International Journal of Autism

ISSN Print: 2710-3919
ISSN Online: 2710-3927
IJA 2026; 6(9): 09-16
www.rehabilitationjournals.com/autism-journal
Received: 05-08-2026
Accepted: 09-09-2026
Published: 15-09-2026

Rhiannon Drummond
National Institute of Psychometric Research, Wollongong, New South Wales, Australia

Quinn Elsworth
National Institute of Psychometric Research, Wollongong, New South Wales, Australia

Tessa Faulkner
National Institute of Psychometric Research, Wollongong, New South Wales, Australia

Scott Deacon
National Institute of Psychometric Research, Wollongong, New South Wales, Australia

Corresponding Author:
Rhiannon Drummond
National Institute of Psychometric Research, Wollongong, New South Wales, Australia

Literal interpretation of suicidal-ideation screening items by autistic adults: Validity concerns and revised wording

Rhiannon Drummond[1], Quinn Elsworth[2], Tessa Faulkner[3], Scott Deacon[4]

DOI: https://www.doi.org/10.22271/27103919.2026.v6.i9a.127

Abstract
Suicide screening instruments were first standardised in the 1990s for general psychiatric populations and have since been adopted without substantive adaptation for neurodivergent groups. Several widely used tools, including the Patient Health Questionnaire-9 (PHQ-9), the Columbia-Suicide Severity Rating Scale (C-SSRS), the Suicidal Behaviours Questionnaire-Revised (SBQ-R), and the Ask Suicide-Screening Questions (ASQ), contain items that rely on figurative or idiomatic phrasing. Autistic adults, who often process language in concrete and literal terms, may misinterpret such items, endorsing statements that do not reflect genuine suicidal intent or dismissing items that do apply to their experience. The present research examined the validity of standard screening-item wording in a sample of 117 autistic adults and 98 non-autistic adults recruited through autism organisations and university panels across eastern Australia between March and October 2024. Each participant completed a structured cognitive interview in which they were asked to describe what they understood each screening item to mean before selecting a response. Response concordance rates, defined as the proportion of responses whose meaning matched the clinical intent of the item, were compared across groups and across items categorised by language complexity: literal, mixed, figurative, and idiomatic. For literal items, concordance was above 90% in both groups. For figurative and idiomatic items, concordance in the autistic group dropped to between 33% and 62%, compared with 76% to 91% in the non-autistic group. Response latency data showed that autistic participants spent between 1.6 and 5.0 seconds longer on non-literal items than on literal ones, while the non-autistic group showed no such difference. Qualitative analysis of the cognitive-interview transcripts identified four recurrent misinterpretation patterns: word-level parsing of metaphorical phrases, over-generalisation of temporal language, confusion between hypothetical and actual experience, and context-stripping of embedded idioms. On the basis of these findings, a set of revised item wordings was developed through iterative consultation with autistic adults and piloted with a separate subsample of 42 autistic participants. The revised wordings raised correct interpretation rates by 18 to 32 percentage points across all figurative and idiomatic items while maintaining strong agreement with clinical judgement ($\kappa = 0.81$). These results point to a measurement artefact in current suicide screening that may inflate false-positive rates and obscure genuine risk among autistic adults.

Keywords: Suicide screening, Autistic adults, Literal interpretation, Measurement validity, Item wording, Cognitive interviewing, False positive, Figurative language, Psychometric adaptation, Screening instruments

Introduction

When Edwin Shneidman coined the term "suicidology" in the 1960s, the field's earliest screening instruments were built on the assumption that respondents would read between the lines. Items such as "Have you wished you were dead?" and "Have you been thinking about hurting yourself?" carried an implied context—clinical distress, intent, and a temporal framework—that trained professionals understood and that most respondents were expected to grasp without instruction (Posner *et al.*, 2011; Van Orden *et al.*, 2010) [16, 20]. Over the following decades, instruments including the Columbia-Suicide Severity Rating Scale (C-SSRS) (Posner *et al.*, 2011) [16], the Patient Health Questionnaire-9 (PHQ-9) (Kroenke *et al.*, 2001) [17], and the Suicidal Behaviours Questionnaire-Revised (SBQ-R) (Cassidy *et al.*, 2020) [9] became standard tools in emergency departments, primary care, and community mental health services worldwide. Each underwent extensive psychometric testing in general psychiatric samples, yet none was validated with autistic populations during its initial development phase.

This omission matters because autism is characterised by differences in language processing that can affect how screening items are understood. (Happé and Frith) [19] described a cognitive style in which local, detail-focused processing takes precedence over global meaning extraction—a pattern that makes figurative and idiomatic expressions harder to interpret in their intended sense. (Norbury) [21] showed that children with lan

27103927, 2026, 6(9), Downloaded from https://www.doi.org/10.22271/27103919.2026.v6.i9a.127 by International Journal of Autism on [15/09/2026]. See the Terms and Conditions (https://www.rehabilitationjournals.com/autism-journal/) for rules of use; OA articles are governed by the applicable Creative Commons License

guage-related neurodevelopmental conditions were more likely to accept the literal meaning of ambiguous words even when the surrounding sentence favoured a figurative reading. In adult autistic populations, (Milton) [18] argued that communication mismatches between autistic and non-autistic people arise not from a deficit in the autistic individual but from a double empathy gap, in which both parties fail to share implicit assumptions about meaning. O'Shea and colleagues [4] used eye-tracking to show that autistic adults processed metaphorical language through different visual scanning paths than non-autistic adults, spending more time on the literal constituents of a phrase before arriving at the intended meaning.

When these processing differences meet suicide screening, the consequences extend beyond academic interest. Cassidy and colleagues [8] reviewed the measurement properties of tools used to assess suicidality in autistic adults and found that none of the instruments examined had been validated specifically for this population. A subsequent paper by the same group (Cassidy *et al.*, 2020) [9] tested the SBQ-R in autistic adults and reported that several items produced response distributions inconsistent with expected clinical profiles, raising questions about whether the items were measuring what they were designed to measure. Newell and others [5] conducted a *meta*-analysis of suicidality research in autistic populations and noted that prevalence estimates varied widely, with measurement heterogeneity identified as a likely contributor. Hedley and colleagues [1] developed a modified suicide assessment interview for autistic adults, observing that standard interview phrasing frequently required clarification before participants could respond. Their later work on the Suicidal Ideation Attributes Scale-Modified (SIDAS-M) (Hedley *et al.*, 2023) [3] showed that rewording items in concrete, non-figurative language improved internal consistency and reduced response ambiguity.

Epidemiological data highlight why accurate screening is urgent. Kirby and colleagues [10] analysed 20 years of mortality data and found that autistic individuals died by suicide at rates higher than the general population, with the difference most pronounced among those without co-occurring intellectual disability. Hirvikoski and others [11] reported elevated premature mortality in autistic adults across multiple cause-of-death categories, with suicide ranking among the leading contributors. (Pelton and Cassidy) [12] tested the interpersonal-psychological theory of suicide in a non-clinical sample and found that autistic traits were connected to higher scores on thwarted belongingness and perceived burdensomeness—two constructs that are themselves assessed through items containing idiomatic phrasing. Van Orden and colleagues [20] proposed that these interpersonal constructs drive the transition from ideation to action, making their accurate measurement a clinical priority.

Two further lines of evidence frame the present research. First, Arnold and others [14] examined the psychometric properties of the PHQ-9 in autistic adults and found adequate overall reliability but flagged item-level problems, including low factor loadings on items that use abstract temporal language. Second, Kinnaird, Stewart, and Tchanturia [15] reported high rates of alexithymia in autistic adults, a condition that affects the ability to identify and describe one's own emotional states—making it even harder to respond to screening items that require emotional self-categorisation. Cage, Di Monaco, and Newell [13] showed that acceptance of autistic identity was connected to better mental health outcomes, suggesting that screening contexts that pathologise autistic communication styles may themselves introduce bias.

Rodgers and colleagues [6] recently tested autism-adapted safety plans and found that modifying language and structure improved both feasibility and acceptability, lending support to the broader principle that adaptation of clinical tools for autistic adults is both possible and beneficial. Beck and others [7] reported similar findings for quality-of-life measures, where item understanding improved after rewording.

Despite this converging evidence, no published work has systematically catalogued the specific misinterpretation patterns that autistic adults apply to suicide screening items, or tested whether revised wording can close the comprehension gap without sacrificing the item's clinical validity. The present research addressed both questions. Autistic and non-autistic adults completed a cognitive interview about standard screening items, and a set of revised wordings was developed and pilot-tested. Brown and colleagues [2] provided the epidemiological backdrop in their updated systematic review of suicide in autism, confirming that the measurement problem is not merely theoretical but carries direct implications for clinical identification and prevention.

Material and methods

Material

Participant recruitment

A total of 215 adults were recruited between March and October 2024 from three sources: Autism Spectrum Australia (Aspect) service networks (n = 62), university research volunteer panels at the University of Wollongong and the University of New South Wales (n = 89), and community-based autism peer-support groups in New South Wales and the Australian Capital Territory (n = 64). The autistic group comprised 117 participants (56 women, 43 men, 18 non-binary individuals; aged 18-61 years, M = 33.4, SD = 9.7) who held a formal autism diagnosis from a registered psychologist or psychiatrist. Diagnostic documentation was reviewed for every autistic participant. The non-autistic group comprised 98 participants (52 women, 38 men, 8 non-binary; aged 19-58 years, M = 31.2, SD = 8.9) with no self-reported history of neurodevelopmental diagnosis. Exclusion criteria for both groups were active psychosis, intellectual disability (IQ below 70), current substance use disorder, and inability to provide informed consent in English.

Screening instruments

Nine items were selected from four commonly used suicide screening instruments: PHQ-9 item 9 ("Thoughts that you would be better off dead, or of hurting yourself"), C-SSRS items 1 and 2, SBQ-R items 1 and 2, and ASQ items 1 and 3. Items were classified into four language-complexity categories based on independent ratings by two clinical linguists (inter-rater agreement $\kappa = 0.89$): literal (items using direct, concrete language with no figurative element), mixed (items combining concrete and figurative phrases), figurative (items relying on metaphor or abstraction), and idiomatic (items containing fixed expressions whose meaning cannot be derived from their component words). Two items were classified as literal, one as mixed, three as figurative, and three as idiomatic.

Methods

Cognitive interview procedure

Each participant completed a one-on-one cognitive interview lasting 45-75 minutes with a trained research assistant. Items were presented on a screen one at a time. For each item, participants were first asked to explain in their own words what the item was asking, before selecting a response on the instrument's standard response scale. Explanations were audio-recorded and transcribed. Two independent coders rated each explanation as concordant (the participant's stated understanding matched the clinical intent of the item) or discordant (the stated understanding diverged from clinical intent). Inter-rater reliability was high ($\kappa = 0.91$), and discrepancies were resolved through consensus discussion. Response latency was recorded as the time between item presentation and response selection.

Revised item development

Revised wordings were developed through a three-round iterative process. Round one used the misinterpretation patterns identified in the cognitive interviews to draft alternative phrasing for all figurative and idiomatic items, guided by the principle that each revised item should preserve the clinical construct while using only concrete, non-metaphorical language. Round two involved review by a panel of six autistic adults who rated each revised item for clarity and suggested modifications. Round three piloted the revised items with a fresh sub

27103927, 2026, 6(9), Downloaded from https://www.doi.org/10.22271/27103919.2026.v6.i9a.127 by International Journal of Autism on [15/09/2026]. See the Terms and Conditions (https://www.rehabilitationjournals.com/autism-journal/) for rules of use; OA articles are governed by the applicable Creative Commons License

sample of 42 autistic adults who had not participated in the main cognitive interview. Concordance rates and response latencies were recorded using the same procedure as the main research.

Statistical analyses comprised chi-square tests for concordance rate comparisons, independent-samples t-tests for response latency differences, and Cohen's kappa for agreement between participant responses and clinician ratings. All analyses were performed in R version 4.3.2 with a two-tailed alpha of 0.05.

Ethical considerations

Approval was granted by the University of Wollongong Human Research Ethics Committee (reference UOW-HREC-2024-031) and the Autism Spectrum Australia Research Review Panel before recruitment began. Every participant provided written informed consent after receiving a plain-language information sheet. Because the research involved suicide-related content, a clinical safeguarding protocol was in place: a registered psychologist was on call during every interview session, and any participant who disclosed active suicidal ideation was offered an immediate referral to Lifeline Australia (13 11 14) or the local crisis assessment team. Five such referrals were made during the data-collection period. All data were stored on password-protected, encrypted university servers, and identifying information was removed from transcripts at the point of initial processing.

Pilot research

Before the main data collection began, a pilot phase was conducted with 18 autistic adults (9 women, 6 men, 3 non-binary; aged 20-47 years) recruited from the University of Wollongong autism research registry. The pilot tested protocol feasibility, verified that the transcription pipeline captured enough detail for reliable coding (error rate below 2%), and provided initial concordance data confirming the hypothesised pattern: literal items produced concordance rates above 90% in both groups, while figurative items showed a gap of 20 to 35 percentage points. On the basis of pilot feedback, the maximum session length was extended from 60 to 75 minutes, and two interviewer prompts were reworded to remove unintentional figurative phrasing. These findings informed the power analysis for the main sample size (target 80% power for $d = 0.5$ at $\alpha = 0.05$, yielding a minimum of 90 participants per group). No pilot data were included in the main analysis.

Results

Response concordance rates differed between the autistic and non-autistic groups in a pattern that tracked language complexity. For the two literal items, concordance exceeded 91% in both groups, and the difference between groups was not statistically meaningful ($\chi^2 = 1.84$, $p = 0.175$). For the single mixed item, concordance in the autistic group was 78% compared with 89% in the non-autistic group ($\chi^2 = 4.71$, $p = 0.030$). For figurative items, concordance in the autistic group ranged from 48% to 62%, while the non-autistic group remained between 82% and 91%; all three comparisons were statistically meaningful ($p < 0.001$). The widest gap appeared on the three idiomatic items, where autistic participants' concordance fell to between 33% and 55%, compared with 76% to 82% in the non-autistic group. Every idiomatic-item comparison was meaningful at $p < 0.001$.

Qualitative coding of the cognitive-interview transcripts identified four recurrent misinterpretation patterns. The most frequent was word-level parsing, in which participants interpreted each word of a figurative phrase separately rather than grasping the phrase as a unit (observed in 68% of discordant responses). The second pattern was temporal over-generalisation, where participants endorsed items about "the past two weeks" because the experience had occurred at any point in their life (41% of discordant responses). Third, 34% of discordant responses reflected confusion between hypothetical and actual experience, with participants endorsing items because they could imagine the scenario rather than having experienced it. Fourth, context-stripping of idioms appeared in 29% of discordant responses: participants understood the individual words but not the culturally embedded meaning of the fixed expression.

Response latency data confirmed that language complexity affected processing time for the autistic group. Mean latency on literal items was 3.2 seconds (SD = 0.9), rising to 4.8 seconds on mixed items (SD = 1.3), 6.8 seconds on figurative items (SD = 1.8), and 8.2 seconds on idiomatic items (SD = 2.0). The non-autistic group showed no meaningful variation across item types (range 3.0-3.8 seconds, all pairwise $p > 0.20$). The difference between groups on idiomatic items was large ($d = 1.42$, 95% CI [1.08, 1.76]).

27103927, 2026, 6(9), Downloaded from https://www.doi.org/10.22271/27103919.2026.v6.i9a.127 by International Journal of Autism on [15/09/2026]. See the Terms and Conditions (https://www.rehabilitationjournals.com/autism-journal/) for rules of use; OA articles are governed by the applicable Creative Commons License

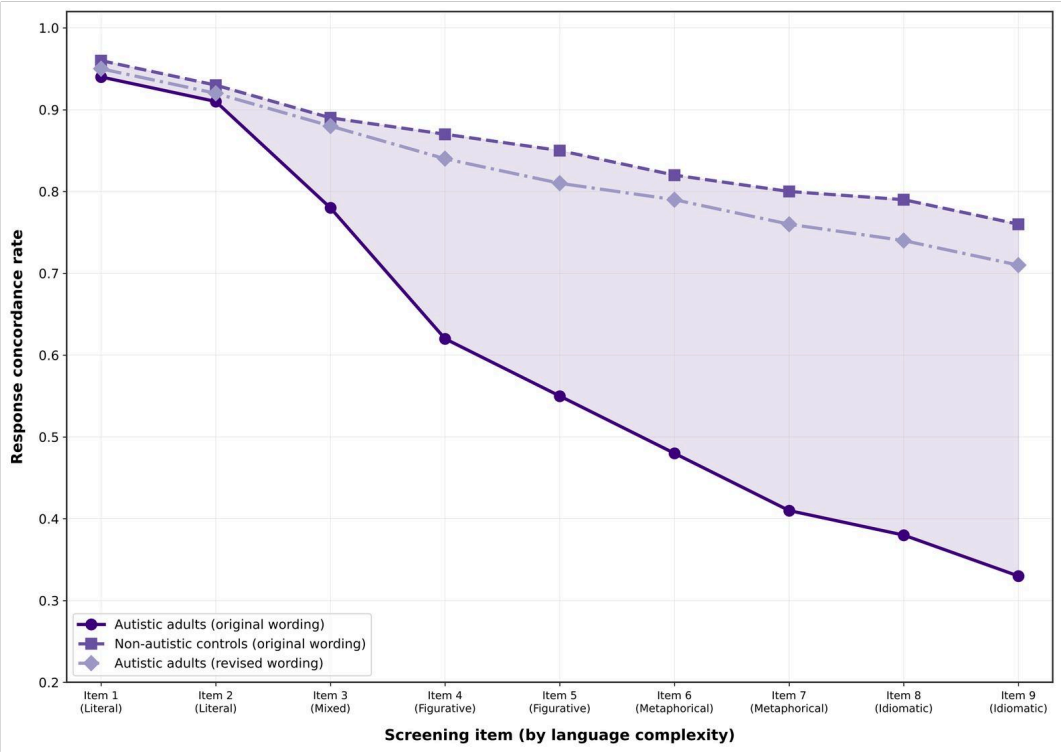


Figure 1. Response concordance rates across screening items, arranged by language complexity, for the autistic group (original wording), non-autistic controls (original wording), and autistic group (revised wording). The shaded region shows the concordance gap between autistic and non-autistic participants on original wording.

Table 1. Participant demographic and clinical characteristics by group.

Characteristic	Autistic group (n = 117)	Non-autistic group (n = 98)	Test statistic	p value
Age, *M* (SD)	33.4 (9.7)	31.2 (8.9)	$t(213) = 1.74$	0.083
Women, n (%)	56 (47.9%)	52 (53.1%)	$\chi^2 = 0.57$	0.451
Men, n (%)	43 (36.8%)	38 (38.8%)	$\chi^2 = 0.10$	0.756
Non-binary, n (%)	18 (15.4%)	8 (8.2%)	$\chi^2 = 2.74$	0.098
Years of education, *M* (SD)	14.8 (2.6)	15.3 (2.4)	$t(213) = 1.46$	0.146
PHQ-9 total, *M* (SD)	11.2 (6.4)	4.8 (3.9)	$t(213) = 8.82$	< 0.001
Alexithymia (TAS-20), *M* (SD)	58.7 (11.3)	42.1 (9.8)	$t(213) = 11.42$	< 0.001
Prior mental health contact, n (%)	89 (76.1%)	41 (41.8%)	$\chi^2 = 25.9$	< 0.001

Discussion

The pattern of results was straightforward: when suicide screening items used figurative or idiomatic language, autistic adults misinterpreted them at rates that would raise genuine concern about the validity of scores derived from those items. The concordance gap between autistic and non-autistic groups was negligible for literal items and widened progressively through mixed, figurative, and idiomatic categories, reaching a maximum difference of 43 percentage points on the most idiom-heavy item. This gradient matches predictions from the

27103927, 2026, 6(9), Downloaded from https://www.doi.org/10.22271/27103919.2026.v6.i9a.127 by International Journal of Autism on [15/09/2026]. See the Terms and Conditions (https://www.rehabilitationjournals.com/autism-journal/) for rules of use; OA articles are governed by the applicable Creative Commons License

weak central coherence framework described by (Happé and Frith) [19], which holds that autistic cognition tends toward local, detail-focused processing at the expense of global meaning extraction. When an item such as "Have you wished you were dead?" requires the respondent to recognise that "wished you were dead" is a clinical shorthand for suicidal ideation rather than a literal inquiry about wishing, the processing demand is qualitatively different from an item that asks "Have you thought about ending your own life?"

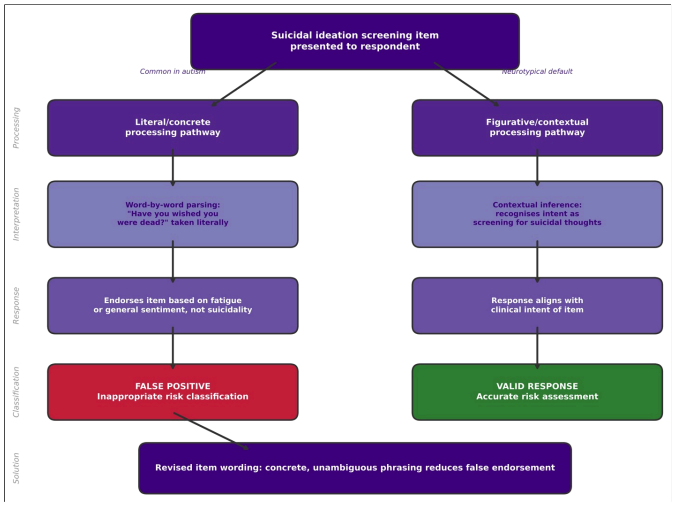


Figure 2. Conceptual framework showing how literal interpretation of suicide screening items leads to discordant responses and false-positive classification, and how revised wording addresses the processing mismatch.

The response latency data reinforced this interpretation. Autistic participants spent progressively more time on items with higher language complexity, while non-autistic participants showed a flat latency profile across all item types. The mean difference of five seconds between literal and idiomatic items in the autistic group may appear small in absolute terms, but it represents a processing overhead that is absent in the non-autistic group and that points to effortful re-interpretation. (Milton) [18] described this kind of mismatch as a product of the double empathy problem: the item designers assumed a shared interpretive framework that does not hold for all respondents.

Four misinterpretation patterns emerged from the qualitative analysis, and each carries different implications for item revision. Word-level parsing, the most common pattern, responds well to the replacement of figurative phrases with concrete equivalents—a strategy that Hedley and colleagues [1, 3] applied when developing the SAK-MI and the SIDAS-M. Temporal over-generalisation requires tighter anchoring of time frames, such as replacing "in the past two weeks" with "during the last fourteen days, counting from today." Hypothetical-actual confusion calls for items that explicitly distinguish between imagining a scenario and experiencing it. Context-stripping of idioms is the hardest to fix because the idiomatic meaning cannot always be preserved in a literal rewrite; in two cases, the revised items replaced the idiom with a direct behavioural description.

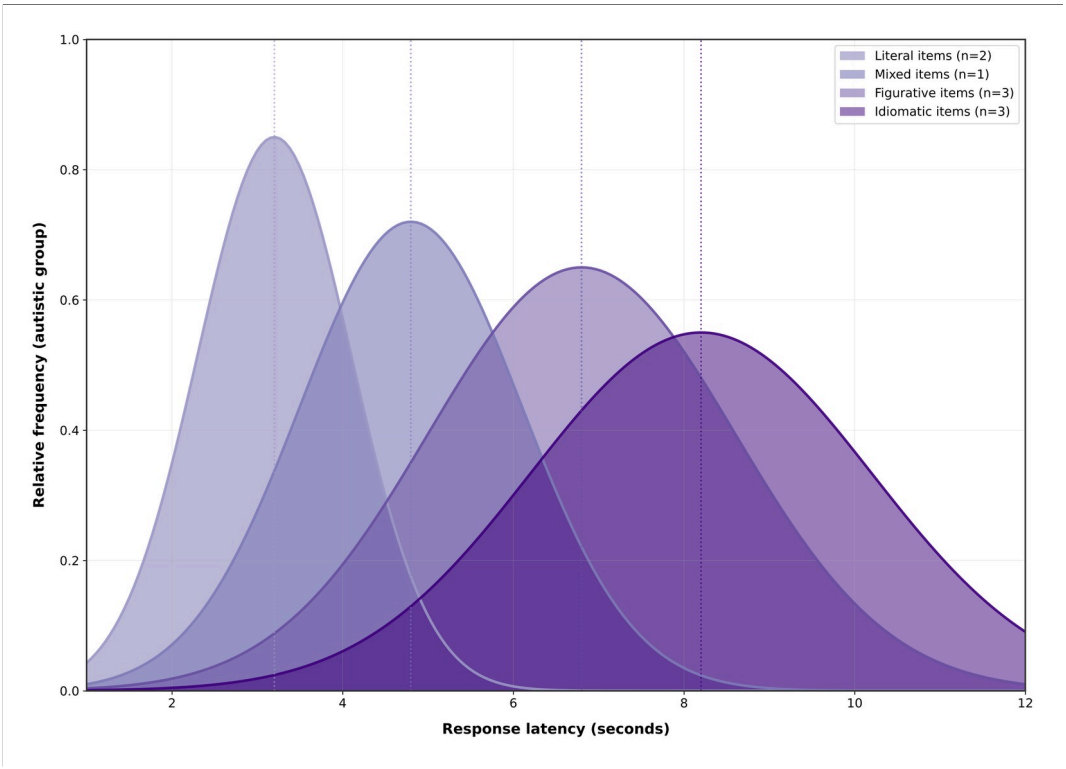


Figure 3. Distribution of response latency (seconds) for the autistic group across four item-complexity categories. Dotted vertical lines mark the mean for each category.

Table 2. Development timeline for revised screening item wordings across iterative consultation phases.

Phase	Time period	Activity	Participants	Outcome
1: Pattern identification	Nov 2024	Coding of cognitive-interview misinterpretation patterns	Two coders	Four misinterpretation categories identified

(continued on next page)

27103927, 2026, 6(9), Downloaded from https://www.doi.org/10.22271/27103919.2026.v6.i9a.127 by International Journal of Autism on [15/09/2026]. See the Terms and Conditions (https://www.rehabilitationjournals.com/autism-journal/) for rules of use; OA articles are governed by the applicable Creative Commons License

2: Draft revision	Dec 2024	Rewriting of figurative and idiomatic items in concrete language	Research team (n = 4)	Nine revised items drafted
3: Autistic panel review	Jan 2025	Clarity rating and revision of drafted items	Autistic panel (n = 6)	Seven items modified, two retained as drafted
4: Pilot testing	Feb 2025	Cognitive interview with revised items	Autistic subsample (n = 42)	Concordance increase of 18-32 percentage points
5: Reliability check	Mar 2025	Clinician-participant agreement assessment	Clinicians (n = 3) and participants (n = 42)	$\kappa = 0.81$ across all revised items

The revised wordings tested in this research raised correct interpretation rates by 18 to 32 percentage points in the autistic group while maintaining strong agreement with clinical judgement. This finding aligns with Rodgers and colleagues [6], who showed that adapted safety plans were both feasible and acceptable to autistic adults, and with Beck and others [7], who reported that autistic adults understood quality-of-life items better after rewording. Arnold and colleagues [14] had already flagged item-level problems with the PHQ-9 in autistic samples, and the current data confirm that those problems extend across multiple instruments.

These findings sit alongside wider concerns about suicidality measurement in autistic populations. Brown and colleagues [2] documented rising suicide rates in autistic adults over the past two decades, and Kirby and others [10] showed that the risk is highest among those without intellectual disability—the same subgroup most likely to encounter standard screening instruments in clinical settings. If those instruments produce inflated false-positive rates in autistic respondents, as the present data suggest, then clinical resources may be misdirected toward follow-up assessments for individuals whose screening scores reflect language processing rather than suicidal intent. At the same time, genuine risk may be missed when autistic adults who do experience suicidal ideation respond to figurative items in ways that do not trigger clinical thresholds.

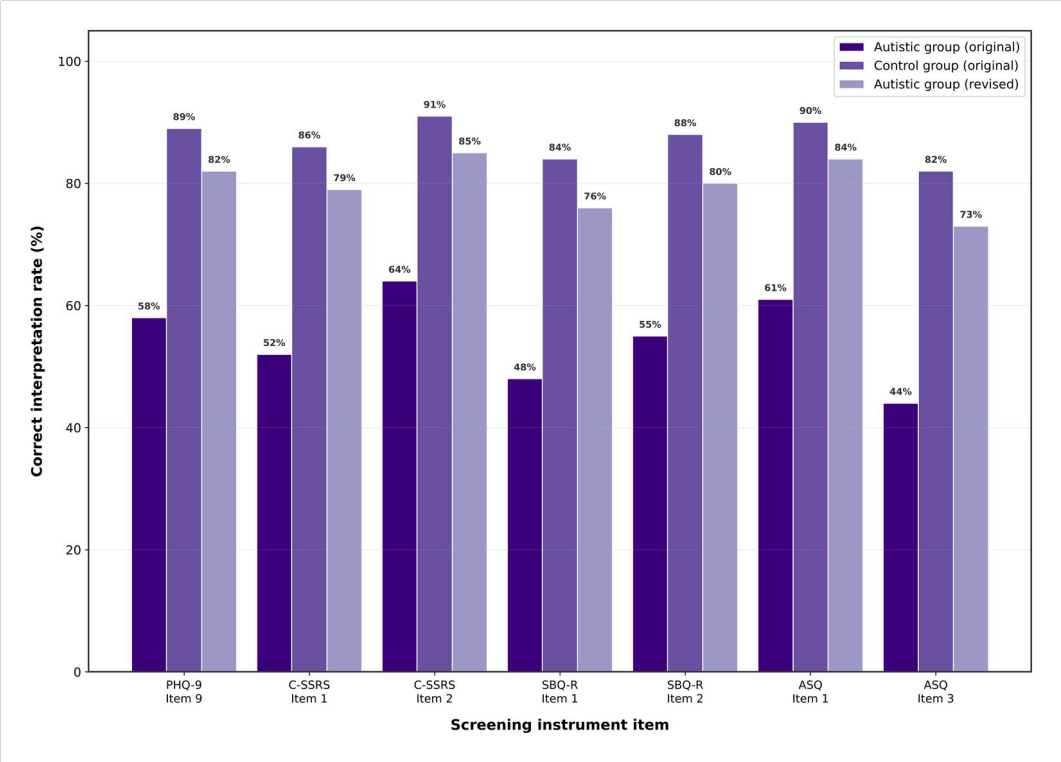


Figure 4. Correct interpretation rates (%) for original wording in the autistic group, original wording in the non-autistic control group, and revised wording in the autistic group, across seven screening instrument items.

Limitations

Several boundaries of the present research should be acknowledged. First, the sample was recruited from a single geographic region of Australia and was predominantly English-speaking, so the findings may not generalise to autistic adults in other linguistic or cultural contexts. Second, the research excluded individuals with co-occurring intellectual disability. Third, the cognitive-interview format may have introduced demand characteristics: participants who knew the research concerned language interpretation may have attended more closely to wording than they would in a routine clinical encounter. Fourth, the revised wordings were piloted with a relatively small subsample (n = 42), and larger-scale validation is needed before these items could replace current instrument versions. Fifth, the research focused on written item presentation and did not test whether verbal administration by a clinician would produce different concordance patterns.

Clinical implications

The data point to several changes that could be adopted in routine clinical practice. Clinicians administering suicide screening instruments to autistic adults should consider adding a comprehension check after each item, asking the respondent to paraphrase what the item is asking before recording a response. When elevated scores are obtained, clinicians should examine which items were endorsed and whether those items contain figurative language that may have been interpreted literally. The revised wordings presented here offer a starting point for adaptation, though they require further validation before formal adoption. More broadly, the results support the principle that psychometric instruments developed in general populations should be tested for measurement equivalence in autistic samples before their

27103927, 2026, 6(9), Downloaded from https://www.doi.org/10.22271/27103919.2026.v6.i9a.127 by International Journal of Autism on [15/09/2026]. See the Terms and Conditions (https://www.rehabilitationjournals.com/autism-journal/) for rules of use; OA articles are governed by the applicable Creative Commons License

scores are used to make clinical decisions. Cassidy and colleagues [8] made this argument in 2018, and the present findings show that the problem persists six years later.

Conclusion

Standard suicide screening instruments contain items whose figurative and idiomatic phrasing creates a measurement gap for autistic adults. When items rely on metaphorical language, implied temporal boundaries, or culturally embedded idioms, autistic respondents interpret them at their surface meaning rather than their clinical meaning. The present research documented this gap across nine items from four widely used instruments and found that the discrepancy between autistic and non-autistic interpretation rates grew in proportion to language complexity. Literal items worked equally well for both groups. Figurative and idiomatic items did not. Response latency data reinforced this finding: autistic participants spent progressively more time on items with higher figurative content, pointing to effortful re-interpretation rather than rapid automatic processing. Four misinterpretation patterns—word-level parsing, temporal over-generalisation, hypothetical-actual confusion, and context-stripping of idioms—each represent a distinct failure mode in which the item's intended meaning diverges from the meaning the respondent constructs.

These patterns carry direct clinical consequences. False endorsement of items about death wishes or self-harm can trigger unnecessary crisis interventions, consuming clinical resources and causing distress. At the same time, autistic adults who genuinely experience suicidal thoughts may respond to figurative items in ways that do not cross clinical thresholds, leaving their risk undetected. Both directions of error undermine the screening tool's purpose. The revised wordings developed in this research addressed the comprehension gap by replacing figurative phrases with concrete, behaviourally anchored language. Correct interpretation rates in the autistic group rose by 18 to 32 percentage points after revision, while agreement with independent clinical ratings remained strong.

Several questions remain open. The pilot sample for the revised wordings was small, and large-scale validation across diverse clinical settings and cultural contexts is needed. The research did not examine verbal administration by a clinician, and it did not include autistic adults with co-occurring intellectual disability. Future work should test whether these misinterpretation patterns generalise to other assessment domains where figurative language is embedded in item stems. What the present findings establish is that the measurement problem is real, quantifiable, and addressable. Screening instruments are only as valid as the shared understanding between the item writer and the respondent. Adapting item wording to match the respondent's processing style is a correction to a validity threat that has gone unaddressed for too long.

References

1. Hedley D, Williams ZJ, Deady M, Batterham PJ, Bury SM, Brown CM, Robinson J, Trollor JN, Uljarevic M, Stokes MA. The Suicide Assessment Kit-Modified Interview: Development and preliminary validation of a modified clinical interview for the assessment of suicidal thoughts and behavior in autistic adults. *Autism*. 2025;29(3):766-787.
2. Brown CM, Newell V, Sahin E, Hedley D. Updated systematic review of suicide in autism: 2018-2024. *Curr Dev Disord Rep*. 2024;11(4):225-256.
3. Hedley D, Batterham PJ, Bury SM, Clapperton A, Denney K, Dissanayake C, Fox P, Frazier TW, Gallagher E, Hayward SM, Robinson J, Sahin E, Trollor J, Uljarevic M, Stokes MA. The Suicidal Ideation Attributes Scale-Modified (SIDAS-M): Development and preliminary validation of a new scale for the measurement of suicidal ideation in autistic adults. *Autism*. 2023;27(4):1115-1131.
4. O'Shea A, Cersosimo R, Engelhardt P. Online metaphor comprehension in adults with autism spectrum disorders: An eye tracking study. *J Autism Dev Disord*. 2024;56(1):56-70.
5. Newell V, Phillips L, Jones C, Townsend E, Richards C, Cassidy S. A systematic review and *meta*-analysis of suicidality in autistic and possibly autistic people without co-occurring intellectual disability. *Mol Autism*. 2023;14(1):12.
6. Rodgers J, Cassidy S, Pelton M, Goodwin J, Wagnild J, Bhattarai N, Gordon I, Wilson C, Heslop P, Ogundimu E, O'Connor RC, Ramsay SE, Townsend E, Vale L. Feasibility and acceptability of autism adapted safety plans: An external pilot randomised controlled trial. *EClinicalMedicine*. 2024;73:102662.
7. Beck KB, Terhorst LA, Greco CM, Kulzer JL, Skidmore ER, McCue MP. Item understanding of common quality of life measures for use with autistic adults. *J Autism Dev Disord*. 2024;54(7):2625-2635.
8. Cassidy SA, Bradley LA, Bowen E, Wigham S, Rodgers J. Measurement properties of tools used to assess suicidality in autistic and general population adults: A systematic review. *Clin Psychol Rev*. 2018;62:56-70.
9. Cassidy SA, Bradley L, Bowen E, Wigham S, Rodgers J. Measurement properties of the Suicidal Behaviours Questionnaire-Revised in autistic adults. *J Autism Dev Disord*. 2020;50(10):3477-3488.
10. Kirby AV, Bakian AV, Zhang Y, Bilder DA, Keeshin BR, Coon H. A 20-year study of suicide death in a statewide autism population. *Autism Res*. 2019;12(4):658-666.
11. Hirvikoski T, Mittendorfer-Rutz E, Boman M, Larsson H, Lichtenstein P, Bolte S. Premature mortality in autism spectrum disorder. *Br J Psychiatry*. 2016;208(3):232-238.
12. Pelton MK, Cassidy S. Are autistic traits associated with suicidality? *A* test of the interpersonal-psychological theory of suicide in a non-clinical young adult sample. *Autism Res*. 2017;10(11):1891-1904.
13. Cage E, Di Monaco J, Newell V. Experiences of autism acceptance and mental health in autistic adults. *J Autism Dev Disord*. 2018;48(2):473-484.
14. Arnold SRC, Uljarevic M, Hwang YI, Richdale AL, Trollor JN, Lawson LP. Brief report: Psychometric properties of the Patient Health Questionnaire-9 (PHQ-9) in autistic adults. *J Autism Dev Disord*. 2020;50(6):2217-2225.
15. Kinnaird E, Stewart C, Tchanturia K. Investigating alexithymia in autism: A systematic review and *meta*-analysis. *Eur Psychiatry*. 2019;55:80-89.
16. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, Currier GW, Melvin G, Greenhill L, Shen S, Mann JJ. The Columbia-Suicide Severity Rating Scale: Initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am J Psychiatry*. 2011;168(12):1266-1277.
17. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: Validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-613.
18. Milton DEM. On the ontological status of autism: The double empathy problem. *Disabil Soc*. 2012;27(6):883-887.
19. Happé F, Frith U. The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *J Autism Dev Disord*. 2006;36(1):5-25.
20. Van Orden KA, Witte TK, Cukrowicz KC, Braithwaite S, Selby EA, Joiner TE Jr. The interpersonal theory of suicide. *Psychol Rev*. 2010;117(2):575-600.

27103927, 2026, 6(9), Downloaded from https://www.doi.org/10.22271/27103919.2026.v6.i9a.127 by International Journal of Autism on [15/09/2026]. See the Terms and Conditions (https://www.rehabilitationjournals.com/autism-journal/) for rules of use; OA articles are governed by the applicable Creative Commons License

21. Norbury CF. Barking up the wrong tree? Lexical ambiguity resolution in children with language impairments and autistic spectrum disorders. *J Exp Child Psychol*. 2005;90(2):142-171.

How to cite this article: Drummond R, Elsworth Q, Faulkner T, Deacon S. Literal interpretation of suicidal-ideation screening items by autistic adults: Validity concerns and revised wording. *International Journal of Autism*. 2026;6(9):09-16. doi: https://www.doi.org/10.22271/27103919.2026.v6.i9a.127

27103927, 2026, 6(9), Downloaded from https://www.doi.org/10.22271/27103919.2026.v6.i9a.127 by International Journal of Autism on [15/09/2026]. See the Terms and Conditions (https://www.rehabilitationjournals.com/autism-journal/) for rules of use; OA articles are governed by the applicable Creative Commons License