Behavioral Intelligence Design: A Framework for AI Systems That Infer Psychological Trajectory

Ben Cotte SSRN Working Paper | April 2026

> **Version:** v4 (1.1) — Three additions: recommendation system distinction (Section 3.1), ABC triad buyer typology (Section 4.1), reflexivity as designed mechanism (Section 5.3). +5 references. **Supersedes:** v3 (submitted to SSRN April 4, 2026, Abstract ID 6518238) **Companion:** `BID-revision-notations.md`

Abstract

For two decades, programmatic advertising has operated a sophisticated inference architecture at planetary scale: capturing behavioral signals, operationalizing them as predictive features, inferring latent states, and acting on the inference in milliseconds. This architecture is commercially proven and scientifically unexamined. It asks one question (will they transact?) and has never asked another.

The stated-versus-revealed preference gap, the empirical regularity that behavior predicts outcomes more reliably than declared preferences (Samuelson, 1938; Kahneman, 2011; Ellis & Huang, 2026), has never been operationalized as the design kernel of a deployed AI system studied under rigorous methodology. The mechanism that exploits this gap exists at commercial scale. The science that grounds it does not.

This paper introduces **Behavioral Intelligence Design (BID)** as the practice and emerging research programme of building AI artifacts that operationalize behavioral signals against behavioral science theory to infer *psychological trajectory* rather than transaction probability. We ground BID in Design Science Research methodology (Hevner et al., 2004; Gregor & Hevner, 2013) and contribute: (1) the first formal definition and vocabulary of BID as a named practice; (2) a behavioral signal taxonomy including the identification of *comparative behavioral signals* as the class that directly encodes the stated-revealed gap; (3) a six-stage inference-to-action pipeline with explicit theoretical grounding requirements; and (4) three deployed artifact demonstrations across consumer, interpersonal, and organizational contexts. We state three falsifiable design propositions and identify the open questions sufficient to sustain cumulative inquiry.

Keywords: Behavioral Intelligence Design, Behavioral Signal Inference, Psychological Trajectory, Design Science Research, Stated vs. Revealed Preferences, Comparative Behavioral Signals, Ecological Validity

---

1. Introduction

1.1 A Practice Without a Name

Thousands of AI systems currently infer human intent from behavioral signals. Real estate platforms route leads based on browsing patterns rather than declared search criteria. Communication platforms analyze response latency to gauge engagement. HR systems track behavioral adoption rates to detect where training is failing. Each shares an architecture: capture behavioral traces, operationalize them into indicators, infer a latent state, act on the inference. Each shares an epistemological commitment: behavior is a more reliable signal than declaration.

None share a name, a theoretical grounding, or a research community. They exist as engineering outputs, proprietary systems optimized for commercial outcomes, not as objects of scientific inquiry. The knowledge embedded in them dies with the deployment context. No journal accumulates findings about what works, under what conditions, and why. You cannot replicate a methodology that has not been articulated. You cannot build on prior work that has not been situated in a shared theoretical frame.

Naming this practice is the precondition for the science to begin.

1.2 The Mechanism and Its Intellectual Abandonment

The architecture at the core of BID did not emerge from academic research. It emerged from commercial necessity. Over roughly two decades, programmatic advertising produced signal capture infrastructure processing billions of daily behavioral events, operationalization pipelines transforming raw signals into predictive features, real-time inference engines modeling intent from behavioral patterns, and action layers serving creative content in milliseconds. The commercial validation is unambiguous: behavioral signals predict transaction probability more reliably than any declared-preference instrument.

What the industry built, the academy never examined. No peer-reviewed methodology paper explains why the architecture works, what theoretical properties of human behavior make behavioral signals more predictive than declarations. No generalized design knowledge was ever extracted. The practitioner community accumulated substantial operational knowledge about signal reliability, operationalization failures, cold-start problems, proxy contamination, and temporal decay. None of it exists in the scientific literature.

The abandonment is structural, not accidental. Commercial incentives reward prediction accuracy, not theoretical understanding. There is no return on investment from publishing the mechanisms of a proprietary system. And so an industry that processes more behavioral inferences per day than all of psychology processes per decade has contributed nothing to the science of behavioral inference.

This paper draws on two decades of practitioner experience with this architecture. The biography is the data.

1.3 The Stated-Revealed Gap as Theoretical Warrant

Samuelson (1938) formalized the insight that consumers' observed choices reveal preference structures their verbal reports cannot capture. Kahneman (2011) provided the cognitive architecture: System 1 governs fast, automatic, behavioral responses; System 2 governs slow, deliberate, declarative output. The two systems frequently disagree, not because subjects are irrational, but because they draw on different information under different cognitive loads. A property seeker browsing listings at midnight on a mobile device operates System 1; the same person completing a search preference form the next morning operates System 2. The behavioral trace and the declared preference measure different things.

Ariely (2008) demonstrated that the gap is systematic, not random: behavioral departures from stated preferences follow stable, identifiable patterns that persist across contexts and populations. Weber and Hsee (1998) showed the gap is culturally modulated, with cross-cultural differences in risk perception larger when measured through behavior than through self-report. For any deployment outside Western WEIRD populations, this is not a caveat. It is a research finding waiting to be generated.

Recent computational work sharpens the stakes. Ellis and Huang (2026) found that revealed preferences predicted choices more accurately than subjects' own written instructions for what they wanted. Himmelstein and Budescu (2023) demonstrated that stated preferences for algorithmic versus human forecasting advice did not predict actual behavioral reliance on those systems. Gu, Wang, and Han (2025) found that even large language models exhibit measurable divergence between stated and revealed preferences, evidence consistent with the interpretation that the gap is structural to any system bridging representation and action.

The gap is not a curiosity. It is a design signal. It identifies what class of AI system would outperform systems built on declared preferences, and therefore what class of system is worth building and studying.

1.4 DSR as the Methodology

Hevner et al. (2004) called for AI artifacts designed from behavioral science theory, evaluated through behavioral science constructs, generating design knowledge that accumulates across deployment contexts. The ISR editorial board (2024) renewed the call, identifying AI-focused design research grounded in operational practice as an underserved pathway. Neither has been answered with a systematic programme. BID closes the loop by coupling three things that have remained separate: behavioral science theory as the warrant, AI artifact construction as the method, and DSR methodology as the knowledge production framework.

1.5 What This Paper Contributes

This paper does not prove that BID generates better outcomes than alternative approaches across all contexts. That is the work of the research programme it proposes. It establishes that the programme is coherent, theoretically grounded, methodologically rigorous, and empirically instantiable:

- The first formal definition of Behavioral Intelligence Design as a named practice and research programme.
- A behavioral signal taxonomy, including the identification of *comparative behavioral signals* as the class encoding the stated-revealed gap.

- A six-stage inference-to-action pipeline with a theoretical warrant requirement at the operationalization stage.
- Three deployed artifact demonstrations generating three falsifiable design propositions.
- A research programme of open questions sufficient to sustain cumulative inquiry.

Founding papers plant stakes, not proof. The programme is open.

---

2. Theoretical Foundations

2.1 The Stated-Revealed Preference Gap: Design Signal, Not Measurement Error

The evidence for the stated-revealed preference gap is among the most replicated in behavioral science. We treat it here not as a literature review but as a theoretical warrant: the gap is the reason BID is possible.

Samuelson (1938) established the foundational proposition: consumers' observed choices reveal preference orderings that introspective reports cannot reliably capture. Kahneman (2011) supplied the cognitive mechanism. System 1 (fast, automatic, pattern-driven) governs the behavioral responses BID captures. System 2 (slow, deliberate, linguistically mediated) governs the declared preferences collected by surveys and forms. They operate in parallel and frequently disagree because they draw on different information under different cognitive loads.

Ariely (2008) established that the disagreement is systematic: subjects consistently act on preferences their declarations misrepresent, and the misrepresentation follows stable patterns. Weber and Hsee (1998) demonstrated cultural modulation: cross-cultural differences in risk perception are larger when measured behaviorally than declaratively, meaning the stated-revealed divergence is itself culturally structured. For deployments in Thai real estate markets, Thai interpersonal communication, and Thai organizational contexts, the cultural dimension of the gap is not a methodological footnote. It is a variable of direct interest.

Recent work confirms persistence in AI interaction. Ellis and Huang (2026) showed that revealed preferences predicted choices more accurately than subjects' own written instructions; when the two conflicted, AI systems defaulted to the less accurate stated signal. Himmelstein and Budescu (2023) demonstrated that stated preferences for algorithmic versus human forecasting advice did not predict actual behavioral reliance. Gu, Wang, and Han (2025) found that large language models exhibit measurable divergence between stated and revealed preferences, evidence that the gap may be structural to any system bridging representation and action.

The design implication is direct: AI systems that capture behavioral signals access a more predictive data stream than those processing declarations. This is not a measurement correction. It is an architectural advantage. BID operationalizes it.

The gap also has boundary conditions that remain underspecified. How does it vary with stakes level? A consumer browsing rental properties on a Tuesday evening may exhibit less stated-revealed divergence than one committing to a purchase representing years of savings. How does it vary across cultural contexts? Weber and Hsee's (1998) finding that cultural differences amplify under behavioral measurement suggests that a BID system deployed in Thailand will encounter a differently structured gap than one deployed in Sydney or Paris. How does it change over longitudinal observation? A lead whose behavioral profile diverges sharply from declared preferences in week one may converge over time as System 2 catches up with System 1, or diverge further as preference crystallization reveals what the initial declaration obscured.

These boundary conditions matter because they determine where and when behavioral inference outperforms declared-preference collection, under what conditions the gap narrows or disappears, and what operationalization adjustments are required. Laboratory conditions, with their compressed timescales, low stakes, and WEIRD subject pools, suppress precisely the variables needed to map these boundaries. Deployed artifacts in naturalistic contexts are the research method positioned to generate this evidence.

2.2 Design Science Research: Artifacts as the Unit of Knowledge

Simon (1996) established that the sciences of the artificial produce knowledge through design, not only through observation of natural phenomena. Hevner et al. (2004) articulated the position for IS research: behavioral science explains and predicts; design science builds and evaluates. The two paradigms complement each other, and the strongest contributions are those where behavioral theory drives design decisions and artifact evaluation generates findings in return.

Three of Hevner's seven guidelines bear the weight for BID. Guideline 1 (Design as Artifact) requires a deployed system operating in live contexts, not a description of what a system might do. Guideline 2 (Problem Relevance) requires that the artifact address problems that are genuinely important and unsolved. Guideline 5 (Research Rigor) requires formal definition, coherent representation, and internal consistency.

Gregor and Hevner's (2013) knowledge contribution framework positions BID in the Exaptation quadrant: mature solution components applied to a new problem class. Behavioral science, AI/ML technology, signal capture infrastructure, and DSR methodology are each individually mature. Their assembly under a common framework directed at psychological trajectory inference across deployment contexts is new. The contribution is the coupling, not the components. This novelty claim is independently grounded in Section 2.5, where we demonstrate that the four preconditions for a research programme are all absent from the existing literature. The integration gap exists prior to BID.

Larsen et al. (2025) provide the evaluative apparatus: criterion validity (does the artifact work?), causal validity (do we understand why?), and context validity (do findings transfer?). Tuunanen, Winter, and vom Brocke (2024) provide the design echelon framework for managing complexity in multi-context DSR projects of the kind BID proposes. All three validity dimensions are addressed in the demonstrations with different degrees of current evidence.

2.3 The Programmatic Advertising Legacy

The inference architecture BID formalizes developed through two decades of commercial pressure, not academic inquiry. That history deserves to be treated as what it is: a massive, replicated engineering experiment that produced a mature architecture without producing any science.

What the industry built: a signal taxonomy covering behavioral traces across digital touchpoints (click patterns, dwell distributions, scroll depth, search refinement sequences, session timing, return visit patterns); an operationalization infrastructure connecting raw signals to predictive features, with techniques for handling sparse data, cold-start problems, proxy contamination, and temporal decay; and an inference-action pipeline proven under extreme conditions of millisecond latency, billions of daily decisions, and adversarial users actively working to defeat the inference (Boerman, Kruikemeier, & Zuiderveen Borgesius, 2017).

What the industry abandoned: theoretical grounding. The industry does not know *why* behavioral signals predict transaction probability. It knows that they do, and it optimizes accordingly. The result is proprietary optimization engines that cannot be generalized. When the deployment context changes, the models fail and must be rebuilt from scratch. No transferable design knowledge was ever extracted.

Programmatic advertising asks "will they click?" BID asks "who are they becoming?" The architecture is shared. The question is not.

BID recovers the mechanism and gives it theory. The operational failures, the cold-start solutions, the temporal decay calibrations: this is practitioner knowledge that behavioral science cannot generate from laboratory settings and that DSR has not yet formalized. I have built commercial implementations of this architecture for twenty years. The knowledge lives in deployed systems, not in journals.

The academic precursors of BID in the author's own work span a decade. Cotte and Haseley (2014) used institutional registration data, not surveys, to predict the timing of student attrition at a Thai international university, demonstrating that behavioral traces from administrative systems predicted departure trajectories that declared satisfaction measures missed. Cotte and Shannon (2018) proposed using advertising platforms as experimental instruments for consumer behavior research, arguing that ad targeting infrastructure generates behavioral data at scale that resolves the sample, segmentation, and ecological validity limitations of laboratory methods. Neither paper named BID. Both instantiated its core commitment: behavioral signals captured from operational systems, operationalized against theoretical constructs, producing inference that declared-preference instruments cannot replicate. The present paper formalizes what those earlier efforts practiced.

2.4 Domain Theories as Construct Libraries

The domain theories in BID's demonstrations are not pillars supporting a separate case for validity. They are construct libraries: the vocabularies the operationalization stage draws on when the pipeline is instantiated in a specific context.

A behavioral signal is theoretically inert until operationalized against a construct. Response latency operationalized against Gottman's (1994) approach-avoidance framework becomes *engagement velocity*, measuring turning-toward versus turning-away behavior. The same response latency operationalized against Lemon and Verhoef's (2016) customer journey framework becomes *decision velocity*, measuring purchase intent crystallization. Same signal class, different construct library. This is the structure of how behavioral inference works, not a limitation.

The **interpersonal library** draws on Gottman (1994, 1999) for the most extensively validated behavioral coding system for relationship dynamics outside clinical settings: the 5:1 positive-to-negative interaction ratio, the Four Horsemen behavioral signatures (criticism, contempt, defensiveness, stonewalling), and the bid-for-connection coding scheme measuring turning-toward versus turning-away. All were developed from behavioral observation without self-report, making them directly applicable to text-based signal analysis. Bowlby (1969) and Ainsworth (1978) provide attachment theory: four attachment styles (secure, anxious-preoccupied, dismissive-avoidant, fearful-avoidant) manifest through distinctive behavioral signatures in response patterns, proximity-seeking, disclosure trajectories, and reactions to perceived threat. Austin (1962) and Searle (1975) provide speech act theory, giving a principled basis for classifying the relational function of messages from linguistic structure alone. Altman and Taylor (1973) provide social penetration theory: disclosure depth and breadth follow predictable trajectories that are themselves behavioral signals. Pennebaker (2011) provides linguistic style matching as an indicator of rapport and alignment.

The **consumer and market library** draws on Khanna and Palepu (1997) for institutional voids, where intermediary institutions are absent, creating conditions where behavioral signal inference is not supplementary but primary. In markets with no MLS, behavioral data from platform interaction is the only market intelligence. Rochet and Tirole (2003) provide two-sided market economics. Lemon and Verhoef (2016) frame each consumer touchpoint as a potential behavioral signal. Kwak, Zhang, and Yu (2021) document how platforms in institutional voids can create or maintain those voids, a finding directly relevant to BID's deployment contexts.

The **organizational library** draws on Cohen and Levinthal (1990) for absorptive capacity and Zahra and George (2002) for the distinction between potential and realized absorption. Seran et al. (2025) document AI-induced cognitive dissonance: the discomfort when AI capability contradicts self-efficacy beliefs. Golgeci et al. (2025) confirm that AI resistance mechanisms remain systematically underexplored relative to adoption enablers, characterizing workplace AI resistance as a distinct phenomenon requiring its own theoretical framework.

2.5 The Integration Gap

The gap BID addresses is not that behavioral inference systems do not exist. They are commercially ubiquitous. The gap is that four conditions necessary for a research programme have all been absent at once.

A name. A search of ISR, MISQ, JAMS, CSCW, and CHI reveals no paper defining "behavioral intelligence design" as a research practice, no paper naming a taxonomy of behavioral signal classes for multi-context deployment. The term appears in scattered industry usage (maritime risk, voice AI, marketing trade press) but has no academic definition and no research community.

A theoretical grounding. Commercial systems use atheoretical machine learning to identify predictive features. They produce systems that work within training distributions but generate no transferable design knowledge.

A methodology. DSR provides one, but has not been applied to behavioral inference systems with behavioral science theory as the design kernel.

A vocabulary. Without shared terminology for signal classification, pipeline stages, and construct mapping, findings from one deployment cannot be compared to another.

Each was absent independently. BID supplies all four. The gap is prior to the definition. That is the independent ground for the novelty claim.

---

3. The BID Framework

3.1 Definition and Scope

Behavioral Intelligence Design is the practice and emerging research programme of building AI artifacts that operationalize behavioral signals against behavioral science theory to infer the *psychological trajectory* of subjects (who they are becoming) rather than the probability of a specific transaction, deployed in operational contexts where stakes are real, data is longitudinal, and the stated-revealed gap is consequential.

Five terms carry the definitional weight.

"*Operationalize against theory*" is the warrant requirement. Every construct must map from a behavioral signal to a theoretical construct via a published, falsifiable warrant. Response latency used as a generic engagement proxy is behavioral analytics. Response latency operationalized as engagement velocity via Gottman's approach-avoidance coding is BID. The difference is epistemological: the warrant makes design knowledge transferable and the inference falsifiable.

"*Psychological trajectory*" distinguishes BID from state-detection and transaction-prediction. A trajectory is directional movement over time: this subject is *becoming* disengaged; this lead's preference is *converging* toward a property category they have never declared; this team member's resistance is *intensifying*. Trajectory requires longitudinal data, sequential signals, and a model of change.

"*Operational contexts*" enforces ecological validity. BID artifacts are deployed, not prototyped. They operate under real stakes in real markets, not under experimental conditions with fictive scenarios.

"*Consequential*" narrows scope. Trivial gaps in low-stakes contexts are not BID problems. BID operates where the inference changes outcomes.

"*Behavioral signals*" specifies the input: the observable residue of behavior (digital traces, interaction patterns, timing distributions, sequences) as distinct from self-report or any form of elicited preference. Captured, not asked for.

BID is distinguished from adjacent concepts. Business intelligence infers business states from business metrics, not psychological states from behavioral signals. Programmatic behavioral targeting shares the architecture but limits the question to transaction probability. Behavioral analytics identifies patterns without inferring *why* and without operationalizing against theory. Artificial Behavior Intelligence (Jo et al., 2025) operates on computer vision (posture, facial expression, embodied interaction); BID operates on text, clicks, journeys, and communication signals.

The distinction from recommendation systems requires specific attention because it is the most likely misreading. Recommendation systems ask "given what this user has done, what should we show them next?" Their output is a ranked list optimized for engagement or conversion. BID asks "given what this user has done, who are they becoming?" Its output is a behavioral vector trajectory that evolves over time; the ranked list, if one exists, is an instrument, not the objective. Five structural divergences separate the two. First, the unit of analysis: recommendation systems model items and item-user affinities; BID models trajectories, time-series of psychological state transitions. Second, the theoretical warrant: recommendation systems cite collaborative filtering and matrix factorization (Koren, Bell, & Volinsky, 2009); BID cites Samuelson (revealed preference), Gottman (approach-avoidance), Kahneman (dual-process), and Hofstede (cultural moderation). Third, the core signal: recommendation systems treat the divergence between stated preferences and observed behavior as noise to be filtered; BID treats it as the primary signal, the comparative signal class formalized in Section 3.2. Fourth, the prior as design surface: recommendation systems treat cold-start as a weakness to be overcome; BID treats the prior, the advertising creative that generates the initial behavioral context, as a first-class design element whose manipulation is itself an experimental intervention. Fifth, the feedback loop: recommendation systems loop within a single platform; BID loops across system boundaries, from advertising (prior setting) through observation (behavioral capture) through inference (model update) through targeting adjustment (prior revision), a cross-system recursive architecture that is designed, not incidental. A reviewer who reads BID as a recommendation system with extra steps has misidentified the category. BID is a behavioral inference architecture that may use recommendation as one instrument among several.

3.2 The Behavioral Signal Taxonomy

The taxonomy operates at the level of mechanism, independent of domain. Domain-specific signal mappings are instantiated in the demonstrations.

Temporal signals capture when and how fast behavior occurs: timing, pacing, rhythm, recency, interval distributions. They are baseline-relative. A four-hour response latency carries different meaning in a fast-paced negotiation and an asynchronous professional exchange. Temporal signals capture urgency, engagement velocity, avoidance patterns, and the pace of preference change.

Sequential signals capture order and progression: paths through an interface, refinement patterns across sessions, the arc from initial contact to inquiry. A property seeker whose search narrows consistently toward a neighbourhood they have never declared interest in exhibits a sequential signal of revealed preference. Sequential signals capture trajectory and intent crystallization.

Intensity signals capture depth and magnitude: dwell time, elaboration, investment in a particular interaction. They capture preference strength and commitment level independent of any declaration.

Comparative signals capture divergence between declared input and behavioral trace, or cross-signal inconsistency within the behavioral record. A subject whose stated budget is 3-5 million baht but whose dwell time and return visits concentrate on 6-8 million baht listings generates a comparative signal. A relationship where stated satisfaction is high but response latency increases and disclosure depth declines generates a comparative signal. An organizational team whose AI satisfaction scores are high but override rates are climbing generates a comparative signal.

The comparative signal class is BID's taxonomic contribution. No existing taxonomy names it explicitly. It is visible only when a system holds both declared input and behavioral trace and is designed to look for divergence. This is the class that directly encodes the stated-revealed preference gap, and it is the one that alternative approaches most systematically ignore.

3.3 The Inference-to-Action Pipeline

The pipeline has six stages. It applies without structural modification across deployment contexts; what changes are the domain theories, the signal sources, and the action repertoire.

Stage 1: Capture. Identify and collect behavioral signals from the operational context. Capture design is theory-informed: collect signals that map to constructs that matter, not everything available. Comparative signals require that declared input be captured alongside behavioral trace, because the comparison requires both channels.

Stage 2: Operationalize. Transform raw signals into measurable psychological constructs using a published, falsifiable theoretical warrant. This is the theoretical heart of BID. Response latency distributions become engagement velocity via Gottman's approach-avoidance coding. Search refinement trajectories become preference crystallization via revealed preference theory. Session timing patterns become decision readiness via customer journey frameworks. Each mapping is explicit, grounded, and open to challenge: a reviewer can evaluate whether the warrant is appropriate and the operationalization faithful.

Stage 3: Infer. Derive the latent psychological trajectory from operationalized constructs. The inferential target is directional movement over time, not a snapshot or a probability. What attachment pattern is emerging? What trust trajectory is forming? What resistance pattern is developing, and which type? The inference combines constructs from multiple signal classes into a composite trajectory estimate. Not all cases require the same inferential depth: some trajectories are resolvable from signal patterns alone, while others require progressively richer reasoning about construct interactions. Production BID systems must manage this complexity gradient as a design variable, balancing inferential resolution against operational cost.

Stage 4: Act. Execute an intervention based on the inferred trajectory. Route a lead by behavioral preference profile rather than declared requirements. Surface an alert when cooling patterns cross a threshold. Trigger differentiated training when resistance signals cluster. Without action, a BID artifact is a monitoring system; with action, it is a behavioral intelligence system.

Stage 5: Evaluate. Assess whether the inference-action pair produced the predicted outcome. The falsifiability condition must be stated before evaluation begins: what outcome pattern would lead to revision of the model? If behavioral matching does not outperform declared-preference matching on conversion, the proposition is falsified. Falsifiability is a design requirement, not a footnote.

Stage 6: Refine. Two types of refinement, and the distinction matters. *Model refinement* adjusts weights within existing construct mappings; the mapping holds, the calibration improves. This is standard ML learning. *Construct revision* changes the theoretical mapping itself when evaluation evidence persistently disconfirms the construct. When response latency operationalized via Gottman fails to predict disengagement in WhatsApp exchanges in Pattaya with the effect sizes the psychotherapy literature predicts, that is a finding about the construct's boundary conditions, not just a calibration failure. This feedback loop from deployed artifact back to behavioral science is the mechanism by which BID advances the discipline it draws on.

3.4 Ecological Validity as Design Requirement

Three properties of deployed artifacts produce knowledge that laboratory conditions systematically suppress. These are stated once as design requirements for any BID contribution.

Real stakes produce qualitatively different behavioral signals than experimental conditions. The stated-revealed gap manifests more strongly when the property purchase is real, the relationship has genuine consequences, and the organizational override affects actual performance. BID's claim is that behavioral signals are more informative under real stakes precisely because System 2's declarative override of System 1 becomes more costly when consequences are real.

Longitudinal accumulation is required to detect trajectory rather than state. Attachment styles manifest through patterns across many interactions, not a single exchange. Preference drift reveals itself across sessions. Resistance develops over weeks of deployment, not a one-time measure.

Naturalistic context means messy, culturally embedded, unpredictable environments where subjects behave for their own reasons. A BID artifact that works only in controlled conditions has failed its design specification.

3.5 Why Three Contexts, Not Three Domain Papers

Mechanism generality can only be established through multi-context instantiation. A single deployment proves the pipeline works once. Three deployments in which everything varies (domain theories, signal sources, cultural context, action repertoire) while the pipeline does not, demonstrate that the mechanism is general rather than context-specific.

Multi-context instantiation also reveals boundary conditions that single-domain work cannot access. The stated-revealed gap varies across consumer, interpersonal, and organizational contexts in ways that generate findings about the gap itself. When comparative signals in a consumer market predict inquiry targets with high accuracy, but the equivalent comparative signal in an organizational context (stated satisfaction diverging from behavioral usage) predicts resistance type with moderate accuracy, that difference is a finding about the mechanism's boundary conditions, not about either domain in isolation. Mapping these variations produces knowledge about the mechanism, not just the domain. No single deployment can generate it.

Three demonstrations by a single practitioner-researcher do not constitute proof of general applicability. They constitute a claim, supported by the theoretical argument, and an invitation to the community. That is what a founding paper does.

---

4. Framework Demonstrations

All three demonstrations satisfy the ecological validity requirements from Section 3.4: deployed artifact, real stakes, longitudinal data, naturalistic context.

4.1 Demonstration A: Behavioral Matching in a Market Without Cooperative Data Infrastructure

Context. Thailand's residential property market exhibits institutional void conditions (Khanna & Palepu, 1997) representative of many Southeast Asian property markets: no multiple listing service, no cooperative database, property information fragmented across LINE groups, Facebook pages, and agent-controlled listings. No shared declared-preference system was ever built. Consumers state preferences to individual agents in ad hoc form; agents match manually based on experience. Systematic stated-revealed divergence is structurally invisible to this process.

The platform operates in the Pattaya/Jomtien corridor of eastern Thailand, where Russian, Chinese, European, and Thai buyers navigate a property market that has never had centralized infrastructure. Agents control their own listings. No shared inventory exists. Price transparency is minimal. In this environment, the only systematic behavioral data about consumer preferences comes from platform interaction. The behavioral trace is not supplementary to institutional data; it is the only data there is.

Pipeline instantiation. Capture collects platform interaction logs: property page views, search filter selections and modifications, inquiry submissions, return visits, cross-listing comparison sequences, and session timing. Declared search filters (budget, type, location, bedrooms) are captured as the comparative baseline.

Operationalization maps signals to constructs: click-to-inquiry timing as *decision velocity* (rate of intent crystallization); progressive filter narrowing across sessions as *preference crystallization* (convergence toward a definable preference profile); dwell time concentration as *revealed preference intensity* (behavioral investment in property types independent of declared criteria);

divergence between declared filter parameters and browsing behavior as the *stated-revealed preference gap score*, the comparative signal class in direct operation.

Inference produces a behavioral preference profile compared against declared parameters. Where divergence exceeds a threshold, the behavioral profile becomes the operative preference for routing.

Action routes leads by behavioral profile, surfaces listings prioritized by behavioral match, and delivers behavioral briefings to agents before first contact. The user-facing interface is itself a design variable: presentation order, visual weight, and cross-reference structure shape the conditions under which the next behavioral signal is produced, making the front-end an instrument in the inference loop rather than a passive display layer.

Evaluation compares inquiry conversion rates and time-to-match for leads routed by behavioral profile versus declared preferences. The falsifiability condition: if behavioral matching does not outperform declared-preference matching on inquiry conversion, Design Proposition 1 is not supported.

Design knowledge. Systematic stated-revealed divergence is detectable in platform logs without additional instrumentation. The comparative signal class, the divergence between declared filters and actual browsing, is the most informative predictor of inquiry target, outperforming any single behavioral or declared signal alone.

A further finding emerges from the convergence dynamics of behavioral vectors. Drawing on the ABC triad from consumer behavior (Fishbein & Ajzen, 1975; Breckler, 1984), the rate at which a subject's behavioral vector converges toward a stable preference profile, *vector convergence rate*, differentiates buyer psychology types without self-report. Affect-led subjects (wide zone exploration, high dwell on lifestyle imagery, low engagement with financial metrics) exhibit slow, broad convergence. Cognition-led subjects (rapid multi-property comparison, high engagement with yield and price-per-square-metre data) exhibit fast, narrow convergence on financial dimensions. Behavior-led subjects (short sessions, urgency signals, repeated return to the same listing, foreign quota and resale price checks) exhibit narrow vectors from session start. The stated-revealed gap varies systematically by type: cognition-led subjects declare lifestyle motivations while dwelling on yield data; affect-led subjects declare casual intent while returning to the same property. Vector convergence rate is already captured in the deployed system's session schema. It differentiates buyer psychology from behavioral trace alone, enriching DP1 from "does behavioral matching outperform declared-preference matching?" to "does BID reveal the psychological structure of buyer decision-making that declared preferences systematically misrepresent?"

Design Proposition 1 (DP1): In markets without cooperative data infrastructure, behavioral signal matching outperforms declared-preference matching on inquiry conversion rate and time-to-match.

Context validity. DP1 is proposed as generalizable, not specific to Thai real estate. The test: does it hold in other institutional void markets where behavioral data from platform interaction is the primary available intelligence? Sub-Saharan residential markets, informal credit markets, emerging economy service platforms are all independent test sites.

4.2 Demonstration B: Psychological Trajectory Inference from Asynchronous Text

Context. An AI system analyzing behavioral signals in asynchronous text messaging (WhatsApp, LINE) to infer relationship dynamics without self-report, clinical infrastructure, or trained-coder annotations. The substrate is naturalistic and unstructured: daily interpersonal exchanges between people with ongoing relationships. No ground-truth labels exist. No subjects are recruited for research purposes. The behavioral trace is the record of actual communication.

Pipeline instantiation. Capture collects message-level behavioral signals: response latency distributions across time-of-day and topic, topic progression trajectories, disclosure depth and breadth, message length and elaboration, and divergence between stated characterizations of the relationship and the behavioral patterns the record reveals.

Operationalization maps signals to constructs from the interpersonal library: response latency as *engagement velocity* via Gottman's (1994) approach-avoidance coding; bid-to-response ratios as *emotional reciprocity* via Gottman's (1999) turning-toward/away framework; linguistic synchrony as *rapport* via Pennebaker (2011); disclosure trajectories as *trust development stage* via Altman and Taylor (1973); speech act distributions as *relational function* via Austin (1962) and Searle (1975).

Inference produces a composite relationship health trajectory: warming, cooling, stabilizing, or turbulence, from combined temporal, sequential, intensity, and comparative signals. The system distinguishes silence patterns with different relational implications: uniform latency increase across all contacts indicates busyness; selective latency increase in one relationship with stable baselines elsewhere indicates withdrawal.

Action surfaces alerts when behavioral indicators cross established trajectory thresholds, providing the subject with information about patterns they may be unable to articulate.

Evaluation examines correspondence between inferred trajectory and subsequent relational developments over six-to-twelve-month observation windows. The falsifiability condition: if inferred cooling trajectories do not precede actual behavioral disengagement at rates above the base rate of relationship dissolution, the model requires revision.

Design knowledge. Multi-signal inference detects relational trajectory that single-signal inference misses. The minimum viable inference set is temporal plus comparative signals: latency combined with divergence between stated relationship characterization and observed behavioral pattern. Single-signal systems generate noisy state estimates; multi-signal systems with comparative signals generate trajectory estimates.

Design Proposition 2 (DP2): Multi-signal behavioral inference (temporal, sequential, intensity, and comparative signals combined) predicts relational trajectory with greater accuracy than single-signal approaches in asynchronous text communication.

The validation challenge, stated frankly. Demonstration B faces a problem the current research cannot fully resolve. Bredgaard, Trinhammer, and Bassignana (2025) validated computational attachment classification against trained-coder labels from psychotherapy transcripts, a gold standard that does not exist for naturalistic messaging. Provisional approaches are available: content validity via expert panel review of signal-to-construct mappings; ecological correlation over twelve-month windows tracking whether flagged trajectories precede disengagement; aggregate comparison against established instruments (ECR-R attachment scales, Gottman's SPAFF coding) where feasible. These are partial solutions. The full validation is an open problem, the most methodologically challenging in the programme. We state it as such. Identifying this gap is itself a research contribution, and we invite the computational social science community to help develop the methodology BID needs for interpersonal inference contexts.

4.3 Demonstration C: Organizational Absorption of Behavioral Intelligence Systems

Context. Three small-to-medium enterprises in Southeast Asia that deployed behavioral intelligence systems across different service domains. In each case, the system's inferences sometimes contradicted the declared expert judgment of organizational members.

The reframing matters. The subject of inference is not the organization as a unit. It is the individual member whose behavioral signals, aggregated, produce organizational patterns. BID studies organizational resistance to behavioral intelligence by applying behavioral intelligence to the resisters.

Pipeline instantiation. Capture collects system usage logs as the behavioral signal stream. The subjects are the system's own users, whose interaction with the artifact generates the data for studying its absorption.

Operationalization maps usage signals to organizational constructs: override frequency as *resistance intensity*; time-to-first-use across features as *adoption velocity*; divergence between stated AI satisfaction (survey data) and actual system usage rate (log data) as *cognitive dissonance index* (Seran et al., 2025), a direct instantiation of the comparative signal class; manual workaround complexity as *resistance investment*, the behavioral cost incurred to avoid the system's inference.

Inference produces a resistance type classification. Three types emerge from behavioral signatures:

Skill-deficit resistance: the user cannot operate the system. Low usage, high errors, frequent help-seeking. Response: targeted training.

Process-level resistance: the system disrupts established workflows. Selective feature avoidance, high workaround complexity, stable usage in non-disrupting features. Response: integration redesign.

Identity-level resistance: the system reveals that stated expertise diverges from the behavioral pattern the system infers from the same data the user claims mastery of. **High stated satisfaction, high override rates, low workaround complexity.** The agent who consistently overrides signal-based lead assignments, explaining "I know my clients better than any algorithm," is not resisting technology failure. They are resisting the revelation that their behavioral pattern of client engagement does not match their self-description of it. The low workaround complexity is diagnostic: the agent knows how to use the system and can use it easily; they choose not to. This is the most consequential resistance type because it is the hardest to detect and the most resistant to standard interventions. Training does not address it (the user is already skilled). Process redesign does not address it (the process is not the problem). Only the behavioral signature, visible in the interaction logs, distinguishes it from the other types. And it is invisible to satisfaction surveys because the defining marker is a high survey score.

Absorption trajectories are tracked over six-to-twelve-month deployment periods, making this longitudinal inference rather than cross-sectional classification. Override rates, workaround complexity, and unprompted system consultation change over time as absorption progresses or stalls.

Action delivers differentiated interventions by type. Training addresses skill-deficit; integration redesign addresses process-level; graduated exposure starting with confirming cases (where the system agrees with the user's judgment) before introducing disconfirming inferences addresses identity-level resistance.

Evaluation examines trajectories over the deployment period: declining overrides, increasing unprompted consultation, and decreasing workaround complexity indicate successful absorption. The falsifiability condition: if resistance type classification based on behavioral signals does not predict differential response to differentiated interventions at rates above chance, DP3 is falsified.

Design Proposition 3 (DP3): Identity-level AI resistance produces a distinctive behavioral signature (high stated satisfaction diverging from actual usage) that is classifiable from behavioral signals and predicts differential response to targeted interventions.

4.4 What the Three Demonstrations Establish Together

Four things, which is what a founding paper needs.

The pipeline instantiates without modification across consumer, interpersonal, and organizational contexts. Domain theories, signal sources, cultural contexts, and action repertoires all change. The six-stage structure does not. This is evidence of mechanism generality, not proof.

The comparative signal class is consistently the most distinctive element. In Demonstration A, the divergence between declared preferences and browsing is the most informative predictor. In Demonstration B, the divergence between stated satisfaction and observed trajectory is the most diagnostic signal. In Demonstration C, the divergence between stated AI satisfaction and actual usage is the signature of the resistance type most requiring intervention. In every context, the class encoding the stated-revealed gap is the one alternative approaches ignore.

Ecological validity produces findings stated-preference approaches miss. In Demonstration A, a survey of property preferences would have reinforced declared parameters; behavioral inference revealed divergence the survey would mask. In Demonstration B, a satisfaction instrument administered during cooling would have returned positive scores. In Demonstration C, the AI satisfaction survey returned high scores from the most resistant users, rendering them invisible to the instrument designed to detect them.

The design propositions are transferable. DP1, DP2, and DP3 are stated in terms that allow independent researchers in different markets, on different platforms, and in different cultural contexts to test them.

---

5. Research Programme

5.1 The Founding Design Propositions

Three falsifiable propositions emerge from the demonstrations. They are starting points, not the programme's only claims.

DP1: In markets without cooperative data infrastructure, behavioral signal matching outperforms declared-preference matching on inquiry conversion rate and time-to-match.

DP2: Multi-signal behavioral inference predicts relational trajectory with greater accuracy than single-signal approaches in asynchronous text communication.

DP3: Identity-level AI resistance produces a distinctive behavioral signature classifiable from behavioral signals, predicting differential response to targeted interventions.

Each states a condition for falsification. Each is testable in independent contexts. The programme accumulates through confirmation, falsification, and boundary condition mapping.

5.2 Open Questions in Operationalization

How do cultural norms affect signal-to-construct mappings? Thai *kreng jai* (the cultural disposition to avoid imposing) produces response latency baselines reflecting consideration rather than disengagement. This adjustment must be theoretically warranted at the design stage, not discovered as a post-hoc anomaly. Which cultural psychology literature provides the construct that modulates the mapping, and what does the modulated mapping predict differently? The question generalizes beyond Thailand: every non-WEIRD deployment will require culturally grounded operationalization adjustments.

What minimum longitudinal duration is required to infer trajectory rather than state? The answer likely varies: preference crystallization in a property search may become interpretable within a week of use; attachment trajectory inference may require months of accumulated exchange. The programme does not yet address this systematically.

Which signal classes carry protected-attribute proxies, and how should proxy audit be structured? Response latency encodes not just engagement but working conditions, device quality, socioeconomic status, and caregiving responsibilities. A system that operationalizes response latency as disengagement when it actually measures a single mother's inability to check messages during working hours is not generating insight. It is reproducing inequality with a theoretical veneer. Linguistic patterns encode education, ethnicity, and language proficiency. Proxy audit, testing signals for correlation with protected attributes before deployment, is both an ethical requirement and a methodological one. Identifying which signals require audit in which contexts is itself a research contribution the programme needs to formalize.

5.3 Open Questions in Inference

Does deploying a BID artifact change the behavioral signals it measures? The question deserves careful treatment because the answer reveals a structural distinction between BID and observational inference systems.

In recommendation systems, reflexivity is a threat to validity. If the system's output influences the user's next action, which the system then observes as input, the feedback loop contaminates the signal. The system optimizes against its own prior output. Goodhart's Law applies: the measure ceases to be a good measure when it becomes a target (Strathern, 1997). This is a genuine problem for systems that claim to observe behavior objectively while shaping it through their own interventions.

BID's architecture treats reflexivity differently. The cross-system feedback loop, from advertising (prior setting) through platform observation (behavioral capture) through inference (model update) through targeting adjustment (prior revision), is not an unintended consequence. It is the designed mechanism. The advertising creative IS the experimental instrument: it sets a prior that shapes the initial behavioral context, and the system measures the trajectory that emerges under that prior. Different creatives delivered to comparable audiences constitute different experimental conditions. A/B testing across these conditions is not a workaround for contamination; it is interventionist design science (Sein, Henfridsson, Purao, Rossi, & Lindgren, 2011).

The distinction is epistemological. Observational systems claim independence between the observer and the observed; reflexivity violates that claim. Interventionist systems acknowledge that they shape the conditions they study and design their methodology accordingly: controlled variation of the intervention (the prior), measurement of the differential behavioral response (the trajectory), and falsifiable comparison of outcomes across conditions. BID does not pretend to observe from outside. It intervenes, measures the effect of the intervention, and calls the recursive loop the contribution.

This does not eliminate all reflexivity concerns. Within a single session, a subject's awareness that listings are being re-ranked may alter browsing behavior. But the cross-system architecture provides a structural defense: the prior is set in the advertising layer (which the subject experiences as a social media ad, not as a system they are interacting with), the observation occurs on the platform (where the subject is navigating for their own purposes), and the inference operates in a layer the subject does not see. The subject cannot simultaneously game signals across all three stages because they experience them as separate contexts. Multi-stage capture across system boundaries is harder to reflexively contaminate than single-platform observation.

How does the stated-revealed gap vary by stakes, domain, and cultural context across deployed settings? The current demonstrations produce three data points: a high-stakes consumer market, an interpersonal communication context, and organizational technology adoption. The gap may be wider in consumer contexts (where System 1 browsing diverges freely from System 2 declared preferences) than in organizational contexts (where professional norms constrain behavioral deviation from stated policy). Or it may be the reverse: organizations may produce larger gaps precisely because professional identity provides stronger motivation to maintain declared positions. The programme needs dozens of data points from independent teams to begin mapping these boundaries.

5.4 Open Questions in Evaluation

When should evaluation evidence trigger construct revision rather than model refinement? The distinction in Stage 6 is real, but the criterion for choosing is underspecified. How persistent must disconfirmation be before the mapping is revised rather than the weights adjusted?

What criterion validity approaches work when ground-truth labels are unavailable? Every naturalistic interpersonal deployment faces this problem. The computational social science community's work on proxy validation is relevant but needs formalization for BID's specific targets.

5.5 Methodological Standards

A contribution to the BID research programme meets these standards operationally, not aspirationally:

- DSR methodology (Peffers et al., 2007) with Larsen et al. (2025) validity framework applied.
- Behavioral science theory as the operationalization kernel, with a published warrant per construct.
- Ecological validity: deployed artifact, real stakes, longitudinal data, naturalistic context.
- Falsifiability condition stated per design proposition before evaluation.
- Proxy audit completed pre-deployment, with results reported.
- Ethical architecture embedded in design (Section 6.3).

These standards define a BID contribution. A practitioner case study that uses behavioral data is not one. A lab experiment that simulates behavioral inference is not one. The standard is deployed, theoretically grounded, falsifiable, and ethically constrained.

---

6. Discussion

6.1 Implications for Design Science Research

BID closes the loop Hevner et al. (2004) called for and the ISR editorial board (2024) renewed: artifacts designed *from* behavioral science theory, evaluated *through* behavioral science constructs, generating design knowledge that feeds back into theory through construct revision.

The IS contribution is a new class of design knowledge: *behavioral signal design principles*. These take the form: "when designing behavioral inference artifacts for [context], operationalize [signal class] against [construct] to infer [trajectory]." Such principles are actionable, transferable, groundable in evidence, and falsifiable, satisfying Gregor and Hevner's (2013) criteria for prescriptive knowledge contributions.

This knowledge was previously unavailable because the systems that could generate it were proprietary and unexamined, and DSR was never applied to behavioral inference with behavioral science as the design kernel. Consider what a behavioral signal design principle looks like in practice: "In institutional void property markets, operationalize search filter divergence against Samuelson's revealed preference framework to infer the stated-revealed gap score; when the gap exceeds a context-calibrated threshold, route by behavioral profile rather than declared preference." That principle is transferable to any platform operating in an institutional void. It does not depend on the specific listing database, the specific cultural context, or the specific ML model. It transfers because the theoretical warrant transfers. This is the class of knowledge BID makes possible. The gap between the two paradigms that Hevner et al. identified in 2004 remains open. BID proposes a bridge.

6.2 Implications for Behavioral Science

Deployed artifacts are a new validation context for behavioral science constructs, one that laboratories cannot provide and the community has not systematically used. When Gottman's approach-avoidance coding fails to predict disengagement in Thai WhatsApp contexts with the effect sizes the psychotherapy literature reports, that is a finding about the construct's boundary conditions, not just a system malfunction. The gap between what a construct predicts in clinically coded sessions and what it predicts in naturalistic asynchronous text is behavioral science evidence.

BID is therefore not merely a consumer of behavioral science. It is a producer of evidence that behavioral science, operating within laboratory and clinical constraints, cannot generate alone. The irony deserves stating: the discipline that discovered the stated-revealed preference gap has never built the systems that exploit it, and the industry that exploits it has never contributed to the discipline. BID sits at the junction and faces both directions.

6.3 Ethical Architecture: Definitional, Not Aspirational

The ethical requirements are design constraints, not discussion-section additions. Systems that do not satisfy them are not BID systems regardless of their stated purposes.

Consent to inference requires, following Wachter and Mittelstadt (2019), that subjects be informed of what psychological states will be inferred from their behavioral data, not merely what data will be collected. This is a higher bar than standard privacy requirements. Systems that capture behavioral data with consent but infer attachment, resistance, or decision readiness without disclosure operate outside the framework.

Proxy audit as pre-deployment gate requires that each behavioral signal be tested for correlation with protected attributes in the deployment context before the system goes live. Not a post-hoc review. A gate. Response latency in a Thai SME context may encode socioeconomic status, caregiving load, or device quality. The audit must establish this before the signal is operationalized.

Domain prohibition follows structurally from the consent requirement. In criminal justice, immigration, insurance underwriting, and state surveillance, the power asymmetry between operator and subject makes meaningful consent to psychological inference structurally impossible (Zuboff, 2019; Ferguson, 2017; Brayne, 2020). The subject cannot freely withhold consent when the consequence is legal jeopardy, denied access, or state action. The prohibition is not an ethical preference added to BID. It is what the consent requirement, applied consistently, produces as a logical consequence.

This boundary gives BID a structural distinction from the darker applications of the same inference architecture. The distinction rests on the definition, not on the practitioner's intentions. Intentions cannot be verified. Definitions can.

6.4 Limitations: Honest and Generative

Single practitioner. All three artifacts were designed by the same practitioner-researcher, in the same cultural region, with overlapping institutional familiarity. Methodological consistency is a strength for internal validity and a real limitation for independence. The specific invitation: test DP1 in a different institutional void market; test DP2 on a different messaging platform in a non-Thai context; test DP3 in a Western enterprise with different organizational culture.

Southeast Asian specificity. Thai markets, SMEs, and communication norms shape signal environments in ways that may not generalize directly. Weber and Hsee (1998) already implies that the gap is culturally modulated. Cross-cultural replication is a first-order priority.

No independent replication. Three demonstrations, no independent tests. The programme is proposed, not established. The design propositions are first claims, not findings. Publishing now, at the SSRN working paper stage, establishes the framework and propositions before replication, so the community can engage with, contest, extend, and replicate them. A founding paper that waited for independent replication would not be a founding paper. It would be a review.

Experimental pipeline in development. The demonstrations draw on operational deployment data, not controlled experiments. Extensions are in design: controlled lab experiments for DP1 with random assignment of matching algorithm, structured validation protocols for Demonstration B, multi-site organizational studies for DP3. Their absence is the natural state of a founding paper preceding the experimental programme it opens.

---

7. Conclusion

What existed before this paper: a mature inference architecture operating for two decades in commercial practice, proven and unexamined; a behavioral science literature documenting with precision why behavioral signals outpredict declarations, without building the systems the documentation implies; a design science methodology capable of producing transferable knowledge from deployed artifacts, never coupled with behavioral science as the theoretical kernel.

The practice was everywhere. The science was nowhere.

This paper names the practice, formalizes the mechanism, grounds the taxonomy, demonstrates instantiation across three contexts, states falsifiable propositions, and opens a research programme.

Behavioral Intelligence Design is not declared a field here. Fields are declared by communities, not by founding papers. What is claimed is more modest and more durable: the practice exists, the mechanism is coherent, the taxonomy is formal, the methodology is rigorous, the design propositions are testable, the open questions are real, and the citation stake is planted.

I have built commercial implementations of this inference architecture for twenty years. I know where the cold-start problems hide, how proxy contamination creeps in, what happens when temporal baselines shift under a market that nobody mapped because no institution existed to map it. That knowledge has lived in deployed systems and in practitioner intuition. It has not lived in the literature. This paper begins the work of putting it there.

The knowledge comes from the practice. That loop, between operational deployment and theoretical contribution, is not a research strategy. It is the field.

The programme is open.

References

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation.* Erlbaum.

Altman, I., & Taylor, D. A. (1973). *Social penetration: The development of interpersonal relationships*. Holt, Rinehart & Winston.

Ariely, D. (2008). *Predictably irrational: The hidden forces that shape our decisions.* Harper.

Austin, J. L. (1962). *How to do things with words.* Oxford University Press.

Breckler, S. J. (1984). Empirical validation of affect, behavior, and cognition as distinct components of attitude. *Journal of Personality and Social Psychology,* 47(6), 1191-1205.

Boerman, S. C., Kruikemeier, S., & Zuiderveen Borgesius, F. J. (2017). Online behavioral advertising: A literature review and research agenda. *Journal of Advertising,* 46(3), 363-376.

Cotte, B., & Haseley, M. (2014). Determinants and timing of student attrition. *Editorial Board,* 113.

Cotte, B., & Shannon, R. (2018). Operationalizing big data for academic research: A case for consumer behavior. In *Proceedings of the 4th ICAMA.* College of Management, Mahidol University, Bangkok.

Bowlby, J. (1969). *Attachment and loss: Vol. 1. Attachment.* Basic Books.

Brayne, S. (2020). *Predict and surveil: Data, discretion, and the future of policing.* Oxford University Press.

Bredgaard, F., Trinhammer, M. L., & Bassignana, E. (2025). The language of attachment: Modeling attachment dynamics in psychotherapy. *arXiv preprint arXiv:2504.16271*.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly,* 35(1), 128-152.

Ellis, K., & Huang, W. (2026). Should I state or should I show? Aligning AI with human preferences. *arXiv preprint arXiv:2603.29317.*

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research.* Addison-Wesley.

Ferguson, A. G. (2017). *The rise of big data policing: Surveillance, race, and the future of law enforcement.* NYU Press.

Golgeci, I., Ritala, P., Arslan, A., McKenna, B., & Ali, I. (2025). Confronting and alleviating AI resistance in the workplace: An integrative review and a process framework. *Human Resource Management Review,* 35(2), 101075.

Gottman, J. M. (1994). *What predicts divorce? The relationship between marital processes and marital outcomes.* Erlbaum.

Gottman, J. M. (1999). *The marriage clinic: A scientifically based marital therapy.* W. W. Norton.

Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly,* 37(2), 337-355.

Gu, Z., Wang, Q., & Han, S. (2025). Alignment revisited: Are large language models consistent in stated and revealed preferences? *arXiv preprint arXiv:2506.00751*.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly,* 28(1), 75-105.

Himmelstein, M., & Budescu, D. V. (2023). Preference for human or algorithmic forecasting advice does not predict if and how it is used. *Journal of Behavioral Decision Making,* 36(1), e2285.

ISR Editorial Board. (2024). Pathways for design research on artificial intelligence. *Information Systems Research,* 35(4), 1199-1207.

Jo, K., Choi, J., Kim, K., Kim, S., Nguyen, D.-L., Vo, X.-T., Priadana, A., & Tran, T.-D. (2025). Artificial behavior intelligence: Technology, challenges, and future directions. *arXiv preprint arXiv:2505.03315.*

Kahneman, D. (2011). *Thinking, fast and slow.* Farrar, Straus and Giroux.

Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer,* 42(8), 30-37.

Khanna, T., & Palepu, K. (1997). Why focused strategies may be wrong for emerging markets. *Harvard Business Review,* 75(4), 41-51.

Kwak, J., Zhang, Y., & Yu, J. (2021). Digital platforms in institutional voids: Maintenance and creation. *World Development,* 148, 105662.

Larsen, K. R., Lukyanenko, R., Mueller, R. M., Storey, V. C., Parsons, J., VanderMeer, D., & Hovorka, D. (2025). Validity in design science. *MIS Quarterly,* 49(4), 1267-1294.

Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing,* 80(6), 69-96.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems,* 24(3), 45-77.

Pennebaker, J. W. (2011). *The secret life of pronouns: What our words say about us.* Bloomsbury.

Rochet, J.-C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association,* 1(4), 990-1029.

Samuelson, P. A. (1938). A note on the pure theory of consumer's behaviour. *Economica,* 5(17), 61-71.

Searle, J. R. (1975). A taxonomy of illocutionary acts. In K. Gunderson (Ed.), *Language, mind, and knowledge* (pp. 344-369). University of Minnesota Press.

Seran, C. E., Tan, M. J. T., Karim, H. A., & AlDahoul, N. (2025). A conceptual exploration of generative AI-induced cognitive dissonance and its emergence in university-level academic writing. *Frontiers in Artificial Intelligence,* 8, 1573368.

Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. *MIS Quarterly,* 35(1), 37-56.

Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.). MIT Press.

Strathern, M. (1997). 'Improving ratings': Audit in the British university system. *European Review,* 5(3), 305-321.

Tuunanen, T., Winter, R., & vom Brocke, J. (2024). Dealing with complexity in design science research: A methodology using design echelons. *MIS Quarterly,* 48(2), 427-458.

Wachter, S., & Mittelstadt, B. (2019). A right to reasonable inferences: Re-thinking data protection law in the age of big data and AI. *Columbia Business Law Review,* 2019(2), 494-620.

Weber, E. U., & Hsee, C. (1998). Cross-cultural differences in risk perception, but cross-cultural similarities in attitudes towards perceived risk. *Management Science,* 44(9), 1205-1217.

Zahra, S. A., & George, G. (2002). Absorptive capacity: A review, reconceptualization, and extension. *Academy of Management Review,* 27(2), 185-203.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power.* PublicAffairs.

Cotte, B. (2026). Behavioral Intelligence Design: A Framework for AI Systems That Infer Psychological Trajectory. SSRN Working Paper. April 2026.