

DEVELOPING AN EVIDENCE-BASED PERFORMANCE MANAGEMENT AND APPRAISAL SYSTEM IN A SMALL INSURANCE BROKERAGE: AN ACTION RESEARCH STUDY

PENGEMBANGAN SISTEM MANAJEMEN DAN PENILAIAN KINERJA BERBASIS BUKTI PADA PERUSAHAAN PIALANG ASURANSI BERSKALA KECIL: STUDI ACTION RESEARCH

Gerald Andrew Renee Pieloor¹, R. Muh Deddy Hanif. S²

^{1,2}Program Studi Magister Manajemen, Sekolah Tinggi Manajemen Asuransi Trisakti, Jakarta^{1,2}
geraldandrew95@gmail.com¹

ABSTRACT

This study developed an evidence-based performance management and appraisal system for the Technical and Claims functions of a small Indonesian insurance brokerage. The organisation previously relied on managerial judgment without documented individual key performance indicators, rating rubrics, evidence requirements, calibration, or systematic development follow-up. A qualitative applied design using two action research cycles was employed. Data were obtained from a semi-structured interview with the director, focused validation discussions with two department heads and two staff members, limited observation, document review, and prototype simulations. Cycle 1 produced an initial performance management flow and appraisal forms but revealed ambiguous indicators, equal weighting, limited evidence traceability, and no mid-year review. Cycle 2 integrated weighted KPIs, operational targets, evidence sources, behavioural rubrics, a mid-year review, calibration, employee responses, and individual development plans, with Finance and General Affairs designated as process owner. The study proposes a minimum performance-governance architecture in which technical measurement devices must be coupled with social calibration and process ownership. The prototype demonstrated initial contextual feasibility, but its effectiveness and inter-rater reliability were not tested longitudinally.

Keywords: performance management; performance appraisal; action research; insurance brokerage; key performance indicators

ABSTRAK

Penelitian ini mengembangkan sistem manajemen dan penilaian kinerja berbasis bukti bagi Bagian Teknik dan Klaim pada perusahaan pialang asuransi berskala kecil di Indonesia. Sebelumnya, perusahaan mengandalkan pertimbangan manajerial tanpa KPI individual terdokumentasi, rubrik penilaian, standar bukti, kalibrasi, dan tindak lanjut pengembangan yang sistematis. Penelitian terapan kualitatif dilakukan melalui dua siklus action research. Data diperoleh dari wawancara semi-terstruktur dengan direktur, diskusi validasi bersama dua kepala bagian dan dua staf, observasi terbatas, telaah dokumen, serta simulasi prototipe. Siklus pertama menghasilkan alur manajemen kinerja dan formulir awal, tetapi masih menunjukkan indikator ambigu, bobot seragam, keterlacakan bukti terbatas, dan belum adanya evaluasi tengah tahun. Siklus kedua mengintegrasikan KPI berbobot, target operasional, sumber bukti, rubrik perilaku, evaluasi tengah tahun, kalibrasi, tanggapan karyawan, dan rencana pengembangan individu, dengan Finance and General Affairs sebagai pemilik proses. Penelitian mengusulkan arsitektur minimum tata kelola kinerja yang memadukan perangkat pengukuran teknis, kalibrasi sosial, dan kepemilikan proses. Prototipe menunjukkan kelayakan kontekstual awal, tetapi efektivitas dan reliabilitas antarpemilai belum diuji secara longitudinal.

Kata kunci: Manajemen Kinerja; Penilaian Kinerja; Action Research; Pialang Asuransi; Indikator Kinerja Utama

PENDAHULUAN

Manajemen kinerja semakin dipahami sebagai proses berkelanjutan yang menghubungkan sasaran organisasi, ekspektasi individual, pemantauan, umpan balik, evaluasi, dan

pengembangan. Armstrong dan Taylor (2023) menempatkan manajemen kinerja sebagai praktik manajerial yang berjalan sepanjang tahun, sedangkan Aguinis (2023) menekankan proses identifikasi, pengukuran, dan pengembangan kinerja

yang selaras dengan tujuan strategis. Pergeseran ini penting karena penilaian tahunan yang berdiri sendiri sering tidak menyediakan informasi tepat waktu untuk memperbaiki kinerja dan mudah berubah menjadi ritual administratif.

Literatur mutakhir menunjukkan bahwa permasalahan sistem kinerja tidak hanya berasal dari instrumen. DeNisi (2024) menunjukkan bahwa kajian penilaian kinerja telah bergerak dari fokus pada akurasi rating menuju perhatian yang lebih luas terhadap konteks, proses sosial, dan penggunaan informasi kinerja. Maley et al. (2024) juga menegaskan bahwa lingkungan kerja yang cepat berubah menuntut sistem yang adaptif dan terhubung dengan pengembangan talenta. Dengan demikian, efektivitas sistem ditentukan oleh hubungan antara desain ukuran, kualitas interaksi manajerial, dan tata kelola pelaksanaan.

Pada tingkat implementasi, manajer lini memadukan sistem formal dan informal sesuai dengan kondisi kerja. Lee et al. (2021) menemukan bahwa praktik formal, seperti sasaran dan review periodik, berjalan bersama percakapan informal dan pemecahan masalah sehari-hari. Pada organisasi kecil dan menengah, sistem pengukuran kinerja juga berkembang secara bertahap mengikuti budaya, sistem informasi, dan gaya kepemimpinan (Sardi et al., 2020). Formalisasi karenanya perlu proporsional: cukup jelas untuk membangun akuntabilitas, tetapi tidak terlalu kompleks sehingga menghambat penggunaan.

Tantangan utama penilaian kinerja adalah membangun kredibilitas keputusan. Pada organisasi kecil, kedekatan antara pimpinan dan karyawan dapat mempercepat komunikasi, tetapi juga meningkatkan ketergantungan pada ingatan dan kesan personal. Belsito dan Reutzel (2020) menunjukkan bahwa

formalisasi appraisal dalam usaha kecil berkaitan dengan dinamika kepercayaan terhadap kepemimpinan. Kivipöld et al. (2020) menegaskan keterkaitan antara penggunaan sistem appraisal, persepsi keadilan, dan efektivitas organisasi. Sementara itu, Micacchi et al. (2024) menunjukkan bahwa desain appraisal yang mendukung keadilan dapat memperkuat keterlibatan kerja.

Konsekuensi appraisal juga bekerja melalui mekanisme psikologis. Ghani et al. (2024) menemukan bahwa pengaruh appraisal terhadap perilaku ekstra-peran dan niat keluar dimediasi oleh komitmen serta rasa berdaya. Lyu et al. (2023) menunjukkan bahwa keadilan appraisal memengaruhi kinerja melalui jalur psikologis yang berbeda. Temuan tersebut memperlihatkan bahwa indikator yang rinci belum cukup apabila karyawan tidak memahami dasar penilaian atau tidak memiliki ruang untuk merespons keputusan.

Penggunaan data dan bukti menjadi bagian penting dari tata kelola kinerja. Peeters et al. (2020) menempatkan efektivitas analitik SDM pada kemampuan organisasi mengubah data menjadi keputusan yang relevan. Pada tingkat operasional, kamus KPI, sumber data, frekuensi pengukuran, dan standar bukti dapat mengurangi variasi interpretasi. Camilleri (2024) menekankan bahwa KPI harus didefinisikan, dapat diukur, dan terhubung dengan tujuan organisasi. Namun, data tidak bersifat netral; pemilihan indikator, bobot, dan ambang penilaian tetap memerlukan pertimbangan manajerial dan proses kalibrasi.

Kebutuhan tersebut semakin relevan pada industri asuransi. Peta Jalan Pengembangan dan Penguatan Perasuransian Indonesia 2023-2027 menempatkan kompetensi SDM, tata kelola, kualitas layanan, dan transformasi

digital sebagai agenda penguatan industri (Otoritas Jasa Keuangan, 2023). POJK Nomor 34 Tahun 2024 menegaskan pengembangan kualitas SDM yang terencana dan berkelanjutan pada sektor perasuransian, penjaminan, dan dana pensiun (Otoritas Jasa Keuangan, 2024). Pada saat yang sama, perubahan teknologi mengubah proses underwriting dan klaim serta meningkatkan kebutuhan atas kompetensi analitis dan pengambilan keputusan (Eling et al., 2022).

Penelitian dilakukan pada PT XYZ, perusahaan pialang asuransi dengan sekitar 30 karyawan. Fokus penelitian adalah Bagian Teknik dan Klaim karena kedua fungsi tersebut membentuk rangkaian layanan yang saling bergantung. Bagian Teknik menangani analisis risiko, quotation, placement, review polis, endorsement, dan renewal. Bagian Klaim menangani pencatatan, notifikasi, pengumpulan dokumen, monitoring, analisis coverage, serta review penyelesaian klaim. Kinerja kedua bagian tidak cukup diukur melalui volume pekerjaan; ketepatan proses, kualitas analisis, dokumentasi, komunikasi, dan integritas juga menentukan kualitas layanan.

Sebelum penelitian, PT XYZ telah melakukan pemantauan operasional melalui rapat produksi dan pembahasan klaim. Namun, perusahaan belum memiliki siklus manajemen kinerja formal, KPI individual, rubrik rating, standar bukti, kalibrasi penilai, serta dokumentasi yang konsisten. Penilaian bergantung pada pengamatan kepala bagian dan diskusi dengan direksi. Kondisi ini menimbulkan dua masalah. Pertama, karyawan tidak selalu memperoleh ekspektasi kinerja yang sama sejak awal periode. Kedua, dasar keputusan sulit ditelusuri karena catatan penilaian tidak dikelola sebagai arsip kinerja.

Penelitian terdahulu lebih banyak menguji hubungan antara appraisal, keadilan, keterlibatan, komitmen, dan kinerja. Kajian tersebut menjelaskan mengapa appraisal penting, tetapi relatif sedikit yang memperlihatkan bagaimana perusahaan jasa profesional berskala kecil membangun sistem dari kondisi informal menuju tata kelola yang dapat digunakan. Gap operasional mencakup penerjemahan pekerjaan ke dalam KPI, penetapan sumber bukti, pembentukan rubrik, penyelarasan penilai, penetapan pemilik proses, serta hubungan hasil appraisal dengan pengembangan.

Penelitian ini bertujuan mengembangkan dan menyempurnakan prototipe manajemen kinerja dan penilaian kinerja bagi Bagian Teknik dan Klaim PT XYZ melalui dua siklus action research. Kontribusi penelitian dirumuskan sebagai arsitektur minimum tata kelola kinerja bagi organisasi kecil. Arsitektur tersebut tidak hanya terdiri atas perangkat teknis berupa KPI, target, bobot, bukti, dan rubrik, tetapi juga perangkat sosial berupa review berkala, diskusi hasil, kalibrasi, mekanisme ketidaksepakatan, dan rencana pengembangan. Penelitian tidak mengklaim telah membuktikan peningkatan kinerja atau reliabilitas statistik; fokusnya adalah kelayakan kontekstual awal dan pembelajaran desain.

METODE PENELITIAN

Penelitian menggunakan desain terapan kualitatif dengan pendekatan action research. Pendekatan ini dipilih karena tujuan penelitian bukan hanya mendeskripsikan kelemahan sistem, tetapi mengembangkan, mencoba secara terbatas, mengevaluasi, dan menyempurnakan prototipe bersama pengguna. Action research menggabungkan tindakan dan penyelidikan melalui siklus yang

berulang (Burkholder et al., 2020). Dalam studi organisasi, pendekatan ini menuntut keterkaitan antara masalah praktis, partisipasi pemangku kepentingan, pengumpulan bukti, refleksi, dan perubahan tindakan (Coghlan & Shani, 2021).

Peneliti merupakan pihak eksternal yang bekerja di industri asuransi, tetapi tidak menjadi karyawan atau pengambil keputusan di PT XYZ. Posisi tersebut memberi jarak profesional sekaligus pengetahuan kontekstual. Risiko bias dikelola melalui reflektivitas, konfirmasi temuan, triangulasi teknik, dokumentasi keputusan revisi, dan pembatasan klaim. Peneliti bertindak sebagai fasilitator diagnosis dan pengembangan prototipe, sedangkan keputusan operasional tetap berada pada pihak perusahaan.

Partisipan terdiri atas seorang direktur sebagai informan utama, Kepala Bagian Teknik dan Kepala Bagian Klaim sebagai peserta validasi manajerial, serta satu staf Teknik dan satu staf Klaim sebagai perwakilan pengguna. Data diperoleh melalui wawancara semi-terstruktur, diskusi validasi, observasi terbatas atas alur dan dokumen kerja, telaah uraian pekerjaan dan dokumen operasional, serta simulasi penggunaan formulir. Struktur partisipan dipilih berdasarkan daya informasi dan keterlibatan langsung dalam proses kinerja, bukan untuk menghasilkan generalisasi statistik.

Prosedur mengacu pada empat tahap, yaitu *planning*, *acting*, *developing*, dan *reflecting*. Siklus pertama berfungsi sebagai eksplorasi dan perumusan masalah. Tahap *planning* memetakan praktik awal, sumber subjektivitas, karakter pekerjaan, serta kebutuhan indikator. Tahap *acting* menghasilkan Prototipe V1 berupa alur manajemen kinerja dan formulir appraisal awal. Tahap *developing* mengevaluasi kejelasan indikator, kelayakan bukti,

bobot, skala, dan alur. Tahap *reflecting* menetapkan prioritas revisi untuk siklus kedua.

Siklus kedua berfungsi sebagai validasi dan penyempurnaan. *Planning* menggunakan hasil refleksi siklus pertama untuk menetapkan kriteria perbaikan. *Acting* menerapkan Prototipe V2 melalui simulasi terbatas dan diskusi pengguna. *Developing* memperbaiki KPI, target, bobot, rubrik, bukti, pembagian peran, review tengah tahun, kalibrasi, dan tindak lanjut pengembangan. *Reflecting* menghasilkan paket prototipe final dan batas implementasinya.

Analisis menggunakan analisis isi yang ditempatkan dalam proses analisis data interaktif. Pernyataan dan catatan yang relevan dipilah menjadi unit makna, diberi kode awal, dikelompokkan ke dalam kategori, dan dibandingkan lintas sumber serta lintas siklus. Kategori utama mencakup ketidakjelasan indikator, subjektivitas penilaian, lemahnya dokumentasi, ketidaktersediaan bukti, perbedaan interpretasi, kebutuhan review, kepemilikan proses, dan pengembangan karyawan. Matriks temuan-revisi digunakan untuk menelusuri hubungan antara bukti Cycle 1 dan perubahan pada Cycle 2.

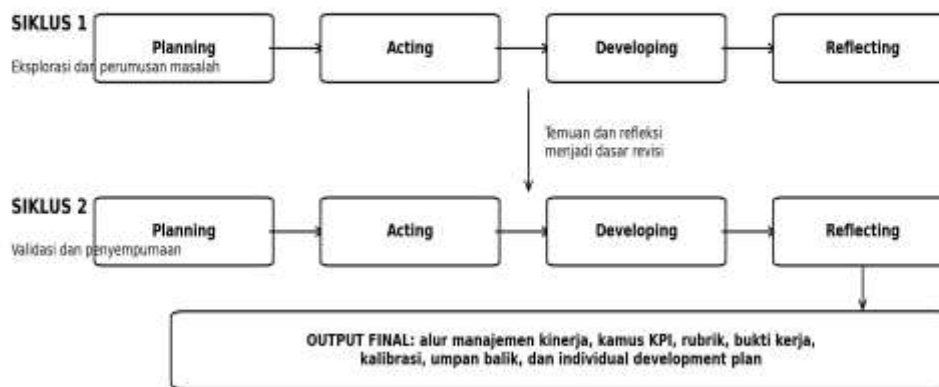
Kredibilitas diperkuat melalui triangulasi sumber terbatas, triangulasi teknik, triangulasi waktu, serta konfirmasi kepada pengguna. Dependabilitas dijaga melalui versi prototipe, notulen, catatan perubahan, dan matriks perbandingan. *Confirmability* dibangun dengan menghubungkan setiap perubahan pada indikator atau alur dengan data, karakter pekerjaan, atau teori. Prinsip kualitas *action research* juga mengikuti kebutuhan transparansi proses, kolaborasi, refleksi, dan kegunaan praktis sebagaimana ditekankan Casey et al. (2023) serta Oberholzer dan Buys (2023).

Validasi prototipe dibatasi pada validasi isi dan kesesuaian kontekstual, uji penggunaan melalui simulasi, serta kalibrasi awal secara kualitatif. Kelayakan awal ditetapkan apabila KPI memiliki definisi, target, bobot, cara ukur, sumber data, dan bukti; formulir dapat digunakan tanpa kendala substantif; alur menjelaskan peran; dan hasil penilaian terhubung dengan umpan balik serta pengembangan. Penelitian tidak melakukan uji Content Validity Index,

Aiken's V, Cohen's Kappa, intraclass correlation, atau evaluasi dampak longitudinal.

Seluruh identitas individu dan perusahaan disamarkan. Informan diberi penjelasan mengenai tujuan, penggunaan data, dan kerahasiaan. Dokumen sensitif tidak disajikan secara rinci. Data digunakan untuk pengembangan sistem dan penelitian akademik, bukan untuk mengambil keputusan individual selama studi.

Gambar 1. Prosedur dua siklus action research



Tabel 1. Sumber data dan fungsi analitis

Sumber	Jumlah/Unit	Fungsi dalam penelitian
Wawancara direktur	1 informan utama	Mendiagnosis sistem awal, kebijakan, dasar keputusan, dan kebutuhan perusahaan.
Diskusi validasi	2 kepala bagian dan 2 staf	Menilai relevansi KPI, kejelasan rubrik, bukti kerja, bobot, dan kemudahan penggunaan.
Dokumen kerja	Uraian pekerjaan, SOP, log dan formulir terkait	Memeriksa kesesuaian indikator serta mengidentifikasi sumber bukti.
Observasi terbatas	Alur kerja Teknik dan Klaim	Memahami proses, keluaran pekerjaan, koordinasi, dan keterbatasan operasional.
Simulasi prototipe	Cycle 1 dan Cycle 2	Mengidentifikasi masalah penggunaan dan menentukan revisi perangkat.

HASIL DAN PEMBAHASAN

Kondisi Awal: Pemantauan Operasional Tanpa Siklus Formal

PT XYZ telah memiliki pemantauan pekerjaan melalui rapat produksi, pembahasan renewal, dan rapat klaim. Namun, pemantauan tersebut berfokus pada status pekerjaan, bukan pada siklus kinerja individual yang dimulai dari kesepakatan sasaran dan berakhir pada pengembangan. Informan utama menjelaskan bahwa monitoring telah berjalan secara operasional, tetapi belum terdapat alur baku yang menghubungkan target, review berkala, penilaian akhir, dan tindak lanjut.

Data kerja sebenarnya tersedia dalam bentuk status renewal, endorsement, placing, claim register, korespondensi, serta perkembangan klaim. Akan tetapi, data tersebut belum dikelola sebagai bukti kinerja individual. Dokumen penilaian juga tidak diarsipkan secara konsisten. Informan menyatakan bahwa penilaian pernah dicatat pada kertas sederhana dan dibuang setelah keputusan dibuat. Kondisi ini menurunkan keterlacakan keputusan dan memperbesar ketergantungan pada ingatan penilai.

Praktik awal menggunakan aspek umum seperti kinerja, disiplin, kerja sama, sikap, dan kontribusi. Kepala bagian kemudian memberikan pertimbangan atau urutan karyawan untuk didiskusikan dengan direksi. Proses ini cepat dan sesuai dengan struktur organisasi yang ramping, tetapi tidak menyediakan definisi kinerja yang sama, target, bobot, atau bukti yang dapat diperiksa. Masalah inti karena itu bukan ketiadaan formulir semata, melainkan ketiadaan arsitektur tata kelola yang menyatukan pengukuran, dialog, dokumentasi, dan tindak lanjut.

Cycle 1: Struktur Awal dan Keterbatasan Prototipe

Cycle 1 menghasilkan dua artefak utama. Pertama, alur manajemen kinerja yang dimulai dari uraian pekerjaan dan tujuan departemen, dilanjutkan diskusi KPI dan target, pelaksanaan kerja, coaching bila diperlukan, penilaian akhir, diskusi hasil, persetujuan, dan eskalasi apabila terjadi ketidaksepakatan. Kedua, formulir appraisal awal untuk Bagian Teknik dan Klaim yang memuat aspek pekerjaan utama dan perilaku kerja.

Pada Bagian Teknik, aspek awal meliputi kecepatan quotation dan placement, kualitas analisis risiko, review polis dan endorsement, dukungan renewal, dokumentasi teknis, dan koordinasi. Pada Bagian Klaim, aspek awal meliputi penerimaan dan pencatatan laporan, notifikasi kepada insurer, kelengkapan dokumen, koordinasi dengan insurer atau loss adjuster, monitoring, analisis coverage, dan settlement.

Diskusi bersama dua staf dan dua kepala bagian menunjukkan bahwa cakupan pekerjaan sudah relevan, tetapi dasar pemberian skor belum jelas. Peserta mempertanyakan perbedaan antara nilai 3, 4, dan 5; bukti yang harus digunakan; dan alasan seluruh indikator memiliki pengaruh yang sama. Nilai akhir berbasis rata-rata sederhana tidak membedakan indikator inti dan perilaku pendukung. Dengan demikian, masalah Cycle 1 terletak pada operasionalisasi, bukan pada daftar pekerjaan.

Refleksi Cycle 1 menghasilkan empat prioritas. Pertama, indikator harus dilengkapi definisi, target, cara ukur, dan bukti. Kedua, bobot perlu menunjukkan kepentingan relatif. Ketiga, review tengah tahun perlu ditambahkan agar sistem tidak bersifat retrospektif. Keempat, proses membutuhkan pemilik yang mengelola jadwal, distribusi, arsip, dan tindak lanjut.

Tabel 2. Perubahan desain dari Cycle 1 ke Cycle 2

Komponen	Cycle 1	Temuan evaluasi	Cycle 2
Alur kinerja	Perencanaan, pelaksanaan, penilaian akhir, persetujuan.	Belum ada titik koreksi selama periode.	Ditambah mid-year review, diskusi kendala, tindak lanjut, dan coaching.
KPI	Aspek kerja umum.	Definisi dan cara ukur ambigu.	KPI operasional dengan target, rumus/cara ukur, sumber data, dan bukti.
Bobot	Relatif seragam/rata-rata.	Tidak mencerminkan prioritas pekerjaan.	Kinerja utama 70% dan kompetensi/perilaku 30%.
Skala	Nilai 1-5 tanpa jangkar rinci.	Perbedaan skor sulit dijelaskan.	Rubrik berbasis tingkat pencapaian dan kualitas bukti.
Bukti	Kolom catatan atasan.	Penilaian tetap bertumpu pada persepsi.	Daftar bukti wajib dan catatan alasan pemberian nilai.
Kalibrasi	Belum tersedia.	Potensi perbedaan standar Head dan Direksi.	Forum Head-Direksi-FGA sebelum finalisasi.
Tindak lanjut	Coaching/pelatihan umum.	Tidak berbasis diagnosis individual.	Individual development plan dengan area, waktu, dan PIC.
Kepemilikan proses	Belum tegas.	Risiko sistem tidak berjalan konsisten.	FGA sebagai process owner administratif.

Cycle 2: Integrasi KPI, Bukti, dan Dialog Kinerja

Cycle 2 menyempurnakan alur menjadi siklus yang lebih lengkap. Sasaran individu diturunkan dari uraian pekerjaan dan tujuan departemen, kemudian dibahas oleh karyawan dan kepala bagian. Mid-year review digunakan untuk menilai capaian sementara, kendala, dan tindakan korektif. Setelah itu, karyawan melaksanakan tindak lanjut dan memperoleh coaching atau pelatihan

sesuai kebutuhan. Penilaian akhir merangkum bukti sepanjang periode, diikuti diskusi hasil, tanggapan karyawan, persetujuan, dan eskalasi apabila terjadi ketidaksepakatan.

Formulir Cycle 2 membagi kinerja utama sebesar 70% dan kompetensi atau perilaku kerja sebesar 30%. Pembagian tersebut mempertahankan fokus pada hasil dan proses. Pada pekerjaan Teknik dan Klaim, kualitas keputusan, kepatuhan, komunikasi, dan integritas tidak dapat dipisahkan dari angka output. Pendekatan gabungan juga sejalan

dengan pandangan Dessler (2020) bahwa standar appraisal harus terkait pekerjaan dan dapat dipertanggungjawabkan, serta dengan kerangka appraisal yang membedakan hasil, perilaku, dan konteks.

KPI Teknik mencakup ketepatan quotation dan placement, kualitas analisis risiko dan kebutuhan pertanggung, akurasi review polis dan endorsement, ketepatan renewal, serta kelengkapan dokumentasi teknis. Bukti yang digunakan antara lain quotation log, email insurer, comparison sheet, risk information sheet, placing slip, policy review checklist, endorsement checklist, renewal register, dan folder klien.

KPI Klaim mencakup ketepatan pencatatan dan notifikasi klaim, kelengkapan dan akurasi dokumen, konsistensi monitoring dan update, kualitas coverage review dan analisis klaim, serta kualitas review settlement atau penolakan. Bukti yang digunakan antara lain claim register, email notifikasi, timestamp laporan, claim document checklist, claim monitoring sheet, coverage review sheet, settlement review, dan draft response.

Setiap KPI dilengkapi target atau standar minimal, persentase pencapaian, nilai tertimbang, dan catatan penilai. Indikator kuantitatif menggunakan indeks pencapaian sesuai karakter ukuran, sedangkan indikator kualitatif menggunakan rubrik yang menghubungkan skor dengan deskripsi kualitas dan bukti. Batas maksimum pencapaian diperlukan agar satu indikator tidak mendominasi nilai total. Kategori akhir menempatkan pemenuhan target sebagai standar, bukan sebagai kinerja luar biasa.

Cycle 2 juga memperjelas peran. Karyawan menyepakati target, menyediakan bukti, merespons hasil, dan menjalankan rencana pengembangan. Kepala bagian menetapkan sasaran

bersama karyawan, memonitor, memberi coaching, dan menyusun penilaian. Direksi memberi keputusan akhir, khususnya ketika terdapat perbedaan. FGA mengelola kalender, distribusi formulir, pengumpulan bukti, arsip, kalibrasi, dan tindak lanjut.

Arsitektur Minimum Tata Kelola Kinerja

Temuan menunjukkan bahwa peningkatan kualitas appraisal memerlukan lebih dari standarisasi metrik. Sistem yang dikembangkan dapat dipahami sebagai arsitektur minimum tata kelola kinerja dengan tiga lapisan. Lapisan pertama adalah **infrastruktur pengukuran**, yang terdiri atas KPI, target, bobot, cara ukur, sumber data, bukti, dan rubrik. Lapisan kedua adalah **infrastruktur dialog**, yang terdiri atas review tengah tahun, feedback, tanggapan karyawan, kalibrasi, dan mekanisme ketidaksepakatan. Lapisan ketiga adalah **infrastruktur keberlanjutan**, yang terdiri atas process owner, arsip, kalender, serta individual development plan.

Lapisan pengukuran mengurangi ruang interpretasi tanpa menghapus pertimbangan profesional. Penilaian pada pekerjaan jasa tidak selalu dapat diubah menjadi angka murni karena kualitas analisis, komunikasi, dan advokasi membutuhkan judgment. Karena itu, objektivitas dalam penelitian ini dipahami sebagai keterlacakan alasan dan konsistensi prosedur, bukan hilangnya seluruh unsur judgment. Barbieri et al. (2023) menunjukkan bahwa kredibilitas appraisal bergantung pada desain dan dukungan implementasi, sedangkan Rubin dan Edwards (2020) memperingatkan bahwa formalisasi tidak otomatis menghilangkan keluhan apabila proses dipersepsikan tidak adil.

Lapisan dialog menjawab keterbatasan sistem yang hanya

mengandalkan formulir. Kalibrasi mempertemukan penilai untuk membahas bukti dan alasan pemberian skor. Forum tersebut tidak bertujuan menyeragamkan seluruh nilai, tetapi menguji konsistensi penerapan standar. Ruang tanggapan karyawan penting karena appraisal memengaruhi komitmen, engagement, dan kepercayaan. Memon et al. (2021) menunjukkan bahwa kepuasan terhadap appraisal berkaitan dengan engagement dan turnover intention, sedangkan Sultana dan Khandakar (2022) menempatkan praktik HRM yang adil sebagai sumber kepercayaan impersonal.

Review tengah tahun menggeser appraisal dari kejutan tahunan menjadi percakapan kinerja. Umpan balik yang terlambat sulit digunakan untuk koreksi, sedangkan umpan balik negatif yang tidak dikelola dapat menghasilkan coping yang merugikan kinerja (Wu et al., 2024). Dalam rancangan PT XYZ, review digunakan untuk mengidentifikasi hambatan, menyesuaikan dukungan, dan mencatat tindakan. Hal ini membuat penilaian akhir lebih merupakan rangkuman bukti yang telah dibahas daripada keputusan yang muncul secara tiba-tiba.

Lapisan keberlanjutan menjelaskan mengapa FGA ditetapkan sebagai process owner. Pada organisasi kecil, sistem sering bergantung pada komitmen individu pimpinan. Tanpa pemilik proses, jadwal mudah terlewat, dokumen tidak terkumpul, dan rencana pengembangan tidak dipantau. Penetapan FGA tidak memindahkan tanggung jawab penilaian dari kepala bagian, tetapi menciptakan akuntabilitas administratif agar siklus berjalan.

Temuan juga memperlihatkan bahwa individual development plan merupakan penghubung antara appraisal dan pengembangan. Sebelumnya, dukungan lebih banyak berupa koreksi ketika masalah terjadi. Prototipe mengubah hasil penilaian menjadi area pengembangan, tindakan, waktu, dan PIC. Dengan demikian, appraisal tidak berhenti sebagai klasifikasi karyawan, tetapi menjadi diagnosis kebutuhan. Siraj dan Hågen (2023) menegaskan peran sistem manajemen kinerja dalam mendukung kinerja pada konteks UKM, tetapi hasil penelitian ini menambahkan bahwa hubungan tersebut perlu dimediasi oleh bukti, dialog, dan tindak lanjut yang terkelola.

Tabel 3. Arsitektur minimum tata kelola kinerja yang dihasilkan

Lapisan	Komponen utama	Risiko yang dikurangi	Batas klaim penelitian
Infrastruktur pengukuran	KPI, target, bobot, rubrik, sumber data, bukti.	Indikator ambigu, recency bias, keputusan tanpa dasar.	Belum diuji validitas isi kuantitatif dan reliabilitas antarpemilai.
Infrastruktur dialog	Mid-year review, feedback, tanggapan karyawan, kalibrasi, eskalasi.	Perbedaan interpretasi, kejutan akhir tahun, rendahnya voice.	Kalibrasi masih berupa simulasi kualitatif terbatas.
Infrastruktur keberlanjutan	FGA process owner, kalender, arsip, IDP,	Sistem berhenti sebagai dokumen dan tidak berulang.	Belum diuji dalam satu periode kinerja penuh.

Lapisan	Komponen utama	Risiko yang dikurangi	Batas klaim penelitian
	monitoring tindak lanjut.		

Kontribusi Teoretis dan Praktis

Secara teoretis, penelitian memperkuat pandangan bahwa sistem kinerja bersifat sosio-teknis. Komponen teknis menyediakan struktur pengukuran, sedangkan komponen sosial menyediakan makna, interpretasi, voice, dan legitimasi. Intervensi yang hanya memperbaiki formulir berisiko menghasilkan kepatuhan administratif tanpa perubahan praktik. Sebaliknya, dialog tanpa standar berisiko kembali pada subjektivitas. Nilai desain muncul ketika kedua komponen dipasang.

Penelitian juga menawarkan tiga prinsip desain kontekstual. Pertama, **keterlacakan bukti harus mendahului klaim objektivitas**. Kedua, **kalibrasi diperlukan ketika indikator tetap membutuhkan judgment profesional**. Ketiga, **kepemilikan proses diperlukan agar siklus bertahan pada organisasi kecil**. Prinsip tersebut tidak diklaim universal, tetapi dapat menjadi proposisi awal untuk diuji pada perusahaan jasa profesional lain.

Secara praktis, prototipe menyediakan alur yang dapat diterapkan bertahap tanpa aplikasi digital. PT XYZ dapat memulai dengan spreadsheet, formulir terstruktur, folder bukti, dan kalender review. Pendekatan ini sesuai dengan kapasitas organisasi dan menghindari desain yang terlalu kompleks. Setelah satu periode penuh, perusahaan dapat menilai waktu administrasi, kualitas bukti, konsistensi rating, penerimaan pengguna, dan dampak terhadap pengembangan.

Pendekatan action research memungkinkan masalah dan solusi

berkembang melalui pembelajaran. Cycle 1 tidak diperlakukan sebagai kegagalan, melainkan sebagai alat untuk memperlihatkan ambiguitas yang tidak terlihat pada tahap konseptual. Cycle 2 kemudian mengubah masukan pengguna menjadi revisi yang dapat ditelusuri. Hal ini sejalan dengan kritik terhadap performance management yang terlalu berfokus pada desain formal dan mengabaikan kualitas implementasi (Pulakos et al., 2019).

Meski demikian, temuan harus dibaca sesuai batasnya. Prototipe belum membuktikan bahwa bias penilai berkurang, engagement meningkat, atau kinerja organisasi membaik. Studi belum menguji validitas statistik, reliabilitas, dan konsekuensi perilaku. Oleh sebab itu, istilah kelayakan awal lebih tepat daripada efektivitas. Penggunaan jangka panjang juga dapat memunculkan risiko gaming, inflasi bukti, beban administrasi, atau ketegangan antara target kuantitatif dan kualitas layanan.

PENUTUP

Kesimpulan

Penelitian menghasilkan prototipe manajemen kinerja dan penilaian kinerja untuk Bagian Teknik dan Klaim PT XYZ melalui dua siklus action research. Kondisi awal menunjukkan pemantauan operasional tanpa sasaran individual terdokumentasi, bukti terstruktur, rubrik, review tengah tahun, kalibrasi, dan tindak lanjut pengembangan. Cycle 1 menghasilkan alur dan formulir awal, tetapi evaluasi mengungkap indikator ambigu, bobot seragam, dan dasar penilaian yang belum dapat ditelusuri.

Cycle 2 mengintegrasikan KPI berbobot, target, sumber bukti, rubrik, mid-year review, feedback, tanggapan karyawan, kalibrasi, individual development plan, dan FGA sebagai process owner. Temuan utama menunjukkan bahwa penguatan objektivitas appraisal membutuhkan hubungan antara infrastruktur pengukuran, infrastruktur dialog, dan infrastruktur keberlanjutan. Formulir yang baik tidak cukup tanpa bukti, kalibrasi, voice, dan pengelolaan proses.

Kontribusi penelitian bersifat kontekstual-terapan. Prototipe menunjukkan kelayakan awal sebagai dasar implementasi bertahap pada organisasi kecil, tetapi belum membuktikan efektivitas, reliabilitas antarpemilai, atau dampak longitudinal. Dengan demikian, hasil penelitian harus digunakan sebagai rancangan awal yang perlu dievaluasi setelah satu periode kinerja penuh.

Saran

PT XYZ disarankan menerapkan prototipe selama minimal satu periode penilaian dengan dua titik review, menyelenggarakan pelatihan penilai, menetapkan batas maksimum indeks pencapaian, dan mengaudit kualitas bukti. FGA perlu memantau kalender, kelengkapan arsip, penyelesaian IDP, dan umpan balik pengguna.

Penelitian selanjutnya perlu menguji validitas isi melalui panel ahli, reliabilitas antarpemilai menggunakan weighted kappa atau intraclass correlation coefficient, serta penerimaan pengguna melalui ukuran kejelasan, waktu pengisian, kelengkapan bukti, dan persepsi keadilan. Desain pra-pasca atau longitudinal diperlukan untuk menilai dampak terhadap kualitas feedback, pencapaian KPI, pengembangan kompetensi, engagement, dan retensi.

Replikasi lintas fungsi dan lintas perusahaan pialang asuransi diperlukan untuk menguji keteralihan prinsip desain. Studi lanjutan juga perlu mengevaluasi risiko perilaku tidak diinginkan, seperti fokus berlebihan pada angka, manipulasi bukti, dan pengabaian pekerjaan penting yang tidak tercantum dalam KPI.

DAFTAR PUSTAKA

- [1] M. Armstrong and S. Taylor. (2023). *Armstrong's Handbook of Human Resource Management Practice: A Guide to the Theory and Practice of People Management* (16th ed.). Kogan Page. London.
- [2] H. Aguinis. (2023). *Performance Management* (5th ed.). SAGE Publications. Thousand Oaks.
- [3] A.S. DeNisi. (2024). A surprising journey through a changing landscape. *Annual Review of Organizational Psychology and Organizational Behavior*, 11: 1-16. <https://doi.org/10.1146/annurev-orgpsych-030822-044158>
- [4] J.F. Maley, M. Dabić, A. Neher, L. Wuersch, L. Martin, and T. Kiessling. (2024). Performance management in a rapidly changing world: Implications for talent management. *Management Decision*, 62(10): 3085-3108. <https://doi.org/10.1108/MD-07-2023-1162>
- [5] Q.Y. Lee, K. Townsend, and A. Wilkinson. (2021). Frontline managers' implementation of the formal and informal performance management systems. *Personnel Review*, 50(1): 379-398. <https://doi.org/10.1108/PR-11-2019-0639>
- [6] A. Sardi, E. Sorano, A. Ferraris, and P. Garengo. (2020). Evolutionary paths of performance measurement and management system: The longitudinal case study of a leading

- SME. *Measuring Business Excellence*, 24(4): 495-510. <https://doi.org/10.1108/MBE-01-2020-0016>
- [7] C.A. Belsito and C.R. Reutzel. (2020). SME employee performance appraisal formalization and trust in leadership change. *International Journal of Organizational Analysis*, 28(2): 434-456. <https://doi.org/10.1108/IJOA-07-2019-1832>
- [8] K. Kivipõld, K. Türk, and L. Kivipõld. (2020). Performance appraisal, justice and organizational effectiveness: A comparison between two universities. *International Journal of Productivity and Performance Management*, 70(1): 87-108. <https://doi.org/10.1108/IJPPM-05-2019-0229>
- [9] L. Micacchi, F. Vidé, G. Giacomelli, and M. Barbieri. (2024). Performance appraisal justice and employees' work engagement in the public sector: Making the most of performance appraisal design. *Public Administration*, 102(3): 815-840. <https://doi.org/10.1111/padm.12952>
- [10] B. Ghani, M.A.R. Malik, and K.R. Memon. (2024). Effects of performance appraisal on employees' extra-role behaviors and turnover intentions: A parallel mediation model. *Personnel Review*, 53(9): 2413-2441. <https://doi.org/10.1108/PR-10-2022-0707>
- [11] B. Lyu, W. Su, Q. Qi, and F. Xiao. (2023). The influence of performance appraisal justice on employee job performance: A dual path model. *SAGE Open*, 13(3). <https://doi.org/10.1177/21582440231194513>
- [12] T. Peeters, J. Paauwe, and K. van de Voorde. (2020). People analytics effectiveness: Developing a framework. *Journal of Organizational Effectiveness: People and Performance*, 7(2): 203-219. <https://doi.org/10.1108/JOEPP-04-2020-0071>
- [13] E. Camilleri. (2024). *Key Performance Indicators: The Complete Guide to KPIs for Business Success*. Routledge. London. <https://doi.org/10.4324/9781032685465>
- [14] Otoritas Jasa Keuangan. (2023). *Peta Jalan Pengembangan dan Penguatan Perasuransian Indonesia 2023-2027*. Otoritas Jasa Keuangan. Jakarta.
- [15] Otoritas Jasa Keuangan. (2024). *Peraturan Otoritas Jasa Keuangan Republik Indonesia Nomor 34 Tahun 2024 tentang Pengembangan Kualitas Sumber Daya Manusia bagi Perusahaan Perasuransian, Lembaga Penjamin, Dana Pensiun, serta Lembaga Khusus Bidang PPDP*. Otoritas Jasa Keuangan. Jakarta.
- [16] M. Eling, D. Nuessle, and J. Staubli. (2022). The impact of artificial intelligence along the insurance value chain and on the insurability of risks. *The Geneva Papers on Risk and Insurance - Issues and Practice*, 47(2): 205-241. <https://doi.org/10.1057/s41288-020-00201-7>
- [17] G. Dessler. (2020). *Human Resource Management* (16th ed.). Pearson. Harlow.
- [18] M. Barbieri, L. Micacchi, F. Vidé, and G. Valotti. (2023). The performance of performance appraisal systems: A theoretical framework for public

- organizations. *Review of Public Personnel Administration*, 43(1): 104-129.
<https://doi.org/10.1177/0734371X211043560>
- [19] E.V. Rubin and A. Edwards. (2020). The performance of performance appraisal systems: Understanding the linkage between appraisal structure and appraisal discrimination complaints. *The International Journal of Human Resource Management*, 31(15): 1938-1957.
<https://doi.org/10.1080/09585192.2018.1424015>
- [20] G.J. Burkholder, K.A. Cox, L.M. Crawford, and J.H. Hitchcock. (2020). *Research Design and Methods: An Applied Guide for the Scholar-Practitioner*. SAGE Publications. Thousand Oaks.
- [21] D. Coghlan and A.B. Shani. (2021). *Conducting Action Research for Business and Management Students*. SAGE Publications. London.
- [22] M. Casey, D. Coghlan, Á. Carroll, and D. Stokes. (2023). Towards a checklist for improving action research quality in healthcare contexts. *Systemic Practice and Action Research*, 36(6): 923-934.
<https://doi.org/10.1007/s11213-023-09635-1>
- [23] M. Oberholzer and P.W. Buys. (2023). A roadmap to business research. In P.W. Buys and M. Oberholzer (Eds.), *Business Research: An Illustrative Guide to Practical Methodological Applications in Selected Case Studies*: 11-40. Palgrave Macmillan. Singapore.
https://doi.org/10.1007/978-981-19-9479-1_2
- [24] M.A. Memon, R. Salleh, M.Z. Mirza, J.-H. Cheah, H. Ting, M.S. Ahmad, and A. Tariq. (2021). Satisfaction matters: The relationships between HRM practices, work engagement and turnover intention. *International Journal of Manpower*, 42(1): 21-50.
<https://doi.org/10.1108/IJM-04-2018-0127>
- [25] S. Sultana and M.S.A. Khandakar. (2022). Do human resource management practices boost up employees' impersonal trust? Evidence from the banking sector of Bangladesh. *Evidence-based HRM*, 10(3): 275-292.
<https://doi.org/10.1108/EBHRM-08-2020-0111>
- [26] S. Wu, Y. Zhang, H. Yang, and T. Tian. (2024). The double-edged sword: Investigating when and why supervisor negative feedback promotes versus inhibits employee performance. *Journal of Managerial Psychology*, 39(4): 452-468.
<https://doi.org/10.1108/JMP-05-2023-0292>
- [27] N. Siraj and I. Hågen. (2023). Performance management system and its role for employee performance: Evidence from Ethiopian SMEs. *Heliyon*, 9(11), e21819.
<https://doi.org/10.1016/j.heliyon.2023.e21819>
- [28] E.D. Pulakos, R. Mueller-Hanson, and S. Arad. (2019). The evolution of performance management: Searching for value. *Annual Review of Organizational Psychology and Organizational Behavior*, 6: 249-271.
<https://doi.org/10.1146/annurev-orgpsych-012218-015009>