



Scoping Review: Situational Judgment Test untuk Pengukuran Kompetensi dan Implikasinya bagi Kompetensi Sosial Guru PAUD

Raisa Rakhmania¹, Florence Yulisinta², Restiana Ertika Latifah³, dan Robiah⁴

^{1,2,3,4} Pendidikan Guru Pendidikan Anak Usia Dini, Sekolah Tinggi Pendidikan Holistik Berbasis Karakter

ABSTRAK. Kompetensi sosial guru Pendidikan Anak Usia Dini (PAUD) menuntut kemampuan memahami situasi interpersonal dan menentukan respons profesional yang tepat, sementara pengukurannya memerlukan pendekatan yang mampu merepresentasikan kompleksitas situasi kerja. Penelitian ini bertujuan memetakan penggunaan dan pengembangan Situational Judgment Test (SJT) pada guru dan pendidik serta menyintesis implikasinya bagi pengembangan SJT kompetensi sosial guru PAUD. Penelitian menggunakan scoping review dengan kerangka Population, Concept, and Context (PCC) dan pelaporan berdasarkan PRISMA-ScR. Pencarian dilakukan melalui Scopus, Crossref, Google Scholar, GARUDA, dan ERIC. Sebanyak 14 artikel yang memenuhi kriteria dianalisis secara deskriptif dan tematik. Sintesis menghasilkan lima tema utama meliputi karakteristik penelitian SJT, konstruk kompetensi terapan, autentisitas skenario dan konteks budaya, format respons dan skoring, serta validasi psikometrik dan fairness. Temuan utama menunjukkan bahwa kemampuan mendiagnosis situasi dan menentukan respons perlu diperlakukan sebagai kemampuan yang berkaitan tetapi berbeda. Kajian ini merumuskan arsitektur diagnose, respond yang mengintegrasikan konstruk spesifik, skenario autentik, opsi respons plausible, skoring parsial, dan validasi multi-bukti sebagai landasan konseptual dan metodologis pengembangan SJT kompetensi sosial guru PAUD.

Kata Kunci : Asesmen; Guru; Kompetensi Sosial; Situational Judgment Test

ABSTRACT. Early childhood teachers' social competence requires the ability to understand interpersonal situations and select appropriate professional responses, while its assessment requires an approach that can represent the complexity of real work situations. This study aimed to map the use and development of Situational Judgment Tests (SJTs) for teachers and educators and to synthesize their implications for developing an SJT of early childhood teachers' social competence. A scoping review was conducted using the Population, Concept, and Context (PCC) framework and reported according to PRISMA-ScR. Searches were conducted through Scopus, Crossref, Google Scholar, GARUDA, and ERIC. Fourteen eligible studies were analyzed descriptively and thematically. The synthesis generated five major themes: characteristics of SJT research, applied competence constructs, authentic scenarios and cultural contexts, response formats and scoring, and psychometric validation and fairness. A key finding indicates that situation diagnosis and response selection should be treated as related but distinct abilities. Accordingly, this review proposes a diagnose, respond architecture integrating specific constructs, authentic scenarios, plausible response options, partial scoring, and multiple sources of validity evidence. This architecture provides a conceptual and methodological foundation for developing a contextual and defensible SJT of early childhood teachers' social competence.

Keyword : Assessment; Teachers; Situational Judgment Test; Social Competence

Copyright (c) 2026 Raisa Rakhmania dkk.

✉ Corresponding author : Raisa Rakhmania

Email Address : rakhmaniaraisa@gmail.com

Received 18 Agustus 2026, Accepted 19 September 2026, Published 19 September 2026

PENDAHULUAN

Guru PAUD menghadapi beragam interaksi sosial yang menuntut kemampuan berkomunikasi, memahami orang lain, membangun hubungan positif, dan mengambil keputusan interpersonal secara tepat [1]. Kompetensi ini penting karena kualitas hubungan guru-anak berkaitan dengan perkembangan sosial anak, sedangkan komunikasi dan kemitraan dengan keluarga turut mendukung kesinambungan pengalaman pendidikan anak [2]. Urgensi ini didukung oleh kajian yang menunjukkan bahwa kompetensi sosial-emosional pendidik berkaitan dengan kesejahteraan guru, hubungan guru-anak, pengelolaan kelas, pembelajaran sosial-emosional, dan keterampilan sosial anak [3]. Keragaman kerangka SEC juga menegaskan pentingnya definisi konstruk yang jelas dalam pengukuran kompetensi sosial guru [4].

Kompetensi sosial guru PAUD bersifat situasional sehingga tidak cukup diukur melalui penilaian diri, tetapi perlu mempertimbangkan pihak yang terlibat dan konteks interaksi [5]. Interaksi dengan keluarga juga membutuhkan komunikasi responsif, empati, kemampuan mendengarkan, dan pembangunan kepercayaan [6]. Kondisi tersebut menunjukkan perlunya asesmen yang menempatkan kompetensi dalam konteks situasi profesional, bukan semata-mata sebagai atribut individual yang dilepaskan dari kondisi penggunaannya [7].

Salah satu pendekatan yang berpotensi memenuhi kebutuhan tersebut adalah *Situational Judgment Test* (SJT), yaitu asesmen berbasis situasi yang meminta peserta memberikan judgment terhadap alternatif tindakan dalam konteks tertentu. SJT telah berkembang pada berbagai bidang dan digunakan untuk mengukur beragam bentuk kompetensi serta *professional judgment*. Namun, perkembangan SJT dan kajian kompetensi sosial-emosional guru belum sepenuhnya terintegrasi dalam konteks kompetensi sosial guru PAUD. Literatur SJT masih tersebar pada berbagai konstruk, tujuan, dan konteks profesi [8], sementara kajian yang secara khusus memetakan prinsip pengembangan SJT untuk diterjemahkan ke dalam asesmen kompetensi sosial guru PAUD masih terbatas.

Penelitian terdahulu menunjukkan bahwa *Situational Judgment Test* (SJT) telah digunakan dalam berbagai konteks pendidikan. Aldrup et al. mengembangkan SJT untuk mengukur kompetensi sosial-emosional guru [9], sedangkan Li et al. menggunakannya untuk membedakan kemampuan guru PAUD dalam menganalisis situasi pedagogis dan menentukan respons [10], [11]. Kajian sebelumnya juga telah memetakan perkembangan SJT secara umum [7] dan penggunaannya untuk mengukur atribut kepribadian dalam seleksi guru [12]. Namun, penelitian tersebut lebih banyak berfokus pada seleksi guru, kompetensi sosial-emosional secara umum, atau kemampuan pedagogis tertentu.

Gap penelitian ini terletak pada belum tersedianya kajian yang mengintegrasikan bukti mengenai konstruk, pengembangan skenario, format respons, sistem skoring, dan validasi psikometrik SJT ke dalam kerangka yang sesuai dengan karakteristik kompetensi sosial guru PAUD, khususnya dalam interaksi dengan anak, keluarga, rekan sejawat, pimpinan satuan pendidikan, dan masyarakat. Untuk mengisi gap tersebut, penelitian ini memetakan penggunaan dan pengembangan SJT pada guru dan pendidik serta

menyintesis implikasinya bagi pengembangan SJT kompetensi sosial guru PAUD. Kebaruan kajian ini terletak pada perumusan arsitektur *diagnose, respond* yang membedakan kemampuan memahami situasi sosial dan menentukan respons profesional sebagai dua kemampuan yang saling berkaitan. Pendekatan *scoping review* dipilih untuk memetakan karakteristik, keragaman, dan kesenjangan bukti pada bidang yang luas dan heterogeny [13], [14].

METODE

Penelitian ini menggunakan *scoping review* untuk memetakan penggunaan dan pengembangan *Situational Judgment Test* (SJT) pada guru dan pendidik serta menyintesis implikasinya bagi pengembangan SJT kompetensi sosial guru PAUD. Kajian mengacu pada tahapan *scoping review* dan dilaporkan berdasarkan *Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews* (PRISMA-ScR)[13], [15]. Pertanyaan kajian mencakup tiga fokus: karakteristik penggunaan dan pengembangan SJT; prinsip pengembangan terkait konstruk, skenario, format respons, skoring, dan validasi; serta implikasinya bagi pengembangan SJT kompetensi sosial guru PAUD.

Kriteria sumber ditetapkan menggunakan kerangka *Population, Concept, and Context* (PCC) [16]. *Population* mencakup guru, calon guru, mahasiswa pendidikan keguruan, dan teacher educator; *Concept* mencakup penggunaan atau pengembangan SJT untuk asesmen, seleksi, maupun evaluasi kompetensi; sedangkan *Context* meliputi sekolah, pendidikan guru, seleksi calon guru, pengembangan profesional, dan PAUD. Kriteria inklusi dan eksklusi disajikan pada Tabel 1.

Tabel 1. Kriteria Inklusi dan Eksklusi Sumber Artikel Ilmiah

Aspek	Kriteria Inklusi	Kriteria Eksklusi
Population	Guru aktif, calon guru, mahasiswa pendidikan keguruan, <i>pre-service teacher</i> , <i>student teacher</i> , <i>teacher candidate</i> , dan <i>teacher educator</i>	Populasi di luar profesi atau pendidikan guru, seperti tenaga medis, pegawai umum, atau profesi lain
Concept	Studi yang secara substantif menggunakan, mengembangkan, atau mengevaluasi <i>Situational Judgment Test/Situational Judgement Test</i>	Studi yang hanya membahas <i>judgment</i> , keputusan situasional, atau skenario kasus tetapi bukan SJT
Context	Sekolah, PAUD, pendidikan guru, seleksi calon guru, dan pengembangan profesional pendidik	Konteks nonpendidikan yang tidak berkaitan dengan profesi guru
Fokus studi	Pengembangan skenario, konstruk, format respons, skoring, validitas, reliabilitas, <i>fairness</i> , atau penggunaan SJT	SJT hanya disebut sebagai referensi atau tidak memberikan informasi yang relevan dengan tujuan <i>review</i>
Jenis publikasi	Artikel jurnal ilmiah dan prosiding lengkap	Editorial, berita, resensi, abstrak konferensi saja, dan publikasi nonilmiah
Bahasa Tahun	Bahasa Inggris atau Indonesia 2016–2026	Bahasa lain Di luar rentang tersebut

Proses penelusuran, identifikasi, deduplikasi, dan penyaringan awal literatur dilaksanakan pada 5–8 Agustus 2026, dengan pencarian terakhir pada seluruh sumber informasi dilakukan pada 8 Agustus 2026. Pencarian dilakukan menggunakan *Publish or Perish* [17] pada Scopus, Crossref, Google Scholar, GARUDA, dan ERIC. Istilah utama

“*Situational Judgment Test*” dan “*Situational Judgement Test*” dikombinasikan dengan istilah yang merujuk pada guru dan pendidikan guru, yaitu *teacher(s)*, *pre-service/preservice teacher*, *student teacher*, *teacher candidate*, *educator(s)*, *teacher education*, serta istilah terkait seleksi dan rekrutmen guru. Hasil pencarian didokumentasikan berdasarkan sumber, kata kunci, jumlah rekod, dan hasil deduplikasi. Skrining judul, abstrak, dan *full-text* dilakukan secara independen oleh dua dari empat penulis untuk setiap artikel, dengan penulis lainnya bertindak sebagai penengah apabila terdapat ketidaksepakatan dalam penentuan kelayakan studi, sebagaimana ditampilkan pada Gambar 1.



Gambar 1. Alur Seleksi Artikel berdasarkan prinsip PRISMA-ScR

Data diekstraksi menggunakan lembar *data charting* yang memuat identitas studi, negara atau konteks, karakteristik partisipan, tujuan dan konstruk SJT, karakteristik skenario, format respons, skoring, pendekatan psikometrik, temuan utama, keterbatasan, dan relevansinya bagi pengembangan SJT kompetensi sosial guru PAUD. Analisis dilakukan dalam dua tahap. Pertama, analisis deskriptif digunakan untuk memetakan karakteristik studi dan desain SJT. Kedua, sintesis tematik dilakukan dengan membandingkan informasi antarstudi, memberikan kode berdasarkan kesamaan makna, dan mengelompokkannya menjadi tema yang menjawab pertanyaan kajian. Temuan lintas studi kemudian diintegrasikan untuk merumuskan prinsip pengembangan SJT kompetensi sosial guru PAUD.

HASIL DAN PEMBAHASAN

Sintesis terhadap 14 artikel yang memenuhi kriteria inklusi menghasilkan lima tema utama: (1) karakteristik penelitian SJT pada guru dan pendidik; (2) konstruk

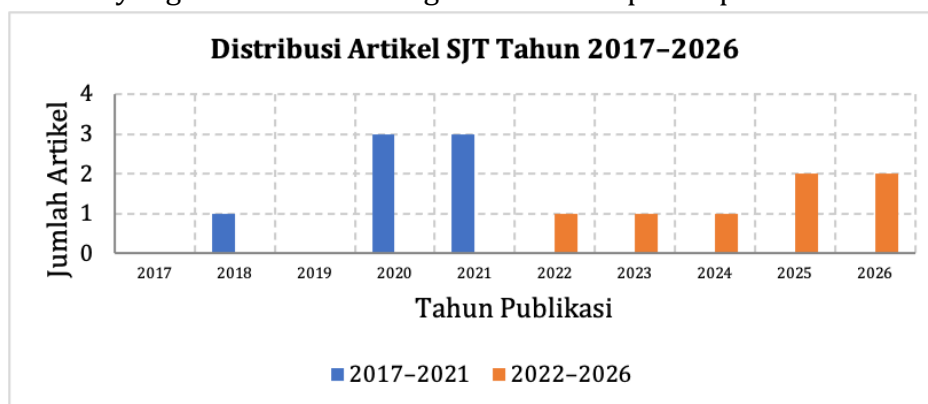
kompetensi terapan dan arsitektur *diagnose-respond*; (3) skenario autentik, konteks budaya, dan opsi respons; (4) format respons dan skoring; serta (5) validasi psikometrik, *fairness*, dan pemanfaatan SJT. Karakteristik studi utama disajikan pada Tabel 2.

Tabel 2. Karakteristik Studi Utama yang Diikutsertakan dalam Scoping Review

Kode	Studi	Asal	Partisipan	Fokus SJT	Format SJT
S1	Kusumangtyas et al. (2025)	Indonesia	1.090 calon guru fisika	<i>Situational judgment</i> dalam penerapan pengetahuan pedagogis dan PCK	SJT dikombinasikan dengan tes PCK; format respons SJT tidak dirinci dalam sintesis
S2	Gupta et al. (2026)	India	1.234 guru aktif	Kompetensi sosial-emosional guru: <i>emotion regulation</i> (ER) dan <i>relationship management</i> (RM)	Komik digital; 13 skenario (7 ER, 6 RM); peserta memberikan rating terhadap strategi respons
S3	Li et al. (2022)	China	1.084 guru PAUD aktif	Kemampuan menganalisis perkembangan/prekonsepsi anak dan menentukan respons pedagogis	16 situasi; mengukur kemampuan analisis dan respons secara terpisah; skor 0–2
S4	Chao et al. (2019)	Taiwan	5.039 mahasiswa; mencakup SMP, SD, dan PAUD	<i>Professional judgment</i> guru yang spesifik menurut jenjang pendidikan	Booklet 30 item; <i>paper-and-pencil</i> ; skoring berbasis SME
S5	Al Hashmi & Klassen (2019)	Oman	guru dan mahasiswa pendidikan guru; analisis akhir ±124 peserta	Seleksi calon guru: komunikasi, organisasi/perencanaan, resiliensi/adaptabilitas, motivasi, dan etika profesional	29 item final; format <i>ranking</i> dan <i>select-best-three</i>
S6	Rofiki et al. (2024)	Indonesia	10 calon subjek PPG PGSD; 4 kasus dianalisis mendalam	Proses kognitif dan intuisi dalam menyelesaikan SJT	Satu soal SJT matematika disertai alasan jawaban dan wawancara
S7	Koschmieder & Neubauer (2021)	Austria	Studi awal 188 peserta; seleksi nyata 3.138 peserta ; studi lanjutan 750 peserta	Regulasi emosi pedagogis: dimensi interpersonal dan intrapersonal	22 item; empat strategi berbasis teori; <i>single-best response</i>
S8	Aldrup et al. (2020)	Jerman	396 Guru aktif, calon guru, serta kelompok <i>face validity</i>	Kompetensi sosial-emosional guru: ER dan RM	15 skenario; peserta menilai efektivitas seluruh alternatif respons; <i>theory-based scoring</i>
S9	Ezra et al. (2026)	Israel	10 guru aktif TK-SMA	Pengetahuan <i>self-regulated learning-problem solving</i> (SRL-PS) dan proses diagnosis situasi	Pilot <i>open-ended</i> ; repository 66 situasi dikembangkan menjadi 10 situasi final
S10	Durksen & Klassen (2017)	Australia	29 pendidik lokal; 99 calon/guru baru pada uji pilot	Kompetensi guru <i>rural/remote</i> : empati/komunikasi,	55 item awal menjadi 32 item pilot; <i>ranking</i> dan <i>choose-three</i> ;

				perencanaan, resiliensi/adaptabilitas, dan konteks budaya	<i>distance scoring</i>
S11	Bardach, Rushby, & Klassen (2021)	Inggris	2.808 calon guru	<i>Fairness</i> dalam seleksi guru; <i>conscientiousness</i> , <i>mindset</i> , dan regulasi emosi	11 item; campuran teks dan video; <i>hybrid scoring</i>
S12	Starčević et al. (2023)	Serbia	245 mahasiswa pendidikan; 45,7% calon guru primary/kindergarten	Kompetensi interkultural, <i>perspective taking</i> , dan <i>isomorphic attribution</i>	9 item; masing-masing 5 opsi; <i>single-best response</i> ; skor 0–1
S13	Li et al. (2025)	China	72 calon guru PAUD dan 322 guru PAUD aktif	<i>Analytical Ability</i> (AA) dan <i>Responsive Ability</i> (RA) dalam situasi pedagogis PAUD	16 skenario; setiap skenario memuat dua pertanyaan AA dan RA; skor 0–2
S14	Bardach et al. (2021)	Inggris	290 calon guru STEM	Pengaruh format penyajian SJT terhadap skor, <i>engagement</i> , dan perbedaan antarkelompok	15 skenario yang sama dalam tiga kondisi: video+teks+audio, video+audio, teks+audio; 4 opsi dinilai <i>appropriate-inappropriate</i> ; <i>distance scoring</i>

Sintesis terhadap 14 artikel yang memenuhi kriteria inklusi menunjukkan bahwa penelitian SJT pada guru dan pendidik telah dipublikasikan sepanjang 2017–2026, dengan distribusi yang berfluktuasi sebagaimana ditampilkan pada Gambar 2.



Gambar 2. Distribusi Artikel SJT Tahun 2017 - 2026

Dalam periode 2017–2026, tujuh artikel diterbitkan pada 2017–2021 dan tujuh artikel lainnya pada 2022–2026. Meskipun jumlah publikasi pada kedua periode sama, pola distribusinya berbeda. Pada periode pertama, publikasi terkonsentrasi pada 2020 dan 2021, sedangkan dalam lima tahun terakhir artikel ditemukan secara lebih berkelanjutan setiap tahun.

Karakteristik dan Lanskap Penelitian SJT pada Guru dan Pendidik. Sebanyak 14 artikel yang memenuhi kriteria inklusi diterbitkan pada 2017–2026 dan berasal dari berbagai konteks di Asia, Eropa, dan Australia. Partisipan mencakup guru aktif, calon guru, mahasiswa pendidikan keguruan, peserta seleksi pendidikan guru, dan guru awal karier, dengan ukuran sampel sekitar 10 hingga lebih dari 5.000 peserta. Studi berskala kecil umumnya digunakan untuk mengeksplorasi proses respons dan pengembangan

skenario, sedangkan studi berskala besar lebih banyak menguji karakteristik psikometrik dan perbedaan antarkelompok. Keragaman ini menunjukkan bahwa SJT telah digunakan dalam berbagai sistem pendidikan dan sangat dipengaruhi oleh karakteristik profesi serta konteks sosial-budaya penggunaannya [8], [16], [18].

Penggunaan SJT juga mencakup beragam tujuan, mulai dari seleksi calon guru hingga pengukuran kompetensi sosial-emosional, regulasi emosi, hubungan guru-peserta didik, kompetensi interkultural, penalaran pedagogis, dan *professional judgment* [9]–[11], [16], [19]–[28]. Bentuk penyajiannya meliputi teks, komik digital, video, audio, maupun kombinasi media, dengan format respons berupa *single-best response*, rating, ranking, dan pemilihan beberapa respons. Sistem skoring juga bervariasi, dari dikotomis hingga *partial*, *distance*, *theory-based*, dan *hybrid scoring*. Variasi tersebut menegaskan bahwa SJT bukan instrumen dengan satu konstruk dan format baku; desain dan interpretasi skornya bergantung pada konstruk serta tujuan asesmen. Namun, bukti yang secara langsung berasal dari konteks PAUD masih terbatas. Hanya sebagian kecil studi secara khusus melibatkan guru atau calon guru prasekolah, sementara mayoritas penelitian berada pada konteks pendidikan guru, sekolah dasar dan menengah, atau seleksi calon guru [7], [12]. Dengan demikian, meskipun SJT telah berkembang sebagai asesmen *professional judgment* guru, penerapannya untuk kompetensi sosial guru PAUD memerlukan pengembangan yang spesifik terhadap karakter interaksi dengan anak, keluarga, rekan sejawat, dan pimpinan satuan pendidikan, bukan sekadar mengadopsi SJT dari jenjang atau konteks lain.

Hakikat Konstruk SJT: Kompetensi Terapan dan Arsitektur *Diagnose* → *Respond*. Hasil sintesis menunjukkan bahwa SJT digunakan untuk menangkap kemampuan menerapkan pengetahuan dan *professional judgment* dalam situasi yang menyerupai konteks kerja, sehingga konstruk yang diukur tidak sepenuhnya identik dengan pengetahuan deklaratif atau persepsi diri. Hal ini terlihat pada Kusumaningtyas et al., yang menemukan korelasi sedang antara *situational judgment* dan PCK ($r \approx 0,46$), tetapi masih terdapat kelemahan dalam penerapan pengetahuan pedagogis. Temuan ini menguatkan bahwa mengetahui prinsip profesional belum tentu diikuti kemampuan menerapkannya secara tepat dalam situasi tertentu, yaitu:

Pemisahan antara memahami situasi dan menentukan tindakan terlihat lebih jelas pada studi yang melibatkan guru PAUD. Li et al. (2022) menemukan korelasi relatif rendah antara kemampuan analisis dan respons ($r = 0,353$), sedangkan Li et al. (2025) secara eksplisit membedakan *Analytical Ability* dan *Responsive Ability* dalam setiap skenario [11], [28]. Temuan serupa juga muncul pada studi kompetensi interkultural yang menekankan *perspective taking* serta pada studi diagnostik yang menunjukkan kecenderungan guru langsung mengusulkan solusi tanpa diagnosis masalah yang lengkap [10], [24]. Secara keseluruhan, bukti tersebut menunjukkan bahwa *situation diagnosis* dan *response judgment* saling berkaitan, tetapi tidak identik.

Berdasarkan pola lintas studi tersebut, kajian ini mengusulkan arsitektur *diagnose* → *respond* sebagai prinsip pengembangan SJT kompetensi sosial guru PAUD. Dalam satu skenario, tahap *diagnose* dapat menilai kemampuan mengidentifikasi masalah, kebutuhan, emosi, perspektif, atau konteks sosial, sedangkan tahap *respond*

menilai pilihan tindakan profesional terhadap situasi yang sama [9], [24], [28]. Arsitektur ini relevan bagi PAUD karena respons sosial yang tepat bergantung pada ketepatan memahami anak, keluarga, dan dinamika hubungan yang melatarbelakangi situasi. Dengan demikian, SJT kompetensi sosial guru PAUD tidak hanya menilai kemampuan memilih tindakan yang dianggap terbaik, tetapi juga proses memahami situasi sosial yang mendasari tindakan tersebut.

Skenario Autentik, Konteks Budaya, dan Pembentukan Opsi Respons. Hasil sintesis menunjukkan bahwa autentisitas dan kekhususan konteks profesi merupakan fondasi penting dalam pengembangan SJT guru. Skenario perlu merepresentasikan situasi yang benar-benar mungkin dihadapi dalam pekerjaan, bukan sekadar persoalan umum. Pengembangannya dapat melibatkan praktisi, *critical incidents*, analisis konteks kerja, dan pengalaman lapangan. Chao et al. mengembangkan SJT yang dibedakan menurut jenjang pendidikan, sedangkan Durksen dan Klassen melibatkan pendidik lokal dalam konteks rural dan Li et al. (2025) mengombinasikan perspektif guru, anak, dan pakar [9], [11], [26]. Temuan tersebut menunjukkan bahwa SJT kompetensi sosial guru PAUD perlu dibangun dari karakteristik interaksi yang spesifik pada pendidikan anak usia dini.

Respons terbuka pada tahap awal juga bermanfaat untuk memperoleh alternatif jawaban yang lebih autentik. Ezra et al. menggunakan situasi yang dihimpun dari pengalaman guru dan respons terbuka untuk mengidentifikasi cara peserta memahami serta menangani masalah [27]. Pendekatan ini dapat membantu menghasilkan *distractor* yang realistis. Alternatif respons sebaiknya sama-sama *plausible* dan merepresentasikan variasi kualitas tindakan, bukan terdiri atas satu jawaban yang jelas benar dan pilihan lain yang mudah dikenali sebagai salah. Dalam konteks PAUD, variasi tersebut dapat mencakup respons optimal, cukup tepat tetapi belum lengkap, terlalu prosedural atau berorientasi kontrol, hingga respons yang berisiko atau bias [9], [16], [26].

Konteks budaya juga perlu menjadi bagian dari desain SJT. Kontekstualisasi tidak cukup melalui penerjemahan bahasa atau penggunaan media yang lebih menarik, tetapi perlu mempertimbangkan aktor, relasi, norma sosial, dan makna respons yang dianggap tepat. Penggunaan video atau komik dapat memperkaya informasi situasional, tetapi tidak otomatis meningkatkan validitas atau menghilangkan perbedaan antarkelompok [23], [26]. Oleh karena itu, media visual lebih tepat digunakan ketika ekspresi, intonasi, gestur, atau informasi nonverbal memang relevan dengan konstruk yang diukur.

Berdasarkan temuan tersebut, pengembangan skenario SJT kompetensi sosial guru PAUD dapat mengombinasikan pendekatan *bottom-up* melalui penggalan *critical incidents* dan pendekatan *top-down* untuk memasukkan situasi penting seperti inklusi, kerahasiaan, perlindungan anak, diskriminasi, dan konflik nilai. *Blueprint* dapat menyilangkan domain kompetensi dengan mitra interaksi, seperti anak, keluarga, rekan guru, pimpinan satuan, dan komunitas. Skenario kemudian diuji melalui respons terbuka atau *cognitive interview* dan dilengkapi opsi yang sama-sama masuk akal, sehingga SJT lebih merepresentasikan kompleksitas *judgment* sosial guru PAUD daripada sekadar mengukur kemampuan mengenali jawaban normatif.

Format Respons dan Sistem Skoring SJT, hasil sintesis menunjukkan bahwa SJT menggunakan beragam format respons, seperti *single-best response*, *rating*, *ranking*, dan pemilihan beberapa respons, dengan sistem skoring dikotomis, *partial credit*, *distance*, maupun *theory-based scoring*. Tidak terdapat satu format yang unggul untuk seluruh konstruk; pemilihannya perlu disesuaikan dengan tujuan asesmen, karakteristik situasi, dan jenis *judgment* yang diukur.

Tabel 3. Perbandingan Format Respons Dan Sistem Skoring SJT

Format respons / skoring	Studi pendukung	Temuan utama	Implikasi untuk SJT kompetensi sosial guru PAUD
<i>Single-best response</i>	S7, S12	Sederhana dan sesuai ketika terdapat satu judgment yang paling defensible; dapat menyederhanakan situasi jika beberapa respons sebenarnya masih plausible	Potensial digunakan pada tahap diagnosis situasi, misalnya memilih interpretasi kebutuhan, emosi, atau perspektif yang paling tepat
Rating seluruh opsi	S2, S8, S14	Memberikan informasi lebih kaya karena peserta menilai kualitas beberapa respons, bukan hanya memilih satu jawaban	Potensial digunakan pada tahap <i>response judgment</i> , terutama ketika beberapa tindakan memiliki tingkat efektivitas berbeda
Ranking	S5, S10	Reliabilitas dapat lebih rendah dan peserta dapat menganggap proses pengurutan membingungkan	Tidak perlu dijadikan pilihan utama; perlu diuji terlebih dahulu pada pilot
<i>Select-best/choose-three</i>	S5, S10	Dapat menjadi alternatif ranking; pada salah satu studi menunjukkan reliabilitas lebih baik dibandingkan ranking	Dapat diuji sebagai format alternatif apabila respons sosial membutuhkan pemilihan beberapa tindakan
Dikotomis (0–1)	S7, S12	Mudah diinterpretasikan, tetapi kurang sensitif terhadap variasi kualitas respons	Lebih sesuai jika terdapat batas substantif yang jelas antara respons tepat dan tidak tepat
<i>Partial credit</i> (misal, 0–2)	S3, S13	Memungkinkan respons dinilai bertingkat dan digunakan pada studi yang langsung melibatkan guru PAUD	Relevan untuk situasi sosial karena respons dapat bersifat optimal, cukup tepat, atau kurang tepat
<i>Distance/theory-based scoring</i>	S8, S10, S14	Skor mempertimbangkan derajat kesesuaian dengan judgment ahli atau kriteria teoretis	Potensial untuk respons sosial yang sama-sama plausible, dengan kunci berbasis teori dan panel ahli

Tabel 3 menunjukkan bahwa *single-best response* lebih sesuai ketika diperlukan satu interpretasi utama, sedangkan *rating* memberikan informasi lebih kaya ketika beberapa tindakan memiliki tingkat efektivitas berbeda [16], [18], [20], [23], [24]. Sebaliknya, *ranking* menunjukkan keterbatasan pada beberapa studi dari sisi reliabilitas maupun kenyamanan peserta [11], [19], [29]. Sistem skoring juga perlu mempertimbangkan sifat respons. Dalam konteks PAUD, penggunaan *partial credit* memungkinkan kualitas *judgment* dinilai secara bertingkat dan tidak direduksi menjadi benar atau salah [9].

Berdasarkan temuan tersebut, arsitektur *diagnose* → *respond* dapat menggunakan *single-best response* atau *multiple-choice* pada tahap diagnosis dan *rating* atau *best–worst* pada tahap respons. Skoring bertingkat, misalnya 0–2 atau 0–3, dapat

dipertimbangkan untuk merepresentasikan derajat kualitas respons dengan kunci berbasis teori, *judgment* ahli, dan bukti empiris. Format akhir tetap perlu diuji pada tahap pilot karena karakteristik psikometriksnya dapat berbeda menurut konstruk dan populasi.

Validasi Psikometrik, Fairness, dan Pemanfaatan SJT, hasil sintesis menunjukkan bahwa kualitas psikometrik SJT tidak dapat dinilai hanya melalui konsistensi internal. Reliabilitas antarstudi bervariasi, sementara pada beberapa instrumen test-retest menunjukkan kestabilan yang lebih baik dibandingkan alpha [11], [16], [19], [24]. Oleh karena itu, evaluasi SJT perlu menggunakan bukti yang beragam, meliputi konsistensi internal, reliabilitas per skenario, test-retest, analisis item, serta struktur internal melalui CFA atau IRT/GRM. Struktur SJT juga tidak selalu homogen; beberapa studi mendukung model multidimensi, sedangkan studi lain menunjukkan struktur yang lebih bergantung pada karakteristik situasi. Dengan demikian, model pengukuran perlu diuji sesuai konstruk yang mendasari instrumen.

Validitas SJT menjadi lebih kuat ketika skor dikaitkan dengan indikator eksternal yang relevan. Pada TRUST, skor SJT berhubungan dengan emotional support dan kualitas hubungan guru-siswa, sedangkan pada SJT kompetensi interkultural ditemukan hubungan dengan tugas interkultural dalam kehidupan nyata [9], [25]. Temuan ini menunjukkan bahwa validasi tidak cukup dilakukan melalui korelasi dengan *self-report*, tetapi perlu melibatkan indikator perilaku atau performa. Dalam konteks PAUD, bukti tersebut dapat diperoleh melalui observasi interaksi guru-anak, simulasi komunikasi dengan keluarga, penilaian multisumber, atau tugas penyelesaian konflik.

Fairness juga merupakan bagian penting dari validasi. Beberapa studi menemukan perbedaan skor berdasarkan gender, etnisitas, maupun format penyajian, tetapi perbedaan rerata tidak dengan sendirinya menunjukkan bias [16], [26]. Karena itu, pengembangan SJT kompetensi sosial guru PAUD perlu mempertimbangkan *measurement invariance*, dan *cognitive interview* untuk memastikan bahwa skenario, bahasa, alternatif respons, serta kunci skoring dipahami secara setara oleh kelompok dengan latar berbeda. Dalam konteks Indonesia, pengujian dapat mempertimbangkan wilayah, budaya, latar pendidikan, jenis satuan PAUD, dan pengalaman mengajar.

Bukti psikometrik selanjutnya perlu disesuaikan dengan tujuan penggunaan skor. SJT berpotensi digunakan untuk diagnosis, refleksi, pembelajaran, dan pengembangan profesional, tetapi penggunaannya untuk seleksi atau keputusan *high-stakes* membutuhkan bukti reliabilitas, validitas, dan *fairness* yang lebih kuat. Oleh karena itu, SJT kompetensi sosial guru PAUD pada tahap awal lebih tepat digunakan sebagai bagian dari asesmen *multimethod* yang bersifat diagnostik dan developmental, bersama observasi, simulasi, atau penilaian lain. Dengan demikian, validasi SJT tidak hanya memastikan bahwa skor reliabel, tetapi juga bahwa interpretasi dan penggunaannya dapat dipertanggungjawabkan.

Kontribusi penelitian ini mencakup aspek teoretis, metodologis, dan praktis. Secara teoretis, kajian ini memperjelas bahwa kemampuan mendiagnosis situasi sosial dan menentukan respons profesional merupakan dua kemampuan yang saling berkaitan, tetapi tidak identik. Secara metodologis, kajian ini mengintegrasikan bukti

yang sebelumnya tersebar mengenai konstruk, pengembangan skenario autentik, pembentukan alternatif respons yang plausible, pemilihan format respons, scoring bertingkat, serta validasi psikometrik dan fairness ke dalam arsitektur *diagnose → respond*. Secara praktis, hasil sintesis memberikan arah bagi pengembangan SJT kompetensi sosial guru PAUD yang mempertimbangkan interaksi dengan anak, keluarga, rekan sejawat, pimpinan satuan pendidikan, dan masyarakat. Pada tahap awal, SJT tersebut lebih tepat digunakan sebagai bagian dari asesmen multimethod yang bersifat diagnostik dan mendukung pengembangan profesional, bersama observasi, simulasi, atau penilaian lain. Penggunaannya untuk seleksi atau keputusan *high-stakes* memerlukan bukti empiris yang lebih kuat mengenai reliabilitas, validitas, dan fairness skor.

Kajian ini memiliki beberapa keterbatasan. Pertama, pencarian dibatasi pada Scopus, Crossref, Google Scholar, GARUDA, dan ERIC serta pada publikasi berbahasa Indonesia atau Inggris, sehingga artikel relevan dalam bahasa lain atau yang terindeks di luar sumber tersebut mungkin belum teridentifikasi. Kedua, kajian hanya menyertakan artikel jurnal dan prosiding lengkap sehingga literatur abu-abu, seperti laporan penelitian, disertasi, dan dokumen pengembangan instrumen yang tidak dipublikasikan, belum tercakup. Ketiga, bukti yang secara langsung melibatkan guru atau calon guru PAUD masih terbatas dan sebagian besar studi berasal dari jenjang pendidikan, negara, serta konteks budaya yang berbeda. Oleh karena itu, implikasi bagi kompetensi sosial guru PAUD Indonesia masih bersifat konseptual dan memerlukan pengujian empiris. Selain itu, pencarian literatur terakhir dilakukan pada 8 Agustus 2026 sehingga publikasi yang tersedia setelah tanggal tersebut tidak tercakup dalam kajian ini. Arsitektur *diagnose → respond* yang dihasilkan perlu diperlakukan sebagai landasan pengembangan instrumen, bukan sebagai model pengukuran yang telah tervalidasi.

KESIMPULAN

Scoping review ini menegaskan bahwa pengembangan *Situational Judgment Test* (SJT) kompetensi sosial guru PAUD perlu mengukur tidak hanya kemampuan menentukan tindakan yang tepat, tetapi juga kemampuan memahami situasi sosial yang melatarbelakanginya. Kebaruan penelitian ini terletak pada perumusan arsitektur *diagnose → respond* yang membedakan kemampuan mendiagnosis situasi sosial dan menentukan respons profesional sebagai dua komponen yang saling berkaitan, tetapi tidak identik. Arsitektur tersebut mengintegrasikan konstruk yang spesifik, skenario autentik dan kontekstual, alternatif respons yang *plausible* dan bertingkat, scoring parsial, serta validasi yang mencakup bukti psikometrik, proses respons, kriteria eksternal, dan fairness. Dengan arsitektur ini, SJT berpotensi memberikan gambaran kompetensi sosial yang lebih dekat dengan tuntutan nyata pekerjaan guru dibandingkan pengukuran yang hanya mengandalkan pengetahuan atau persepsi diri. Kontribusi penelitian ini adalah menyediakan landasan konseptual dan metodologis bagi pengembangan SJT kompetensi sosial guru PAUD yang autentik, kontekstual, dan dapat dipertanggungjawabkan. Instrumen tersebut terutama diarahkan sebagai bagian dari

asesmen *multimethod* untuk diagnosis dan pengembangan profesional sebelum digunakan dalam pengambilan keputusan yang bersifat *high-stakes*.

PENGHARGAAN

Terima kasih disampaikan kepada Kementerian Pendidikan Tinggi, Sains, dan Teknologi Republik Indonesia atas pendanaan melalui Program Penelitian Dosen Pemula (PDP) BIMA Tahun 2026, berdasarkan Kontrak Induk 283/C3/DT.05.00/PL-BARU/2026 serta Kontrak Turunan 1519/LL4/PG/2026 dan 010/KP/UPPM/STPHBK/V/2026.

REFERENSI

- [1] R. Rakhmania, B. P. D. Riyanti, dan M. Purwanti, "Construct Validity and Reliability: Perception of Teacher Competency Scale," *JP3I (Jurnal Pengukuran Psikol. dan Pendidik. Indones.*, vol. 14, no. 2, hal. 46–74, Nov 2025, doi: 10.15408/jp3i.v14i2.41460.
- [2] T. G. Hummel, F. Cohen, J. Gessulat, dan Y. Anders, "Measuring interaction quality between parents and professionals and its relation to preschool characteristics," *Int. J. Educ. Res. Open*, vol. 3, hal. 100195, 2022, doi: 10.1016/j.ijedro.2022.100195.
- [3] G. Olaru, J. Burrus, C. MacCann, F. M. Zaromb, O. Wilhelm, dan R. D. Roberts, "Situational Judgment Tests as a method for measuring personality: Development and validity evidence for a test of Dependability," *PLoS One*, vol. 14, no. 2, hal. e0211884, Feb 2019, doi: 10.1371/journal.pone.0211884.
- [4] G. Lozano-Peña, F. Sáez-Delgado, Y. López-Angulo, dan J. Mella-Norambuena, "Teachers' Social-Emotional Competence: History, Concept, Models, Instruments, and Recommendations for Educational Quality," *Sustainability*, vol. 13, no. 21, hal. 12142, Nov 2021, doi: 10.3390/su132112142.
- [5] N. Rahminawati, A. Hakim, A. Sobarna, dan M. Masnipal, "Teaching Competency Development for Early Childhood Education," in *Proceedings of the 2nd Social and Humaniora Research Symposium (SoRes 2019)*, 2020. doi: 10.2991/assehr.k.200225.063.
- [6] L. Halimah, M. S. Yuliariatiningsih, dan Mirawati, "Strategies for optimizing partnership: workshop on teamwork for teachers and parents in early childhood education," *Int. J. Early Years Educ.*, vol. 34, no. 1, hal. 24–41, Jan 2026, doi: 10.1080/09669760.2023.2299239.
- [7] S. Kepes, S. K. Keener, F. Lievens, dan M. A. McDaniel, "An Integrative, Systematic Review of the Situational Judgment Test Literature," *J. Manage.*, vol. 51, no. 6, hal. 2278–2319, Jul 2025, doi: 10.1177/01492063241288545.
- [8] P. Schäpers, S. Krumm, F. Lievens, J. Freudenstein, J. Schulze, dan C. J. König, "Situation Descriptions in Situational Judgment Tests: A Matter of Including Trait-Relevant Situational Cues?," *Int. J. Sel. Assess.*, vol. 34, no. 1, hal. 1–17, Feb 2026, doi: 10.1111/ijsa.70032.
- [9] K. Aldrup, B. Carstensen, M. M. Köller, dan U. Klusmann, "Measuring Teachers' Social-Emotional Competence: Development and Validation of a Situational Judgment Test," *Front. Psychol.*, vol. 11, no. May, hal. 1–20, Mei 2020, doi: 10.3389/fpsyg.2020.00892.
- [10] L. Li, Y. Xi, B. Yu, Z. BeiBei, dan X. Lü, "Pre-service and in-service preschool teachers' abilities to analyze and respond to children's preconceptions: evidence

- from situational judgement tests,” *Early Years*, hal. 1–17, Nov 2025, doi: 10.1080/09575146.2025.2589471.
- [11] L. Li, X. Lv, Y. Xi, dan L. Guo, “Can Preschool Teachers’ Accurate Analysis of the Development Trajectories of Children’s Preconceptions Ensure Their Effective Response? Evidence from Situational Judgement Tests,” *Sustainability*, vol. 14, no. 18, hal. 11725, Sep 2022, doi: 10.3390/su141811725.
 - [12] A. I. Nadmilail, M. E. E. Mohd Matore, S. M. Maat, dan L. Sheridan, “Broad vs. narrow traits: a scoping review of measuring personality traits in teacher selection using the situational judgment test,” *Front. Psychol.*, vol. 14, Sep 2023, doi: 10.3389/fpsyg.2023.1217321.
 - [13] A. C. Tricco *et al.*, “PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation,” *Ann. Intern. Med.*, vol. 169, no. 7, hal. 467–473, Okt 2018, doi: 10.7326/M18-0850.
 - [14] M. D. J. Peters *et al.*, “Updated methodological guidance for the conduct of scoping reviews,” *JBIM Evid. Synth.*, vol. 18, no. 10, hal. 2119–2126, Okt 2020, doi: 10.11124/JBIES-20-00167.
 - [15] E. du Plooy, D. Casteleijn, dan D. Franzsen, “Personalized adaptive learning in higher education: A scoping review of key characteristics and impact on academic performance and engagement,” *Heliyon*, vol. 10, no. 21, hal. e39630, Nov 2024, doi: 10.1016/j.heliyon.2024.e39630.
 - [16] L. Bardach, J. V. Rushby, dan R. M. Klassen, “The selection gap in teacher education: Adverse effects of ethnicity, gender, and socio-economic status on situational judgement test performance,” *Br. J. Educ. Psychol.*, vol. 91, no. 3, hal. 1015–1034, Sep 2021, doi: 10.1111/bjep.12405.
 - [17] Harzing.com, “Publish or Perish.” Tarma Software Research Ltd, 2025. [Daring]. Tersedia pada: <https://harzing.com/resources/publish-or-perish/windows>
 - [18] M. A. Sorrel, J. Olea, F. J. Abad, J. de la Torre, D. Aguado, dan F. Lievens, “Validity and Reliability of Situational Judgement Test Scores,” *Organ. Res. Methods*, vol. 19, no. 3, hal. 506–532, Jul 2016, doi: 10.1177/1094428116630065.
 - [19] D. A. Kusumaningtyas, Ishafit, M. Fitriawanati, dan E. Kurniasari, “Analysis of Future Physics Teachers Abilities Using STJ and Pedagogical Content Knowledge Approaches,” *J. Phys. Conf. Ser.*, vol. 3139, no. 1, hal. 012098, Nov 2025, doi: 10.1088/1742-6596/3139/1/012098.
 - [20] C. Koschmieder dan A. C. Neubauer, “Measuring emotion regulation for preservice teacher selection: A theory-driven development of a situational judgment test,” *Pers. Individ. Dif.*, vol. 168, hal. 110363, Jan 2021, doi: 10.1016/j.paid.2020.110363.
 - [21] T. L. Durksen dan R. M. Klassen, “The development of a situational judgement test of personal attributes for quality teaching in rural and remote Australia,” *Aust. Educ. Res.*, vol. 45, no. 2, hal. 255–276, Apr 2018, doi: 10.1007/s13384-017-0248-5.
 - [22] W. Al Hashmi dan R. M. Klassen, “Developing a situational judgement test for admission into initial teacher education in Oman: An exploratory study,” *Int. J. Sch. Educ. Psychol.*, vol. 8, no. sup1, hal. 187–198, Des 2020, doi: 10.1080/21683603.2019.1630042.
 - [23] H. Gupta, B. Carstensen, U. Klusmann, dan N. C. Singh, “Bringing situations to life: Validating the TRUST comic version, a situational judgement test of teachers’ social-emotional competence,” *Soc. Emot. Learn. Res. Pract. Policy*, vol. 8, hal. 100218, Des 2026, doi: 10.1016/j.sel.2026.100218.

- [24] I. Rofiki *et al.*, "Intuisi Logis Mahasiswa PPG Universitas Negeri Malang dalam Menyelesaikan Situational Judgement Test (SJT)," *Kogn. J. Ris. HOTS Pendidik. Mat.*, vol. 4, no. 3, Sep 2024, doi: 10.51574/kognitif.v4i3.1936.
- [25] J. Starčević, A. Altaras Dimitrijević, dan Z. Jolić Marjanović, "Validating a situational judgment test of intercultural competence for use with education professionals," *Int. J. Sch. Educ. Psychol.*, vol. 11, no. 2, hal. 180–192, Apr 2023, doi: 10.1080/21683603.2022.2069619.
- [26] L. Bardach, J. V. Rushby, L. E. Kim, dan R. M. Klassen, "Using video- and text-based situational judgement tests for teacher selection: a quasi-experiment exploring the relations between test format, subgroup differences, and applicant reactions," *Eur. J. Work Organ. Psychol.*, vol. 30, no. 2, hal. 251–264, Mar 2021, doi: 10.1080/1359432X.2020.1736619.
- [27] O. Ezra, D. Avidov, G. Cohen, A. Cohen, dan A. Bronshtein, "Teachers' self-regulated learning: Situational judgment test finds cognitive skills outperform motivational skills," *Teach. Teach. Educ.*, vol. 176, hal. 105506, Jul 2026, doi: 10.1016/j.tate.2026.105506.
- [28] T.-Y. Chao, Y.-T. Sung, dan J.-L. Huang, "Construction of the situational judgment tests for teachers," *Asia-Pacific J. Teach. Educ.*, vol. 48, no. 4, hal. 355–374, Agu 2020, doi: 10.1080/1359866X.2019.1633621.
- [29] A. A. Grant, L. Jeon, dan C. K. Buettner, "Chaos and commitment in the early childhood education classroom: Direct and indirect associations through teaching efficacy," *Teach. Teach. Educ.*, vol. 81, hal. 50–60, Mei 2019, doi: 10.1016/j.tate.2019.02.010.