorn's (2023) account of contestable systems across their lifecycle and Kinchin's (2024) critique of procedural voice reduced to fairness perceptions. Bujold, Gaudet and Parent-Rocheleau (2026) similarly treat "correctability", the perceived opportunity to appeal and correct erroneous algorithmic procedures, as a distinct cue in workers' justice evaluations. The present extension is an evidence-to-outcome discipline for platform work: each claimed governance capability must be supported by evidence that actually observes that capability. The framework is not a sequential mediation model, a maturity score or a claim that these distinctions have never appeared in earlier scholarship.

Intelligibility concerns the worker's ability to identify the basis of the decision and the information relevant to it. It is a practical achievement, not a count of explanatory sentences. For an illustrative dispatch case, a proximity-based reason may clarify the allocation while leaving open whether location data were accurate or the rule was appropriate. The first issue concerns understanding, the second evidential accuracy and the third justification. An interface that combines them into a single reassurance message prevents both the worker and the evaluator from identifying which question has been answered.

Contestability concerns the practical opportunity to raise those questions and introduce relevant information. It includes a discoverable channel, an understandable process and a route to someone able to consider the challenge. The capability should not depend on a worker already possessing a technical account of the algorithm: an accessible process can begin with a request for clarification. Explanations and challenge can therefore be reciprocal. The framework does not require every worker to understand a complete model before a concern may be heard. Nor does it assume that an initial explanation covers everything the reviewer may need to investigate.

Effective redress concerns what follows a warranted challenge. The responsible organisation must be able to examine evidence, identify the relevant decision or process defect, and implement an appropriate response. Where a defect is established, responses can include reversal, data correction, compensation for a verified loss or revision of a defective rule. Otherwise, a reasoned affirmation may be appropriate. These examples specify possible evaluation targets; they are not descriptions of remedies currently offered by a particular platform. A service that closes every ticket quickly may still fail this capability when closure is unrelated to the merits of the challenge.

4.2 Functional complementarity is not statistical synergy

Two capabilities can complement each other functionally without producing a positive statistical interaction on a rating scale. An explanation may make a concern easier to articulate; a review route may give the explanation a practical use. This is an institutional argument about their roles. Statistical synergy is a different claim: the effect of offering both features exceeds the sum of their separate effects on a specified outcome and scale. The former cannot be established or rejected solely by the latter, and evidence of a marginal treatment effect does not establish an interaction.

The absence of a supported interaction in Yurrita et al.'s (2023) loan experiment makes this distinction consequential rather than semantic. It is possible that separate capabilities matter without amplifying one another's immediate perceptual effects. It is also possible that effects differ across decisions, populations or measures. Choosing among those possibilities requires evidence; a negative or null interaction should not be edited out of a synthesis because a complementary governance architecture is normatively attractive. Table 3 describes four interface configurations while deliberately withholding any claim that their effects follow a universal rank order.

Table 3. Explanation–appeal configurations under a separate redress condition.

Configuration	Capability made visible	Unresolved institutional question
Neither explanation nor appeal	Outcome notification only.	How can the worker discover or challenge an error?
Explanation without appeal	A reason may be intelligible.	Can an inaccurate premise or inappropriate rule be questioned?
Appeal without explanation	A challenge can begin through a review route.	Can the worker obtain enough information to formulate the concern?
Explanation and appeal	Understanding and challenge can support each other.	Can authorised review deliver a justified response or correction?

Author synthesis, not an empirical treatment ranking. Redress authority and implementation must be examined in every configuration; displaying both features does not establish them.

4.3 Avoiding outcome substitution

The principal cross-literature risk is outcome substitution: evidence for one endpoint is used to support a different institutional claim. An informational-fairness effect becomes a claim that the allocation process is fair. A preferred human-review option becomes a claim that review works, and continued participation becomes a claim that the worker has endorsed the rules. These substitutions are tempting because the proxy is easier to observe than the outcome of interest. They are not repaired by using a larger sample unless the study also measures the missing outcome.

The framework does not devalue perceptions. Workers' accounts of dignity, clarity and procedural treatment remain important even when the original decision is substantively correct. Rather, it requires parallel evaluation. Perceptions answer how the process is experienced. Operational records answer whether the channel can be used, whether review addresses the concern and whether a justified correction follows. Independent standards are still needed to judge the allocation rule itself. Agreement among these sources would strengthen a governance claim; disagreement would identify a problem worth investigating rather than an inconvenience to average away.

This approach also clarifies the contribution beyond existing reviews. Broad algorithmic-management reviews organise practices and worker responses; fairness-perception reviews map predictors and outcomes; contestability frameworks describe opportunities for intervention (Alfrink et al., 2023; Kadolkar et al., 2025; Starke et al., 2022). The present synthesis connects those perspectives by specifying the evidential transition from communicated feature to perceived experience, practical challenge and justified remedy. It makes the transition visible without claiming that the literature has already validated every link. That is a narrower, but more usable, contribution than declaring one interface feature universally superior. Recent panel evidence points in the same direction. Phan Khanh and Le (2026) report a non-linear relationship between the coherence of algorithmic human-resource signals and workers' withdrawal, with a turning point that depends on whether workers can contest algorithmic decisions. If this finding replicates, contestability is a condition that changes how explanations are read, rather than an add-on to them.

5. Managerial Implications and Research Agenda

5.1 Evaluate the procedure, not just the notice

For platform managers, the framework supports a diagnostic audit rather than a promise that a particular message will improve retention. Begin with a consequential decision type and map what the worker is told, what can be disputed, who considers the dispute and what that person can change. Table 4 proposes indicators for each stage. They are evaluative recommendations derived from the synthesis, not validated thresholds or a claim about legal compliance. A platform

should report denominators and reasons for non-use so that a low appeal rate is not automatically interpreted as satisfaction.

Table 4. Proposed evaluation matrix for platform decision governance.

Capability or process	Illustrative evidence to collect	Interpretive safeguard
Intelligibility	Comprehension of the specific reason; ability to identify disputed data; accessibility by language.	Do not substitute message length or satisfaction for understanding.
Access to challenge	Eligible decisions, channel discovery, attempts, completed requests, time and other burdens.	Non-use may reflect barriers rather than agreement.
Review quality and authority	Evidence considered, reviewer remit, reasoned responses, escalation routes and conflict safeguards.	Human involvement alone does not demonstrate competence or independence.
Justified redress	Verified defects, appropriate corrections, implementation time and consistency across comparable cases.	Raw reversal rates do not measure accuracy or fairness.
Organisational learning	Repeated complaint categories, data repairs and changes to defective rules or processes.	Closing individual tickets can leave a recurring cause unchanged.
Worker consequences	Subsequent work access and losses, alongside fairness and dignity measures over a stated period.	Attitudinal improvement alone does not establish material benefit.

Author-proposed indicators. Practical implementation requires context-specific definitions, permissions and safeguards; no platform audit was conducted for this article.

The audit should distinguish a challenge to the data from a challenge to the rule. A worker may accept a proximity rule but dispute the recorded location; another may question whether that rule fairly distributes desirable tasks. A review team authorised to correct coordinates may lack authority to reconsider allocation policy. Treating both concerns as identical customer-service tickets would conceal this distinction. The required escalation route depends on what is disputed, not simply on whether a human answered the request.

There are also trade-offs to examine rather than assume away. More detail can assist understanding but can impose reading burden; faster review can reduce waiting but may leave less time for investigation. A higher reversal rate can indicate responsive correction or a worse initial decision process. Managers should therefore pair efficiency measures with evidence of decision quality and worker experience. The aim is a documented, reasoned response to a relevant concern, not an optimisation exercise in producing favourable survey scores.

5.2 Separate interface experiments from operational validation

A focused next step is a factorial experiment crossing a concrete explanation with an accessible review opportunity while holding an adverse allocation constant. This follows an established design logic rather than introducing the first such comparison. The study should distinguish marginal effects from the interaction and retain separate informational, procedural and dignity outcomes. A communicated appeal is a treatment package whose wording may combine accessibility, a human reviewer and response time. It should not be read as an isolated test of any one component unless the design separates them.

A subsequent operational study must answer a different question. It should observe actual attempts to challenge decisions, the grounds raised, the review undertaken and the implemented response. Where randomisation is feasible and ethically appropriate, improvements to existing procedures can be compared without removing established protections. Otherwise, a carefully specified longitudinal comparison can document change while acknowledging confounding. Neither design should equate a simulation’s stated intention with actual retention, and neither should infer mediation merely because fairness ratings correlate with later participation.

Three design choices are especially consequential. First, the decision type must be fixed or explicitly stratified: task allocation, pay calculation and account restriction need not have the same stakes or remedies. Second, follow-up should be long enough to distinguish an initial response to a promise from experience of its delivery. Third, estimates should remain interpretable under missing outcomes and selective appeal use. Those who appeal are not a random subset of affected workers; comparisons between appellants and non-appellants need to address that selection rather than attribute all differences to the review process.

5.3 Test collective and contextual conditions directly

Worker representatives, peer groups and advisers may contribute information that an individual lacks or reveal patterns invisible in a single dispute. That possibility motivates comparisons between individual-only and supported contestation routes, with attention to who can participate and whose concerns remain unheard. The Malaysian evidence on collective action makes this an especially relevant extension, but does not determine the result in advance. Evaluation should distinguish help with understanding, help with submitting a challenge and influence over the rules themselves.

For Malaysian delivery and ride-hailing settings, recruitment should be grounded in actual work experience and the languages used in the decision process. Cognitive testing should examine the credibility of the scenario and the meaning of a promised review. No language, occupational category or national label should be treated as a proxy for common expectations. The resulting research should report the population reached and the populations not reached, rather than imply national representativeness from the location alone. These are requirements for testing transportability, not additional findings about Malaysian workers. The commencement of the Gig Workers Act 2025 also allows a before-and-after design. For comparable decision types, researchers could compare the volume, grounds and outcomes of internal grievances, conciliation referrals and Tribunal awards. Workers' accounts of whether the notices they received were intelligible would complete the picture. Such a design would test redress directly rather than inferring it from fairness ratings.

6. Limitations of the Review

This review has a purposive rather than exhaustive corpus. Search-engine retrieval, selection from an existing protocol's references and access to available original material can favour visible, accessible and already prominent work. The English-language focus may under-represent locally published research and worker perspectives expressed elsewhere. The source registry makes the retained evidence inspectable, but does not remove those selection effects. Findings describe the selected literature and should not be interpreted as the prevalence of particular effects across all algorithmic-management research.

Access depth also limits the analysis. Seventeen publications were used at abstract level. Such sources contextualise the problem but cannot support a detailed risk-of-bias judgement or unreported subgroup result. The principal experimental comparisons received closer original-text inspection, yet the review did not independently reanalyse their data. AI-assisted extraction and a consistency pass do not replace independent human screening, and no inter-rater agreement or certainty score is asserted.

The studies differ in populations, decision stakes, treatment content and measurement. This prevents a defensible ranking of explanation against appeal or a single pooled effect. The framework itself is conceptual: its capabilities are separately testable, but their proposed relationships are not established causal pathways. Malaysian material supports contextual relevance and collective-contestation questions, not local experimental effectiveness. The description of the Gig Workers Act relies on official news reporting and legal commentary rather than a doctrinal analysis of the statute. Finally, the article does not assess the current legal adequacy of any platform's procedures. Those boundaries leave the central conceptual result intact while identifying what empirical and institutional work remains necessary.

7. Conclusion

The evidence supports separating explanation, contestability and redress rather than treating them as interchangeable forms of transparency. Explanations can improve fairness perceptions under particular conditions, but their effects depend on what is explained and what is measured. Review preferences and communicated appeal opportunities supply different evidence from demonstrated correction. Worker-centred studies further locate these encounters within continuing relationships of control, dependence and collective voice.

The review's contribution is a three-capability framework and an evidence-to-outcome matrix for algorithmic platform work. Intelligibility should be tested through understanding, contestability through usable challenge, and redress through authorised, justified responses to defects. Functional complementarity among these capabilities should not be confused with a proven positive statistical interaction. Fairness ratings remain part of the evaluation, but cannot substitute for the missing institutional outcome. The practical question is whether workers can move from a relevant concern to a meaningful response. Giving a reason or displaying an appeal button does not by itself settle it.

Acknowledgement

Acknowledgements, funding sources and sponsor roles require confirmation by the responsible authors before submission.

Research Ethics and Data Availability

This article synthesises published research. No new participants were recruited and no individual-level worker data were collected. The accompanying source registry and retrieval notes document the materials used for the synthesis. No ethics approval, exemption determination, preregistration or public repository deposit is claimed.

AI-Assisted Preparation

AI assistance was used for source discovery, evidence organisation, critical synthesis, language editing and document formatting. Bibliographic identifiers were re-resolved against Crossref on 8 October and 22 November 2026. Source-access limits are reported in the methodology and accompanying registry. Responsible human authors must verify the content, approve authorship and declarations, and accept accountability for the final submission.

Conflict of Interest

Conflict-of-interest declarations have not been supplied. The responsible authors must confirm their individual disclosures before submission; an absence of conflicts is not assumed.

Appendix A: Retrieval and Extraction Record

The bounded retrieval included the following searches on 8 October 2026: “algorithmic management” “fairness” “transparency”; “algorithmic decisions” “contestability” “explanations”; “platform workers” “human review” “appeal”; and “Malaysia” “platform” “food delivery” “blackouts”. These were supplemented by exact-title and DOI look-ups for the retained publications, including the direct experimental comparators and review-methodology sources. The searches were conducted through web retrieval, not represented as Scopus, Web of Science or other subscription-database searches.

The registry records citation identity, publication year, DOI where available, source-access route, inspection depth, original-text locator, supported proposition and inference limit. Methods and findings were compared by decision setting, sample or documentary unit, exposure, outcome and the distinction between perception and operational remedy. Inclusion in the registry means that the source was used within its recorded limits; it does not certify exhaustive full-text appraisal, independent replication or human-author approval.

References

Alfrink, K., Keller, I., Kortuem, G., & Doorn, N. (2023). Contestable AI by design: Towards a framework. Minds and Machines, 33(4), 613–639. https://doi.org/10.1007/s11023-022-09611-z

Bernama. (2026, March 31). Act 872 provides greater protection to 1.64 million gig workers. *Bernama*. Retrieved from https://bernama.com/en/news.php?id=2539439

Binns, R., Van Kleek, M., Veale, M., Lyngs, U., Zhao, J., & Shadbolt, N. (2018). “It’s reducing a human being to a percentage”: Perceptions of justice in algorithmic decisions. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (Paper 377). New York, NY: ACM. https://doi.org/10.1145/3173574.3173951

Bujold, A., Gaudet, M.-C., & Parent-Rocheleau, X. (2026). When algorithmic management plays an intermediary role in worker–organization relationships: A social exchange perspective. *Organizational Psychology Review*. Advance online publication. https://doi.org/10.1177/20413866261461258

Bujold, A., Parent-Rocheleau, X., & Gaudet, M.-C. (2022). Opacity behind the wheel: The relationship between transparency of algorithmic management, justice perception, and intention to quit among truck drivers. *Computers in Human Behavior Reports, 8*, 100245. https://doi.org/10.1016/j.chbr.2022.100245

Cameron, L. D. (2024). The making of the “good bad” job: How algorithmic management manufactures consent through constant and confined choices. Administrative Science Quarterly, 69(2), 458–514. https://doi.org/10.1177/00018392241236163

Colquitt, J. A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. Journal of Applied Psychology, 86(3), 386–400. https://doi.org/10.1037/0021-9010.86.3.386

Dasgupta, P., Carbery, R., McDonnell, A., & Jooss, S. (2026). Algorithmic management and worker agency: The platform perspective on algoactivism. *Human Relations, 79*(12), 1539–1574. https://doi.org/10.1177/00187267261445807

Duggan, J., Sherman, U., Carbery, R., & McDonnell, A. (2020). Algorithmic management and app-work in the gig economy: A research agenda for employment relations and HRM. Human Resource Management Journal, 30(1), 114–132. https://doi.org/10.1111/1748-8583.12258

Grimmelikhuijsen, S. (2023). Explaining why the computer says no: Algorithmic transparency affects the perceived trustworthiness of automated decision-making. *Public Administration Review, 83*(2), 241–262. https://doi.org/10.1111/puar.13483

Henning, A., & Langenbach, P. (2026). Bridging the human–AI fairness gap: How providing reasons enhances the perceived fairness of public decision-making. Journal of Empirical Legal Studies, 23(1), 39–59. https://doi.org/10.1111/jels.70019

Hu, P., Zeng, Y., Wang, D., & Teng, H. (2024). Too much light blinds: The transparency–resistance paradox in algorithmic management. *Computers in Human Behavior, 161*, 108403. https://doi.org/10.1016/j.chb.2024.108403

Ius Laboris. (2026, March 11). *Malaysia's landmark Gig Workers Act coming into effect*. Retrieved from https://iuslaboris.com/insights/malaysias-landmark-gig-workers-act-coming-into-effect/

Jabagi, N., Croteau, A.-M., Audebrand, L. K., & Marsan, J. (2025). Do algorithms play fair? Analysing the perceived fairness of HR-decisions made by algorithms and their impacts on gig-workers. The International Journal of Human Resource Management, 36(2), 235–274. https://doi.org/10.1080/09585192.2024.2441448

Kadolkar, I., Kepes, S., & Subramony, M. (2025). Algorithmic management in the gig economy: A systematic review and research integration. Journal of Organizational Behavior, 46(7), 1057–1080. https://doi.org/10.1002/job.2831

Keegan, A., & Meijerink, J. (2025). Algorithmic management in organizations? From edge case to center stage. Annual Review of Organizational Psychology and Organizational Behavior, 12, 395–422. https://doi.org/10.1146/annurev-orgpsych-110622-070928

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. Academy of Management Annals, 14(1), 366–410. https://doi.org/10.5465/annals.2018.0174

Kinchin, N. (2024). "Voiceless": The procedural gap in algorithmic justice. International Journal of Law and Information Technology, 32(1), eaae024. https://doi.org/10.1093/ijlit/eaae024

Lee, M. K., Kusbit, D., Metsky, E., & Dabbish, L. (2015). Working with machines: The impact of algorithmic and data-driven management on human workers. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (pp. 1603–1612). New York, NY: ACM. https://doi.org/10.1145/2702123.2702548

Lyons, H., Velloso, E., & Miller, T. (2021). Conceptualising contestability: Perspectives on contesting algorithmic decisions. Proceedings of the ACM on Human-Computer Interaction, 5(CSCW1), Article 106. https://doi.org/10.1145/3449180

Lyons, H., Wijenayake, S., Miller, T., & Velloso, E. (2022). What's the appeal? Perceptions of review processes for algorithmic decisions. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (pp. 1–15). New York, NY: ACM. https://doi.org/10.1145/3491102.3517606

Mirbabaie, M., Langer, M., Rieskamp, J., & Hofeditz, L. (2026). Transparency fallacy: Perceived fairness in algorithmic management. Business & Information Systems Engineering, 68(4), 829–850. https://doi.org/10.1007/s12599-025-00963-1

Möhlmann, M., Zalmanson, L., Henfridsson, O., & Gregory, R. W. (2021). Algorithmic management of work on online labor platforms: When matching meets control. MIS Quarterly, 45(4), 1999–2022. https://doi.org/10.25300/MISQ/2021/15333

Muldoon, J., & Raekstad, P. (2023). Algorithmic domination in the gig economy. *European Journal of Political Theory*, *22*(4), 587–607. https://doi.org/10.1177/14748851221082078

Ochmann, J., Michels, L., Tiefenbeck, V., Maier, C., & Laumer, S. (2024). Perceived algorithmic fairness: An empirical study of transparency and anthropomorphism in algorithmic recruiting. Information Systems Journal, 34(2), 384–414. https://doi.org/10.1111/isj.12482

Phan Khanh, D., & Le, T. T. (2026). When algorithmic HRM signals backfire: Contestability thresholds in gig worker sensemaking. *International Journal of Manpower*. Advance online publication. https://doi.org/10.1108/IJM-02-2026-0138

Rani, U., & Furrer, M. (2021). Digital labour platforms and new forms of flexible work in developing countries: Algorithmic management of work and workers. Competition & Change, 25(2), 212–236. https://doi.org/10.1177/1024529420905187

Rosenblat, A., & Stark, L. (2016). Algorithmic labor and information asymmetries: A case study of Uber's drivers. International Journal of Communication, 10, 3758–3784. Retrieved from https://ijoc.org/index.php/ijoc/article/view/4892

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. Journal of Business Research, 104, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Starke, C., Baleis, J., Keller, B., & Marcinkowski, F. (2022). Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature. Big Data & Society, 9(2), 20539517221115189. https://doi.org/10.1177/20539517221115189

Torraco, R. J. (2016). Writing integrative reviews of the literature: Methods and purposes. International Journal of Adult Vocational Education and Technology, 7(3), 62–70. https://doi.org/10.4018/IJAVET.2016070106

Uchiyama, Y., Furuoka, F., & Omar, S. A. (2025). The rise and contestation of platform capitalism: Evidence from two food delivery blackouts in Malaysia. Journal of Industrial Relations, 67(3), 481–508. https://doi.org/10.1177/00221856241294107

van Zoonen, W., von Bonsdorff, M. E., & van der Heijden, B. I. J. M. (2026). Algorithmic surveillance and workers' compliance: The role of trust, privacy concerns, and fairness in online crowdwork. *Human Relations, 79*(7), 795–824. https://doi.org/10.1177/00187267251379698

Wesche, J. S., Hennig, F., Kollhed, C. S., Quade, J., Kluge, S., & Sonderegger, A. (2024). People's reactions to decisions by human vs. algorithmic decision-makers: The role of explanations and type of selection tests. *European Journal of Work and Organizational Psychology, 33*(2), 146–157. https://doi.org/10.1080/1359432X.2022.2132940

Wood, A. J., Graham, M., Lehdonvirta, V., & Hjorth, I. (2019). Good gig, bad gig: Autonomy and algorithmic control in the global gig economy. Work, Employment and Society, 33(1), 56–75. https://doi.org/10.1177/0950017018785616

Yurrita, M., Draws, T., Balayn, A., Murray-Rust, D., Tintarev, N., & Bozzon, A. (2023). Disentangling fairness perceptions in algorithmic decision-making: The effects of explanations, human oversight, and contestability. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (Article 134). New York, NY: ACM. https://doi.org/10.1145/3544548.3581161

Yurrita, M., Verma, H., Balayn, A., Alfrink, K., Gadiraju, U., & Bozzon, A. (2025). Identifying algorithmic decision subjects' needs for meaningful contestability. Proceedings of the ACM on Human-Computer Interaction, 9(7), Article CSCW234. https://doi.org/10.1145/3757415

Zhou, L., Lei, X., Cooke, F. L., Huang, X., & Zhang, J. (2025). The dual effects of algorithmic management on platform workers: An attribution perspective. *Human Resource Management, 64*(6), 1687–1707. https://doi.org/10.1002/hrm.70009