

## Artikel

# Beritahu Kami Apa yang Sebenarnya Anda Pikirkan: Analisis Protokol Think Aloud dari Tes Refleksi Kognitif Verbal

Nick Byrd <sup>1,\*</sup> , Brianna Joseph <sup>2</sup>, Gabriela Gongora <sup>3</sup> dan Miroslav Sirota <sup>4</sup><sup>1</sup> Program Beasiswa Penelitian Pascadoktoral Komunitas Intelijen, Stevens Institute of Technology, Hoboken, NJ 07030, AS<sup>2</sup> AIOps, IBM, Atlanta, GA 30319, Amerika Serikat<sup>3</sup> College of Business, Florida State University, Tallahassee, FL 32306, AS<sup>4</sup> Departemen Psikologi, Universitas Essex, Colchester CO4 3SQ, Inggris; [msirota@essex.ac.uk](mailto:msirota@essex.ac.uk)\* Korespondensi: [nick.a.byrd@gmail.com](mailto:nick.a.byrd@gmail.com)

**Abstrak:** Interpretasi standar dari tes refleksi kognitif mengasumsikan bahwa respons yang benar adalah reflektif dan respons yang terpancing adalah tidak reflektif. Namun, penelusuran proses tes refleksi *matematis* sebelumnya telah meragukan interpretasi ini. Dalam dua penelitian (N = 201), kami menggunakan protokol think-aloud yang telah divalidasi secara langsung dan online untuk menguji bagaimana asumsi ini dipenuhi oleh Tes Refleksi Kognitif Verbal (vCRT) yang baru, tervalidasi, kurang familiar, dan non-matematis. Pikiran yang diverbalisasikan dalam kedua penelitian mengungkapkan bahwa sebagian besar (tapi tidak semua) respons yang benar melibatkan refleksi dan sebagian besar (tapi tidak semua) respons yang dipancing tidak memiliki refleksi. Protokol think-aloud tampaknya mencerminkan kinerja bisnis seperti biasa: berpikir keras tidak mengganggu kinerja tes dibandingkan dengan kelompok kontrol. Data-data ini menunjukkan bahwa vCRT biasanya memenuhi interpretasi standar dari tes refleksi (meskipun bukan tanpa pengecualian) dan bahwa vCRT dapat menjadi ukuran yang baik untuk konstruk yang diteorikan oleh penjelasan dua faktor dari 'refleksi' (seperti disengaja dan sadar).

**Kata kunci:** tes refleksi kognitif; analisis protokol think-aloud; psikometri; penilaian dan pengambilan keputusan; heuristik dan bias



**Kutipan:** Byrd, Nick, Brianna Joseph, Gabriela Gongora, dan Miroslav Sirota. 2023. Beritahu Kami Apa yang Sebenarnya Anda Pikirkan: Analisis Protokol Think Aloud terhadap Tes Refleksi Kognitif Verbal. *Journal of Intelligence* 11: 76.  
<https://doi.org/10.3390/jintelligence11040076>

Diterima: 20 Juli 2022

Revisi: 27 Maret 2023

Diterima: 17 April 2023

Diterbitkan: 21 April 2023



**Hak Cipta:** © 2023 oleh penulis. Pemegang lisensi MDPI, Basel, Swiss. Artikel ini adalah artikel akses terbuka yang didistribusikan di bawah syarat dan ketentuan lisensi Creative Commons Atribusi (CC BY) (<https://creativecommons.org/licenses/by/>

4.0/).

## 1. Pendahuluan

Jika Anda sedang mengikuti perlombaan lari dan Anda melewati orang yang berada di posisi kedua, di posisi manakah Anda sekarang? Interpretasi standar untuk masalah seperti ini mengasumsikan bahwa jawaban yang muncul dengan cepat dan mudah di benak banyak orang adalah "juara pertama". Namun, setelah direnungkan, banyak orang menyadari bahwa

jawaban yang benar adalah "tempat kedua". Soal ini dianggap sebagai tes refleksi karena dirancang untuk memancing kita ke arah jawaban tertentu yang, setelah direnungkan, dapat kita sadari bahwa jawaban tersebut tidak benar (Byrd 2022b, 2022c). Dengan demikian, interpretasi standar dari label refleksi memikat respons yang salah dengan "tidak reflektif" dan respons yang benar dengan "reflektif" (Pennycook et al. 2015a).

Sejak diperkenalkannya Tes Refleksi Kognitif (Kahneman dan Frederick 2002), teori-teori tentang refleksi telah berkembang (Evans dan Stanovich 2013). Di tengah kemajuan ini, beberapa ahli teori menyaring lusinan perbedaan reflektif-tidak reflektif (lihat Frankish 2010, Tabel 1) menjadi hanya dua perbedaan yang agak ortogonal: pemrosesan otomatis versus pemrosesan yang disengaja dan representasi sadar versus representasi tidak sadar (Shea dan Frith 2016). Menurut penjelasan dua faktor dari 'refleksi' ini, pemikiran *reflektif* melibatkan lebih banyak representasi kesadaran (atau kesadaran) dari penalaran yang relevan dan penghambatan yang lebih disengaja (berlawanan dengan penerimaan langsung dari dorongan awal seseorang) dan pemikiran tidak *reflektif* kurang terwakili secara sadar dan lebih otomatis (Byrd 2022c). Dengan kata lain, *isi* dari penalaran reflektif lebih mudah diakses (misalnya, untuk menjelaskan secara verbal) dan *proses* berpikir reflektif melibatkan lebih banyak pertimbangan ulang (misalnya, meragukan atau mengecek ulang) dibandingkan berpikir tidak reflektif (Byrd 2019). Para ahli teori juga berpendapat bahwa

kebutuhan akan refleksi mungkin bergantung pada konteks: dalam domain penalaran yang sudah dikenal, penalaran yang tidak reflektif mungkin dapat mencapai hasil yang diinginkan, tetapi dalam domain yang tidak dikenal, mengatasi kesalahan atau bias mungkin memerlukan beberapa refleksi (Pennycook et al. 2015a).

Seiring dengan berkembangnya teori-teori refleksi, begitu pula pemahaman kita tentang tes refleksi. Beberapa peneliti telah menantang interpretasi standar dari tes refleksi matematis (disingkat mCRT) dan penjelasan dua faktor tentang 'refleksi' (Stanovich 2018). Memang, beberapa peneliti menemukan bahwa sebagian besar (67%) dari mereka yang menjawab tes refleksi dengan benar setelah musyawarah telah menjawab dengan benar di bawah tekanan waktu atau beban kognitif sebelum musyawarah (Bago dan De Neys 2019). Meskipun paradigma dua respons tersebut telah membantu menguji teori intervensi standar tentang refleksi-yang menyatakan bahwa respons yang benar melibatkan intervensi terhadap respons standar (alias, terpancing) (Evans 2007)-paradigma ini mengabaikan banyak informasi yang berguna tentang proses pemecahan masalah tes refleksi. Jadi, peneliti lain telah mendengarkan setiap kata yang diucapkan oleh peserta tes ketika berpikir keras selama tes refleksi. Salah satu studi berpikir keras juga "menimbulkan keraguan" tentang interpretasi standar mCRT: *sebagian besar* (77%) respons yang benar pada mCRT "dimulai dengan jawaban yang benar" dan *banyak* (39%) respons yang memikat pada mCRT yang melibatkan "merefleksikan respons pertama" (Szasz dkk., 2017). Beberapa investigasi psikometrik terhadap tes refleksi menunjukkan bahwa respons yang benar tapi tidak reflektif dan respons yang terpicat tapi reflektif ini dapat dijelaskan oleh keakraban domain (Purcell et al. 2021), kecerdasan (Thompson dan Johnson 2014), atau strategi (Markovits et al. 2021). Meskipun nilai prediktif mCRT tetap ada setelah mengulang tes (Bialek dan Pennycook 2018; Stagnaro dkk. 2018), prediktor terbaik untuk performa mCRT seringkali adalah performa tes matematika secara umum (Attali dan Bar-Hillel 2020; Erceg dkk. 2020). Jadi, mCRT tidak hanya melacak refleksi, tapi juga kompetensi matematika (alias kemampuan berhitung) dan faktor lainnya.

### 1.1. Tes Refleksi Kognitif Verbal

Sirota dan koleganya (2021) mengembangkan dan memvalidasi varian baru, 10 butir soal, *non-matematis* dari tes refleksi kognitif (mCRT) yang terkenal (matematika) dari Shane Frederick dan koleganya (Frederick 2005; Kahneman dan Frederick 2002) untuk mengatasi masalah keakraban dan kemampuan berhitung (Byrd 2022d). Salah satu soal dari tes refleksi kognitif verbal Sirota dan rekan-rekannya (atau disingkat vCRT) adalah contoh soal pembuka: "Jika Anda sedang mengikuti sebuah perlombaan, dan Anda melewati orang yang berada di posisi kedua, di posisi manakah Anda sekarang?" Beberapa penelitian menemukan bahwa vCRT memiliki konsistensi internal yang tinggi, reliabilitas tes-retes yang tinggi, dan lebih sedikit keterkaitannya dengan kemampuan matematika secara umum dibandingkan dengan mCRT-bahkan ketika diterjemahkan ke dalam bahasa dan konteks yang berbeda (Sobkow dkk. 2023). Hal ini menunjukkan bahwa vCRT merupakan pelengkap atau pengganti mCRT yang menjanjikan dalam banyak konteks penelitian.

### 1.2. Analisis Protokol Think-Aloud

Para peneliti telah lama menyerukan untuk melakukan investigasi terhadap isi dan proses refleksi, bukan hanya hasilnya (Stromer-Galley 2007). Untungnya, Ericsson dan rekan-rekannya telah mengembangkan dan memvalidasi protokol think-aloud bersamaan (Ericsson 2003; Ericsson dan Simon 1993) yang telah terbukti dapat mengatasi beberapa masalah yang terkenal pada protokol laporan verbal awal seperti konfabulasi (contohnya, Wilson dan Nisbett 1978) atau gangguan kinerja pada masalah-masalah yang sepertinya tidak dapat dipecahkan tanpa adanya pemahaman (contohnya, Schooler dkk. 1993). Sebagai contoh, meminta peserta untuk *memverbalisasikan* atau *mengingat kembali* (daripada menjelaskan atau membenarkan) pemikiran mereka tidak serta merta mengganggu kinerja tugas atau menghasilkan laporan verbal yang tidak konsisten dengan kinerja yang mereka amati (Fox dkk. 2011; Petitmengin dkk. 2013). Verbalisasi bersamaan bahkan dapat *membantu* orang untuk memikirkan sebuah kata yang membentuk kata majemuk yang umum dengan tiga kata lain-misalnya, memikirkan

---

'pohon' ketika diberikan 'apel', 'keluarga', dan 'rumah' ([Ball dan Stevens 2009](#); lihat juga [Blech et al. 2020](#)). Jadi, meminta peserta untuk berpikir keras saat memecahkan masalah (yaitu, secara bersamaan) dapat membantu peserta memberi tahu kami apa yang sebenarnya mereka pikirkan lebih baik daripada permintaan retrospektif untuk "menjelaskan", yang dapat menyiratkan bahwa peserta harus memverbalisasikan pembenaran (daripada ingatan) tentang apa yang mereka pikirkan ([Ericsson dan Simon 1980](#)). Evans dan rekan-rekannya

dengan sukses menerapkan kesadaran tentang laporan verbal yang bersamaan ini pada tes refleksi logis untuk membedakan antara teori penalaran "rasionalis" dan "dua faktor" atau "proses ganda" (Evans dkk. 1983). Szaszi dan koleganya juga menggunakan metode think-aloud seperti itu untuk menyelidiki proses kognitif yang terlibat dalam menyelesaikan CRT asli (2017) (Szaszi et al. 2017). Jadi, protokol think-aloud yang bersamaan mungkin juga berguna untuk menelusuri proses berpikir peserta tes vCRT dan menentukan apakah berpikir dengan suara keras memiliki efek "reaktif" pada kinerja tes (Ericsson 2003).

### 1.3. Penelitian Saat Ini

Tujuan utama kami adalah (a) untuk menguji apakah berpikir keras mengubah kinerja tes refleksi, (b) menguji beta platform berpikir keras online, (c) mengukur deviasi antara interpretasi standar tes refleksi dan penjelasan dua faktor refleksi, (d) untuk menilai bagaimana kinerja vCRT bergantung pada keakraban vCRT, dan (e) menilai akun intervensi default dari respons tes refleksi. Kami telah mendaftarkan dua hipotesis. Pertama, berpikir dengan keras selama vCRT akan memberikan bukti respon yang benar tapi tidak reflektif dan respon yang terpancing tapi reflektif. Kedua, berpikir dengan suara keras tidak akan secara signifikan menghalangi kinerja vCRT-yaitu, tidak akan berdampak atau meningkatkan kinerja tes refleksi. Hasil eksperimen dan studi lanjutan menghasilkan hasil yang dihipotesiskan. Mereka juga mendeteksi bahwa penjelasan dua faktor dari 'refleksi' sangat kuat, meskipun tidak sempurna, berhubungan dengan interpretasi standar kinerja tes refleksi. Semua manipulasi, tindakan, dan pengecualian dilaporkan. Semua pedoman etika APA dan IRB diikuti. Hipotesis, metode, strategi analisis, data, dan skrip R yang telah didaftarkan sebelumnya ada di Open Science Framework: <https://osf.io/rk3jq>.

## 2. Belajar 1

Penelitian pertama bertujuan untuk menguji efek berpikir keras pada jawaban akhir dari pertanyaan-pertanyaan vCRT. Tujuan sekundernya adalah untuk menguji perbedaan performa vCRT antara peserta yang sudah terbiasa dan yang masih naif, korelasi antara interpretasi standar tes refleksi dan penjelasan yang lebih baru tentang 'refleksi', serta tingkat respons yang benar namun tidak reflektif dan respons yang terpancing namun reflektif.

### 2.1. Metode

*Peserta.* Orang-orang direkrut dari ruang publik di sebuah kampus universitas di Amerika Serikat bagian tenggara. Kami mendaftarkan terlebih dahulu jumlah sampel target sebanyak 100 peserta-50 peserta per kondisi (Simmons et al. 2013). Setelah berbulan-bulan perekrutan, mencapai jumlah sampel yang sudah terdaftar sebelumnya dengan protokol tatap muka menjadi tidak mungkin dilakukan secara etis dan praktis karena Organisasi Kesehatan Dunia mengumumkan pandemi global (Ghebreyesus 2020), kampus universitas ditutup, dan IRB universitas mengumumkan bahwa semua pengumpulan data tatap muka harus dihentikan hingga pemberitahuan lebih lanjut (Office for Human Subjects Protection 2020). Karena protokol tidak dapat direplikasi secara online, kami harus menghentikan pengumpulan data setelah merekrut hanya 99 peserta (usia rata-rata = 23,41 tahun; 48 diidentifikasi sebagai perempuan, 50 sebagai laki-laki, dan 1 tidak memilih jenis kelamin; 51 diidentifikasi sebagai kulit putih, 15 sebagai kulit hitam, 16 sebagai Hispanik atau Latin, 1 sebagai Kepulauan Pasifik, dan 16 sebagai etnis lainnya).

### 2.2. Prosedur dan Bahan

*Manipulasi.* Setelah menyetujui untuk berpartisipasi, partisipan diarahkan ke survei Qualtrics menggunakan kode QR di mana mereka secara acak ditugaskan ke kondisi think-aloud atau kondisi kontrol. Untuk memastikan bahwa partisipan dalam kedua kondisi tersebut menyelesaikan survei di depan peneliti, mereka diminta untuk tetap berada di meja penelitian hingga akhir survei untuk menerima kompensasi mereka: kesempatan untuk memenangkan speaker pintar, botol air, atau buku yang ada di atas meja.

*Protokol berpikir keras.* Peserta yang secara acak ditugaskan untuk kondisi think-aloud

---

diminta untuk meminta instruksi dari peneliti. Setelah peneliti menjelaskan protokol think-aloud kepada peserta, peserta memiliki kesempatan untuk meminta klarifikasi dan

persetujuan dengan memilih tombol berlabel "Saya menerima dan memahami instruksi dari peneliti". Kemudian peneliti memulai rekaman audio di ponsel pintar dan partisipan berlatih berpikir dengan suara keras pada tugas pra-survei, "Untuk berlatih berpikir dengan suara keras, silakan ucapkan kalimat ini dengan keras, diikuti dengan nomor berikut ini." Nomor masing-masing

Peserta yang membaca dengan suara keras dibuat secara acak dan digunakan untuk memasang jawaban survei secara anonim dengan setiap rekaman think-aloud yang sesuai. Para peserta diingatkan untuk berpikir keras sesuai kebutuhan selama survei berlangsung.

*Tes Refleksi Kognitif Verbal.* Peserta menyelesaikan Tes Refleksi Kognitif Verbal atau vCRT yang terdiri dari 10 butir soal (Sirota et al. 2021). Urutan item ditetapkan untuk semua peserta-tidak ada efek latihan yang diamati (Lampiran A Gambar A1) - dan tanggapan diketik ke dalam kotak teks. Mengikuti interpretasi standar tes refleksi, skor reflektif dihitung dengan menjumlahkan respons yang benar (misalnya, peringkat kedua) dan skor tidak reflektif dihitung dengan menjumlahkan respons yang diiming-imingi (misalnya, peringkat pertama) pada soal-soal penalaran verbal. Respons salah yang tidak dipancing hanya 1,67% dari semua respons, lebih rendah dari apa yang telah diamati untuk mCRT yang digunakan secara luas (Stupple et al. 2017, Tabel 5).

*Pertanyaan tentang iming-iming.* Untuk menguji apakah jawaban yang benar mengikuti jawaban yang diiming-imingi-seperti yang biasa dilakukan dalam penelitian tes refleksi (Howarth dan Handley 2016)-peserta melaporkan apakah jawaban yang diiming-imingi itu muncul dalam benak mereka setelah mereka memberikan jawaban dari setiap pertanyaan tes refleksi. Misalnya, setelah menjawab pertanyaan yang disebutkan di atas tentang meloloskan pembalap di posisi ke-2, para peserta ditanya, "Pernahkah Anda berpikir bahwa 'posisi pertama' bisa menjadi jawabannya?"

*Kesengajaan dan kesadaran dalam rekaman think-aloud.* Penjelasan dua faktor dari 'refleksi' menyatakan bahwa penalaran bersifat reflektif ketika penalaran tersebut lebih disengaja dan lebih sadar diwakili (Shea dan Frith 2016). Penalaran dikatakan disengaja ketika tidak hanya menerima respons awal yang otomatis dan dikatakan sadar ketika peserta dapat mengartikulasikan bagian-bagian dari penalaran mereka (Byrd 2019, 2022c). Jadi, dua penilai mahasiswa (yang tidak mengetahui hipotesis proyek), dan kemudian penulis pertama, menilai setiap respons. Sebelum para penilai mengetahui jawaban akhir partisipan, para penilai menilai kesengajaan dari setiap respons-yaitu, apakah partisipan secara verbal mempertimbangkan kembali respons awal mereka (Light's  $K_{avg} = 0.59$ )-dan representasi sadar-yaitu, apakah partisipan mengutarakan alasan untuk atau menentang respons apa pun (Light's  $K_{avg} = 0.43$ ). Opsi penilaian termasuk "ya", "tidak", dan "tidak tentu", dan kesepakatan "moderat" antara penilai memungkinkan peringkat "ya" dari setiap peserta untuk dirata-ratakan (Landis dan Koch 1977, hlm. 165).

*Keakraban tes.* Penelitian sebelumnya menemukan bahwa banyak peserta yang sudah terbiasa dengan pertanyaan-pertanyaan tes refleksi dan bahwa keakraban tersebut dapat menjadi prediktor terbaik untuk performa tes refleksi (Byrd 2022d; Stieger dan Reips 2016). Jadi, para penilai juga menilai apakah setiap peserta menyebutkan bahwa mereka sudah terbiasa dengan soal-soal vCRT (Light's  $K = 0.82$ ). Penentuan diberi label "ya", "tidak", dan "tidak pasti". Parameter keakraban adalah jumlah rata-rata penilaian "ya" per peserta.

### 2.3. Hasil

Kami menguji tingkat respons yang benar namun tidak reflektif dan respons yang terpancing namun reflektif pada vCRT, efek berpikir keras pada performa tes refleksi, korelasi antara skor tes refleksi standar dan penjelasan terbaru tentang 'refleksi', serta korelasi antara keakraban tes dan performa tes.

*Respons yang benar namun tidak reflektif dan respons yang terpancing namun reflektif.* Laporan verbal Think Aloud terkadang menyimpang dari interpretasi standar tes refleksi. Kami memperkirakan beberapa orang akan mendapatkan jawaban yang benar sebelum melakukan refleksi dan tanpa terlebih dahulu memikirkan jawaban yang diiming-imingi, dan terpancing untuk menerima jawaban yang salah meskipun telah melakukan refleksi. Tabel 1 menegaskan ekspektasi yang sudah terdaftar sebelumnya: interpretasi standar dari jawaban yang benar dan jawaban yang dipancing biasanya tetapi tidak

---

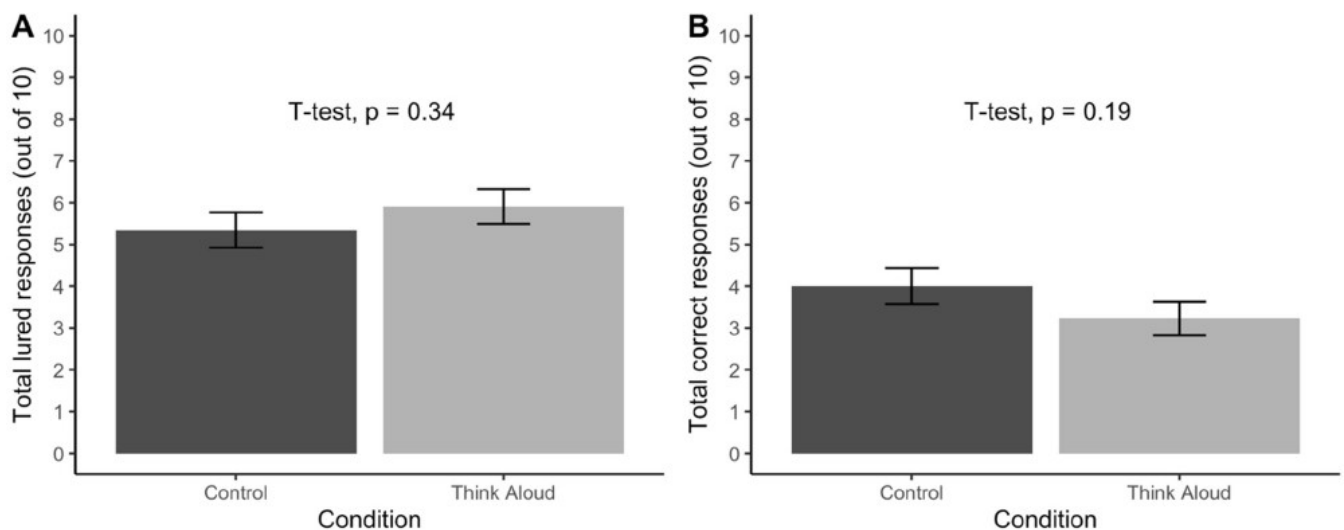
sempurna sesuai dengan penjelasan dua faktor dari 'refleksi'.



**Tabel 1.** Kategorisasi standar dan dua faktor dari respons tes refleksi berdasarkan analisis protokol think-aloud terhadap respons tes refleksi verbal dalam Studi 1. Contoh verbalisasi berdasarkan pertanyaan tes refleksi berikut: "Jika Anda sedang mengikuti perlombaan lari, dan Anda melewati orang yang berada di posisi ke-2, di posisi manakah Anda sekarang?".

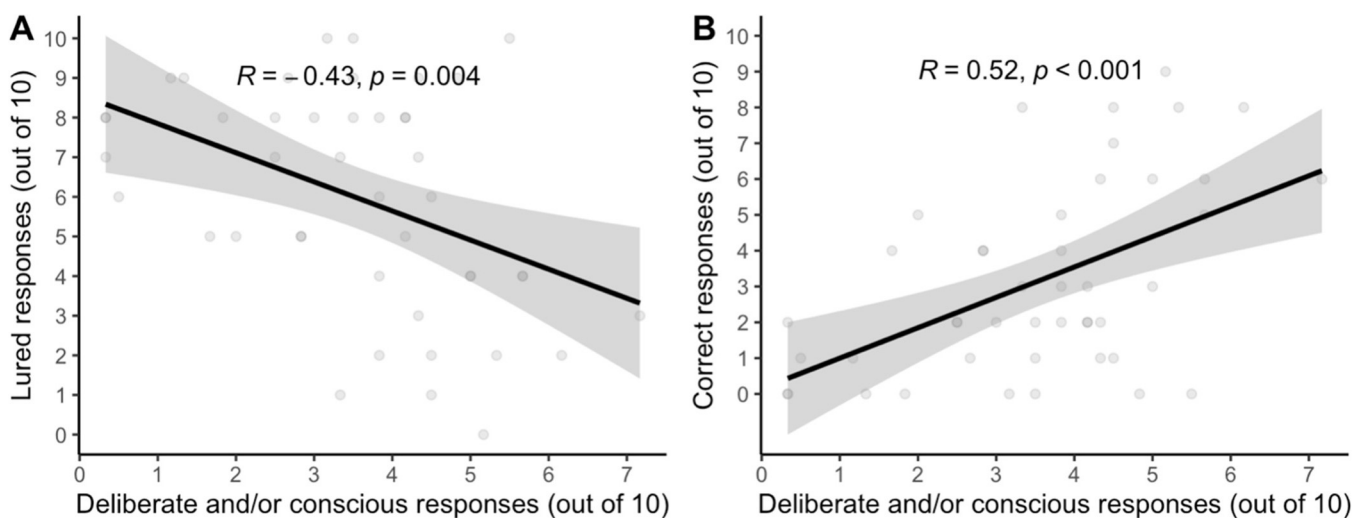
Kategori	Contoh Verbalisasi	Jawaban	Standar	Dua Faktor	Nilai
Benar dan reflektif	"1 jelas. Sebenarnya tidak. . . Kedua."	Benar.	Reflektif	Reflektif	80.2%
Benar-tetapi tidak reflektif	"Kedua"	Benar.	Reflektif	Tidak reflektif	19.8%
Terpikat dan tidak reflektif	"1"	Terpikat	Tidak reflektif	Tidak reflektif	71.5%
Memikat namun reflektif	"Saya ingin mengatakan 1, tapi, ya, 1."	Terpikat	Tidak reflektif	Reflektif	28.5%

*Berpikir dengan suara keras tidak memengaruhi kinerja.* Gambar 1 menegaskan hipotesis kami yang sudah terdaftar sebelumnya dan penelitian meta-analisis sebelumnya (Fox et al. 2011): kami tidak mendeteksi efek interferensi dari berpikir keras pada jumlah respons yang terpancing atau benar pada vCRT.



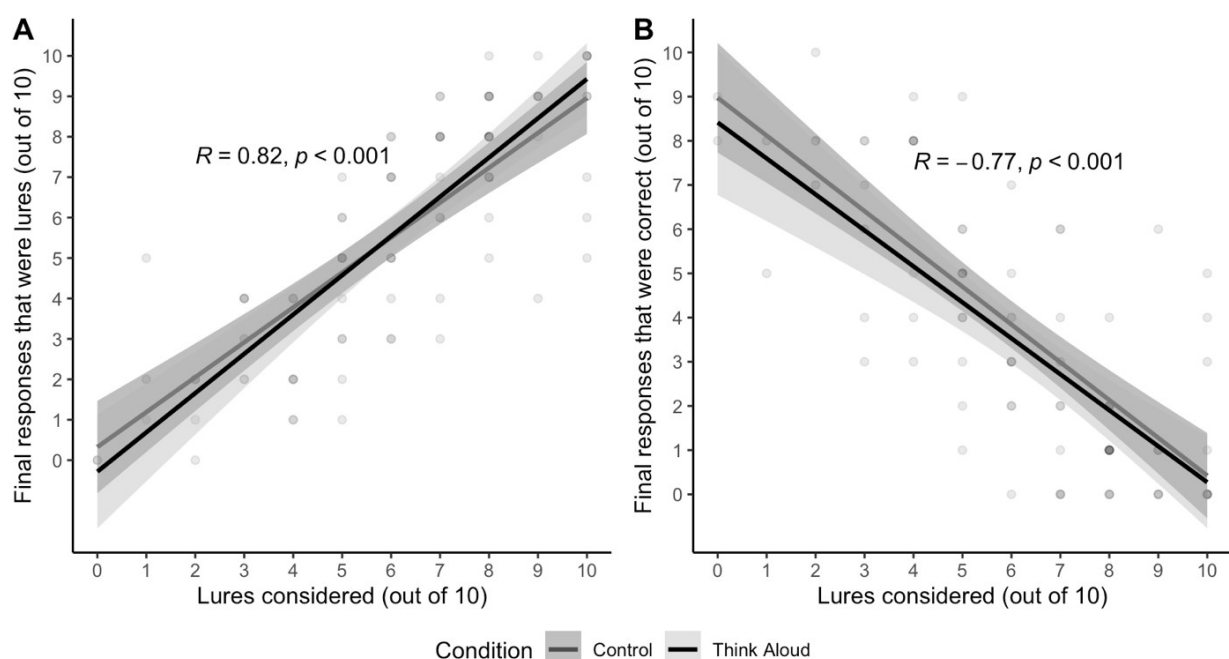
**Gambar 1.** Pengaruh berpikir dengan suara keras pada (A) jumlah respons yang terpancing dan (B) jumlah respons yang benar pada tes refleksi kognitif verbal (vCRT) dalam Studi 1 (N = 99). Bilah galat mewakili kesalahan standar.

*Interpretasi dua faktor memprediksi interpretasi standar.* Analisis regresi digunakan untuk memahami seberapa baik interpretasi standar dari tes refleksi selaras dengan penjelasan dua faktor dari para ahli teori proses ganda tentang 'refleksi'. Gambar 2 menunjukkan bahwa kedua faktor tersebut selaras: semakin banyak verbalisasi partisipan yang melibatkan pertimbangan ulang yang disengaja atas respons apa pun atau artikulasi sadar atas alasan respons apa pun, semakin kecil kemungkinan partisipan untuk menerima respons yang dipancing dan semakin besar kemungkinan partisipan untuk menerima respons yang benar. Dengan kata lain, sejauh mana *pemikiran* partisipan menunjukkan dua faktor teoretis refleksi berhubungan erat dengan sejauh mana *jawaban akhir* partisipan dianggap reflektif atau tidak reflektif berdasarkan interpretasi standar dari respons tes.



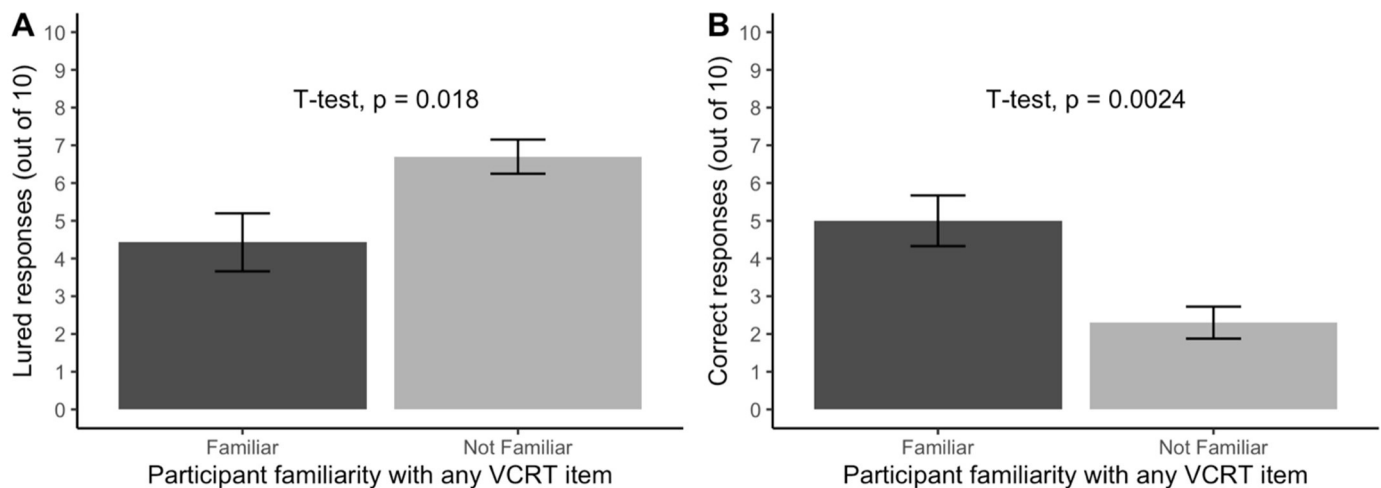
**Gambar 2.** Korelasi antara pengkodean "dua faktor" dari respons reflektif (disengaja dan/atau disadari) dan (A) jumlah respons yang dipancing-atau yang disebut "tidak reflektif"-dan (B) jumlah respons yang benar-atau yang disebut "reflektif" pada tes refleksi verbal (vCRT) pada kondisi think-aloud Studi 1 ( $N = 47$ ) dengan pita kesalahan standar abu-abu.

Mempertimbangkan respons yang dipancing. Frederick (2005) mengamati bahwa respons mCRT yang benar sering kali melibatkan pertimbangan respons yang diiming-imingi. Beberapa orang berhipotesis bahwa iming-iming menarik karena mereka cenderung merasa benar (Thompson et al. 2013). Jika itu benar, maka orang seharusnya tidak hanya cenderung mempertimbangkan respons yang diiming-imingi, tetapi mereka yang mempertimbangkan iming-iming juga seharusnya sangat mungkin menerima iming-iming sebagai jawaban akhir mereka. Gambar 3 menegaskan hal ini: pertimbangan terhadap respons yang diiming-imingi pada suatu titik dalam setiap keputusan relatif tinggi (rata-rata = 6,13, rentang = 0-10,  $SD = 2,59$ ), dan hanya dengan mempertimbangkan iming-iming hampir secara sempurna memprediksi apakah seseorang menerima iming-iming tersebut sebagai jawaban akhir.



**Gambar 3.** Korelasi antara pertimbangan umpan dan (A) penerimaan umpan dan (B) penerimaan respons yang benar pada tes refleksi verbal (vCRT) dalam Studi 1 ( $N = 99$ ). Pita abu-abu mewakili kesalahan standar.

*Keakraban dengan soal memprediksi performa tes.* Pada sekitar 27% rekaman think-aloud, peserta menyebutkan keakraban mereka dengan setidaknya satu soal di vCRT-misalnya, "Saya pernah melihat soal-soal ini di TikTok". Gambar 4 menunjukkan perbedaan besar dalam performa vCRT antara peserta yang sudah terbiasa dan yang naif pada respons yang dipancing ( $d = -0.87$ ) dan respons yang benar ( $d = 1.13$ ).



**Gambar 4.** Jumlah (A) respons yang dipancing dan (B) respons yang benar pada tes refleksi verbal (vCRT) di antara para peserta dalam kondisi think-aloud Studi 1 ( $N = 47$ ) tergantung pada laporan diri mereka yang tidak diminta tentang keakraban dengan vCRT dengan bilah kesalahan standar.

#### 2.4. Diskusi

Data ini menunjukkan keselarasan yang signifikan antara penjelasan dua faktor 'refleksi' dari para ahli teori proses ganda dan interpretasi standar dari jawaban tes refleksi. Data tersebut juga menunjukkan bahwa peserta universitas sebagian besar naif terhadap vCRT meskipun keakraban yang dilaporkan sendiri tetap menjadi prediktor yang kuat untuk interpretasi standar kinerja vCRT.

Orang mungkin bertanya-tanya apakah hasil ini akan tereplikasi dalam validasi think-aloud yang lebih besar dari vCRT. Untungnya, hasil awal kami menunjukkan bahwa protokol think-aloud tidak akan mempengaruhi kinerja vCRT secara signifikan. Jadi, replikasi yang lebih besar secara metodologis dimungkinkan. Sayangnya, protokol think-aloud berskala besar yang dilakukan secara langsung sangat memakan waktu, membosankan, atau di masa pandemi-tidak etis. Untuk mengatasi tantangan ini, kami bermitra dengan sebuah perusahaan rintisan untuk mengembangkan platform untuk survei think-aloud berskala besar dan online.

### 3. Belajar 2

Studi 2 bertujuan untuk mereplikasi temuan dari Studi 1 pada sampel peserta yang baru dan menguji kelayakan metodologi survei think-aloud online. Untuk melakukan hal ini, kami mereproduksi semua instruksi dan langkah-langkah dalam kondisi think-aloud pada Studi 1 dalam platform survei audio daring, Phonic ([Perrodin dan Todd 2021](#); [Phonic Inc. 2020](#)).

#### 3.1. Metode

**Peserta.** Partisipan monolingual berbahasa Inggris direkrut dari Prolific ([Palan dan Schitter 2018](#); [Peer et al. 2017](#)) dengan bayaran \$9,85/jam berdasarkan waktu penyelesaian rata-rata dari kondisi think-aloud pada Studi 1. Untuk memastikan kualitas data, Prolific memberi tahu calon partisipan bahwa kompensasi akan bergantung pada persetujuan dan kemampuan mereka untuk memberikan rekaman pemikiran mereka yang dapat digunakan selama survei. Kami bertujuan untuk menggandakan ukuran sampel yang telah terdaftar sebelumnya dari kondisi think-aloud Studi 1 ( $N = 47$ ), merekrut 102 partisipan (usia rata-rata = 30,38; 57 diidentifikasi sebagai perempuan, 38 sebagai laki-laki, dan 7

---

tidak memilih jenis kelamin; 85 diidentifikasi sebagai kulit putih, 3 sebagai kulit hitam, 3 sebagai Hispanik atau Latin, dan 11 sebagai etnis lainnya).

### 3.2. Prosedur dan Bahan

*Platform survei audio Phonic.* Kami menggunakan platform survei audio online (Phonic) untuk merekam verbalisasi secara bersamaan; layanan ini disediakan sebagai imbalan untuk pengujian beta platform survei yang baru.

*Materi dari Studi 1.* Semua materi dari kondisi berpikir keras pada Studi 1 disertakan dalam Studi 2. Peserta berlatih berpikir keras sebelum survei dan kemudian berpikir keras sambil menyelesaikan vCRT 10 item yang sama dengan pengingat untuk memverbalisasikan semua pemikiran mereka selama tes - sekali lagi, tidak ada efek latihan yang teramati (Lampiran A Gambar A2). Setelah setiap jawaban disampaikan, peserta menjawab pertanyaan lanjutan yang sama tentang apakah mereka mempertimbangkan respons yang dipancing. Kesengajaan, kesadaran, dan keakraban tes dari rekaman think-aloud para peserta untuk setiap pertanyaan dinilai oleh peneliti utama dengan menggunakan opsi penilaian yang sama dengan Studi 1. Mereplikasi hasil dari Studi 1, tingkat jawaban salah yang tidak terpancing sangat rendah (1,72%).

### 3.3. Hasil

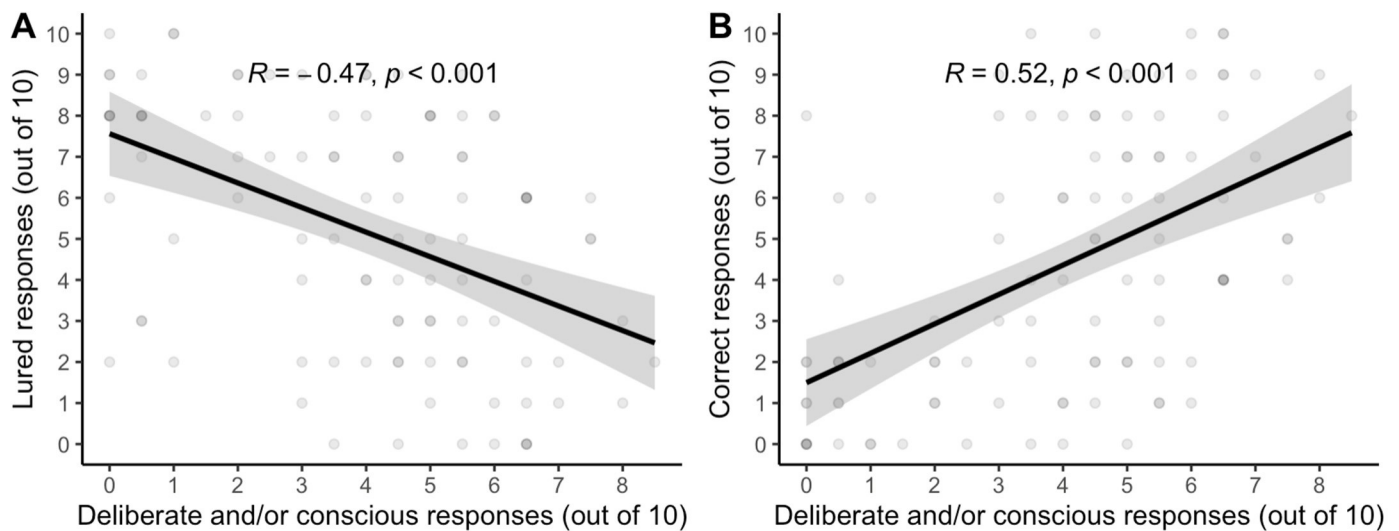
Kami menguji korelasi antara keakraban tes dan kinerja tes, korelasi antara skor tes refleksi standar dan penjelasan terbaru dari 'refleksi', serta tingkat respons yang benar namun tidak reflektif dan respons yang terpancing namun reflektif pada vCRT.

*Respons yang benar tetapi tidak reflektif dan respons yang terpancing namun reflektif.* Untuk menguji kesepakatan antara interpretasi standar dari tes refleksi dan penjelasan dua faktor yang lebih baru tentang 'refleksi', tingkat respons yang benar-tetapi-tidak reflektif dan respons yang terpancing namun reflektif ditentukan oleh rekaman think-aloud dari para peserta Prolific. Tabel 2 menunjukkan replikasi dari kesepakatan yang dominan namun tidak sempurna antara interpretasi standar dan penjelasan dua faktor dari jawaban tes refleksi.

**Tabel 2.** Kategorisasi standar dan dua faktor dari respons tes refleksi berdasarkan analisis protokol think-aloud terhadap respons tes refleksi verbal dalam Studi 2. Contoh verbalisasi berdasarkan pertanyaan tes refleksi berikut: "Jika Anda sedang mengikuti perlombaan lari, dan Anda melewati orang yang berada di posisi ke-2, di posisi manakah Anda sekarang?".

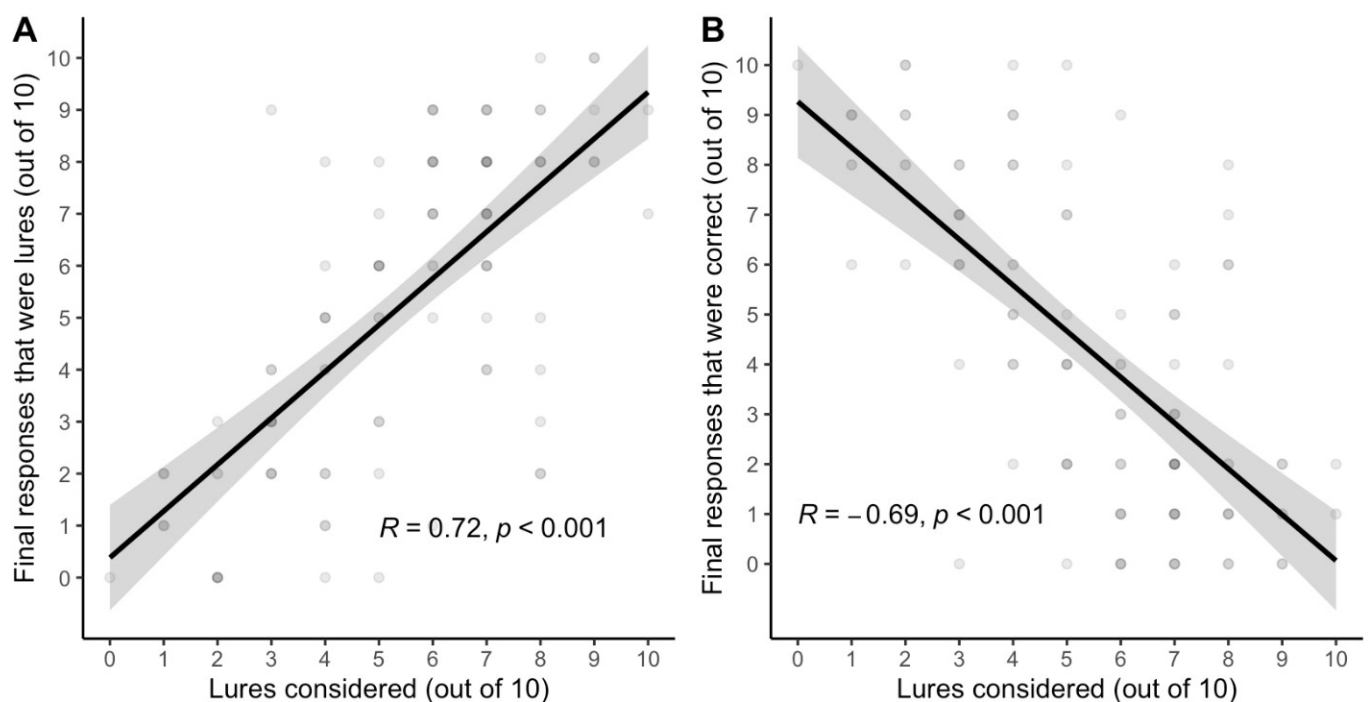
Kategorisasi	Contoh Verbalisasi	Jawaban	Standar	Dua Faktor	Nilai
Benar dan reflektif	"1. Tidak. . . 2. Whoopsie"	Benar.	Reflektif	Reflektif	68.5%
Benar-tetapi tidak reflektif	"Kedua"	Benar.	Reflektif	<i>Tidak reflektif</i>	31.5%
Terpikat dan tidak reflektif	"1"	Terpikat	Tidak reflektif	Tidak reflektif	75.8%
Memikat namun reflektif	"1. . . atau itu tipuan? . . . Saya akan mengatakan 1."	Terpikat	Tidak reflektif	<i>Reflektif</i>	24.2%

*Interpretasi dua faktor memprediksi interpretasi standar.* Analisis regresi lain digunakan untuk menguji bagaimana interpretasi standar dari tes refleksi sesuai dengan penjelasan dua faktor dari para ahli teori proses ganda tentang 'refleksi'. Gambar 5 menunjukkan replikasi dari korelasi mereka: peningkatan jumlah respons peserta yang melibatkan pemikiran yang disengaja atau disadari-sebagaimana ditentukan oleh rekaman think-aloud-berkorelasi dengan penurunan yang signifikan dalam jumlah respons yang terpancing dan peningkatan yang signifikan dalam jumlah respons yang benar.



**Gambar 5.** Korelasi antara pengkodean "dua faktor" dari respons reflektif (disengaja dan/atau disadari) dan (A) jumlah respons yang dipancing-atau yang disebut "tidak reflektif" dan (B) jumlah respons yang benar-atau yang disebut "reflektif"-pada kinerja tes refleksi verbal (vCRT) dalam Studi 2 (N = 102). Pita abu-abu mewakili kesalahan standar.

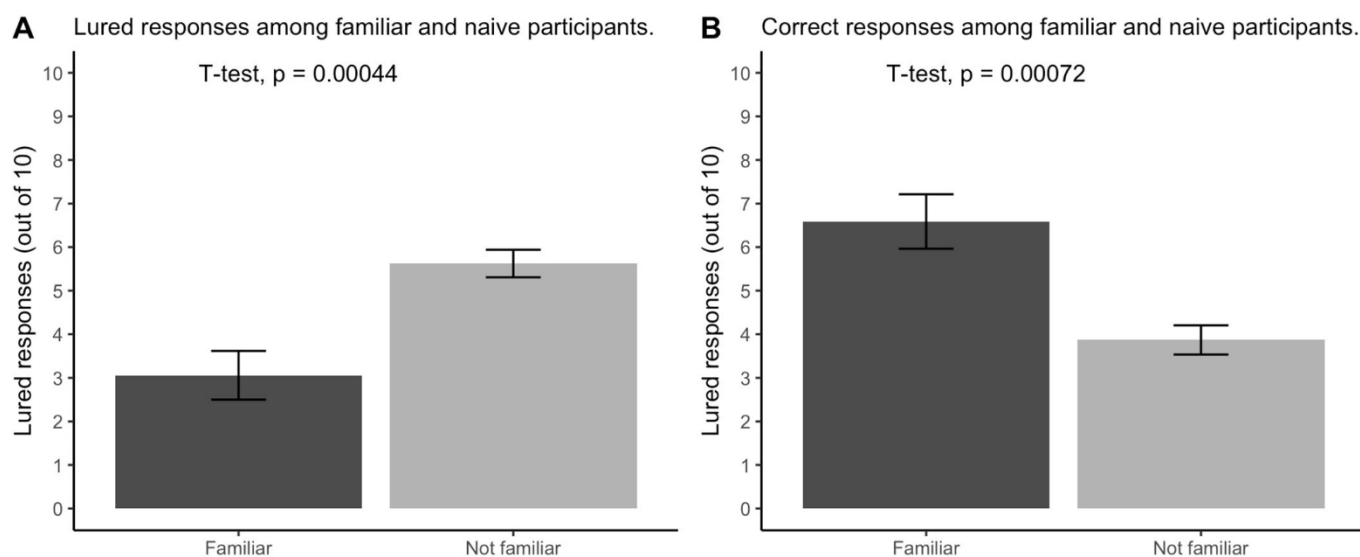
*Pertimbangan tanggapan* yang dipancing. Gambar 6 menunjukkan replikasi dari perasaan yang kuat akan kebenaran tanggapan yang diiming-imingi. Memang, *pertimbangan* iming-iming tidak hanya relatif tinggi (rata-rata = 5.37, rentang = 0-10, SD = 2.38), tetap menjadi prediktor terbaik untuk *menerima* respons yang dipancing dan yang benar pada vCRT.



**Gambar 6.** Korelasi antara pertimbangan umpan dan (A) *penerimaan* umpan dan (B) *penerimaan* respons yang benar pada kinerja tes refleksi verbal (vCRT) Studi 2 (N = 102). Pita abu-abu mewakili kesalahan standar.

*Keakraban tes memprediksi kinerja tes.* Pada sekitar 17% rekaman think-aloud, peserta Prolific menyebutkan keakraban sebelumnya dengan setidaknya satu item pada vCRT-secara signifikan

kurang familiar dibandingkan dengan 27% familiaritas di antara partisipan universitas kami,  $t = -2.9$ , 95% CI [0.09, 0.24],  $p = 0.005$ . Gambar 7 menunjukkan replikasi perbedaan besar dalam kinerja vCRT antara peserta yang akrab dan naif untuk respons yang dipancing ( $d = -0.91$ ) dan respons yang benar ( $d = 0.90$ ).



**Gambar 7.** Jumlah (A) respons yang dipancing dan (B) respons yang benar pada tes refleksi verbal (vCRT) di antara para peserta dalam Studi 2 ( $N = 102$ ) tergantung pada laporan diri mereka yang tidak diminta tentang keakraban mereka dengan vCRT. Bilah galat mewakili kesalahan standar.

#### 4. Diskusi Umum

Dalam dua studi protokol think-aloud, kami menemukan bahwa respons yang benar dari partisipan kami pada tes refleksi kognitif verbal terjadi setelah memikirkan respons awal yang secara intuitif menarik tapi salah (yaitu, iming-iming). Ketika menerima iming-iming sebagai jawaban *akhir* mereka, pemikiran partisipan kami biasanya tidak memiliki upaya yang disengaja untuk mengoreksi respons awal yang menarik dan salah (iming-iming). Pengecualian dari dua lintasan pemikiran yang diharapkan pada vCRT ini hanya terdiri dari 19–31% respons - jauh lebih sedikit daripada 39–77% respons luar biasa yang terdeteksi pada mCRT yang digunakan secara luas (Szasz et al. 2017). Dengan demikian, temuan kami sebagian besar selaras dengan interpretasi standar respons vCRT (Pennycook et al. 2015a) dan penjelasan dua faktor yang lebih baru tentang penalaran reflektif (Byrd 2019, 2022a, 2022c; Shea dan Frith 2016). Meskipun demikian, minoritas substansial dari respons yang benar tapi tidak reflektif dan terpicat namun reflektif menunjukkan peluang untuk mengurangi kesalahan pengukuran tes refleksi (Machery 2021) untuk menghindari pelabelan yang salah terhadap beberapa respons yang benar sebagai reflektif dan beberapa respons yang terpicat sebagai tidak reflektif atau kikir (Toplak dkk. 2014).

Lebih lanjut, kami tidak menemukan bahwa berpikir dengan suara keras mengganggu kinerja pada tes refleksi verbal. Studi ini juga menegaskan bahwa sebagian besar peserta universitas dan peserta Prolific tidak terbiasa dengan vCRT. Meskipun demikian, laporan diri yang tidak diminta tentang keakraban dengan vCRT merupakan prediktor yang kuat untuk interpretasi standar dari kinerja vCRT dalam kedua studi tersebut. Bukti ini mereplikasi dan memperluas beberapa fitur yang menjanjikan dari vCRT (Sirota et al. 2021).

##### 4.1. Implikasi Metodologis

Penelitian ini juga menunjukkan bahwa protokol think-aloud dapat mengungkapkan nuansa yang berharga dan tidak terdeteksi dalam kinerja tes refleksi kognitif. Sebagai contoh, rekaman think-aloud mengungkapkan bahwa interpretasi standar terhadap respons tes refleksi salah memberi label 19–31% respons sebagai respons reflektif atau tidak reflektif.

---

Wawasan ini tampaknya meningkatkan beban pembenaran untuk menggunakan interpretasi standar tes refleksi atau *tidak* menggunakan protokol think-aloud ([Byrd 2022b](#))



yang lebih akurat mendeteksi ciri-ciri yang disengaja dan disadari dari pemikiran reflektif (Shea dan Frith 2016). Bahkan jika para peneliti tidak memikirkan kembali interpretasi atau ketergantungan mereka terhadap tes refleksi, mereka tetap perlu membenarkan status quo.

Selain implikasi untuk *mengukur* refleksi, mungkin juga ada implikasi untuk *memanipulasi* refleksi. Ingatlah paradigma dua respons dalam mengukur refleksi (misalnya, Bago dan De Neys 2019). Paradigma ini mungkin tidak memberikan wawasan tentang proses yang mengubah respons awal peserta, tetapi paradigma ini menunjukkan apakah respons awal ini sudah benar. Jadi, paradigma dua respons bisa menjadi kompromi antara kemudahan interpretasi standar tes refleksi dan kebosanan pengkodean transkrip tes refleksi (misalnya, Buric dan Šrol 2020). Namun, idealnya, respons kedua ini akan muncul setelah semacam tugas reflektif - sebagai lawan dari masa tunggu yang dipaksakan di mana peserta dapat melakukan apa pun yang mereka inginkan - untuk memastikan bahwa perubahan yang diamati adalah hasil dari refleksi dan bukan faktor lain (Isler et al. 2020).

#### 4.2. Implikasi Teoritis

Hasil lain dari penelitian ini adalah dukungan yang dapat diandalkan untuk penjelasan "perasaan benar" dari kinerja tes refleksi (Pennycook et al. 2015b; Thompson et al. 2011). Sebagian besar respons melibatkan *pertimbangan* iming-iming dan *pertimbangan* iming-iming merupakan prediktor terbaik untuk *menerima* respons yang terpicat dan benar pada vCRT. Jika iming-iming tidak secara signifikan lebih menarik daripada respons lain yang mungkin, maka akan sulit untuk menjelaskan dominannya *pertimbangan* iming-iming, *penerimaan* iming-iming, dan korelasinya yang kuat. Hal ini mungkin berimplikasi pada perdebatan antara akun intervensi standar tentang refleksi dan alternatifnya (Howarth dan Handley 2016). Mereka yang sampai pada respons yang benar sangat tidak mungkin mempertimbangkan respons yang diiming-imingi. Dengan kata lain, apa yang disebut sebagai respons reflektif (yaitu, benar) pada tes refleksi biasanya tidak melibatkan intervensi pada respons default (terpancing). Tentu saja, data ini mengonfirmasi bahwa ada *beberapa* kasus intervensi reflektif default. Jadi, bukti yang ada saat ini mungkin tidak memalsukan akun intervensi default, melainkan menunjukkan bahwa ini bukanlah penjelasan yang lengkap tentang kinerja uji refleksi.

#### 4.3. Keterbatasan

Penelitian saat ini dibatasi oleh sumber daya untuk mendengarkan dan mengkodekan laporan verbal yang diucapkan dengan keras. Hal ini menghasilkan ukuran sampel yang minimal untuk pertanyaan penelitian yang dibahas dalam makalah ini (Simmons et al. 2013). Meskipun efek yang diharapkan terdeteksi-beberapa di beberapa populasi, baik secara langsung maupun online-masih ada peluang bagi peneliti dengan sumber daya yang lebih banyak untuk melakukan replikasi dan perluasan skala yang lebih besar dari penelitian yang sudah ada. Sebagai contoh, para peneliti dapat membuat anotasi aspek-aspek tambahan dari transkrip partisipan dari tes refleksi-misalnya, apakah mengatasi intuisi yang salah bervariasi tergantung pada *jenis* alasan yang dipertimbangkan partisipan: alasan untuk respons awal mereka, alasan untuk respons selanjutnya, alasan yang menentang respons awal mereka, atau alasan yang menentang respons selanjutnya (Cullen dkk., 2022). Kami berharap dapat melihat apa yang ditemukan oleh peneliti lain dengan dataset kami dan peneliti lain (terbuka). Ada juga peluang bagi platform survei think aloud online untuk meningkatkan transkripsi ucapan, analisis sentimen, dan fitur lain yang cukup untuk mengotomatisasi dan oleh karena itu meningkatkan penelitian protokol think aloud.

### 5. Kesimpulan

Penelitian ini mereplikasi sebagian dan mengklarifikasi validasi yang ada dari tes refleksi kognitif verbal, sebagian berkat teknologi survei audio online yang baru. Sebagian besar peserta naif terhadap tes ini dan interpretasi standar dari tes refleksi sebagian besar selaras dengan penjelasan yang lebih lanjut tentang penalaran reflektif. Jika digabungkan dengan penelitian yang sudah ada yang menunjukkan bahwa tes

---

refleksi verbal memiliki konsistensi internal yang tinggi, reliabilitas tes-retes yang tinggi, dan lebih sedikit keterkaitannya dengan kemampuan matematika atau jenis kelamin, maka bukti-bukti yang ada menunjukkan bahwa vCRT dapat menjadi pelengkap yang menjanjikan untuk tes refleksi yang sudah digunakan secara luas. Meskipun demikian, mungkin masih ada peluang untuk meningkatkan

pemahaman tentang refleksi dengan menerapkan kembali protokol think-aloud online untuk penelitian berskala besar. Dengan demikian, baik tes refleksi verbal maupun protokol think-aloud online merupakan alat yang menjanjikan untuk memajukan pemahaman kita tentang penalaran reflektif dan alternatifnya.

**Kontribusi Penulis:** Konseptualisasi, N.B. dan M.S.; Metodologi, N.B.; Analisis formal, N.B.; Investigasi, N.B.; Kurasi data, N.B., B.J. dan G.G.; Penulisan-draf awal, N.B.; Penulisan-tinjauan & penyuntingan, N.B. dan M.S.; Visualisasi, N.B.; Pengawasan, N.B.; Administrasi proyek, N.B.; Perolehan dana, N.B. Semua penulis telah membaca dan menyetujui versi naskah yang diterbitkan.

**Pendanaan:** Penelitian ini mendapatkan dana sebesar \$250 dan dua asisten peneliti (B. Joseph & G. Gongora) melalui Program Kesempatan Penelitian Sarjana di Florida State University. APC untuk artikel ini dibebaskan.

**Pernyataan Dewan Peninjau Institusi:** Penelitian ini dianggap "dikecualikan" oleh Kantor Perlindungan Subjek Manusia di Florida State University (STUDY00000697, 2019-12-13: <https://osf.io/z7bua>).

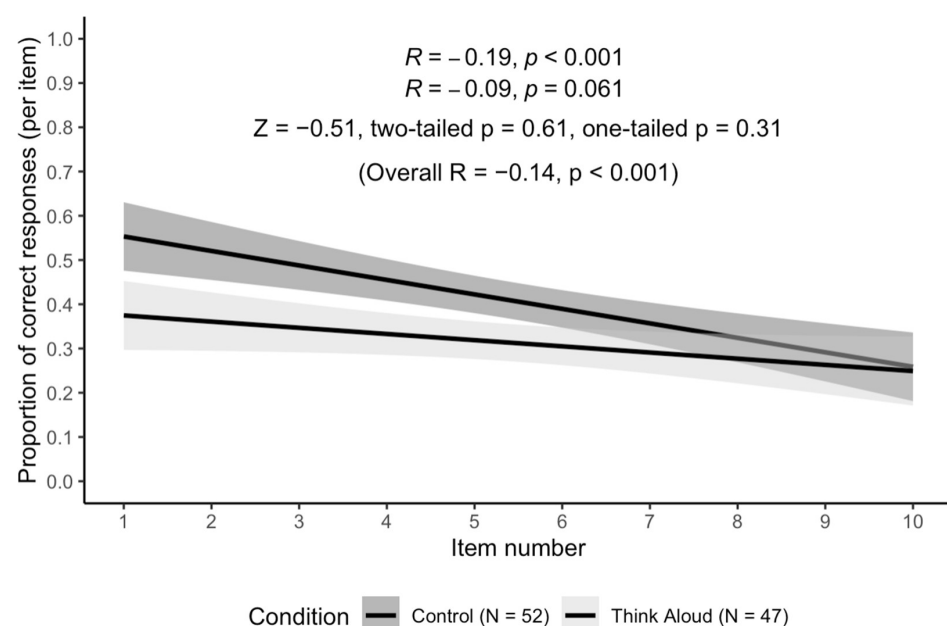
**Pernyataan Persetujuan Berdasarkan Informasi:** "Lembar Informasi" direkomendasikan sebagai pengganti formulir persetujuan oleh Kantor Perlindungan Subjek Manusia di Florida State University (STUDY00000697, 2019-12-13: <https://osf.io/ax85v>).

**Pernyataan Ketersediaan Data:** Semua file data dan analisis tersedia di <https://osf.io/rk3jq/>.

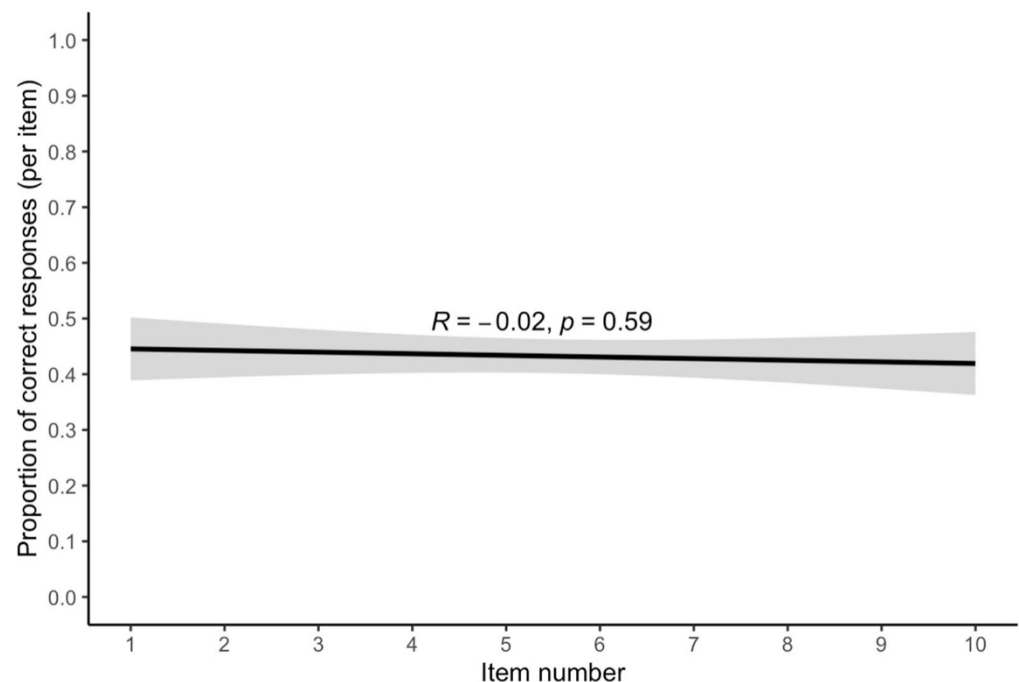
**Ucapan terima kasih:** Terima kasih kepada almarhum Anders Ericsson yang telah memberikan saran dalam pengembangan instruksi peserta serta para mahasiswa dalam seminar akhir *Analisis Keahlian & Protokol* Ericsson atas umpan baliknya terhadap presentasi awal penelitian ini. Terima kasih juga kepada Jeff Larsen, Sunny Khemlani, Megan Peters, Kelompok Penelitian Kognitif di University of Surrey, Pusat Penelitian tentang Diri dan Identitas di University of Southampton, Departemen Ilmu Sosial dan Ilmu Pengambilan Keputusan di Carnegie Mellon University, dan peserta dari Masyarakat 2022 untuk Psikologi Kepribadian dan Sosial, Masyarakat 2022 untuk Penghakiman dan Pengambilan Keputusan, Penelitian Perilaku Daring 2021, dan konferensi Asosiasi Ilmu Psikologi 2021 atas umpan baliknya terhadap presentasi proyek ini.

**Konflik Kepentingan:** Layanan survei audio telepon disediakan sebagai imbalan untuk pengujian beta platform dan fitur-fiturnya (Studi 2).

## Lampiran A



**Gambar A1.** Tes efek latihan dalam Studi 1 ( $N = 99$ ). Performa tidak meningkat tetapi menurun selama survei. Besarnya penurunan ini tidak berbeda secara dramatis di antara kondisi-kondisi ( $Z$ -score). Pita kesalahan menunjukkan kesalahan standar.



**Gambar A2.** Uji efek latihan dalam Studi 2 (N = 102). Performa tidak meningkat atau menurun selama survei. Pita kesalahan menunjukkan kesalahan standar.

#### *Petunjuk Pikirkan dengan Keras*

Dalam percobaan ini kami tertarik dengan apa yang Anda pikirkan ketika Anda menemukan jawaban atas beberapa pertanyaan yang akan saya minta untuk Anda jawab. Untuk melakukan hal ini, kami meminta Anda untuk BERPIKIR KERAS saat mengerjakan soal yang diberikan.

Yang saya maksud dengan 'berpikir dengan keras' adalah saya ingin Anda memberi tahu saya SEMUA yang Anda pikirkan sejak pertama kali Anda melihat pertanyaan sampai Anda memberikan jawaban. Saya ingin Anda berbicara dengan suara lantang secara KONSTAN sejak Anda memulai survei hingga Anda menyelesaikannya.

Tolong jangan rencanakan apa yang Anda katakan atau coba jelaskan apa yang Anda katakan. Bersikaplah seolah-olah Anda sedang sendirian di dalam ruangan dan berbicara kepada diri sendiri.

Yang paling penting adalah Anda harus terus berbicara. Jika Anda diam dalam jangka waktu yang lama, saya akan meminta Anda untuk berbicara.

Apakah Anda memahami apa yang harus Anda lakukan?

Bagus, sekarang Anda bisa berlatih berpikir dengan suara keras pada contoh soal ini.

[Pastikan mereka membaca dengan suara keras (termasuk kode) dan terus berpikir dengan suara keras saat berlatih] Baik. Sekarang Anda akan memulai survei.

[Gunakan frasa seperti 'ingatlah untuk mengucapkan pikiran Anda dengan lantang' jika mereka terdiam selama beberapa detik.] [Ucapkan terima kasih kepada mereka setelah mereka selesai.]

#### **Tes Refleksi Kognitif Verbal yang Dimodifikasi (Sirota et al. 2021)**

- (1) Ayah Mary memiliki 5 anak perempuan tetapi tidak memiliki anak laki-laki-Nana, Nene, Nini, Nono. Siapa nama anak perempuan kelima yang mungkin?

[jawaban benar: Mary, jawaban terpancing:

Nunu] [jeda halaman]

"Apakah Anda ingat pernah berpikir bahwa 'Nunu' bisa menjadi jawabannya?"

Ya Tidak

- (2) Jika Anda sedang mengikuti perlombaan lari, dan Anda melewati orang yang berada di posisi ke-2, di posisi manakah Anda sekarang?

[jawaban benar: ke-2, jawaban terpancing: ke-1]

[Page break]

"Apakah Anda ingat pernah berpikir bahwa '1' bisa menjadi jawabannya?"

Ya Tidak

- (3) Saat itu adalah malam yang penuh badai dan sebuah pesawat lepas landas dari bandara JFK di New York. Badai semakin parah, dan pesawat jatuh-separuh mendarat di Amerika Serikat, separuhnya lagi mendarat di Kanada. Di negara manakah Anda menguburkan para korban yang selamat?  
[jawaban yang benar: jangan mengubur korban, jawaban yang memikat: jawaban tentang lokasi penguburan] [jeda halaman]  
"Apakah Anda ingat pernah berpikir bahwa penguburan korban adalah sebuah pilihan?"  
Ya Tidak
- (4) Seekor monyet, tupai, dan burung berlomba untuk mencapai puncak pohon kelapa. Siapa yang akan mendapatkan pisang lebih dulu, monyet, tupai, atau burung?  
[jawaban benar: tidak ada pisang di pohon kelapa, jawaban pancingan: salah satu hewan] [jeda halaman]  
"Apakah Anda pernah berpikir bahwa 'burung', 'tupai', atau 'monyet' bisa menjadi jawabannya?"  
Ya Tidak
- (5) Di sebuah rumah satu lantai berwarna merah muda, ada orang merah muda, kucing merah muda, ikan merah muda, komputer merah muda, kursi merah muda, meja merah muda, telepon merah muda, kamar mandi merah muda-semuanya berwarna merah muda! Apa warna tangga yang mungkin ada di sana?  
[jawaban yang benar: rumah satu lantai mungkin tidak memiliki tangga, jawaban yang memikat: warna merah muda] [Jeda halaman]  
"Apakah Anda ingat pernah berpikir bahwa 'merah muda' bisa menjadi jawabannya?"  
Ya Tidak
- (6) Berapa banyak dari setiap binatang yang Musa taruh di atas bahtera? [jawaban yang benar: tidak ada; jawaban yang terpancing: dua]  
[Page break]  
"Apakah Anda ingat pernah berpikir bahwa 'dua' bisa menjadi jawabannya?"  
Ya Tidak
- (7) Angin berhembus ke arah barat. Sebuah kereta listrik berjalan ke arah timur. Ke arah mata angin manakah asap dari lokomotif berhembus?  
[jawaban yang benar: tidak ada asap dari kereta listrik, jawaban pancingan: barat] [jeda halaman]  
"Apakah Anda ingat pernah berpikir kapan saja lokomotif akan mengeluarkan asap?"  
Ya Tidak
- (8) Jika Anda hanya memiliki satu korek api dan Anda masuk ke dalam ruangan gelap di mana terdapat lampu minyak, koran dan kayu-yang mana yang akan Anda nyalakan terlebih dahulu?  
[jawaban benar: korek api, jawaban terpancing: lampu minyak] [Jeda halaman]  
"Apakah Anda pernah berpikir bahwa 'lampu minyak', 'koran', atau 'kayu' bisa menjadi jawabannya?"  
Ya Tidak
- (9) Apakah etis jika seorang pria menikahi saudara perempuan dari jandanya? [jawaban yang benar: tidak mungkin, jawaban pancingan: ya, tidak]  
[Page break]  
"Apakah Anda ingat pernah berpikir bahwa seorang pria dapat menikahi saudara perempuan dari jandanya?"  
Ya Tidak
- (10) Kalimat mana yang benar: (a) "kuning telur berwarna putih" atau (b) "kuning telur berwarna kuning"?  
[jawaban yang benar: kuning telur berwarna kuning, jawaban terpancing: b] [Jeda halaman]  
"Apakah Anda ingat pernah berpikir bahwa 'a' atau 'b' bisa menjadi jawabannya?" Ya Tidak

### Penilaian dan Pengkodean

- Kondisi: Kondisi kontrol = 0. Kondisi Think Aloud = 1.
- Pengkodean standar: Jumlahkan jawaban yang benar dan jawaban yang dipancing untuk parameter reflektif dan tidak reflektif.
- Dari transkrip kondisi think-aloud, buat variabel untuk (a) apakah setiap peserta mempertimbangkan kembali setiap respons awal mereka, (b) apakah setiap peserta menyampaikan alasan untuk atau menentang respons apa pun, dan (c) apakah peserta menyebutkan bahwa mereka sudah terbiasa dengan pertanyaan tes refleksi verbal.

### Referensi

- Attali, Yigal, dan Maya Bar -Hillel. 2020. Daya Tarik Palsu dari Iming-iming Cepat. *Penghakiman dan Pengambilan Keputusan* 15: 93 - 111. [CrossRef]
- Bago, Bence, dan Wim De Neys. 2019. Sistem Cerdas 1: Bukti untuk Sifat Intuitif dari Respon yang Benar pada Masalah Pemukul dan Bola. *Berpikir & Penalaran* 25: 257 - 99. [CrossRef]
- Ball, Linden J., dan Alexandra Stevens. 2009. Bukti untuk Komponen Analitik Berbasis Verbal untuk Pemecahan Masalah Wawasan. *Prosiding dari Konferensi Tahunan Ketiga Puluh Satu Masyarakat Ilmu Pengetahuan Kognitif* 31: 1060-65.
- Białek, Michał, dan Gordon Pennycook. 2018. Tes Refleksi Kognitif Kuat terhadap Berbagai Paparan. *Metode Penelitian Perilaku* 50: 1953-59. [CrossRef]
- Blech, Christine, Robert Gaschler, dan Merim Bilalic'. 2020. Mengapa Orang Gagal Melihat Solusi Sederhana? Menggunakan Protokol Think-Aloud untuk Mengungkap Mekanisme di Balik Efek Einstellung (Mental Set). *Berpikir & Penalaran* 26: 552 - 80. [CrossRef]
- Buric', Roman, dan Jakub Šrol. 2020. Perbedaan Individu dalam Intuisi Logis pada Masalah Penalaran yang Disajikan di bawah Paradigma Dua Respon. *Jurnal Psikologi Kognitif* 32: 460 - 47. [CrossRef]
- Byrd, Nick. 2019. Apa yang Dapat (dan Tidak Dapat) Kita Simpulkan tentang Bias Implisit dari Eksperimen Debiasing. *Synthese* 198: 1427 - 55. [CrossRef]
- Byrd, Nick. 2022a. Penjelasan Dua Faktor dari "Refleksi": Menyatukan, Memahami, dan Memandu Filsafat dan Ilmu Pengetahuan Penalaran Reflektif. hlm. 1 - 21. Tersedia secara online: <https://researchgate.net/publication/370131881> (diakses pada 21 Maret 2023).
- Byrd, Nick. 2022b. Semua Ukuran Tidak Diciptakan Sama: Tes Refleksi, Berpikir Keras, dan Protokol Disosiasi Proses. Tersedia online: <https://researchgate.net/publication/344207716> (diakses pada 21 Maret 2023).
- Byrd, Nick. 2022c. Reflektivisme terikat & Identitas Epistemik. *Metafilosofi* 53: 53 - 69. [CrossRef]
- Byrd, Nick. 2022d. Pemikir Besar Tidak Berpikir Sama: Pandangan Para Filsuf Diprediksi oleh Refleksi, Pendidikan, Kepribadian, dan Perbedaan Demografis Lainnya. *Tinjauan Filsafat dan Psikologi. Jurnal Psikologi*, Vol. 2, No. 2, Desember 2018, hlm. 1 - 4. [CrossRef]
- Cullen, Simon, Nick Byrd, dan Philipp Chapkovski. 2022. Eksperimen dalam Keseimbangan Reflektif Menggunakan Platform Socrates. Makalah dipresentasikan di Remotely to Reflection on Intelligent Systems: Menuju Definisi Lintas Disiplin, Stuttgart, Jerman, Oktober 20-21; Tersedia secara online: <https://researchgate.net/publication/370132037> (diakses pada 21 Maret 2023).
- Erceg, Nikola, Zvonimir Galic, dan Mitja Ružojc'ic'. 2020. Sebuah Refleksi tentang Refleksi Kognitif-Menguji Validitas Konvergen dari Dua Versi Tes Refleksi Kognitif. *Penghakiman & Pengambilan Keputusan* 15: 741 - 55. [CrossRef]
- Ericsson, K. Anders. 2003. Verbalisasi Pikiran yang Valid dan Non-Reaktif Selama Pelaksanaan Tugas Menuju Solusi terhadap Masalah Sentral Introspeksi sebagai Sumber Data Ilmiah. *Jurnal Studi Kesadaran* 10: 1-18.
- Ericsson, K. Anders, dan Herbert A. Simon. 1980. Laporan Verbal sebagai Data. *Tinjauan Psikologis* 87: 215 - 51. [CrossRef]
- Ericsson, K. Anders, dan Herbert A. Simon. 1993. *Analisis Protokol: Laporan Verbal sebagai Data*, edisi revisi. Cambridge: Bradford Books/MIT Press.
- Evans, Jonathan. 2007. Tentang Resolusi Konflik dalam Teori Proses Ganda Penalaran. *Berpikir & Penalaran* 13: 321 - 39. [CrossRef]
- Evans, Jonathan, dan Keith E. Stanovich. 2013. Teori Proses Ganda Kognisi Tinggi Memajukan Perdebatan. *Perspektif tentang Ilmu Psikologi* 8: 223 - 41. [CrossRef] [PubMed] [PubMed]
- Evans, Jonathan, Julie L. Barston, dan Paul Pollard. 1983. Tentang Konflik antara Logika dan Kepercayaan dalam Penalaran Silogistik. *Memori & Kognisi* 11: 295 - 306. [CrossRef]
- Fox, Mark C., K. Anders Ericsson, dan Ryan Best. 2011. Apakah Prosedur Pelaporan Pemikiran Verbal Harus Bersifat Reaktif? Sebuah Meta-Analisis dan Rekomendasi untuk Metode Pelaporan Terbaik. *Buletin Psikologi* 137: 316 - 44. [CrossRef]
- Frankish, Keith. 2010. Teori Penalaran Proses Ganda dan Sistem Ganda. *Kompas Filsafat* 5: 914-26. [CrossRef]
- Frederick, Shane. 2005. Refleksi Kognitif dan Pengambilan Keputusan. *Jurnal Perspektif Ekonomi* 19: 25 - 42. [CrossRef]
- Ghebreyesus, Tedros Adhanom. 2020. Sambutan Pembukaan Direktur Jenderal WHO pada Media Briefing tentang COVID-19 pada 11 Maret. Tersedia secara daring: [who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19-11-maret-2020](https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19-11-maret-2020) (diakses pada tanggal 21 Maret 2023).
- Howarth, Stephanie, dan Simon Handley. 2016. Bias Kepercayaan, Tingkat Dasar dan Penilaian Moral: Mengevaluasi Kembali Akun Proses Ganda Interversionis Default. Dalam *Berpikir Pikiran*. Diedit oleh Niall Galbraith, Erica Lucas, dan David Over. New York: Taylor & Francis, pp. 97-111.
- Isler, Ozan, Onurcan Yilmaz, dan Burak Dogruyol. 2020. Mengaktifkan Pemikiran Reflektif dengan Pelatihan Pembeneran Keputusan dan Debiasing. *Penghakiman dan Pengambilan Keputusan* 15: 926 - 38. [CrossRef]



- Kahneman, Daniel, dan Shane Frederick. 2002. Keterwakilan Ditinjau Kembali: Substitusi Atribut dalam Penilaian Intuitif. Dalam *Heuristik dan Bias: Psikologi Penilaian Intuitif*. Diedit oleh Thomas Gilovich, Dale W. Griffin dan Daniel Kahneman. New York: Cambridge University Press, hal. 49-81.
- Landis, J. Richard, dan Gary G. Koch. 1977. Pengukuran Kesepakatan Pengamat untuk Data Kategorikal. *Biometrika* 33: 159 - 174. [CrossRef]
- Machery, Edouard. 2021. Keyakinan yang keliru pada data. *Jurnal Eropa untuk Filsafat Ilmu Pengetahuan* 11: 34. [CrossRef]
- Markovits, Henry, Pier-Luc de Chantal, Janie Brisson, Éloise Dubé, Valerie Thompson, dan Ian Newman. 2021. Strategi Penalaran Memprediksi Penggunaan Penalaran Logis yang Sangat Cepat. *Memori & Kognisi* 49: 532 - 43. [CrossRef]
- Kantor Perlindungan Subjek Manusia. 2020. Penghentian Sementara Beberapa Penelitian Subjek Manusia FSU. *Berita Universitas Negeri Florida (blog)*. Maret 23. Tersedia secara daring: <https://news.fsu.edu/announcements/covid-19/2020/03/23/temporary-cessation-to-some-fsu-human-subjects-research/> (diakses pada 23 Maret 2020).
- Palan, Stefan, dan Christian Schitter. 2018. Prolific.Ac-Sebuah Kumpulan Subjek untuk Eksperimen Online. *Jurnal Perilaku dan Eksperimental Keuangan* 17: 22 - 27. [CrossRef]
- Peer, Eyal, Laura Brandimarte, Sonam Samat, dan Alessandro Acquisti. 2017. Beyond the Turk: Platform Alternatif untuk Crowdsourcing Penelitian Perilaku. *Jurnal Psikologi Sosial Eksperimental* 70: 153 - 163. [CrossRef]
- Pennycook, Gordon, James Allan Cheyne, Derek J. Koehler, dan Jonathan A. Fugelsang. 2015a. Apakah Tes Refleksi Kognitif adalah Ukuran Refleksi dan Intuisi? *Metode Penelitian Perilaku* 48: 341 - 48. [CrossRef]
- Pennycook, Gordon, Jonathan A. Fugelsang, dan Derek J. Koehler. 2015b. Apa yang Membuat Kita Berpikir? Model Proses Ganda Tiga Tahap dari Keterlibatan Analitik. *Psikologi Kognitif* 80: 34 - 72. [CrossRef]
- Perrodin, David D, dan Richard Watson Todd. 2021. Pilihan dalam Mengumpulkan Data Kualitatif Secara Asinkron: Beralih dari Respons Tertulis ke Respons Lisan untuk Pertanyaan Terbuka. *DRAL4* 2021: 11. Tersedia secara online: <https://sola.pr.kmutt.ac.th/dral2021/wp-content/uploads/2022/06/3.pdf> (diakses pada 21 Maret 2023).
- Petitmengin, Claire, Anne Remillieux, Béatrice Cahour, dan Shirley Carter-Thomas. 2013. Kesenjangan dalam Temuan Nisbett dan Wilson? A Akses Orang Pertama ke Proses Kognitif Kita. *Kesadaran dan Kognisi* 22: 654 - 69. [CrossRef]
- Phonic Inc. 2020. Survei yang Dapat Anda Jawab dengan Suara Anda. Tersedia secara online: [phonic.ai](https://phonic.ai) (diakses pada 21 Februari 2020).
- Purcell, Zoe A., Colin A. Wastell, dan Naomi Sweller. 2021. Pengalaman Spesifik Domain dan Pemikiran Proses Ganda. *Berpikir & Penalaran* 27: 239 - 277. [CrossRef]
- Schooler, Jonathan W., Stellan Ohlsson, dan Kevin Brooks. 1993. Pemikiran di Balik Kata-kata: Ketika Bahasa Membayangi Wawasan. *Jurnal Psikologi Eksperimental: Umum* 122: 166 - 83. [CrossRef]
- Shea, Nicholas, dan Chris D. Frith. 2016. Teori Proses Ganda dan Kesadaran: Kasus untuk Kognisi "Tipe Nol". *Ilmu Saraf Kesadaran* 2016: niw005. [CrossRef] [PubMed]
- Simmons, Joseph P., Leif D. Nelson, dan Uri Simonsohn. 2013. Kehidupan setelah Peretasan-P. Makalah dipresentasikan pada Pertemuan Masyarakat Psikologi Kepribadian dan Sosial, New Orleans, LA, AS, 17-19 Januari; hal. 38. Tersedia secara online: <https://papers.ssrn.com/abstract=2205186> (diakses pada 21 Maret 2023).
- Sirota, Miroslav, Lenka Kostovicová, Marie Juanchich, Chris Dewberry, dan Amanda Claire Marshall. 2021. Mengukur Refleksi Kognitif tanpa Matematika: Mengembangkan dan Memvalidasi Tes Refleksi Kognitif Verbal. *Jurnal Pengambilan Keputusan Perilaku* 34: 322 - 43. [CrossRef]
- Sobkow, Agata, Angelika Olszewska, dan Miroslav Sirota. 2023. Struktur Faktor Refleksi Kognitif, Numerasi, dan Kecerdasan: Bukti dari Adaptasi Polandia dari CRT Verbal. *Jurnal Pengambilan Keputusan Perilaku* 36: e2297. [CrossRef]
- Stagnaro, Michael N., Gordon Pennycook, dan David G. Rand. 2018. Kinerja pada Tes Refleksi Kognitif Stabil di seluruh Waktu. *Penghakiman dan Pengambilan Keputusan* 13: 260 - 67. [CrossRef]
- Stanovich, Keith E. 2018. Kekikiran dalam Kognisi Manusia: Interaksi Deteksi, Pengabaian, dan Mindware. *Pemikiran & Penalaran* 24: 423-44. [CrossRef]
- Stieger, Stefan, dan Ulf-Dietrich Reips. 2016. Batasan Tes Refleksi Kognitif: Keakraban. *PeerJ* 4: e2395. [CrossRef]
- Stromer-Galley, Jennifer. 2007. Mengukur Konten Musyawarah: Sebuah Skema Pengkodean. *Jurnal Musyawarah Publik* 3. [CrossRef]
- Stuppel, Edward JN, Melanie Pitchford, Linden J. Ball, Thomas E. Hunt, dan Richard Steel. 2017. Lebih lambat tidak selalu lebih baik: Bukti Waktu Respons Memperjelas Peran Terbatas Pemrosesan Informasi yang Kikir dalam Tes Refleksi Kognitif. *PLoS ONE* 12: e0186404. [CrossRef]
- Szaszi, Barnabas, Aba Szollosi, Bence Palfi, dan Balazs Aczel. 2017. Tes Refleksi Kognitif Ditinjau Kembali: Menjelajahi Cara Individu Menyelesaikan Tes. *Berpikir & Penalaran* 23: 207 - 34. [CrossRef]
- Thompson, Valerie A., dan Stephen C. Johnson. 2014. Konflik, Metakognisi, dan Pemikiran Analitik. *Berpikir & Penalaran* 20: 215 - 214. [CrossRef]
- Thompson, Valerie A., Jamie A. Prowse Turner, dan Gordon Pennycook. 2011. Intuisi, Nalar, dan Metakognisi. *Kognitif Psikologi* 63: 107 - 40. [CrossRef] [PubMed] [Beasiswa]
- Thompson, Valerie A., Jonathan Evans, dan Jamie ID Campbell. 2013. Bias Pencocokan pada Tugas Seleksi: Cepat dan Terasa Nyaman. *Berpikir & Penalaran* 19: 431 - 52. [CrossRef]

Toplak, Maggie E., Richard F. West, dan Keith E. Stanovich. 2014. Menilai Pemrosesan Informasi Kikir: Perluasan dari Tes Refleksi Kognitif. *Berpikir & Penalaran* 20: 147 - 168. [[CrossRef](#)]

Wilson, Timothy, dan Richard E. Nisbett. 1978. Keakuratan Laporan Verbal Tentang Efek Rangsangan pada Evaluasi dan Perilaku. *Psikologi Sosial* 41: 118 - 31. [[CrossRef](#)]

**Penafian/Catatan Penerbit:** Pernyataan, opini, dan data yang terkandung dalam semua publikasi adalah milik masing-masing penulis dan kontributor dan bukan milik MDPI dan/atau editor. MDPI dan/atau editor tidak bertanggung jawab atas segala cedera pada orang atau properti yang diakibatkan oleh ide, metode, instruksi, atau produk apa pun yang dirujuk dalam konten.