**Journal of Innovative and Creativity**
Journal Homepage: https://joecy.org/index.php/joecy
e-mail: journaljoecy@gmail.com

Faculty of Education
University of Pahlawan Tuanku Tambusai
ISSN 2776-771X (Online)
ISSN 2962-570X (Printed)
6 (2) 2026, Page: 33341-33354

# Pengembangan dan Validasi Situasional Judgement Test untuk mengukur Efikasi Diri pada Mahasiswa

**Alvira Andrea Jati[1], Yassina Haehta Ratu[2], Mia Lukluul Milah[3], Teguh Mulya Sandi[4], Nabila Lutfi Fauzan[5], Tasya Augustiya[6], Novita Sari[7]**

**[1,2,3,4,5,6,7]Fakultas Sosial dan Humaniora, Universitas Muhammadiyah Bandung, Indonesia**

Corresponding Author: Name, Surname, e-mail: alviraajati@gmail.com[1] , yhaehta@gmail.com[2], mialuluulmilah@gmail.com[3], themulya06@gmail.com[4], nabilalutfifauzan@gmail.com[5] , Tasya.augustiya@umbandung.ac.id[6], novita.sari@umbandung.ac.id[7]

Published: Juni , 2026

## ABSTRAK

Efikasi diri berperan penting dalam keberhasilan akademik mahasiswa, namun sebagian besar instrumen yang tersedia masih berbasis self-report skala likert yang rentan terhadap bias sosial dan kurang merepresentasikan respons individu dalam konteks situasional akademik. Penelitian ini bertujuan mengembangkan dan memvalidasi Situasional Judgment Test (SJT) untuk mengukur efikasi diri mahasiswa berdasarkan teori Bandura (1997) dengan melibatkan 204 mahsiswa aktif di indonesia berpartisipasi melalui teknik *convenience sampling*, dengan instrumen terdiri dari 18 skenario dan tiga opsi jawaban. Validitas isi menunjukan indeks Aiken sebesar 0,80-1,00, reabilitas memperoleh koefisien Alpha Cronbach sebesar 0,65, dan hasil *Confirmatory Factor Analysis* menunjukkan model memiliki kecocokan yang cukup baik. Instrumen SJT ini menunjukan karateristik psikometrik yang memadai sebagai alternatif pengukuran efikasi diri mahasiswa yang lebih kontekstual.

## ABSTRACT

Self-efficacy plays an important role in students' academic success; however, most of the available instruments are still based on self-report Likert scale measures, which are susceptible to social bias and do not adequately represent individual responses in the academic situational context. This study aims to develop and validate a Situational Judgment Test (SJT) to measure students' self-efficacy based on Bandura's theory (1997) by involving 204 active students in Indonesia who participated through the convenience sampling technique. The instrument consists of 18 scenarios and three response options. Content validity indicates an Aiken's index of 0.80–1.00, reliability yields a Cronbach's Alpha coefficient of 0.65, and the results of Confirmatory Factor Analysis show that the model has a reasonably good fit. This SJT instrument demonstrates adequate psychometric characteristics as an alternative way to measure students' self-efficacy in a more contextual manner.

**Keywords:** *Self Efficacy*, Mahasiswa, Validitas, Reliabilitas

© The Author(s). 2021 Open Access This article is licensed under a Creative Commons Atibution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0

## PENDAHULUAN

Fenomena mahasiswa drop out (DO) masih menjadi perhatian dalam sistem pendidikan tinggi di Indonesia. Data statistik pendidikan tinggi yang dikemukakan oleh kementrian pendidikan, kebudayaan, riset, dan teknologi pada tahun 2022 menunjukan bahwa terdapat sekitar 375.134 mahasiswa yang mengalami putus studi dari total 9,32 juta mahasiswa terdaftar, atau setara dengan kurang lebih 4,02% dari total keseluruhan mahasiswa di Indonesia, angka ini terjadi penurunan 7% dari tahun 2020 (Kemendikbudristek, 2023). Akan tetapi jumlah tersebut menunjukan permasalahan yang signifikan bagi keberlangsungan studi mahasiswa salah satu faktor yang membuat mahasiswa DO adalah kondisi mental mahasiswa. Telah ditunjukan oleh survei nasional kesehatan mental, seperti kecemasan serta depresi terbilang cukup tinggi dikangalangan usia mahasiswa (Kementerian

Kesehatan Republik Indonesia, 2022). Sehingga hal ini akan sangat berpengaruh pada motivasi belajar, kepercayaan diri, konsentrasi, serta kemampuan dalam menyelesaikan tantangan yang dihadapi di lingkungan akademik. Berdasarkan fenomena tersebut dapat dipahami bahwa permasalahan mahasiswa tidak hanya dipengaruhi oleh faktor akademik saja, akan tetapi berkaitan erat dengan kemampuan efikasi diri pada mahasiswa dalam menghadapi setiap permasalahan dalam lingkungan akademik.

Efikasi diri adalah suatu keyakinan pada diri sendiri yang akan berpengaruh dalam usahanya juga ketekunanya ketika menyelesaikan permasalahan yang sedang dihadapi (Santrock 2008). Efikasi diri sangat berpengaruh karena rendahnya efikasi diri pada mahasiswa dapat menyebabkan kurang nya kemampuan mahasiswa dalam menghadapi permasalahan di perguruan tinggi (Salim,F., & Fakhrurrozi, M. M. 2020). Efikasi diri sebagai keyakinan seseorang dalam meyakini kemampuannya ketika melakukan suatu hal guna mencapai hasil tertentu (Albert Bandura 1977). Ketika mahasiswa memiliki keyakinan pada kemampuan yang dimilikinya maka hal ini akan berdampak pada usaha yang dilakukan menjadi lebih optimal.

Efikasi diri pada mahasiswa sangat diperlukan karena hal ini telah terbukti bahwa tingginya efikasi pada mahasiswa berperan dalam pencapaian akademik mahasiswa. Mahasiswa yang percaya diri dalam kemampuannya untuk belajar serta menyelesaikan tugas akan menunjukan motivasi intrinsik yang lebih tinggi, keterlibatan dalam pembelajaran, serta memiliki ketekunan dalam menghadapi kesulitan belajar (Schunk & Pajares, 2002). Individu dengan efikasi diri tinggi umumnya menunjukan kepercayaan diri yang kuat, ketahanan terhadap stres, kemauan untuk mencoba hal baru, serta kecenderungan untuk menyelesaikan tugas-tugas menantang dengan strategi adaptif dan pemikiran positif. Mereka lebih gigih dan tahan banting ketika menghadapi kegagalan. Sebaliknya individu dengan tingkat efikasi diri rendah cenderung meragukan kemampuannya sendiri, mudah merasa cemas cenderung menghindari tantangan dan memiliki resiko lebih tinggi terhadap munculnya perilaku pasif. Sehingga adanya permasalahan yang sering dijumpai seperti DO di lingkungan mahasiswa menunjukan bahwa efikasi diri berperan sangat penting terhadap ketahanan mahasiswa dalam menghadapi setiap tantangan serta kesulitan selama berproses di lingkungan akademik, sehingga efikasi pada mahasiswa harus diidentifikasi secara akurat sehingga membutuhkan alat ukur yang valid untuk mengukur efikasi diri pada mahasiswa.

Telah banyak alat ukur efikasi telah dikembangkan sebelumnya, oleh Tiara et.al (2022) Dengan judul Validitas Konstruk Alat Ukur *College Academic Self-Efficacy Scale* pada Mahasiswa, instrumen ini telah diadaptasi ke bahasa Indonesia pada tahun 2022. Penelitian menggunakan metode kuantitatif dengan 116 responden mahasiswa aktif, pertanyaan menggunakan skala likert dengan 33 pertanyaan. Berdasarkan hasil analisis CFA, adaptasi *College Academic Self-Efficacy Scale* (CASES) dalam konteks Indonesia belum memenuhi kriteria fit secara konstruk karena nilai CFI (0,888) belum memenuhi kriteria ≥ 0,90. Meskipun nilai GFI telah memenuhi kriteria fit dan RMSEA berada pada kategori moderate.

Pengembangan alat ukur self-efficacy lainnya dilakukan oleh Pratiwi, R (2024) menggunakan skala likert, hasil dari penelitian telah diujikan pada 315 mahasiswa serta menunjukan bahwa dari 42 aitem yang telah disusun yang memenuhi kelayakan 32 aitem yang memenuhi kriteria kelayakan berdasarkan indeks beda. Berdasarkan hasil penelitian, alat ukur skala efikasi diri pada mahasiswa belum dapat dikatakan layak dalam konteks Indonesia. Meskipun telah melalui uji validitas isi (CVI), hasil uji daya beda menunjukan sebagian besar item memiliki indeks diskriminasi dibawah 0,30. Serta nilai koefisien reliabilitas Cronbach's Alpha sebesar 0,55 menunjukan konsistensi internal yang rendah dan belum memenuhi standar reliabilitas yang direkomendasikan (≥ 0,70). Maka alat ukur ini masih memerlukan perbaikan pada aspek penyusunan indikator dan item.

Selain itu Febrian Solikhin (2020) juga mengembangkan skala self efficacy pada siswa dalam pembelajaran kimia, menggunakan skala likert dengan 22 butir pertanyaan, melibatkan 300 responden siswa SMA. melakukan model perhitungan menggunakan *RASCH* model, serta dalam penelitian ini pengujian validitas dan reliabilitas instrumen menggunakan program *Quest* langsung diketahui besar validitas dan reliabilitas. Setiap pertanyaan dinilai valid karena nilai infit MNSQ berada dalam rentang 0,77-1,30. Sehingga instrumen ini dinilai cukup reliabel untuk digunakan dalam pengukuran efikasi diri pada mahasiswa.

Dari ketiga alat ukur efikasi diri yang telah dibuat sebelumnya belum terdapat alat ukur dengan bentuk Situasional Judgement Test (SJT) dan Instrumen efikasi diri pada mahasiswa di Indonesia masih terbatas dalam metode, dan mayoritas menggunakan skala likert Sehingga dengan demikian, terdapat kebutuhan pengembangan instrumen yang lebih kontekstual dan berbasis situasi nyata untuk meminimalkan bias, sehingga belum ada pengembangan berbentuk *Situational Judgement Test* (SJT) dalam konteks efikasi diri pada mahasiswa, maka untuk memperkaya alat ukur efikasi diri pada mahasiswa pada penelitian ini akan dibuat alat efikasi diri pada mahasiswa dengan alat ukur *Situational Judgement Test* (SJT).

# LANDASAN TEORI

*Self-efficacy* atau efikasi diri merupakan aspek penting dalam psikologi yang memengaruhi bagaimana individu merasakan, berpikir, bertindak, dan memotivasi diri mereka dalam menghadapi berbagai tantangan hidup. Konsep ini pertama kali dikembangkan oleh Albert Bandura (1977) dan didefinisikan sebagai keyakinan individu terhadap kemampuannya dalam mengorganisasi dan melaksanakan tindakan yang diperlukan guna mencapai hasil tertentu (Warner & Schwarzer, 2017).

Keyakinan ini tidak hanya berkaitan dengan persepsi kemampuan, tetapi juga mempengaruhi bagaimana seseorang berfikir, merasakan, memotivasi diri, dan bertindak dalam menghadapi tekanan atau hambatan. *Self-efficacy* memengaruhi berbagai aspek kehidupan manusia. Individu dengan efikasi diri tinggi umumnya menunjukkan kepercayaan diri yang kuat, ketahanan terhadap stres, kemauan untuk mencoba hal baru, serta kecenderungan untuk menyelesaikan tugas-tugas yang menantang dengan strategi adaptif dan pemikiran positif (Garmezy, 1991). Mereka lebih gigih dan tahan banting ketika menghadapi kegagalan. Sebaliknya, individu dengan tingkat *self-efficacy* rendah cenderung meragukan kemampuannya sendiri, mudah merasa cemas, cenderung menghindari tantangan, dan memiliki risiko lebih tinggi terhadap munculnya perilaku pasif atau learned helpness.

Dalam bidang pendidikan, *self-efficacy* telah terbukti berperan dalam pencapaian akademik siswa. Siswa yang percaya diri dalam kemampuannya untuk belajar dan menyelesaikan tugas akan menunjukkan motivasi intrinsik yang lebih tinggi, keterlibatan dalam pembelajaran, serta ketekunan dalam menghadapi kesulitan belajar (Schunk & Pajares, 2002). Dalam bidang psikologi klinis, *self-efficacy* juga telah dikaitkan dengan kemampuan menghadapi gangguan mental, seperti depresi dan kecemasan, karena individu yang memiliki efikasi diri tinggi cenderung lebih optimis dan percaya bahwa mereka mampu mengontrol kehidupannya.

Bandura (1997) menyebutkan bahwa *self-efficacy* terdiri dari tiga dimensi utama: magnitude (tingkat kesulitan tugas yang diyakini dapat diselesaikan), strength (kekuatan keyakinan dalam menghadapi rintangan), dan generality (cakupan generalisasi keyakinan dalam berbagai situasi). Ketiga dimensi ini penting untuk dipertimbangkan dalam penyusunan alat ukur yang valid dan reliabel. Alat ukur yang baik harus mampu merepresentasikan keyakinan individu dalam menghadapi tugas-tugas dengan berbagai tingkat kesulitan, keteguhan mereka untuk bertahan saat mengalami hambatan, serta sejauh mana keyakinan tersebut dapat diterapkan dalam berbagai konteks kehidupan. Mengingat pentingnya peran *self-efficacy* dalam berbagai domain kehidupan dan dampaknya terhadap keberfungsian individu, maka diperlukan alat ukur yang mampu menangkap konstruk ini secara utuh dan akurat.

Pengembangan instrumen yang didasarkan pada dimensi-dimensi yang telah dijelaskan Bandura tidak hanya memberikan landasan teoritis yang kuat, tetapi juga memungkinkan pengukuran yang spesifik, kontekstual, dan bermakna. Dengan demikian, penelitian ini bertujuan untuk menyusun dan menguji validitas serta reliabilitas alat ukur *self-efficacy* berdasarkan dimensi magnitude, strength, dan generality, agar dapat digunakan sebagai alat asesmen psikologis dalam berbagai konteks, khususnya dalam bidang pendidikan dan pengembangan pribadi.

# METODE PENELITIAN

## Subjek dan Teknik Pengambilan Sampel

Subjek penelitian berjumlah 204 Mahasiswa aktif di seluruh perguruan tinggi di Indonesia Mahasiswa tersebut 35,3% adalah laki-laki dan 64,7% adalah perempuan. Teknik pengambilan sampel menggunakan Convenience sampling yaitu pengambilan sample berdasarkan kemudahan akses

peneliti terhadap responden (Ackoff, 1953; Taherdoost,2016). Pengumpulan data dilakukan melalui penyebaran angket menggunakan *Google From* kepada Mahasiswa dengan kriteria responden adalah mahasiswa aktif di perguruan tinggi di Indonesia, Mahasiswa semester 1-8, Mahasiswa usia 17 tahun keatas, Kuesioner disebarkan melalui media sosial.

**Instrumen Penelitian**

Instrumen yang digunakan adalah skala self-efficacy yang dikembangkan berdasarkan teori Bandura dalam bentuk *Situational Judgement Test* (Bandura, 1997). *Yang* memuat 18 aitem unidimensional efikasi diri. Dapat dipahami melalui model proses respon yang didasari oleh teori Weekley & Ployhart (2013).

Merupakan gambaran suatu situasi yang melalui beberapa tahapan kognitif, yaitu :

memahami situasi, mengidentifikasi tujuan yang relevan, serta menghasilkan respons, yang pada akhirnya akan mengevaluasi efektivitas setiap respon sebelum memilih jawaban. Maka skor SJT mencerminkan kombinasi antara memahami kemampuan situasi, pengalaman sebelumnya, serta pengetahuan implisit terhadap perilaku yang efektif.

Dalam alat ukur terdapat suatu pernyataan dan beberapa pilihan untuk menyelesaikan persoalan tersebut tetapi hanya satu jawaban yang dianggap paling benar (Moss, 1997). Terdapat 3 pilihan jawaban pada setiap pernyataan dengan metode skoring, mendapat poin 1 jika efikasi diri rendah menunjukan keraguan pada kemampuan diri, memiliki keyakinan bahwa tidak mampu mengatasi situasi yang dihadapi, mendapat poin 2 jika respons menunjukkan adanya usaha ketika menghadapi suatu situasi, akan tetapi masih disertai keraguan, ketergantungan pada bantuan orang lain serta keyakinan diri yang belum sepenuhnya kuat, mendapat poin 3 jika respons menunjukan adanya keyakinan diri yang kuat terhadap kemampuan diri, keberanian ketika mengambil tindakan, memiliki inisiatif, serta memiliki rasa optimis terhadap keberhasilan ketika menghadapi suatu situasi.

**Tabel 1. Blue Print Alat Ukur**

| Dimensi | Indikator | Nomor Aitem | Jumlah Aitem |
|---|---|---|---|
| *Self Efficacy* | Mahasiswa meyakini usaha akan berpengaruh pada hasil | 1,2 | 2 |
| | Meyakini pendapat yang disampaikan layak untuk disampaikan | 3,4 | 2 |
| | Meyakini bahwa setiap hasil karena usaha bukan kemampuan | 5,6 | 2 |
| | Meyakini usaha akan menghasilkan apa yang diinginkan | 7 | 1 |
| | Meyakini setiap kegiatan akan memberikan dampak positif | 8,9 | 2 |
| | Mampu menenangkan diri ketika menghadapi situasi yang tidak sesuai harapan | 10,11 | 2 |
| | Percaya diri dalam menghadapi situasi yang biasanya menimbulkan rasa cemas | 12,13 | 2 |
| | Memilih menantang diri sendiri agar lebih berani | 14,15 | 2 |
| | Memilih berinteraksi dengan orang-orang yang selalu positif | 16,17 | 2 |
| | Memilih untuk mencoba hal baru | 18 | 1 |
| | Total | | 18 |

**Prosedur Penelitian**

Prosedur penelitian dilaksanakan melalui tujuh tahapan yaitu (Azwar, 2024) :

a) Identifikasi Tujuan Ukur : Menetapkan teori konstruk yang akan diukur. Teori *Self Efficacy* Albert Bandura (1997) yang mengukur *Self Efficacy* pada mahasiswa.
b) Pembatasan Domain Ukur : Menguraikan konsep teori dengan memfokuskan pada aspek keperilakuan berdasarkan teori *self-Efficacy* yang terdiri dari empat dimensi yaitu *cognitive*, *affective*, *selection*, serta *motivation*. Pemfokusan dilakukan karena aitem yang diukur memungkinkan mengukur hal lain secara bersamaan dan bukan hanya dari segi c*ognitive*,

*affective*, *selection*, serta *motivation* saja. Sehingga alat ukur yang dikembangkan tidak hanya mengukur empat aspek akan tetapi mengukur *Self-efficacy* secara langsung.

c) Operasionalisasi Dimensi : Sebelum melaksanakan proses pengukuran, peneliti terlebih dahulu mengkaji teori *self-efficacy* dari Albert Bandura sebagai sebagai dasar konseptual dalam pengembangan alat ukur. berdasarkan teori tersebut, dimensi yang diukur yang dapat dijadikan menjadi indikator keperilakuan dalam terdiri dari: 1) *Cognitive* (kognitif) perilaku yang menunjukan keyakinan dan melibatkan proses berpikir pada seseorang, 2) *Affective* (afektif) perasaan pada saat menghadapi suatu situasi, 3) *Motivation* (motivasi) motivasi seseorang dalam mencapai suatu hal dan *Selection* (seleksi) perilaku individu dengan menyeleksi pilihan dalam pengambilan keputusan, namun struktur alat ukur yang semula terdiri dari 4 aspek difokuskan menjadi 1 agar model sesuai dan memenuhi kriteria pengembangan alat ukur *self-efficacy* ini menggunakan metode *Situational Judgement Test* (SJT), Uji coba ini bertujuan untuk mengukur dimensi *self-efficacy* dalam konteks situasional akademik.

d) Penyusunan Kisi-Kisi : Peneliti memetakan dimensi *self-efficacy* yang telah dipisahkan menjadi 4 aspek keperilakuan lalu dituangkan dalam kisi-kisi (*blueprint*) yang menghasilkan 10 indikator.

e) Validasi Isi Indikator : Indikator tercantum melalui proses tahapan penilaian *expert judgement* yang dilakukan oleh ahli. Penilaian dari para ahli digunakan untuk menentukan apakah aitem telah merepresentasikan setiap indikator konstruk *Self Efficacy*. Item yang kurang sesuai direvisi sesuai dengan saran ahli. setelah itu hasil dari penilaian para ahli dilakukan analisis menggunakan koefisien Aiken's V untuk memperoleh validitas isi.

f) Penulisan Item: Penulisan item disesuaikan dengan indikator keperilakuan berdasarkan hasil validasi yang menghasilkan 19 item. Dilakukan uji keterbacaan oleh tim peneliti setelah penulisan item dilakukan dengan tujuan memperbaiki redaksi item yang dapat menimbulkan interpretasi ganda dan merubah redaksi item yang sulit untuk dipahami.

g) Validasi Isi Item : Menyesuaikan kembali item dengan dimensi indikator keperilakuan dalam *self-efficacy*. Penskalaan dilakukan dengan pemberian bobot skor pada item yang memiliki 3 alternatif pilihan jawaban di setiap item, pemberian skor menggunakan skor yang berdasar pada kualitas response karakteristik skala *Situational Judgement Test*. Hasilnya disusun menjadi *field test* sebagai data yang laik diujicobakan kepada subjek partisipan.

h) Uji Coba Instrumen : *Situational Judgement Test Self Efficacy* diujicobakan kepada 100 orang mahasiswa aktif dari berbagai universitas dan program studi. Uji coba bertujuan untuk memperoleh data empiris sebagai dasar analisis psikometri instrumen

i) Uji Coba Instrumen : *Situational Judgement Test Self Efficacy* diujicobakan kepada 100 orang mahasiswa aktif dari berbagai universitas dan program studi. Uji coba bertujuan untuk memperoleh data empiris sebagai dasar analisis psikometri instrumen

j) berbasis *Situational Judgement Test* memungkinkan satu butir merepresentasikan lebih dari satu aspek secara bersamaan sehingga pemisahan aspek tidak muncul secara jelas. Oleh karena itu, peneliti memfokuskan pengukuran *self-efficacy* sebagai satu konstruk menyeluruh dan mempertahankan 18 item setelah satu item dieliminasi. Berdasarkan penetapan konstruk tersebut, analisis kemudian dialihkan ke *Confirmatory Factor Analysis* untuk menguji kesesuaian model satu faktor. analisis statistik dilakukan di JAMOVI didasarkan pada kemampuan analisis dan pengolahan data yang lebih sesuai dengan karakteristik *Situational Judgement Test*.

k) Estimasi reliabilitas Rulon = 0,52 dan Feld = 0,54. Serta cronbach alpha 0,65

l) Validasi Konstrak dilakukan dengan CFA *(Confirmatory factor analysis*) hasil menunjukan 0,01

m) Penyusunan Instrumen Final : Berdasarkan hasil analisis diperoleh 18 aitem valid dari total 19 aitem awal. Aitem-aitem tersebut disusun menjadi *Situational Judgement Test* final yang digunakan dalam pengumpulan data penelitian utama.

## HASIL DAN PEMBAHASAN

Penelitian ini melibatkan 204 Mahasiswa aktif di seluruh Indonesia sebagai responden 35,3% laki-laki dan 64,7% perempuan mulai dari semester 1-8 dengan usia diatas 17 tahun. Mayoritas partisipan berdomisili di wilayah Jawa Barat, serta partisipan memiliki kriteria yang sesuai untuk menjadi responden dan partisipan dengan sukarela untuk berpartisipasi. Penelitian ini bertujuan untuk mengembangkan alat ukur efikasi diri. pada Mahasiswa yang dikembangkan melalui teori Albert Bandura, yang mencangkup 4 dimensi yaitu kognitif, afektif, motivasi dan seleksi (Albert Bandura 1997).

Uji validitas isi dilakukan pada alat ukur ini untuk mengetahui bahwa butir-butir dalam alat ukur telah mempresentasikan konstruk yang akan diukur. Penilaian dilakukan oleh dua magister psikologi dan satu sarjana psikologi yang expert di bidang psikologi menggunakan skala penilaian 1 tidak sesuai hingga 5 sangat sesuai mewakili konstruk yang akan diukur. Analisis validitas isi dihitung menggunakan indeks Aiken's V.

Berdasarkan dari hasil perhitungan, nilai Aiken's V pada seluruh butir berkisar antara 0,83 hingga 1,00 . Menurut kriteria Aiken, nilai V ≥ 0,80 menunjukan bahwa butir memiliki validitas yang baik. Pada penelitian ini seluruh butir memiliki nilai ≥ 0,80 sehingga dinyatakan valid secara isi. Hasil menunjukan bahwa setiap butir telah representatif mengukur konstruk efikasi diri pada mahasiswa, sehingga telah memenuhi persyaratan untuk melanjutkan tahap uji konstruk.

Uji validitas konstruk dilakukan menggunakan CFA (*Confirmatory Factor Analysis*) hal ini dilakukan untuk menguji kesesuaian model teoritis dengan data empiris. Analisis dilakukan menggunakan JAMOVI. Berdasarkan hasil analisis, diperoleh nilai indeks kelayakan sebagai model berikut Chi-Square 203, p- value 0,01, CFI 0,797, TLI 0,770, RMSEA 0,0496 dan RMSR 0,0496. Hasil ini menunjukan model kecocokan parsial dengan beberapa indeks fit yang belum memenuhi kriteria ideal.

Dilakukan uji reliabilitas untuk mengetahui konsistensi internal instrumen. Pada penelitian ini digunakan tiga pendekatan, yaitu Cronbach's Alpha, Rulon, dan Feldt. Hasil analisis menunjukan nilai Cronbach Alpha sebesar 0.659 menurut kriteria umum, nilai Alpha ≥ 0,70 menunjukkan reliabilitas yang kurang baik. Maka dengan demikian instrumen ini kurang dalam memiliki tingkat konsistensi internal yang baik. Hasil perhitungan reliabilitas menggunakan metode Rulon diperoleh koefisien sebesar 0,519 yang menunjukkan reliabilitas sedang menuju rendah, sehingga instrumen ini memiliki konsistensi internal yang belum optimal sehingga hasil pengukuran belum dapat dikatakan sepenuhnya kuat. Sementara itu nilai, koefisien feldt 0,541 yang sedikit lebih tinggi dari rulon menunjukkan bahwa estimasi reliabilitas yang relatif serupa yaitu berada pada kategori sedang menuju rendah, sehingga hal ini menunjukkan bahwa instrumen masih memerlukan evaluasi lebih lanjut, seperti revisi penambahan atau eliminasi jumlah butir item yang lebih representatif terhadap konstruk. Upaya ini menjadi sangat penting dilakukan untuk dapat meningkatkan konsistensi internal instrumen pada penelitian selanjutnya.

**Uji Reliabilitas**

Uji reliabilitas dilakukan dengan pendekatan menggunakan formula *Alpha Cronbach, Rulon,* dan *Feldt.* Uji reliabilitas menunjukkan nilai *Alpha Cronbach* sebesar 0,65. Menurut Anggraini et al., (2022) jika nilai *Alpha Cronbach* menunjukkan nilai > 0,60 variabel tersebut dapat dikategorikan reliabel. Koefisien *Alpha Cronbach* dari penelitian adalah 0,65 dimana instrumen ini tetap dianggap reliabel karena melebihi 0,60.

**Tabel. 2 *Frequentist Individual Item Reliability Statistics***

| Item | Item-rest Correlation |
|---|---|
| C1 | 0.639 |
| C2 | 0.650 |
| C3 | 0.629 |
| C4 | 0.634 |
| C5 | 0.663 |
| C6 | 0.669 |
| M1 | 0.618 |

| | |
|---|---|
| M2 | 0.628 |
| M3 | 0.625 |
| A1 | 0.635 |
| A2 | 0.669 |
| A3 | 0.669 |
| A4 | 0.638 |
| A5 | 0.637 |
| S1 | 0.612 |
| S2 | 0.671 |
| S3 | 0.685 |
| S5 | 0.634 |

Berdasarkan analisis item-rest corelation, item ini dapat dinyatakan reliabel, dikarenakan nilai item >0,3 sehingga seluruh item dapat dinyatakan reliabel karena tidak ada item yang dibawah 0,3.

**Hasil Uji Rulon**

Hasil Uji Rulon Uji Rulon menggunakan formula :

$$rxx' = 1 - \frac{S^2 d}{S^2 d}$$

Diketahui :

$S^2$ d =9.975

$S^2$ x =20.734

Maka :

$$rxx' = 1 - \frac{9.975}{20.734} = 0{,}519$$

Hasil pengujian felt Rulon mendapatkan hasil 0,519 dengan pembelahan item ganjil genap

**Hasil Uji Feldt**

Hasil Uji Rulon Uji Feldt menggunakan formula :

$$\frac{4(Sy1y2)}{\left[S2\,x - \left\{\frac{S2\,y1 - S2\,y2}{S_x}\right\}\right]\ 2}$$

Diketahui :

Sy1y2 = 2,690

$S^2_{y1}$ = 9,764

$S^2_{y2}$ =5,590

$S^2 x$ =20,734

$S_x$ = 4,553

Maka :

$$\frac{4(2{,}690)}{\left[20{,}734 - \left\{\frac{9{,}764 - 5{,}590}{4{,}553}\right\}\right]} = 0{,}541$$

Hasil pengujian Rulon mendapatkan hasil 0,519 dengan pembelahan item ganjil genap.

**Tabel 3. Hasil Statistik Reliabilitas**

| Scale Reliability | Statistics |
|---|---|
| Coefficient | Estimate |
| Coefficient α | 0.659 |
| Rulon | 0,519 |
| Feldt | 0,541 |

Hasil uji reliabilitas menunjukkan bahwa setiap instrumen telah memiliki konsistensi internal yang cukup baik melalui pendekatan *frequentist*. Ketika dilakukan evaluasi setiap item, nilai koefisien α

apabila item dihapus (α if item dropped) berada pada kisaran 0.612 sampai 0.669 seluruh nilai tersebut lebih rendah atau sangat mendekati nilai total α yaitu sebesar 0.659. pada temuan ini mengindika sikan bahwa tidak terdapat item yang secara signifikan yang dapat menurunkan reliabilitas keseluruhan data.

Pada nilai item-rest correlation yang berada pada rentang 0.612 hingga 0.685 sehingga hal ini memperlihatkan bahwa semua item telah memiliki korelasi positif dengan skor total konstruk yang diukur. Sehingga rentang nilai ini memberikan informasi bahwa setiap item berkorelasi seimbang terhadap konsistensi internal, dengan adanya item yang berkorelasi tinggi seperti (S3) 0.685, hal ini menunjukan kontribusi yang lebih kuat dalam merepresentasikan konstruk yang diukur, sedangkan item (S1) 0.612 tetap tergolong baik karena masih diatas batas minimal yang umumnya direkomendasikan dalam analisis reliabilitas. Sehingga dengan demikian tidak ada item yang menunjukan korelasi bermasalah ataupun rendah sehingga seluruh aitem belum layak untuk dipertahankan.

Pengujian reliabilitas menggunakan metode *split-half* dilakukan melalui uji Rulon sehingga menghasilkan nilai reliabilitas sebesar 0,519 hal ini didapat dari pembelahan item ganjil – genap. Pada nilai Rulon masih diambang batas minimum yang sering digunakan yaitu 0,70 sehingga hal ini mengindikasikan bahwa belahan tes tersebut belum optimal dan memiliki kestabilan yang belum cukup kuat. Selanjutnya pada perhitungan Feldt menghasilkan nilai sebesar 0,541 hal ini lebih tinggi dari sebelumnya akan tetapi nilai tersebut tetap masih berada dalam kategori reliabilitas yang belum dalam standar ideal. Sehingga perbedaan Rulon dan Feldt yang relatif kecil menunjukan bahwa estimasi reliabilitas cenderung konsisten, akan tetapi masih harus ditingkatkan dalam konsistensi internal instrumen secara keseluruhan. Secara umum, hasil pada seluruh instrumen telah menunjukan reliabilitas tingkat sedang, serta belum cukup kuat untuk dinyatakan stabil.

**Uji Validitas Konstruk *(Confirmatory Factor Analysis)***

Setelah itu dilakukan *Confirmatory Factor Analysis (CFA)* hal ini dilakukan untuk dapat menguji kesesuaian indikator-indikator yang telah dibuat serta disusun beanr-benar merepresentasikan faktor yang telah ditetapkan berdasarkan teori yang telah ditentukan.

**Tabel 4. Hasil *Confirmatory Factor Analysis***

| Model | Estimasi | Keterangan |
|---|---|---|
| Chi'square | 203 | Tidak Fit |
| P | ,001 | Tidak Fit |
| TLI | 0,770 | Tidak Fit |
| CFI | 0,797 | Tidak Fit |
| RMSEA | 0,0496 | Fit |
| SRMR | 0,0496 | Fit |

Hasil *Confirmatory Factor Analysis* (CFA) menunjukan bahwa kesesuaian indikator-indikator yang telah dibuat serta disusun belum benar-benar merepresentasikan faktor yang telah ditetapkan berdasarkan teori Bandura.

Uji chi-Square pada faktor model menghasilkan nilai $X^2$ = 203 dengan nilai p=0,001. Secara statistik, nilai <0,05 menunjukan adanya perbedaan yang signifikan antara model teoritis dengan data empiris. Akan tetapi, uji chi-square diketahui sangat sensitif terhadap ukuran sampel, sehingga pada jumlah sampel yang banyak model yang memadai dapat tetap menunjukan hasil signifikan (Kline, 2016).

Nilai *Tucker-Lewis Index* (TLI) sebesar 0,770 dan *Comparative Fit Index* (CFI) sebesar 0,797 hal ini menunjukan bahwa model belum mencapai kriteria *good fit*. Secara umum nilai CFI dan TLI ≥0,90 menunjukan nilai yang memadai, sedangkan nilai ≥95 menunjukan model yang sangat baik (Hu&Bentler, 1999; Hair et al., 2022). Sehingga berdasarkan kedua indeks ini, model belum berada pada kategori fit yang memerlukan modifikasi lebih lanjut. Selanjutnya, nilai *Root Mean Square Error of Approximation* (RMSEA) sebesar 0,0496 berada di bawah batas 0,05 sehingga menunjukan kategori *close fit*. Nilai RMSEA ≤0,05 telah mengindikasikan kesesuaian model yang sangat baik, sedangkan nilai di rentang 0,05-0,08 masih dapat diterima. Maka dengan demikian berdasarkan RMSEA, model menunjukan tingkat kecocokan yang sangat baik dengan data empiris.

Nilai *Standardized Root Mean Square Residual* (SRMR) sebesar 0,0634 berada di bawah ambang batas ≤ 0,08 sehingga telah menunjukan residual antara matriks kovarians yang diobservasi dan yang diprediksi relatif kecil, maka dari itu model masih dapat diterima (Hu & Bentler, 1999). Secara keseluruhan meskipun pada nilai CFI dan TLI masih diambang batas kriteria ideal, akan tetapi indeks RMSEA dan SRMR telah menunjukan nilai kesesuaian yang cukup baik. model menunjukkan kecocokan parsial dengan beberapa indeks fit yang belum memenuhi kriteria ideal. Sehingga memerlukan perbaikan pada butir aitem, agar dapat memenuhi kriteria ideal

**Uji Validitas Konten *(Aiken's V)***

Validitas isi adalah validitas yang dinilai dengan cara menguji kesesuaian serta kelayakan isi tes dengan melalui analisis rasional atau expert *judgement*. Validitas isi ini memastikan bahwa alat ukur *Situational Judgement Test* untuk mengukur efikasi diri pada mahasiswa memiliki item yang memadai dan representatif mengungkapkan konsep tersebut. Semakin item mencerminkan area ataupun konsep yang diukur, maka semakin valid. Aiken (1985) memformulasikan Aiken's V untuk menghitung koefisien validitas isi berdasarkan hasil penilaian panitia ahli, sekurang kurangnya n per orang sebagai proporsi sejauh ana item mewakili konstruk yang diukur.

Indeks validitas item dari alat ukur *Situational Judgement Test* efikasi diri pada mahasiswa yang dinilai oleh 3 rater untuk menguji kesesuaian butir soal dengan indikator dari efikasi diri pada mahasiswa menggunakan kartu CVI. Berdasarkan kriteria validitas AIKEN item yang layak digunakan berada di rentang 0,80 sampai 1,00. Maka seluruh item bernilai valid.

*Situational Judgement Test* efikasi diri 18 item ini menunjukan nilai minimal validitasnya 0,80 sampai dengan 1,00. Validitas alat ukur tergolong sangat tinggi (Guillford 1956). Hal ini dapat dikatakan bahwa pertanyaan *Situational Judgement Test* efikasi diri pada mahasiswa memiliki tingkat validitas yang tinggi. Alat ukur yang valid adalah alat ukur yang memiliki ketepatan serta kecermatan alat ukur dalam melakukan fungsinya, yaitu mengukur apa yang seharusnya diukur (Azwar, 2021)

**Tabel 5. Hasil Penilaian Expert Judgement**

| utir Aitem | Ekspert Judgement | | |
|---|---|---|---|
| | *ndeks V* | riteria | Kesimpulan |
| . | ,83 | 0,80 | Valid |
| . | ,00 | 0,80 | Valid |
| . | ,00 | 0,80 | Valid |
| . | ,00 | 0,80 | Valid |
| . | ,83 | 0,80 | Valid |
| . | ,00 | 0,80 | Valid |
| . | ,00 | 0,80 | Valid |
| . | ,00 | 0,80 | Valid |
| . | ,00 | 0,80 | Valid |
| 0. | ,00 | 0,80 | Valid |
| 1. | ,83 | 0,80 | Valid |

| | | | |
|---|---|---|---|
| 2. | ,83 | 0 ,80 | Valid |
| 3. | ,00 | 0 ,80 | Valid |
| 4. | ,00 | 0 ,80 | Valid |
| 5. | ,92 | 0 ,80 | Valid |
| 6. | ,00 | 0 ,80 | Valid |
| 7. | ,00 | 0 ,80 | Valid |
| 8. | ,00 | 0 ,80 | Valid |

**Tabel 6. *Situational Judgement Test Self Efficacy***

| Aspek | Aitem | ode |
|---|---|---|
| *Self efficacy* | 1. Anda belajar dengan bersungguh-sungguh terlebih dahulu sebelum akan melaksanakan ujian, ketika ujian akan dilaksanakan anda merasa ragu dalam menjawab soal ujian dengan benar, hanya satu hingga dua soal yang anda yakini jawabannya dengan benar | 1 |
| | 2. Ketika akan presentasi anda menyiapkan beberapa hari sebelumnya dan anda telah yakin bahwa dapat menyampaikan dengan jelas setiap materi yang akan disampaikan, tetapi pada saat anda menyampaikan materi di depan kelas anda merasa gugup dan lupa ketika menyampaikannya | 2 |
| | 3. Anda sedang mengikuti perkuliahan, keadaan kelas sedang ramai dikarenakan dosen membuka sesi diskusi dengan meminta mahasiswa untuk memberikan contoh kasus dari materi yang sebelumnya sudah disampaikan. Anda merasa contoh yang akan disampaikan itu relevan | 3 |
| | 4. Ketika anda telah selesai presentasi dan menutup presentasi dengan sesi diskusi, teman-teman di kelas anda banyak yang ingin bertanya kepada anda karena tertarik dengan pembahasan yang anda sampaikan | 4 |
| | 5. Hari ini dosen anda mengumumkan akan membagikan hasil kuis di kelas. Anda diinstruksikan untuk menghampiri dosen anda yang memiliki lembar hasil kuis. Anda mendapati nilai kuis anda dibawah rata-rata, padahal anda merasa bahwa anda sudah belajar dengan sungguh-sungguh | 5 |
| | 6. Hasil dari ujian telah dibagikan dan anda merasa sudah cukup dalam belajar, sebelum ujian ketika lembar hasil ujian sudah anda lihat ternyata hasil ujian rendah | 6 |
| | 7. Anda sedang dalam perkuliahan dan ingin menyampaikan pendapat karena termotivasi melihat lingkungan kelas sedang ramai dan semuanya ikut aktif diskusi kita ingin ikut serta dalam menyampaikan pendapat | 1 |
| | 8. Ada rekrutmen untuk mengikuti organisasi dan anda ingin mencoba untuk mengikuti organisasi dan teman-teman anda pun mengikuti organisasi | 2 |
| | 9. ketika mengikuti seminar yang diadakan oleh kampus dapat memberikan dampak positif bagi mahasiswa yang mengikutinya | 3 |
| | 10. Ketika kamu sedang melakukan presentasi di depan kelas, kamu sudah mempersiapkan materi nya dengan matang namun teman -teman di | 1 |

| | |
|---|---|
| kelas terlihat sibuk sendiri dan tidak menyimak saat kamu presentasi | |
| 11. Ketika anda presentasi tetapi terjadi kegugupan selama presentasi sehingga presentasi tidak berjalan sesuai dengan rencana yang telah dipelajari | 2 |
| 12. Ketika anda sedang berada di kelas terlihat tidak memperhatikan , tiba-tiba dosen memintamu menjelaskan ulang materi yang baru saja disampaikan | 3 |
| 13. Rapat organisasi meminta anda menjelaskan kepada anggota lain untuk menyampaikan hasil yang sudah dibahas dengan divisi masing masing | 4 |
| 14. Ketika anda dihadapkan dengan anggota kelompok yang tidak berkontribusi dalam mengerjakan tugas yang telah dibagikan anda merasa marah karena anda harus mengambil alih tugas teman anda | 5 |
| 15. Ketika di kelas dosen menawarkan kesempatan untuk mempresentasikan materi lebih awal namun Sebagian teman anda ragu tetapi anda merasa bahwa ini melatih keberanian | 1 |
| 16. Ketika anda berada di posisi mengerjakan skripsi dan 3 hari lagi akan ada bimbingan bersama dosen, teman-teman anda mengajak anda untuk berkemah 2 hari satu malam | 2 |
| 17. Anda memiliki kesempatan untuk memilih sendiri teman untuk tugas proyek kelompok. Beberapa teman anda memiliki gaya belajar yang santai, mengikuti alur ada juga teman yang aktif, suka memberi masukan, dan mendukung kerja sama tim | 3 |
| 18. Anda ingin mencoba hal baru ketika menjadi mahasiswa yaitu mengikuti kegiatan volunteer yang diadakan oleh kampus tetapi tidak ada teman yang kenal untuk melakukan kegiatan volunteer | 5 |

## PEMBAHASAN

Hasil pengujian menunjukkan bahwa alat ukur efikasi diri yang dikembangkan menggunakan metode *Situational Judgement Test* memiliki beberapa keterbatasan psikometrik yang perlu diperbaiki. Nilai koefisien Alpha Cronbach sebesar 0,659, meskipun dapat diterima menurut beberapa literatur (Anggraini et al., 2022). Masih berada di bawah standar ideal 0,70 yang umumnya direkomendasikan untuk instrumen psikologi yang siap digunakan dalam penelitian atau praktik ( Nunnally & Bernstein, 1994). Hal ini mengindikasikan bahwa konsistensi internal instrumen belum optimal dan memerlukan perbaikan, baik melalui penambahan item maupun revisi kualitas item yang ada.

Hasil *Confirmatory Factor Analysis* (CFA) juga menunjukkan hasil yang marginal. Nilai CFI sebesar 0,797 dan TLI sebesar 0,770 berada di bawah nilai ideal ≥0,90 yang direkomendasikan untuk model yang baik (Hu & Bentler, 1999). Meskipun nilai RMSEA sebesar 0,0496 dan SRMR sebesar 0,0634 menunjukkan hasil yang dapat diterima (Hair et al., 2022), rendahnya nilai CFI dan TLI mengindikasikan bahwa struktur faktor yang diusulkan belum sepenuhnya didukung oleh data empiris. Hal ini menunjukkan adanya misfit antara model teoritis dengan data yang diperoleh, yang dapat disebabkan oleh kompleksitas item SJT yang mengukur multiple konstruk secara simultan, atau kemungkinan bahwa struktur satu faktor belum sepenuhnya merepresentasikan kompleksitas konstruk efikasi diri.

Lebih lanjut, nilai reliabilitas split-half melalui metode Rulon (0,519) dan Feldt (0,541) menunjukkan hasil yang belum sesuai, berada dibawah standar minimum yaitu 0,70 (Nunnally & Bernstein, 1994). Rendahnya nilai ini mengindikasikan bahwa kedua belahan tes belum konsisten dalam mengukur konstruk yang sama, sehingga alat ukur ini dapat dikembangkan kembali agar lebih optimal. Hasil ini memperkuat kebutuhan mendesak untuk melakukan revisi substansial terhadap instrumen, termasuk kemungkinan penambahan item untuk meningkatkan cakupan konstruk dan konsistensi pengukuran.

Perbandingan dengan penelitian terdahulu menunjukkan posisi instrumen ini di antara hasil yang beragam. Beberapa penelitian pengembangan instrumen efikasi diri di Indonesia juga mengalami tantangan serupa dalam mencapai standar psikometrik yang ideal. Penelitian Pratiwi et al. (2023) yang

mengembangkan skala efikasi diri pada mahasiswa menggunakan teori Bandura (1997) dengan 42 item menghasilkan hanya 32 item yang memenuhi kriteria kelayakan berdasarkan indeks daya beda, dengan koefisien reliabilitas Cronbach's Alpha sebesar 0,55 yang berada di bawah standar yang direkomendasikan.

Hal ini menunjukkan bahwa pengukuran efikasi diri dalam konteks Indonesia memang menantang dan memerlukan perbaikan dalam penyusunan indikator dan item. Sementara itu, penelitian Solikhin (2020) yang mengembangkan alat ukur efikasi diri siswa dalam pembelajaran kimia menggunakan model RASCH dengan 22 butir pertanyaan menghasilkan reliabilitas sebesar 0,73, menunjukkan bahwa dengan metode dan konteks yang tepat, pengembangan instrumen efikasi diri yang memadai adalah mungkin. Hal ini menegaskan bahwa hasil marginal dalam penelitian ini bukan sesuatu yang tidak dapat diperbaiki, namun memerlukan upaya pengembangan lebih lanjut yang sistematis. Penelitian juga menunjukkan bahwa mahasiswa yang percaya diri dalam kemampuannya untuk belajar dan menyelesaikan tugas akan menunjukkan motivasi intrinsik yang lebih tinggi, keterlibatan dalam pembelajaran, serta ketekunan dalam menghadapi kesulitan (Schunk & Pajares, 2002), yang menegaskan pentingnya pengembangan instrumen efikasi diri yang valid dan reliabel.

Dari sisi teoretis, hasil penelitian ini memberikan pembelajaran penting tentang kompleksitas pengukuran efikasi diri menggunakan metode SJT. Hasil yang marginal menunjukkan bahwa asumsi awal tentang struktur konstruks perlu disusun kembali. Dimensi kognitif, afektif, motivasi, dan seleksi yang semula direncanakan mungkin memerlukan operasionalisasi yang lebih jelas dan terpisah, atau sebaliknya, konseptualisasi sebagai satu konstruk perlu didukung dengan landasan teoritis yang lebih kuat dari awal pengembangan (Taherdoost, 2016). Proses pengembangan instrumen memerlukan pendekatan yang sistematis dan iteratif untuk menghasilkan alat ukur yang memenuhi standar psikometrik (Azwar, 2024). Individu dengan efikasi diri tinggi umumnya menunjukkan kepercayaan diri yang kuat, ketahanan terhadap stres, kemauan untuk mencoba hal baru, serta kecenderungan untuk menyelesaikan tugas-tugas yang menantang dengan strategi adaptif dan pemikiran positif sehingga pengembangan alat ukur yang mampu mengidentifikasi tingkat efikasi diri dengan akurat menjadi sangat penting. Meskipun demikian, penelitian ini tetap memberikan kontribusi dalam eksplorasi penggunaan metode SJT untuk mengukur efikasi diri, yang masih jarang dilakukan. Instrumen ini dapat menjadi prototipe untuk pengembangan lebih lanjut, namun dalam kondisi saat ini belum siap digunakan untuk keperluan asesmen atau intervensi tanpa perbaikan substansial. Revisi yang diperlukan mencakup penambahan item untuk meningkatkan reliabilitas, perbaikan kualitas item untuk meningkatkan daya beda, pertimbangan ulang struktur faktor, dan pengujian pada sampel yang lebih besar dan beragam.

Penelitian ini memiliki beberapa keterbatasan serius yang perlu diakui. Pertama, penggunaan convenience sampling membatasi generalisabilitas hasil (Taherdoost, 2016). Kedua, dominasi responden dari Jawa Barat mengurangi keragaman geografis dan sosial-budaya. Ketiga, ukuran sampel 204 responden, meskipun memadai untuk analisis dasar, mungkin belum optimal untuk CFA dengan 18 item (Kline, 2015). Keempat, belum dilakukannya uji test-retest membuat stabilitas temporal instrumen tidak diketahui (Azwar, 2024). Kelima, dan yang paling kritis, adalah hasil psikometrik yang marginal menunjukkan bahwa instrumen ini masih dalam tahap pengembangan awal dan memerlukan iterasi lebih lanjut sebelum dapat digunakan secara luas.

Penelitian selanjutnya sangat disarankan untuk melakukan revisi menyeluruh terhadap instrumen, termasuk penambahan item, perbaikan kualitas item, dan pengujian ulang pada sampel yang lebih besar (minimal 300-400 responden untuk CFA yang lebih stabil) (Kline, 2015). Uji test-retest perlu dilakukan untuk memastikan stabilitas temporal. Eksplorasi struktur faktor alternatif juga perlu dipertimbangkan, termasuk kemungkinan model multidimensional yang lebih kompleks. Validitas konkuren dan prediktif dengan variabel eksternal seperti prestasi akademik juga perlu diuji untuk memberikan bukti validitas yang lebih komprehensif. Secara keseluruhan, penelitian ini menghadirkan pembelajaran berharga dalam pengembangan instrumen efikasi diri berbasis metode SJT. Meskipun properti psikometrik yang dihasilkan belum mencapai standar ideal, proses pengembangan ini memberikan pemahaman mendalam tentang kompleksitas pengukuran efikasi diri dalam konteks situasional. Dengan perbaikan yang sistematis berdasarkan temuan penelitian ini, instrumen memiliki potensi untuk dikembangkan menjadi alat ukur yang valid dan reliabel.

## KESIMPULAN

Instrumen efikasi diri berbasis *Situational Judgement Test* (SJT) ini pada awalnya terdiri dari 19 aitem, kemudian setelah melalui proses validasi isi, uji coba, dan analisis psikometrik,satu aitem dieliminasi sehingga diperoleh 18 aitem dalam bentuk final. Validitas isi menunjukkan indeks Aiken's V berada pada rentang 0,83–1,00 yang menandakan representativitas konten yang baik. Namun, hasil validitas konstruk melalui CFA menunjukkan kecocokan model yang masih parsial (CFI = 0,797; TLI = 0,770; RMSEA = 0,0496; SRMR = 0,0496). Dari sisi reliabilitas, nilai Cronbach's Alpha sebesar 0,659 serta koefisien Rulon (0,519) dan Feldt (0,541) menunjukkan konsistensi internal berada pada kategori sedang dan belum mencapai standar ideal ≥ 0,70.

Dari aspek reliabilitas, nilai Cronbach's Alpha sebesar 0,659 menunjukkan konsistensi internal pada kategori sedang, sementara estimasi split-half melalui Rulon (0,519) dan Feldt (0,541) juga berada pada tingkat sedang dan belum mencapai standar ideal ≥ 0,70. Secara keseluruhan, instrumen ini menunjukkan karakteristik psikometrik awal yang cukup menjanjikan sebagai prototipe alat ukur efikasi diri berbasis situasional, tetapi masih memerlukan pengembangan lanjutan, khususnya dalam peningkatan konsistensi internal dan pemantapan struktur konstruk sebelum dapat digunakan secara luas untuk keperluan asesmen maupun penelitian skala besar.

## REFERENCES


Anggraini, F. D. P., Aprianti, & Setyawati, V. A. V. (2022). Pembelajaran Statistika Menggunakan Software SPSS untuk Uji Validitas dan Reliabilitas. Jurnal BASICEDU, 6(4), 6491-6504. https://doi.org/10.31004/basicedu.v6i4.3206

Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. *Educational and psychological measurement,* 40(4), 955-959.

Azwar, S. (2021). Skala Psikologi. Yogyakarta: Pustaka

Azwar, S. (2024). Penyusunan skala psikologi (Edisi III). Pustaka Pelajar.

Bandura. A. (1977). *Self-efficacy: Toward a unifying theory of behavioral change.* Psychological Review, 84, 191-215

Garmezy, N. (1991). Resilience and vulnerability to adverse developmental outcomes associated with poverty. *American Behavioral Scientist,* 34, 416–430

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2022). *Multivariate data analysis* (8th ed.). Cengage.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jum C. Nunnally, I. H. B. (1994). *Psychometric Theory.*

Kementerian Kesehatan Republik Indonesia. (2022). *Indonesia National Adolescent Mental Health Survey (I-NAMHS): Laporan nasional.*

Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi. (2023). *Statistik pendidikan tinggi tahun 2022.*

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.

Moss, F.A. (1926). Do you know how to get along with people? *Scientific American*, July 192

Musyafirah, T. D., Tira, U. A., Rosalia, V., Fajriyah, Z. A., & Nurmalitasari, F. (2024). *Validitas Konstruk Alat Ukur College Academic Self-efficacy Scale pada Mahasiswa.*

Norman Garmezy, P. (1991). Resilience in Children's Adaptation to Negative Life Events and Stressed Environments. *Pediatric Annals, 20*(9), 459–466.

Pratiwi, R., Sari, Y. R., Rahmah, N., Renita, E. P., Husardi, R. W., Fitriani, R. A., Khotimah, D. M., Adila, W. S., & Fauzan, D. A. (2024). *Pengembangan Skala Efikasi Diri Pada Mahasiswa. 2.*

Salim, F., & Fakhrurrozi, M. M. (2020). Efikasi diri akademik dan resiliensi pada mahasiswa. *Jurnal Psikologi*, *16*(2), 175-187.

Santrock, J. (2008). Physical development and biological aging. *A Topical Approach to Life-Span Development. New York: McGraw-Hill Companies, Inc*, 129-132.

Schunk, D. H., & Pajares, F. (2002). The development of academic self-efficacy. In *Development of achievement motivation* (pp. 15-31). Academic Press.

Solikhin, F. (2020). Pengembangan Alat Ukur Tingkat Efikasi Diri Siswa dalam Pembelajaran Kimia. *Jurnal Pengukuran Psikologi dan Pendidikan Indonesia (JP3I)*, *9*(1), 11–18. https://doi.org/10.15408/jp3i.v9i1.14491

Taherdoost, H. (2016). Sampling methods in research methodology: How to choose a sampling technique for research. *International Journal of Academic Research in Management*, *5*, 18–27.

Warner, L. M., & Schwarzer, R. (2017). Self-efficacy. *The Wiley encyclopedia of health psychology*, 2.

Weekley, J. A., & Ployhart, R. E. (Eds. ). (2006). (2013). *Situational Judgement Tests: Theory, measurement, and application. Mahwah, NJ: Lawrence Erlbaum Associates.* https://doi.org/doi.org/10.4324/9781410617903

Weekley, J. A., & Ployhart, R. E. (Eds.). (2013). *Situational Judgement Tests: Theory, measurement, and application*. Routledge.