



DOI: <https://doi.org/10.38035/jpsn.v4i3.705>
<https://creativecommons.org/licenses/by/4.0/>

Pengembangan Model Evaluasi Pembelajaran Kirkpatrick di PT PLN (Persero) dengan Menggunakan Structural Equation Modeling

Rachmad Ariyadi¹, Sri Gunani Partiw², Erwin Widodo³

¹Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia, ariyadi.rachmad@gmail.com

²Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia, sg.partiwi@its.ac.id

³Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia, erwin@ie.its.ac.id

Corresponding Author: ariyadi.rachmad@gmail.com¹

Abstract: PT PLN (Persero) has adopted the Kirkpatrick model to evaluate learning effectiveness, yet the results of each level are still commonly reported separately. This condition limits empirical evidence on the causal relationships among Reaction, Learning, Behavior, and Result. This study aims to analyze the causal relationships across levels in the Kirkpatrick evaluation model at PT PLN (Persero) using Partial Least Squares Structural Equation Modeling. The data were obtained from the 2025 learning evaluation. Of 775 participant records, 151 were identified as complete, and 100 randomly selected records were used as the final sample based on the minimum sample size calculation using G*Power. Reaction and Learning were assessed by participants, while Behavior and Result were measured using a multisource approach involving both participants and supervisors. The results show that Reaction has a positive and significant effect on Learning. Learning has no significant effect on Behavior. Behavior has a positive and significant effect on Result. These findings indicate that positive learning experiences support learning outcomes, but changes in workplace behavior require organizational support, monitoring, supervisor involvement, and required drivers. This model provides a more integrated and data-driven basis for learning evaluation.

Keyword: Learning Evaluation, Kirkpatrick, Multisource Assessment, PLN, PLS-SEM.

Abstrak: PT PLN (Persero) telah menggunakan model Kirkpatrick untuk mengevaluasi efektivitas pembelajaran, tetapi hasil tiap level masih cenderung dilaporkan secara terpisah. Kondisi ini membatasi bukti empiris mengenai hubungan kausal antara Reaction, Learning, Behavior, dan Result. Penelitian ini bertujuan menganalisis hubungan kausal antarlevel dalam model evaluasi Kirkpatrick di PT PLN (Persero) dengan Partial Least Squares Structural Equation Modeling. Data berasal dari evaluasi pembelajaran tahun 2025. Dari 775 data peserta, 151 data dinyatakan lengkap, dan 100 data acak digunakan sebagai sampel akhir berdasarkan perhitungan kebutuhan sampel minimum dengan G*Power. Evaluasi Reaction dan Learning diisi oleh peserta, sedangkan Behavior dan Result menggunakan pendekatan multisource yang

melibatkan peserta dan atasan. Hasil menunjukkan bahwa Reaction berpengaruh positif dan signifikan terhadap Learning. Learning tidak berpengaruh signifikan terhadap Behavior. Behavior berpengaruh positif dan signifikan terhadap Result. Temuan ini menunjukkan bahwa pengalaman belajar yang positif mendukung hasil belajar, tetapi perubahan perilaku kerja membutuhkan dukungan organisasi, monitoring, peran atasan, dan required drivers. Model ini memberikan dasar evaluasi pembelajaran yang lebih terintegrasi dan berbasis data.

Kata Kunci: Evaluasi Pembelajaran, Kirkpatrick, Multisource Assessment, PLN, PLS-SEM.

PENDAHULUAN

Perubahan lingkungan bisnis, digitalisasi, dan tuntutan akuntabilitas mendorong perusahaan untuk mengembangkan kompetensi sumber daya manusia secara berkelanjutan. Selain itu, George dan Schillebeeckx, (2022) menyampaikan bahwa transformasi digital juga berkaitan dengan tujuan organisasi, tata kelola, transparansi, akuntabilitas, perubahan struktur organisasi, serta kemampuan perusahaan dalam merespon dinamika geopolitik dan tuntutan pasar. Kondisi ini relevan bagi PT PLN (Persero) (selanjutnya dalam penelitian ini disebut PLN) karena perusahaan memiliki peran strategis dalam penyediaan tenaga listrik nasional. Peningkatan kompetensi pegawai menjadi bagian penting untuk mendukung inovasi serta pencapaian kinerja perusahaan (Nguyen & Dao, 2023).

Pelatihan tidak dapat dipahami hanya sebagai aktivitas administratif. Literatur pengembangan sumber daya manusia menunjukkan bahwa pelatihan dapat mendukung peningkatan kompetensi, inovasi, produktivitas, dan kinerja organisasi (Dachner dkk., 2021; Martins, 2021; Nguyen & Dao, 2023). Namun, manfaat pelatihan perlu dibuktikan melalui sistem evaluasi yang mampu menjelaskan hubungan antara pengalaman belajar, hasil belajar, perubahan perilaku, dan dampak pada organisasi (D. L. Kirkpatrick, 1970; J. D. Kirkpatrick & Kirkpatrick, 2016).

Model *Kirkpatrick* banyak digunakan karena membagi evaluasi pelatihan ke dalam empat level, yaitu *Reaction*, *Learning*, *Behavior* dan *Result* (Kirkpatrick 1970; Kirkpatrick and Kirkpatrick 2016). Model ini sederhana dan mudah diterapkan, tetapi mendapat kritik karena hubungan antarlevel sering diasumsikan, bukan diuji secara empiris (Bates, 2004). Skor *Reaction* yang tinggi sering dibaca sebagai tanda keberhasilan pembelajaran, walaupun belum tentu diikuti oleh peningkatan hasil belajar, perubahan perilaku di tempat kerja, atau dampak pada unit kerja.

Permasalahan tersebut juga terjadi dalam evaluasi pembelajaran di PLN Pusdiklat. Evaluasi telah mengadopsi empat level *Kirkpatrick* (Badrul S dkk, 2020), tetapi hasil tiap level masih cenderung dibaca secara parsial. Kondisi ini berpotensi menimbulkan bias ketika organisasi hendak menyimpulkan efektivitas pembelajaran dan kontribusinya terhadap unit kerja. Karena itu, penelitian ini berfokus pada rumusan masalah tentang bagaimana hubungan kausal antarlevel dalam model evaluasi pembelajaran *Kirkpatrick*.

Research gap penelitian ini berada pada terbatasnya bukti empiris yang secara kuat menjelaskan hubungan kausal antarlevel dalam model *Kirkpatrick*, terutama pada evaluasi pembelajaran organisasi yang memanfaatkan data *multisource*. PLN telah menerapkan evaluasi berdasarkan empat level *Kirkpatrick*. Namun, hasil evaluasi pada setiap level masih lebih banyak dipahami secara terpisah. Berdasarkan hal tersebut, penelitian ini menguji keterkaitan antara *Reaction*, *Learning*, *Behavior*, dan *Result* dengan menggunakan *PLS-SEM*, sehingga hasil evaluasi pembelajaran dapat dianalisis secara lebih terpadu dan didukung oleh data empiris.

Sejumlah studi terdahulu telah menerapkan *SEM* untuk menganalisis keterkaitan antarlevel dalam model *Kirkpatrick*. Sebagai contoh, Faisal-E-Alam dan Islam (2025)

membuktikan bahwa susunan hierarkis dalam model *Kirkpatrick* dapat diuji secara empiris melalui hubungan antara *Reaction*, *Learning*, *Behavior*, dan *Result*. Meskipun demikian, penelitian *SEM* sebelumnya masih menyisakan keterbatasan karena fokus utamanya lebih banyak diarahkan pada pengujian relasi antarlevel, sedangkan mutu sumber data pada level *Behavior* dan *Result* belum memperoleh perhatian yang memadai. *Behavior* dan *Result* merupakan tingkat evaluasi yang berhubungan dengan penerapan hasil pembelajaran serta pengaruhnya terhadap unit kerja. Oleh karena itu, penilaiannya tidak dapat hanya bertumpu pada persepsi peserta. Penelitian ini berbeda dari studi terdahulu karena menerapkan *multisource assessment* dengan melibatkan peserta dan atasan langsung dalam pengukuran konstruk *Behavior* dan *Result*. Pendekatan tersebut memperkuat proses evaluasi karena perubahan perilaku dan dampak pembelajaran dapat diamati melalui dua sudut pandang. Dengan demikian, unsur kebaruan penelitian ini terletak pada pengujian *chain of evidence Kirkpatrick* menggunakan *PLS-SEM* melalui *higher-order construct*, *two-stage approach*, dan *multisource assessment* dalam konteks pembelajaran korporat di PLN.

Kontribusi pokok penelitian ini terletak pada penyusunan model evaluasi pembelajaran yang lebih terpadu dan berbasis data untuk menilai apakah hubungan antarlevel *Kirkpatrick* benar-benar terbentuk dalam lingkungan pembelajaran korporat.

Penelitian ini bertujuan menganalisis hubungan kausal antarlevel pada penerapan evaluasi *Kirkpatrick* di PLN. Fokus penelitian dibatasi pada tiga jalur utama dalam *chain of evidence Kirkpatrick*, yaitu *Reaction* terhadap *Learning*, *Learning* terhadap *Behavior*, dan *Behavior* terhadap *Result*. Pembatasan ini dilakukan agar penelitian lebih fokus pada kontribusi empiris utama, yaitu pengujian jalur kausal antarlevel evaluasi.

Pendekatan *Partial Least Squares Structural Equation Modeling (PLS-SEM)* digunakan karena dapat menguji hubungan antarvariabel laten, mengevaluasi model pengukuran, dan menguji model struktural secara simultan (Hair dkk., 2022). *PLS-SEM* juga sesuai untuk model yang memiliki *higher-order construct* dan jumlah sampel yang relatif terbatas. Dalam penelitian ini, pendekatan *two-stage* digunakan untuk mengakomodasi konstruk bertingkat, sedangkan pendekatan *multisource* digunakan pada *Behavior* dan *Result* agar penilaian tidak hanya bergantung pada persepsi peserta.

TINJAUAN PUSTAKA DAN PENGEMBANGAN HIPOTESIS

Model Kirkpatrick

Model *Kirkpatrick* memaparkan evaluasi pelatihan ke dalam empat tingkatan. *Reaction* digunakan untuk menilai tanggapan peserta terhadap proses pembelajaran. *Learning* berfokus pada pengukuran capaian pengetahuan, keterampilan, sikap, kepercayaan diri, dan komitmen. *Behavior* menilai sejauh mana hasil pembelajaran diterapkan oleh peserta setelah kembali ke lingkungan kerja. *Result* mengukur dampak pembelajaran terhadap hasil yang menjadi sasaran organisasi (J. D. Kirkpatrick dan Kirkpatrick, 2016). Model ini banyak diadopsi karena memiliki struktur yang sederhana dan relatif mudah digunakan. Namun, keterkaitan antarlevel dalam model tersebut tidak dapat dimaknai sebagai proses yang berlangsung secara otomatis. (Bates, 2004) menyatakan bahwa model *Kirkpatrick* kerap memperoleh kritik karena hubungan antarlevel sering diasumsikan berjalan secara linear.

New World Kirkpatrick Model (NWKM) mengembangkan model *Kirkpatrick* dengan menekankan pada *critical behaviors* dan *required drivers*. *Critical behaviors* mengacu pada perilaku kerja inti yang diharapkan muncul setelah peserta mengikuti pembelajaran. Sedangkan *required drivers* meliputi dukungan, penguatan, pemantauan, *coaching*, *reinforcement*, dan *accountability* yang dibutuhkan untuk mendorong terwujudnya perilaku tersebut (J. D. Kirkpatrick & Kirkpatrick, 2016). Oleh sebab itu, *Learning* tidak secara otomatis berlanjut menjadi *Behavior*. Dalam sudut pandang transfer *knowledge*, capaian belajar perlu direpresentasikan dalam konteks pekerjaan melalui dukungan setelah pembelajaran, peluang

untuk mempraktikkan hasil belajar, serta lingkungan kerja yang kondusif. Pemahaman ini selaras dengan (Dachner dkk., 2021) yang menegaskan bahwa pengembangan pegawai perlu diarahkan pada penerapan hasil pembelajaran di tempat kerja.

Keterkaitan antara *Behavior* dan *Result* dapat dimaknai melalui pandangan bahwa pelatihan akan memberi nilai bagi organisasi apabila kompetensi yang didapat peserta benar-benar digunakan dalam pelaksanaan pekerjaan. (Paul dkk., 2024) menyampaikan bahwa evaluasi pelatihan perlu memperhatikan hubungan antarlevel agar organisasi dapat mengetahui kaitan antara capaian belajar, perilaku kerja, dan dampak pelatihan. Dalam konteks *corporate training*, (Capatina dkk., 2024) memperlihatkan bahwa pelatihan mampu mendukung *knowledge retention*, *knowledge sharing*, dan *job performance*. Selain itu, Martins (2021), Nguyen dan Dao (2023), serta Qalati dkk. (2022) menunjukkan bahwa pelatihan, kapasitas sumber daya manusia, dan perilaku kerja pegawai memiliki hubungan dengan kinerja organisasi. Oleh karena itu, *Behavior* berperan sebagai penghubung penting antara *Learning* dan *Result* karena manfaat pembelajaran baru dapat terlihat ketika hasil belajar diwujudkan dalam perilaku kerja.

Konteks Evaluasi Pembelajaran di PLN

PLN Pusdiklat menjalankan evaluasi pembelajaran sebagai bagian dari siklus pengembangan kompetensi (Badrul S dkk, 2020). Siklus ini dimulai dari analisis kebutuhan pembelajaran, penyusunan desain pembelajaran, pengembangan materi, pelaksanaan pembelajaran, dan evaluasi. Dalam praktiknya, evaluasi Level 1 dan Level 2 dilakukan pada saat atau setelah proses pembelajaran, sedangkan Level 3 dan Level 4 dilakukan setelah peserta kembali ke unit kerja. Pola ini sesuai dengan prinsip *Kirkpatrick* yang membedakan hasil jangka pendek dan dampak setelah pembelajaran (J. D. Kirkpatrick & Kirkpatrick, 2016).

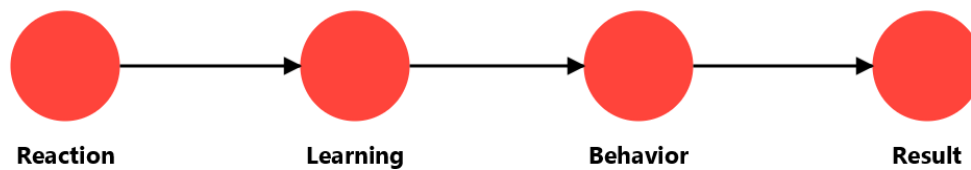
Meskipun struktur evaluasi telah mengacu pada *Kirkpatrick*, cara pembacaan hasil evaluasi masih perlu dikembangkan. Nilai tiap level dapat memberikan informasi penting, tetapi belum cukup untuk menjelaskan hubungan sebab-akibat. Misalnya, nilai kepuasan yang tinggi belum tentu menjadi bukti bahwa peserta memperoleh hasil belajar yang baik. Demikian pula, skor *post-test* yang baik belum tentu menjamin peserta menerapkan hasil belajar dalam pekerjaan. Oleh sebab itu organisasi perlu untuk melakukan evaluasi setelah pelatihan dan tidak cukup hanya dari reaksi dan hasil pembelajaran saja (Nielsen & Shepherd, 2022).

Kondisi tersebut penting karena PLN Pusdiklat membutuhkan evaluasi yang dapat menjelaskan kontribusi pembelajaran terhadap unit kerja. Pembelajaran berbasis kinerja yang dilaksanakan oleh PLN tidak hanya diarahkan untuk meningkatkan pengetahuan peserta. Pembelajaran juga diharapkan mampu memperbaiki cara kerja, meningkatkan penerapan prosedur, memperkuat pemecahan masalah, dan mendukung pencapaian sasaran unit. Oleh karena itu, evaluasi perlu menghubungkan hasil belajar di kelas dengan perilaku kerja dan dampak unit.

Pendekatan *multisource* menjadi relevan dalam konteks ini karena keterlibatan atasan sangat mempengaruhi pelaksanaan hasil pelatihan tersebut (Nielsen & Shepherd, 2022). Evaluasi *Behavior* dan *Result* tidak sebaiknya hanya dinilai oleh peserta, karena peserta dapat memiliki bias penilaian diri. Karena itu, penelitian ini menggunakan peserta dan atasan sebagai sumber penilaian untuk memperkuat kualitas informasi pada level *Behavior* dan *Result*.

Dengan konteks tersebut, pengembangan model evaluasi dalam penelitian ini diarahkan untuk memberikan bukti empiris atas hubungan antarlevel. Model tidak dimaksudkan untuk mengganti *Kirkpatrick*, tetapi untuk memperkuat cara penggunaan *Kirkpatrick* melalui analisis statistik. Hasil yang diperoleh diharapkan dapat membantu organisasi menentukan bagian mana dari proses pembelajaran yang perlu diperbaiki, apakah pada pengalaman belajar, hasil belajar, dukungan pascapembelajaran, atau instrumen pengukuran dampak.

Hipotesis Penelitian



Gambar 1. Struktur hirarkis Kirkpatrick (Faisal-E-Alam & Islam, 2025)

Berdasarkan struktur hierarkis *Kirkpatrick*, evaluasi pembelajaran dapat dipahami sebagai proses bertahap yang dimulai dari *Reaction*, kemudian berlanjut ke *Learning*, *Behavior*, dan *Result*. *Reaction* menggambarkan respons peserta terhadap proses pembelajaran, seperti keterikatan, relevansi materi, dan kepuasan terhadap penyelenggaraan pembelajaran. Respons positif peserta diharapkan dapat menciptakan kondisi belajar yang lebih baik, sehingga peserta lebih mudah memahami materi, membangun komitmen, serta memiliki kesiapan untuk menerapkan hasil pembelajaran. Oleh karena itu, hipotesis pertama dalam penelitian ini menyatakan bahwa *Reaction* berpengaruh positif terhadap *Learning*. Hipotesis ini penting untuk menguji apakah pengalaman peserta selama pembelajaran benar-benar berhubungan dengan hasil belajar yang diperoleh.

Selanjutnya, model *Kirkpatrick* menjelaskan bahwa hasil belajar seharusnya mendorong perubahan perilaku peserta ketika kembali ke unit kerja. Dalam konteks ini, *Learning* tidak hanya dipahami sebagai nilai tes, tetapi juga mencakup komitmen dan kepercayaan diri peserta untuk menerapkan hasil pembelajaran. Jika hasil belajar tersebut kuat, maka peserta diharapkan mampu menerapkan pengetahuan dan keterampilan dalam pekerjaan. Oleh karena itu, hipotesis kedua menyatakan bahwa *Learning* berpengaruh positif terhadap *Behavior*. Pada tahap berikutnya, perubahan perilaku kerja peserta diharapkan dapat memberikan dampak terhadap unit kerja, sehingga hipotesis ketiga menyatakan bahwa *Behavior* berpengaruh positif terhadap *Result*. Ketiga hipotesis ini digunakan untuk menguji jalur utama dalam *chain of evidence* *Kirkpatrick*, yaitu apakah proses pembelajaran dapat bergerak dari pengalaman belajar, hasil belajar, penerapan di tempat kerja, hingga dampak pada unit kerja.

Tabel 1. Hipotesis penelitian

Hipotesis	Penjelasan
H1	Reaction berpengaruh positif terhadap Learning
H2	Learning berpengaruh positif terhadap Behavior
H3	Behavior berpengaruh positif terhadap Result

METODE PENELITIAN

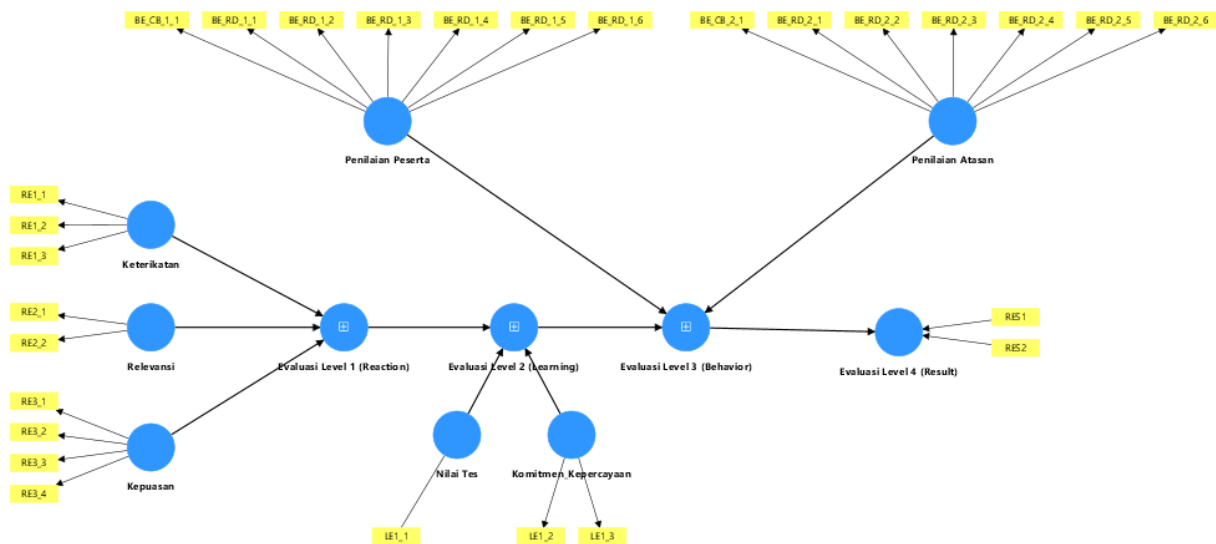
Penelitian ini menggunakan desain kuantitatif dengan menggunakan pendekatan *PLS-SEM*. Unit analisis adalah pegawai PT PLN (Persero) yang mengikuti Pembelajaran Berbasis Kinerja pada tahun 2025. Data berasal dari hasil evaluasi pembelajaran yang dilakukan oleh PLN Pusdiklat. Data untuk evaluasi Level 1 (*Reaction*) dan evaluasi Level 2 (*Learning*) menggunakan isian kuesioner yang dilakukan oleh Peserta. Sedangkan untuk evaluasi Level 3 (*Behavior*) dan Level 4 (*Result*) menggunakan pendekatan *multisource* dengan melibatkan peserta dan atasan langsung.

Dari total 775 data peserta yang tersedia, terdapat 151 data yang tergolong lengkap dan memenuhi syarat untuk dianalisis karena memuat informasi pada seluruh level evaluasi, yaitu *Reaction*, *Learning*, *Behavior*, dan *Result*. Data yang tidak mencakup salah satu unsur evaluasi tersebut tidak disertakan dalam analisis karena tidak mampu menggambarkan hubungan antarlevel dalam model *Kirkpatrick* secara menyeluruh. Penentuan jumlah sampel minimum dilakukan dengan bantuan *G*Power* menggunakan parameter *effect size* (f^2) sebesar 0,15, tingkat signifikansi (α) sebesar 0,02, *statistical power* sebesar 0,95, serta jumlah prediktor

sebanyak 28. Berdasarkan hasil perhitungan tersebut, jumlah sampel minimum yang dibutuhkan adalah 95 responden. Atas dasar itu, penelitian ini menetapkan 100 data sebagai sampel akhir untuk pengujian model PLS-SEM. Sampel akhir diperoleh melalui pemilihan acak dari 151 data lengkap sehingga setiap data yang memenuhi kriteria memiliki kesempatan yang sama untuk masuk sebagai sampel penelitian. Pemilihan sampel tidak ditentukan oleh nilai *Reaction*, *Learning*, *Behavior*, *Result*, unit kerja, jenis pembelajaran, ataupun karakteristik tertentu dari hasil evaluasi. Proses pengambilan sampel secara acak dilakukan untuk memenuhi batas minimum kebutuhan analisis serta menekan kemungkinan terjadinya selection bias dalam pemilihan data. Dengan demikian, 151 data menunjukkan jumlah data lengkap yang tersedia, sedangkan 100 data merupakan sampel akhir yang digunakan dalam pengujian model PLS-SEM. Penggunaan 100 data dari 151 data lengkap tetap dinyatakan sebagai keterbatasan penelitian karena penggunaan seluruh data lengkap berpotensi menghasilkan estimasi model yang lebih stabil dan memperkuat generalisasi temuan.

Model penelitian terdiri atas empat konstruk utama, yaitu *Reaction*, *Learning*, *Behavior* dan *Result*. *Reaction* dibentuk oleh keterikatan, relevansi, dan kepuasan. *Learning* direpresentasikan melalui nilai tes dan komitmen-kepercayaan. *Behavior* diukur melalui penilaian peserta dan penilaian atasan yang mencerminkan penerapan hasil belajar dan *required drivers*. *Result* mengukur persepsi dampak positif pembelajaran terhadap unit kerja.

Model penelitian ini menggunakan *higher-order construct* untuk menggambarkan empat level evaluasi Kirkpatrick. Dimensi yang membentuk setiap level evaluasi dijelaskan melalui *lower-order construct*, sedangkan *higher-order construct* digunakan untuk menganalisis keterkaitan antarlevel dalam model Kirkpatrick. Arah pengukuran ditetapkan berdasarkan landasan konseptual masing-masing konstruk. *Reaction* diposisikan sebagai konstruk multidimensi yang direfleksikan oleh keterikatan, relevansi, dan kepuasan. Sementara itu, *Behavior* dijelaskan melalui dua sumber penilaian, yaitu peserta dan atasan, agar penerapan hasil belajar dapat ditangkap secara lebih menyeluruh.



Gambar 2. Desain SEM PLS - NWKM

Analisis dilakukan dalam dua tahap. Tahap pertama mengevaluasi outer model untuk memperoleh *latent variable score* dari *lower-order construct*. Tahap kedua menggunakan skor yang telah didapat pada tahap pertama sebagai indikator untuk menguji hubungan antarlevel Kirkpatrick. Evaluasi mencakup *outer loading*, reliabilitas, validitas konvergen, validitas diskriminan, kolinearitas, koefisien jalur, *R-square* (R^2), *effect size* (f^2), model fit, dan uji signifikansi berbasis *bootstrapping*. Konstruk *Result* diperlakukan sebagai *composite score* karena indikator RES_1 dan RES_2 menunjukkan variansi sangat rendah dan menimbulkan

masalah *near-singular covariance* pada saat *bootstrapping*. *Composite score* dibentuk dengan menghitung nilai rata-rata dari RES_1 dan RES_2. Hasil rata-rata tersebut selanjutnya digunakan sebagai indikator tunggal untuk konstruk *Result* dalam pengujian model struktural. Keputusan ini ditempuh sebagai langkah teknis untuk menjaga kestabilan estimasi model, bukan sebagai upaya untuk meningkatkan signifikansi hubungan struktural. Dengan pendekatan tersebut, makna substantif konstruk *Result* tetap dipertahankan (Warmbrod, 2014). Validasi model dilakukan dengan melihat hasil pengukuran dan hasil struktural secara berurutan. *Outer model* dievaluasi lebih dahulu untuk memastikan bahwa indikator masih dapat menjelaskan konstruk. Setelah itu, *inner model* dievaluasi untuk membaca kekuatan hubungan antarlevel. Urutan ini penting agar interpretasi hubungan struktural tidak dilepaskan dari kualitas pengukuran konstruk.

Tabel 2. Konstruk dan dimensi pembentuk model penelitian

Konstruk	Dimensi/indikator pembentuk	Makna pengukuran
<i>Reaction</i>	Keterikatan, relevansi, kepuasan	Respons peserta terhadap proses pembelajaran
<i>Learning</i>	Nilai tes, komitmen-kepercayaan	Hasil belajar dan kesiapan menerapkan pembelajaran
<i>Behavior</i>	Penilaian peserta dan penilaian atasan	Penerapan hasil belajar di unit kerja
<i>Result</i>	Composite score dampak pembelajaran	Dampak pembelajaran pada unit kerja

Tabel 3. Indikator penelitian yang digunakan

Konstruk	Kode	Deskripsi indikator
<i>Reaction</i>	RE1_1	Informasi kegiatan pembelajaran sebelum pembelajaran dimulai
<i>Reaction</i>	RE1_2	Keterlibatan peserta selama kegiatan pembelajaran
<i>Reaction</i>	RE1_3	Suasana kelas membantu proses belajar
<i>Reaction</i>	RE2_1	Materi sesuai dengan bidang pekerjaan peserta
<i>Reaction</i>	RE2_2	Materi dapat diterapkan di unit kerja
<i>Reaction</i>	RE3_1	Kepuasan terhadap cara mengajar instruktur
<i>Reaction</i>	RE3_2	Kepuasan terhadap pelayanan pembelajaran
<i>Reaction</i>	RE3_3	Ketepatan waktu pelaksanaan
<i>Reaction</i>	RE3_4	Kepuasan terhadap akomodasi dan infrastruktur
<i>Learning</i>	LE1_1	Skor post-test sebagai ukuran pengetahuan, keterampilan, dan sikap
<i>Learning</i>	LE1_2	Komitmen peserta untuk menerapkan hasil belajar
<i>Learning</i>	LE1_3	Kepercayaan diri peserta untuk menerapkan hasil belajar
<i>Behavior</i>	BE_CB_1_1	Penilaian peserta atas penerapan keterampilan di unit kerja
<i>Behavior</i>	BE_RD_1_1	Bimbingan atasan menurut peserta
<i>Behavior</i>	BE_RD_1_2	Dukungan atau penyemangat menurut peserta
<i>Behavior</i>	BE_RD_1_3	Sistem yang akuntabel dan responsif menurut peserta
<i>Behavior</i>	BE_RD_1_4	Keyakinan bahwa ilmu membantu pekerjaan menurut peserta
<i>Behavior</i>	BE_RD_1_5	Pembelajaran berkelanjutan setelah pembelajaran awal menurut peserta
<i>Behavior</i>	BE_RD_1_6	Penghargaan karena menerapkan pengetahuan menurut peserta
<i>Behavior</i>	BE_CB_2_1	Penilaian atasan atas penerapan keterampilan peserta
<i>Behavior</i>	BE_RD_2_1	Bimbingan atasan menurut atasan
<i>Behavior</i>	BE_RD_2_2	Dukungan atau penyemangat menurut atasan
<i>Behavior</i>	BE_RD_2_3	Sistem yang akuntabel dan responsif menurut atasan
<i>Behavior</i>	BE_RD_2_4	Keyakinan bahwa ilmu membantu pekerjaan menurut atasan
<i>Behavior</i>	BE_RD_2_5	Pembelajaran berkelanjutan setelah pembelajaran awal menurut atasan
<i>Behavior</i>	BE_RD_2_6	Penghargaan karena menerapkan pengetahuan menurut atasan
<i>Result</i>	RES_1	Persepsi peserta atas dampak positif pembelajaran pada unit
<i>Result</i>	RES_2	Persepsi atasan atas dampak positif pembelajaran pada unit

HASIL DAN PEMBAHASAN

Evaluasi Measurement Model

Hasil pengujian outer loading menunjukkan bahwa sebagian besar indikator reflektif memperoleh nilai lebih dari 0,70. Temuan ini mengindikasikan bahwa indikator-indikator tersebut telah mampu merepresentasikan konstruknya secara memadai. Terdapat dua indikator

yang memiliki nilai di bawah 0,70, yaitu RE3_4 sebesar 0,445 dan BE_CB_1_1 sebesar 0,645. Indikator BE_CB_1_1 tetap dipertahankan karena konstruk yang diukurnya masih menunjukkan reliabilitas dan validitas konvergen yang baik. Sebaliknya, RE3_4 dihapus karena memiliki nilai loading yang rendah, meskipun konstruk kepuasan tetap memenuhi batas minimum reliabilitas dan AVE. Selain itu, penghapusan RE3_4 juga dipertimbangkan karena konstruk tersebut masih memiliki indikator reflektif lain yang dapat merepresentasikan konstruknya, sedangkan BE_CB_1_1 dipertahankan karena berperan sebagai unsur critical behavior yang tidak dapat digantikan.

Hasil reliabilitas dan validitas konvergen menunjukkan bahwa *Cronbach alpha*, *composite reliability*, dan *AVE* berada di atas ambang minimum yang direkomendasikan. Dengan demikian, konstruk dalam *outer model* memiliki konsistensi internal dan validitas konvergen yang memadai. Nilai *HTMT* antara Nilai Tes dan Komitmen-Kepercayaan sebesar 0,038 serta antara Penilaian Peserta dan Penilaian Atasan sebesar 0,453 menunjukkan validitas diskriminan yang baik. Namun, nilai *HTMT* antara keterikatan, relevansi, dan kepuasan berada di atas 0,90. Kondisi ini menunjukkan adanya potensi kemiripan konseptual antar dimensi *Reaction*. Nilai *HTMT* yang tinggi menunjukkan bahwa peserta cenderung menilai keterikatan, relevansi, dan kepuasan sebagai pengalaman belajar yang saling berkaitan. Oleh karena itu, dimensi *Reaction* perlu ditafsirkan sebagai aspek yang berdekatan secara konseptual, bukan sebagai konstruk yang sepenuhnya terpisah. Ketiga dimensi tetap dipertahankan karena secara konseptual merepresentasikan *Reaction* dalam *NWKM*. Namun, penelitian berikutnya perlu meninjau ulang redaksi indikator *Reaction* agar setiap dimensi memiliki batas konseptual yang lebih jelas

Tabel 4. Hasil uji outer loading

Hubungan Indikator dan Laten	Outer loadings	Intepretasi
RE1_1 <- Keterikatan	0,84	Valid
RE1_2 <- Keterikatan	0,921	Valid
RE1_3 <- Keterikatan	0,838	Valid
RE2_1 <- Relevansi	0,974	Valid
RE2_2 <- Relevansi	0,974	Valid
RE3_1 <- Kepuasan	0,889	Valid
RE3_2 <- Kepuasan	0,875	Valid
RE3_3 <- Kepuasan	0,911	Valid
RE3_4 <- Kepuasan	0,445	Jauh dari batas valid, sangat kuat dipertimbangkan untuk dieliminasi
LE1_1 <- Nilai Tes	1	Indikator tunggal
LE1_2 <- Komitmen_Kepercayaan	0,966	Valid
LE1_3 <- Komitmen_Kepercayaan	0,966	Valid
BE_CB_1_1 <- Penilaian Peserta	0,645	Dekat dekat batas valid, dapat dipertimbangkan untuk di pertahankan jika reliabilitas dan AVE memenuhi kriteria
BE_CB_2_1 <- Penilaian Atasan	0,726	Valid
BE_RD_1_1 <- Penilaian Peserta	0,962	Valid
BE_RD_1_2 <- Penilaian Peserta	0,938	Valid
BE_RD_1_3 <- Penilaian Peserta	0,883	Valid
BE_RD_1_4 <- Penilaian Peserta	0,928	Valid
BE_RD_1_5 <- Penilaian Peserta	0,923	Valid
BE_RD_1_6 <- Penilaian Peserta	0,826	Valid

Hubungan Indikator dan Laten	Outer loadings	Intepretasi
BE_RD_2_1 <- Penilaian Atasan	0,839	Valid
BE_RD_2_2 <- Penilaian Atasan	0,921	Valid
BE_RD_2_3 <- Penilaian Atasan	0,93	Valid
BE_RD_2_4 <- Penilaian Atasan	0,895	Valid
BE_RD_2_5 <- Penilaian Atasan	0,862	Valid
BE_RD_2_6 <- Penilaian Atasan	0,869	Valid

Tabel 5. Hasil reliabilitas dan validitas konvergen

Konstruk	Cronbach alpha	Composite reliability	AVE
Kepuasan	0,801	0,873	0,646
Keterikatan	0,834	0,901	0,752
Relevansi	0,946	0,974	0,949
Komitmen-Kepercayaan	0,927	0,965	0,932
Penilaian Atasan	0,943	0,954	0,749
Penilaian Peserta	0,948	0,959	0,771

Tabel 6. Hasil uji validitas diskriminan

Hubungan konstruk	Nilai Heterotrait-monotrait ratio (HTMT)
Relevansi - Keterikatan	0.964
Keterikatan - Kepuasan	0.953
Relevansi - Kepuasan	0.919
Nilai Tes - Komitmen-Kepercayaan	0.038
Penilaian Peserta - Penilaian Atasan	0.453

Evaluasi Structural Model

Hasil model struktural menunjukkan bahwa *Reaction* memiliki hubungan positif yang kuat terhadap *Learning* dengan koefisien jalur sebesar 0,667. Sedangkan *Learning* memiliki hubungan positif yang sangat lemah terhadap *Behavior* dengan koefisien jalur sebesar 0,051. Hal berbeda untuk *Behavior* yang memiliki hubungan positif terhadap *Result* dengan koefisien jalur sebesar 0,305. Adapun Nilai SRMR sebesar 0,047 menunjukkan *model fit* yang memadai karena berada di bawah batas 0,08.

Nilai R^2 sebagai *explanatory power* pada *Learning* sebesar 0,446 menunjukkan kemampuan penjelasan yang moderat. Artinya, *Reaction* mampu menjelaskan 44,6% variasi *Learning*. Sedangkan Nilai R^2 *Behavior* sebesar 0,003 menunjukkan bahwa *Learning* hampir tidak mampu menjelaskan variasi *Behavior*. Nilai R^2 *Result* sebesar 0,093 menunjukkan kemampuan penjelasan yang lemah. Hasil ini memperlihatkan bahwa hubungan antarlevel Kirkpatrick tidak selalu berjalan otomatis secara berurutan.

Hasil yang didapatkan pada *Effect size* memperkuat pola tersebut. Pengaruh *Reaction* terhadap *Learning* memiliki f^2 sebesar 0,804 yang termasuk besar. Akan tetapi Pengaruh *Learning* terhadap *Behavior* memiliki f^2 sebesar 0,003 yang sangat kecil. Dan Pengaruh *Behavior* terhadap *Result* memiliki f^2 sebesar 0,103 yang dapat dipahami sebagai kecil menuju sedang. Hasil ini menunjukkan bahwa pengalaman belajar memiliki kontribusi besar terhadap hasil belajar, sedangkan perubahan perilaku memerlukan faktor lain di luar faktor hasil belajar.

Tabel 7. Hasil uji struktural model

Pengujian	Konstruk/jalur	Nilai	Interpretasi
R^2	<i>Learning</i>	0,446	Moderat
	<i>Behavior</i>	0,003	Sangat lemah
	<i>Result</i>	0,093	Lemah
f^2	<i>Reaction</i> -> <i>Learning</i>	0,804	Besar
	<i>Learning</i> -> <i>Behavior</i>	0,003	Sangat kecil

	<i>Behavior -> Result</i>	0,103	Kecil menuju sedang
<i>Model fit</i>	<i>SRMR</i>	0,047	Model fit memadai

Pengujian Hipotesis

Pengujian signifikansi dilakukan melalui prosedur *bootstrapping* dengan 5.000 subsamples, *two-tailed test*, dan tingkat signifikansi 5%. Signifikansi hubungan dinilai berdasarkan *p-value* dan 95% *confidence interval*. Hubungan dinyatakan signifikan apabila *p-value* < 0,05 dan *confidence interval* tidak melewati nilai nol.

Tabel 8. Hasil uji hipotesis

Hipotesis	Koefisien	T-statistic	p-value	25% CI	95% CI	Keputusan
<i>Reaction -> Learning</i>	0,667	9,622	0,000	0,547	0,780	Diterima
<i>Learning -> Behavior</i>	0,051	0,422	0,673	-0,180	0,260	Tidak diterima
<i>Behavior -> Result</i>	0,305	2,272	0,023	0,040	0,560	Diterima

Pembahasan

Hasil H1 menunjukkan bahwa *Reaction* berpengaruh positif dan signifikan terhadap *Learning*. Temuan ini berarti bahwa pengalaman peserta selama pembelajaran berhubungan dengan hasil belajar yang diperoleh. Hasil ini didukung oleh Faisal-E-Alam dan Islam (2025) yang membuktikan bahwa *Reaction* memiliki pengaruh signifikan terhadap *Learning* dalam pengujian antarlevel *Kirkpatrick*. Keterikatan, relevansi, dan kepuasan bukan sekadar indikator administratif, melainkan bagian dari kondisi belajar yang membantu peserta memahami materi, membangun komitmen, dan meningkatkan kepercayaan diri. Hal ini sejalan dengan *The New World Kirkpatrick Model* dimana *Reaction* menjadi tahapan awal dalam proses pembelajaran yang efektif (J. D. Kirkpatrick & Kirkpatrick, 2016). Bagi PLN Pusdiklat, temuan ini menegaskan pentingnya desain pembelajaran yang relevan dengan pekerjaan peserta, metode yang mendorong partisipasi, serta kualitas layanan pembelajaran yang stabil.

Hasil H2 menunjukkan bahwa *Learning* tidak berpengaruh signifikan terhadap *Behavior*. Temuan ini penting karena menunjukkan bahwa hasil belajar di kelas belum otomatis berubah menjadi perilaku kerja di unit. Hal ini sejalan dengan kritik yang disampaikan oleh Bates (2004) terhadap model *Kirkpatrick* dimana hubungan antarlevel tidak selalu terjadi otomatis dan linier. Peserta dapat memperoleh skor belajar yang baik, tetapi belum tentu memiliki kesempatan, dukungan, atau penugasan yang memadai untuk menerapkan pembelajaran. Hal ini sejalan dengan NWKM yang menekankan bahwa transfer pembelajaran tidak hanya membutuhkan *critical behaviors* tetapi juga *required drivers* (J. D. Kirkpatrick & Kirkpatrick, 2016). Dalam konteks PLN, perubahan perilaku dapat dipengaruhi oleh arahan atasan, dukungan organisasi, sistem yang responsif, keyakinan peserta dalam menerapkan hasil belajar, pembelajaran yang berkelanjutan, dan penghargaan atas penerapan pengetahuan.

Hasil H3 menunjukkan bahwa *Behavior* berpengaruh positif dan signifikan terhadap *Result*. Temuan ini menunjukkan bahwa perilaku kerja menjadi penghubung penting antara pembelajaran dan dampak pada unit kerja. Meskipun koefisien jalurnya tidak besar, pengaruhnya signifikan. Artinya, ketika peserta menerapkan hasil pembelajaran dalam pekerjaan, pembelajaran berpeluang memberikan kontribusi terhadap hasil unit. Temuan ini sejalan dengan Qalati dkk. (2022) yang menyatakan bahwa perilaku kerja pegawai berpengaruh pada pencapaian kinerja organisasi. Hal ini memperkuat posisi evaluasi *Behavior* sebagai level kunci dalam *Kirkpatrick*.

Dalam menafsirkan pengaruh *Behavior* terhadap *Result*, perlakuan *Result* sebagai *composite score* perlu diperhatikan. Pendekatan ini digunakan untuk menjaga stabilitas estimasi model karena indikator RES_1 dan RES_2 memiliki variansi yang sangat rendah serta menimbulkan masalah *near-singular covariance* pada proses *bootstrapping*. Namun, penggunaan *composite score* membatasi kemampuan model dalam menguji *Result* sebagai

konstruk laten multidimensi. Oleh karena itu, *Result* dalam penelitian ini dipahami sebagai skor gabungan yang merepresentasikan persepsi terhadap dampak pembelajaran pada unit kerja. Dengan demikian, hubungan *Behavior* terhadap *Result* perlu diartikan sebagai keterkaitan antara penerapan perilaku kerja dan persepsi umum mengenai dampak pembelajaran, bukan sebagai bukti langsung atas pencapaian kinerja objektif unit kerja. Rendahnya variansi indikator *Result* juga menunjukkan bahwa instrumen *Result* perlu dikembangkan agar lebih sensitif dalam membedakan dampak pembelajaran antar peserta maupun antar unit kerja.

Secara keseluruhan, hasil penelitian menunjukkan bahwa hubungan antarlevel *Kirkpatrick* tidak selalu berjalan secara otomatis. *Reaction* terbukti mendukung *Learning*, tetapi *Learning* belum cukup untuk menghasilkan perubahan *Behavior* tanpa dukungan pascapembelajaran. Sementara itu, *Behavior* menjadi penghubung penting menuju *Result*, meskipun interpretasi *Result* perlu memperhatikan penggunaan *composite score*. Temuan ini menjadi dasar bagi perumusan implikasi teoretis dan praktis pada bagian berikutnya.

IMPLIKASI PENELITIAN

Implikasi Teoritis

Penelitian ini memberikan implikasi teoretis terhadap pengembangan *New World Kirkpatrick Model* dalam evaluasi pembelajaran organisasi. Hasil penelitian menunjukkan bahwa hubungan antarlevel *Kirkpatrick* tidak selalu berjalan secara otomatis. *Reaction* berpengaruh signifikan terhadap *Learning*, tetapi *Learning* tidak berpengaruh signifikan terhadap *Behavior*. Temuan ini memperkuat pandangan bahwa hasil belajar di kelas belum cukup untuk menghasilkan perubahan perilaku kerja tanpa dukungan *required drivers*, seperti dukungan atasan, *coaching*, *monitoring*, *reinforcement*, dan *accountability*.

Penelitian ini juga menunjukkan bahwa *PLS-SEM* dapat digunakan untuk menguji *chain of evidence* dalam model *Kirkpatrick*. Melalui *SEM*, evaluasi pembelajaran tidak hanya menampilkan skor pada setiap level, tetapi juga menguji keterkaitan empiris antara pengalaman belajar, hasil belajar, perilaku kerja, dan dampak pada organisasi. Selain itu, penggunaan *multisource assessment* pada *Behavior* dan *Result* memperkuat pengukuran evaluasi karena perubahan perilaku dan dampak pembelajaran dinilai dari perspektif peserta dan atasan langsung.

Implikasi Praktikal

Secara praktis, hasil penelitian ini menunjukkan bahwa evaluasi pembelajaran organisasi tidak cukup dilakukan dengan membaca skor *Reaction*, *Learning*, *Behavior*, dan *Result* secara terpisah. Dengan pendekatan *SEM*, PLN Pusdiklat dapat memantau apakah pengalaman belajar mendukung hasil belajar, apakah hasil belajar berubah menjadi perilaku kerja, dan apakah perilaku kerja berdampak pada unit kerja.

Temuan *Reaction* yang berpengaruh terhadap *Learning* menunjukkan bahwa kualitas pengalaman belajar tetap perlu dikelola melalui relevansi materi, keterikatan peserta, dan kepuasan terhadap proses pembelajaran. Namun, *Reaction* perlu diposisikan sebagai indikator awal, bukan sebagai bukti akhir keberhasilan pelatihan. Temuan *Learning* yang tidak berpengaruh signifikan terhadap *Behavior* menunjukkan perlunya penguatan mekanisme pascapembelajaran, seperti dukungan atasan, *coaching*, *monitoring*, *action learning*, umpan balik, dan sistem akuntabilitas.

Pendekatan *multisource assessment* juga membantu pengambilan keputusan karena evaluasi tidak hanya bergantung pada *self-assessment* peserta. Penilaian atasan memberikan informasi tambahan tentang penerapan perilaku kerja di unit. Dengan demikian, organisasi dapat memperoleh gambaran evaluasi yang lebih seimbang. Kerangka *SEM* juga dapat digunakan sebagai alat monitoring untuk mengidentifikasi titik lemah dalam rantai evaluasi pembelajaran dan menentukan intervensi yang lebih tepat.

KETERBATASAN PENELITIAN

Penelitian ini memiliki beberapa keterbatasan. Pertama, Penelitian ini menggunakan 100 data sebagai sampel akhir dari 151 data lengkap yang tersedia. Meskipun jumlah tersebut telah memenuhi kebutuhan minimum sampel berdasarkan G*Power, pemanfaatan seluruh data lengkap pada penelitian selanjutnya dapat memberikan estimasi model yang lebih stabil dan meningkatkan kekuatan generalisasi hasil.

Kedua, konstruk *Result* diperlakukan sebagai *composite score* karena indikator memiliki variansi yang sangat rendah dan menimbulkan masalah *near-singular covariance*. Perlakuan ini membantu stabilitas model, tetapi juga menunjukkan bahwa pengukuran *Result* perlu diperbaiki. Variansi yang rendah dapat terjadi ketika responden memberikan jawaban yang terlalu seragam, sehingga indikator kurang mampu membedakan tingkat dampak pembelajaran antar peserta atau unit.

Ketiga, beberapa dimensi *Reaction* memiliki nilai *HTMT* diatas 0,90. Hal ini menunjukkan adanya potensi tumpang tindih konseptual antara keterikatan, relevansi, dan kepuasan. Walaupun ketiga dimensi tetap dipertahankan berdasarkan dasar konseptual *NWKM*, penelitian berikutnya perlu meninjau kembali redaksi indikator agar tiap dimensi dapat dibedakan secara lebih jelas.

Keempat, data penelitian berasal dari lingkungan PT PLN (Persero), sehingga hasil tidak dapat langsung digeneralisasi ke organisasi lain. Organisasi lain dapat memiliki sistem pembelajaran, budaya kerja, desain evaluasi, dan mekanisme dukungan pascapembelajaran yang berbeda. Oleh karena itu, pengujian ulang pada konteks lain diperlukan untuk memperkuat validitas eksternal model.

KESIMPULAN

Penelitian ini menganalisis hubungan kausal antarlevel dalam model evaluasi pembelajaran *Kirkpatrick* di PT PLN (Persero) dengan menggunakan *PLS-SEM*. Hasil penelitian menunjukkan bahwa *Reaction* berpengaruh positif dan signifikan terhadap *Learning*, *Learning* tidak berpengaruh signifikan terhadap *Behavior*, dan *Behavior* berpengaruh positif dan signifikan terhadap *Result*.

Secara teoretis, penelitian ini memberikan kontribusi terhadap pengembangan *New World Kirkpatrick Model* dengan menunjukkan bahwa hubungan antarlevel tidak selalu berjalan otomatis. Temuan *Learning* yang tidak signifikan terhadap *Behavior* menegaskan bahwa hasil belajar di kelas belum cukup untuk menghasilkan perubahan perilaku kerja tanpa dukungan pascapembelajaran, peran atasan, monitoring, *coaching*, *reinforcement*, *accountability*, dan kesempatan praktik di tempat kerja. Dengan demikian, model *Kirkpatrick* perlu dipahami sebagai *chain of evidence* yang perlu diuji secara empiris, bukan hanya sebagai urutan evaluasi yang diasumsikan linear.

Penelitian ini juga memberikan kontribusi melalui penggunaan *multisource assessment* pada *Behavior* dan *Result*. Pendekatan ini memperkuat evaluasi karena penerapan hasil pembelajaran dan dampaknya tidak hanya dinilai oleh peserta, tetapi juga oleh atasan langsung. Selain itu, penelitian ini menunjukkan bahwa *PLS-SEM* dapat digunakan sebagai kerangka monitoring evaluasi pembelajaran untuk melihat keterkaitan antara pengalaman belajar, hasil belajar, perilaku kerja, dan dampak pada unit kerja.

Penelitian selanjutnya disarankan untuk menggunakan data lengkap yang lebih besar, menguji ulang struktur dimensi *Reaction*, mengembangkan indikator *Result* berbasis *outcome* objektif, dan memasukkan *required drivers* sebagai konstruk eksplisit agar hubungan *Learning* terhadap *Behavior* dapat dijelaskan secara lebih mendalam.

REFERENSI

- Badrul S dkk. (2020). *Pedoman Pelaksanaan Evaluasi Pembelajaran*.
- Bates, R. (2004). A critical analysis of evaluation practice: The Kirkpatrick model and the principle of beneficence. *Evaluation and Program Planning*, 27(3), 341–347. <https://doi.org/10.1016/j.evalprogplan.2004.04.011>
- Bhatia, M., Stewart, A. E., Wallace, A., Kumar, A., & Malhotra, A. (2021). Evaluation of an In-Situ Neonatal Resuscitation Simulation Program Using the New World Kirkpatrick Model. *Clinical Simulation in Nursing*, 50, 27–37. <https://doi.org/10.1016/j.ecns.2020.09.006>
- Dachner, A. M., Ellingson, J. E., Noe, R. A., & Saxton, B. M. (2021). The future of employee development. *Human Resource Management Review*, 31(2). <https://doi.org/10.1016/j.hrmr.2019.100732>
- Faisal-E-Alam, M., & Islam, A. R. M. T. (2025). Exploring the hierarchical framework of the Kirkpatrick model in training evaluation from a developing country. *Evaluation and Program Planning*, 113. <https://doi.org/10.1016/j.evalprogplan.2025.102696>
- George, G., & Schillebeeckx, S. J. D. (2022). Digital transformation, sustainability, and purpose in the multinational enterprise. In *Journal of World Business* (Vol. 57, Number 3). Elsevier Inc. <https://doi.org/10.1016/j.jwb.2022.101326>
- Hair, J. F. ., Hult, G. T. M. ., Ringle, C. M. ., & Sarstedt, Marko. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (third). SAGE Publications, Inc.
- Kirkpatrick, D. L. (1970). *EVALUATION OF TRAINING*.
- Kirkpatrick, J. D., & Kirkpatrick, W. K. (2016). *FOUR LEVELS OF TRAINING EVALUATION*. www.copyright.com,
- Martins, P. S. (2021). Employee training and firm performance: Evidence from ESF grant applications. *Labour Economics*, 72. <https://doi.org/10.1016/j.labeco.2021.102056>
- Nguyen, D. T., & Dao, T. K. (2023). The mediating role of innovation in the relationship between high-performance human resource management practices and firm performance. *Heliyon*, 9(12). <https://doi.org/10.1016/j.heliyon.2023.e22720>
- Nielsen, K., & Shepherd, R. (2022). Understanding the outcomes of training to improve employee mental health: A novel framework for training transfer and effectiveness evaluation. *Work and Stress*, 36(4), 377–391. <https://doi.org/10.1080/02678373.2022.2028318>
- Qalati, S. A., Zafar, Z., Fan, M., Sánchez Limón, M. L., & Khaskheli, M. B. (2022). Employee performance under transformational leadership and organizational citizenship behavior: A mediated model. *Heliyon*, 8(11). <https://doi.org/10.1016/j.heliyon.2022.e11374>
- Sathishwaran R, Helen S, Bonny, B. P., & S, D. (2025). Structural equation modelling analysis of the Kirkpatrick training evaluation model. *Journal of Agricultural Education and Extension*. <https://doi.org/10.1080/1389224X.2025.2573498>
- Warmbrod, J. R. (2014). Reporting and Interpreting Scores Derived from Likert-type Scales. *Journal of Agricultural Education*, 55(5), 30–47. <https://doi.org/10.5032/jae.2014.05030>