



Initial Validation for the Assessment of Pragmatic Abilities and Cognitive Substrates (APACS) Hebrew Battery in Adolescents and Young Adults With Typical Development

Shanie Fussman and Nira Mashal*

Faculty of Education, Bar-Ilan University, Ramat Gan, Israel

OPEN ACCESS

Edited by:

Marcela Pena,
Pontificia Universidad Católica de
Chile, Chile

Reviewed by:

Eliseo Díez-Itza,
University of Oviedo, Spain
Valentina Bambini,
University Institute of Higher Studies in
Pavia, Italy

*Correspondence:

Nira Mashal
mashaln@mail.biu.ac.il
Nira.Mashal@biu.ac.il

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Communication

Received: 13 August 2021

Accepted: 21 December 2021

Published: 20 January 2022

Citation:

Fussman S and Mashal N (2022) Initial
Validation for the Assessment of
Pragmatic Abilities and Cognitive
Substrates (APACS) Hebrew Battery in
Adolescents and Young Adults With
Typical Development.
Front. Commun. 6:758384.
doi: 10.3389/fcomm.2021.758384

Currently there is no validated battery to assess pragmatic abilities in Hebrew. The use of such battery has great importance, as it may provide norms to the assessment of impaired pragmatic skills across several populations, such as ASD, schizophrenia, specific learning disorders and intellectual disabilities. In order to validate the battery, the Assessment of Pragmatic Abilities and Cognitive Substrates (APACS) was translated to Hebrew. The APACS battery was previously validated and found high reliability and validity for ages 19–89 years. The battery includes six tasks, focusing on two main domains: pragmatic production and pragmatic comprehension. The assessment of pragmatic production will be conducted by the use of interview and description tasks, whereas pragmatic comprehension will be assessed by narratives, two figurative language, and humor tasks. The translated battery, APACS-Heb is currently the most comprehensive and the first validated battery for pragmatic tests in Hebrew. Forty Hebrew-speaking adolescents ages 16–20 participated in the study. All participants performed screening tests assessing vocabulary, Theory of Mind and social responsiveness. In addition, the validity and test-retest reliability of APACS-Heb were assessed. Furthermore, the effect of vocabulary, Theory of Mind and social responsiveness on performance was evaluated. High internal consistency, content validity and test-retest reliability was found for most APACS-Heb tasks and all composite scores. Furthermore, an effect of age and gender was found for most tasks with females outperformed males. In addition, a contribution of Theory of Mind to pragmatic production, pragmatic comprehension and APACS total scores was found. Lastly, a factor analysis revealed two factors, in which the first factor correlates with most tasks, and the second factor correlates only to humor. The results thus suggest that humor is a separate skill among the other pragmatic skills. In conclusion, normative data was collected for the APACS-Heb battery, and it was found that it is a valid and reliable measure of pragmatic skills. Since APACS is a comprehensive battery assessing the various aspects of figurative language, it can identify the specific deficits in figurative language and therefore may pinpoint the appropriate intervention program for each individual.

Keywords: pragmatics, theory of mind, assessment, humor, figurative language

1 INTRODUCTION

Pragmatics is among the main components in language, that is being used for human communication. When being used for communicative purposes, pragmatics allows richer meanings, beyond the explicit ones (Martin and McDonald, 2003). These broader meanings, that derive both from language and context, are mainly a product of the social context. Thus, the ability to communicate relies not only upon lingual abilities, but also upon the context, the people involved and general knowledge. The use of pragmatic in language is mostly subconscious, and awareness to pragmatic rules often happens when such rules are broken. Pragmatic skills enable the speaker to bridge the gap between literal meanings and non-literal communicative intentions in context (Vicente and Falkum, 2021). One of the main components of pragmatics is figurative language. Figurative language is common in daily social discourse, in routine classroom activity (Kerbel and Grunwell, 1997), and in electronic and written media (Reyes et al., 2012). In figurative language, the listener is unable to perform interpretation solely by linking the isolated literal meanings of the word, phrase or sentence components (Martin and McDonald, 2003; Rapp and Wild, 2011). Rather, in order to interpret the listener must understand the other's intention from the given context and the social-communicative meaning. Figurative language refers to all language components that cannot be interpreted literally, all of which are key elements in social communication (Vulchanova et al., 2015). Indeed, individuals with pragmatics deficits report fewer satisfying friendships and relationships, and a higher sense of loneliness when compared to typically developing peers (Tierney et al., 2014).

The current study aims to validate a battery of tests focusing on figurative language and narrative comprehension. The figurative language comprehension scales include metaphors, humor and idioms. Common to these aspects of figurative language is the need to derive the non-literal meaning of the expression and to go "beyond" the literal interpretation in order to grasp the speaker's intention in a given context (Giora, 1997). Metaphoric expressions includes both familiar and non-familiar (novel) metaphors. Different cognitive processes underly the processing of these two types of metaphors: The interpretation of a familiar metaphor is stored in the mental lexicon, and the process of comprehending it is conditioned by the ability to access and retrieve an existing knowledge (Glucksberg et al., 2001; Kasirer and Mashal, 2016). Unlike familiar metaphors, the ability to comprehend novel metaphors requires other cognitive resources such as working memory, selective attention, divergent thinking, non-verbal intelligence, and mental flexibility (Chiappe and Chiappe, 2007; Beaty and Silvia, 2012; Mashal et al., 2013; Kasirer and Mashal, 2016; Menashe et al., 2020). Another component of pragmatics is the ability to comprehend humor. There are different types of humor, and in the current study semantic jokes are used. In semantic jokes, there is usually a deviation against lexical semantic rules, or a communicative-pragmatic violation, in which ambiguity in the interpretation is presented (Vrticka et al., 2013). According to Suls (1972), understanding a joke

requires both a surprise element and coherence. The main cognitive component required to establish coherence is flexibility. Thus, a deficit in one or more of these cognitive abilities may predict a difficulty in the comprehension of figurative language.

Comprehending narratives is another important component in pragmatics (Arcara and Bambini, 2016). This skill is linked to hearing comprehension, and includes the comprehension of narrative texts, photograph narrative and conversations (Tompkins et al., 2013; Arcara and Bambini, 2016). The narrative ability has an important role in human communication, since the ability to tell experiences and stories is a main component of social interactions in everyday life (Duinmeijer et al., 2012). Thus, different aspects of figurative language have an important role in understanding and maintaining social relations.

Adolescence and early childhood are marked by a significant improvement in the development of pragmatic skills, that reaches a plateau in adulthood (Nippold et al., 1997). Firstly, children acquire the ability to comprehend narratives prior to reading, and most first graders are experts in both understanding and creating stories (Skarakis-Doyle and Dempsey, 2008; Tompkins et al., 2013). Many verbal behaviors in childhood are based on narratives, and by using them children conclude on links between events and acquire knowledge on the world around them (Botting, 2002; Duinmeijer et al., 2012). The more lingual skills typically developed (TD) children acquire, the higher syntax complexity, vocabulary and length appear in their narratives (Botting, 2002). In addition to narrative comprehension, metaphor comprehension also seems to be achieved during childhood. It was found that 10, 11, and 12 year-olds did not differ significantly from one another in metaphor comprehension, whereas 9 year olds performed lower than them when presented with mental metaphors (Lecce et al., 2019). Furthermore, a similar pattern was found for Theory of Mind, in which 9 year olds performed significantly lower than other age groups. That is, around age ten, there is a significant improvement both in the comprehension of mental metaphors and in Theory of Mind performance. Therefore, there might be a relationship between the two. These findings are consistent with previous studies, in which a bidirectional longitudinal relationship between metaphor comprehension and Theory of Mind was found for eight and 9 year olds (Del Sette et al., 2020). Furthermore, it was also found that early Theory of Mind predicted better metaphor comprehension later. Interestingly, idiom comprehension also seems to improve significantly around age ten (Vulchanova et al., 2011). It was found that third graders display advanced linguistic skills in the comprehension and interpretation of idioms. Finally, the comprehension of figurative language seems to be achieved at nine or 10 years and continues to improve significantly during adolescence (Nippold et al., 1997; Vulchanova et al., 2015).

Currently, there is no validated comprehensive test to assess pragmatic abilities in adolescents and adults that are native

Hebrew speakers. While some questionnaires in Hebrew are aiming at assessing pragmatic skills, there are several limitations that should be noted. The Children's Communication Checklist (CCC) aims at identifying potential language impairment, predominantly pragmatic abnormalities, among children with communication difficulties (Bishop, 1998). The 70-item questionnaire is filled by parents or teachers. While being used for both clinical and research purposes, it is important to note few limitations of the current tool. Firstly, while assessing pragmatic production, the CCC does not identify pragmatic comprehension difficulties. In addition, as opposed to questionnaires performed by the individual himself, checklist ratings run the risk of subjective bias. Furthermore, the CCC is restricted to children aged 4–16 years and cannot be used to assess adolescents' and adults' pragmatic skills. As far as pragmatic comprehension, there are few questionnaires assessing the ability to infer non-literal meanings to idioms, proverbs, novel and familiar metaphors (Mashal and Kasirer, 2011; Saban-Bezalel and Mashal, 2019). However, these questionnaires do not assess pragmatic production nor narrative or humor comprehension. An additional instrument in Hebrew that is used to assess the pragmatic components of natural peer conversation, is The Pragmatic Rating Scale-Young (PRS-Y; Bauminger-Zviely et al., 2014). This instrument was modified from the original PRS that was originally developed for parents of children with ASD (Landa et al., 1992). However, it is important to note, that while PRS-Y is a detailed scale addressing many pragmatic components in natural social interactions, it is restricted to preschoolers with HFASD, and therefore it cannot be used for adolescents and adults (Bauminger-Zviely et al., 2014). Furthermore, PRS-Y requires videotaping the social interactions, and the coding of behaviors every 2 min over 10 min interactions, by experienced speech therapists with expertise in children with ASD. In other languages, some tools expend the evaluation of pragmatics to non-verbal abilities, including the Assessment Battery for Communication (ABaCo; Angeleri et al., 2012). ABaCo is a validated battery developed for the assessment of pragmatic abilities in patients with brain injuries and neuropsychological disorders. The battery includes five evaluation scales: linguistic, extralinguistic, paralinguistic, context and conversational, addressing both comprehension and production. The battery may be used for Italian speaking adults, ranging from 16 to 73 years.

In the current study, a battery of pragmatic tests—the Assessment of Pragmatic Abilities and Cognitive Substrates (APACS) test will be tested. The APACS is a test that was developed in Italy and translated to Hebrew in order to standardize various pragmatic skills among participants from 19 to 89 years (Arcara and Bambini, 2016). The test includes six tasks that are divided to two sections: pragmatic production and pragmatic comprehension. The test has significant internal validity for all APACS tasks, a significant construct and content validity and a test-retest reliability for most tasks, excluding narratives. In addition, a consistent pattern of education and age effects were found for most APACS tasks and scores. Age and education showed

some general effects, whereas gender did not. Furthermore, a factor analysis was performed, and suggested a two-factor solution in which the first factor was mostly correlated to *figurative language 1*, *figurative language 2*, and *narratives*, whereas the second factor had a very strong correlation for *Humor*, and a moderate correlation with *narratives* as well. The *Description* task was not included in the factor analysis due to a ceiling effect. Additionally, a practice effect was found solely for the *Narratives* task. Lastly, a low variance was observed across all tasks (Arcara and Bambini, 2016).

There is great importance in using and validating this battery, since it is the most comprehensive battery for pragmatic tests in Hebrew, and currently there are no validated alternatives. Due to the developmental characteristics of figurative language, it was decided to perform the research on adolescent participants. Furthermore, the minimum age for APACS was 19 years (Arcara and Bambini, 2016), whereas the Hebrew version (APACS-Heb) will be assessed for participants with age range of 16–20 years.

Creating norms is especially important for populations such as ASD, schizophrenia, specific learning disorders and intellectual disabilities as the ability to comprehend pragmatic language among these populations is often delayed or lacking (Bruce et al., 2006; Vulchanova et al., 2015; Cappelli et al., 2018). While adolescence and young adulthood are periods of social change and new social demands, and the use of pragmatic language is essential to fulfill those demands, individuals with ASD experience many social challenges as they face delay or impairment in the development of figurative language (Tantam, 2000). In addition, after creating pragmatic norms, the APACS battery could be used to assess not only pragmatic comprehension but also production abilities of individuals with pragmatic deficits. Since APACS is a comprehensive battery assessing the various aspects of figurative language, it can identify the specific deficits in figurative language and therefore may pinpoint the appropriate intervention program for each individual.

The aim of the present study is creating for the first time pragmatic norms for Hebrew speakers adolescents from 16 to 20 years, while using a translation of the APACS test. In addition, the current study aims at offering the adaptation of the APACS test in Hebrew and evaluating its use in the adolescent and young adult population. The present study expands from previous studies by adding screening tests, assessing the participants' vocabulary, Theory of Mind and social responsiveness, and addressing the link between these tests and the participants' performance on APACS-Heb. In addition, the present study will be used to assess the battery's validity and test-retest reliability. In accordance with previous findings regarding APACS psychometric properties, we expect APACS-Heb to have high internal consistency, test-retest reliability and content validity (Arcara and Bambini, 2016). In addition, we expect age to predict performance on APACS-Heb. Furthermore, we expect that the factor analysis would reveal two factors, one of which will have high loading mostly for the *Humor* task.

2 METHOD

2.1 Participants

Forty native Hebrew-speaking TD participants (21 woman and 19 man) were recruited from high schools and youth movements. The age of the participants ranged from 16 to 20 years ($M = 17.4$, $SD = 0.99$). All participants had intact or corrected vision and reported no neurological problems. For minor participants, parents received an introductory letter about the experiment and signed an informed consent prior to the beginning of this study. Individuals over the age of 18 signed an informed consent. The study was approved by the Israeli Ministry of Education.

2.2 Materials

All participants performed screening tests assessing vocabulary, ToM and social responsiveness. In addition, their pragmatic abilities were assessed by performing the APACS test.

Screening Tests

Vocabulary

Vocabulary was assessed by a Hebrew translated version (version A) of PPVT-5: Peabody Picture Vocabulary Test Fifth Edition (Dunn, 2019). In this test, participants were presented with four pictures to choose from when instructed. The test includes 240 words, including nouns, verbs, and adjectives, that are divided into categories from different areas of life. The questionnaire is suitable for ages 2.5–90 years. The raw grade was transformed to a scaled score with an average score of 100.

Theory of Mind

Theory of Mind (ToM) skills were measured by a Hebrew translated version of the Hinting test (Saban-Bezalel and Mashal, 2019). The test includes ten brief stories describing social interaction between two characters and a social “hint” by one of them. The participant was then required to make a judgement about the intention of the character, based on the given hint. If the participant responded correctly, he was given two points. If the participant failed to respond correctly, an explicit hint was given. If the participant responded correctly following the explicit hint, he was given one point. If the participant failed to give the correct response following both hints, he was not given any points. The score ranges from 0 to 20 points, and a higher score indicated on better ToM and comprehension of intentions skills. The Hinting test was previously used to assess ToM skills, and it has good psychometric properties (Greig et al., 2004; Marjoram et al., 2005).

Social Responsiveness

To assess the participants’ social impairment in naturalistic social settings, they have filled the questionnaire SRS-2: Social Responsiveness Scale (Constantino and Gruber, 2012). The questionnaire comprises 65 items using a “1” (not true) to “4” (almost always true) point Likert scale, generating one total score (For example: *I avoid initiating social contact with other adults*).

Scores range from 65 to 260. Higher scores indicated greater social impairment.

Assessment of Pragmatic Abilities

In order to assess pragmatic skills, the participants have performed the APACS test: The Assessment of Pragmatic Abilities and Cognitive Substrates (Arcara and Bambini, 2016). The APACS test was recently translated by a professional translator from Italian to Hebrew (APACS-Heb). A number of alterations have been made, such as changing names and places, so that the test will better fit Hebrew-speaking population. Figurative phrases (idioms, metaphors, and proverbs) were derived from previous study (Mashal and Kasirer, 2011). The test includes six tasks, while focusing on two main domains in pragmatics-pragmatic production and pragmatic comprehension.

Task 1: Interview

This task is a semi-constructed interview, in which the participant’s ability to be involved in a conversation was measured. The interviewer assessed the discourse produced by the participant, while focusing on pragmatic aspects such as informativeness, coherence and information flow. Each item was given a score of 0 (never), 1 (sometimes) or 2 (always). Maximal score: 44.

Task 2: Description

This task assessed the participant’s ability to produce and share information from everyday life situations, based on the description of photographs and the narrative from those situations. For example: a woman buying flowers in a flower store. The description of each salient element in the photograph was scored 0 (missed identification), 1 (partial identification) or 2 (correct identification). Maximal score: 48.

Task 3: Narratives

This task assessed the ability to comprehend the main aspects of a narrative text. Six stories, that are based on newspaper, radio, and television news articles, were read to the participant. For example: an article about a dog returned to its owner after it got lost. The participant was then presented with questions that rely on comprehending the text explicitly and implicitly. The implicit questions were based on the ability to draw conclusions from the text. Every answer was scored according to its accuracy (0/1 or 0/1/2). Maximal score: 56.

Task 4: Figurative Language 1

This task assessed the ability to infer non-literal meanings using multiple choice questions, including idioms, novel and familiar metaphors, and proverbs from previous studies (Mashal and Kasirer 2011). A sentence and its three possible interpretations were read to the participant (for example: *do not cry over spilled milk*), and he was then required to choose the correct interpretation (it is a shame to regret something that happened and cannot be changed). Every item was scored from 0 (incorrect) to 1 (correct). Maximal score: 15.

Task 5: Humor

This task assesses the ability to comprehend humor using multiple choice questions. After being presented with a short story, the participant is required to choose the most suitable punchline. A joke and three possible endings are read to the participant (for example: when my grandmother turned 50, she started walking 1 km a day), and he is then required to choose the funniest ending (now she is 97 years old and no one knows where she is). Every item is scored from 0 (incorrect) to 1 (correct). Maximal score: 7.

Task 6: Figurative Language 2

This task assesses the ability to infer non-literal meanings by providing a verbal explanation to idioms, novel and familiar metaphors and proverbs. Stimuli were selected from previous studies (Mashal and Kasirer, 2011; Saban-Bezael and Mashal, 2019). After each sentence is read to the participant, he is required to explain it (for example: *wolf in sheep's clothing*). An item is scored 2 points when the participant provides a good description to the meaning of the figurative sentence; An item is scored 1 point when a partial explanation is provided, such as a literal meaning rather than an abstract one; an item is scored 0 points when the participant provides a literal interpretation, rephrases the figurative sentence or is not familiar with it. Maximal score: 30.

Composite Scores

For each participant, three composite scores are computed using the scores of the six tasks. The score for pragmatic production will be computed according to the scores of the Interview and Description tasks. The score for pragmatic comprehension will be computed according to the scores of the Narrative, Figurative Language 1, Humor and Figurative Language 2 tasks. Every composite score is obtained after transforming the original score for each task in proportion to the other tasks and averaging them. That is, every task has an equal contribution to the composite score. In addition, the APACS composite score was calculated by averaging the scores of pragmatic production and pragmatic comprehension.

2.3 Procedure

All participants performed the screening tests and the APACS test in a single session of approximately 70–80 min. The Test-Retest reliability of APACS was assessed in a subset of 20 participants (mean age = 17.04 years, SD = 0.84), that were tested at two separate times with a 2-week interval.

2.4 Statistical Analysis

Firstly, internal consistency, construct and content validity were assessed for all tasks and composite scores. Internal consistency was calculated by means of Cronbach's alpha on all items in each task, construct validity was measured by the correlation between task scores and content validity was evaluated by averaging the means of the 5-point Likert scale ratings made by judges. In addition, Test-Retest

TABLE 1 | Distribution of Age and Gender for the 40 healthy participants of APACS-Heb normative data.

Socioeconomic status	Age					Total m/f
	16	17	18	19	20	
1	8	13	11	3	1	17/20
2	0	0	2	0	0	0/1
3	0	0	1	0	0	1/0
7	1	0	0	0	0	1/0
Total M/F	5/4	6/7	4/10	3/0	1/0	19/21

Reliability was assessed with Pearson correlation coefficients, and paired *t* tests were used to evaluate the practice effect. In addition, a factor analysis was performed by performing an exploratory factorial analysis using a varimax rotation.

Furthermore, the effect of demographic variables on participants' scores was assessed by a series of multiple regression analyses. Moreover, Pearson correlation coefficients were calculated between all screening tests and tasks. Lastly, a series of multiple regressions was carried in order to assess the contribution of the screening tests to the explained variance of the participants' scores.

3 RESULTS

Details on the distribution of participants' demographic variables are reported in **Table 1**. Socioeconomic status was obtained from Israel's Department of Education's 'Transparency in Education' site. In the mentioned site, each high school receives a score from 1 to 10, according to the averaged socioeconomic information of the students attending the school. Such information includes parents' education, calculated income per family member, school's peripherality, and immigration from developing countries. 1 refers to high socioeconomic status, whereas 10 refers to low socioeconomic status. Raw results on APACS-Heb for the 40 participants are reported in **Table 2**.

The results of **Table 2** show that participants have high scores, including ceiling scores for Interview, Description, and Humor.

Descriptive statistics on the individual difference measure are presented in **Supplementary Appendix SA**. The **Supplementary Appendix SA** shows that participants scored the mean \pm one SD (or higher) as follows: 80% for Pragmatic Production, 87.5% for Pragmatic Comprehension and 80% for APACS Total score. As far as participants scoring two SDs (or lower) below mean, the following percentages were obtained: 10% for Pragmatic Production, 5% for Pragmatic Comprehension and for APACS Total scores.

3.1 Internal Consistency

The internal consistency of APACS-Heb (the Hebrew version of APACS) was calculated by means of Cronbach's alpha on all items in each APACS task on the whole sample of 40 participants. Results indicate that as expected all APACS-Heb tasks have

TABLE 2 | Descriptive statistics of APACS-Heb results.

Task or composite score	Mean	SD	Min	Max	Skewness	Kurtosis
Interview	43.35	1.35	37	44	-3.32	12.87
Description	47.20	1.22	43	48	-1.82	3.17
Narratives	50.95	4.87	32	56	-2.09	5.40
Figurative Language 1	13.87	0.72	12	15	-0.66	0.97
Humor	6.68	0.47	6	7	-0.78	-1.47
Figurative Language 2	23.45	5.15	8	30	-1.15	0.84
Pragmatic Production	0.99	0.02	0.90	1	-2.13	4.82
Pragmatic Comprehension	0.89	0.07	0.66	0.97	-1.40	2.33
APACS Total	0.94	0.04	0.78	0.99	-1.54	3.27

TABLE 3 | Test-Retest reliability and practice effect of APACS.

Task or composite score	Test-retest Reliability	Score difference (Retest minus test)	t-test (Test vs. Retest)
Interview	0.95 ^b	0.15	$t_{(19)} = -1.371$, $p = 0.19$
Description	0.83 ^b	0.05	$t_{(19)} = -0.27$, $p = 0.79$
Narratives	0.81 ^b	1.70	$t_{(19)} = -2.11$, $p = 0.05^*$
Figurative Language 1	0.50 ^a	0.10	$t_{(19)} = -0.57$, $p = 0.58$
Figurative Language 2	0.92 ^b	-0.30	$t_{(19)} = 0.60$, $p = 0.56$
Pragmatic Production	0.90 ^b	0.00	$t_{(19)} = -0.78$, $p = 0.45$
Pragmatic Comprehension	0.90 ^b	0.01	$t_{(19)} = -1.54$, $p = 0.14$
APACS Total	0.93 ^b	0.01	$t_{(19)} = -2.03$, $p = 0.06$

^aSignificance at the 0.05 level (2-tailed).

^bSignificance at the 0.01 level (2-tailed). The Humor task was excluded due to inconsistency as a result of low number of items, leading to an impossible calculation. Results of the paired t-tests comparing the scores at Test and Retest with the corresponding p values are presented.

acceptable internal consistency: 0.77 for Interview; 0.59 for Description; 0.78 for Narratives (excluding two narrative texts due to a ceiling effect); 0.98 for Figurative Language 1; 0.78 for Humor; 0.82 for Figurative Language 2.

3.2 Test-Retest Reliability and Practice Effect

The Test-Retest reliability was assessed for all APACS-Heb tasks, using a subset of 20 participants (mean age = 17.04, SD = 0.84, 10 males and 10 females). The participants were tested at two separate times with a 2-week interval, by the same examiner. **Table 3** includes number of test-retest indices, including Test-Retest reliability, practice effect and t-test scores. Test-Retest reliability was assessed with Pearson correlation coefficients. Most APACS-Heb scores demonstrated good to excellent correlations ($r > 0.8$), excluding *Figurative Language 1* (0.50).

Practice effect was calculated as a mean of the difference in scores across time (Retest minus Test). In addition, Paired t tests were used to evaluate group differences in repeated test scores, and therefore to assess the practice effect.

As can be seen in **Table 3** the results of pairwise comparisons revealed a significant practice effect only for the *Narratives* task, in which participants scored higher in the second measurement. No significant increases were seen in other tasks or composite scores.

3.3 Factorial Structure and Construct Validity

The different pragmatic domains assessed in the APACS-Heb battery are possibly associated to different cognitive processes, and therefore it is important to assess their underlying factor structure. In order to examine the factor structure of APACS-Heb, an exploratory factorial analysis was performed using a varimax rotation, without using a fixed number of factors. The *Description* task was excluded due to a ceiling effect and due to its exclusion from the factor analysis for the APACS battery (Arcara and Bambini, 2016). The factor analysis suggested a two-factor solution which accounted for 49.50 and 22.18% of the variance, respectively. The correlation between APACS-Heb task scores is reported in **Table 4**, and the factor analysis results are reported in **Table 5**.

As can be seen in **Table 4**, the *Humor* task was the only task not demonstrating significant correlations with other tasks.

Loading inspection reveals high loadings in all production and comprehension tasks, excluding *Humor*, for the first factor. The second factor has the highest loadings in *Humor*.

3.4 Content Validity

Content validity refers to the extent to which the items in a test are relevant to, and representative of, the construct it intends to measure. To assess content validity, we asked five judges with a bachelor in psychology, ages 28–30 years, to rate on a 5-point Likert scale the degree to which each task or composite score

TABLE 4 | Correlations between APACS-Heb task scores.

	Interview	Description	Narratives	Figurative language 1	Humor	Figurative language 2
Interview	—	—	—	—	—	—
Description	0.38 ^a	—0	—	—	—	—
Narratives	0.48 ^b	0.39 ^a	—	—	—	—
Figurative Language 1	0.28	0.06	0.49 ^b	—	—	—
Humor	−0.06	0.115	0.09	−0.012	—	—
Figurative Language 2	0.49 ^b	0.60 ^b	0.76 ^b	0.40 ^a	0.19	—

^aCorrelation is significant at the 0.05 level (2-tailed).

^bCorrelation is significant at the 0.01 level (2-tailed).

TABLE 5 | Results of factor analysis on APACS-Heb tasks.

Task	Factor 1-loadings	Factor 2-loadings
Interview	0.70	−0.10
Narratives	0.89	0.13
Figurative Language 1	0.68	−0.30
Humor	0.02	0.96
Figurative Language 2	0.86	0.27

TABLE 6 | Content validity of APACS-Heb.

Task or composite score	Appropriateness
Interview	4.72
Description	4.88
Narratives	4.87
Figurative Language 1	4.88
Humor	4.86
Figurative Language 2	4.87
Pragmatic Production	5.00
Pragmatic Comprehension	5.00
APACS Total	5.00

The table reports item evaluation by raters, computed on a 5-point Likert scale. Mean (across experts and items) is reported for each subscale.

measures the construct it is aimed to measure (Sacco et al., 2008). In each task, the experts were presented with a written statement regarding the extent to which the item measures the targeted domain, and they were required to rate the statement for each item in the task (e.g., “This item evaluates the ability to comprehend humor”). In addition, they were required to rate the quality of APACS-Heb composite scores. A score of 1 in the Likert scale indicated “I strongly disagree”, whereas a score of 5 indicated “I strongly agree”. Item scores were averaged, obtaining a score for each task and composite score.

Results of **Table 6** indicate that the mean values across raters are very high (all above 4.70), indicating that the raters judged the items and composite scores to be appropriate.

3.5 Effect of Demographic Variables on APACS-Heb Tasks and Composite Score

In order to establish the effect of age and gender on APACS-Heb scores, a series of multiple regression analyses was carried out for

each APACS-Heb task and composite score as dependent variables. Age was included as a continuous variable whereas gender was included as a factor with two levels. Each analysis initially included all predictors, and was followed by backward elimination, in which non-significant items were removed. The data in **Table 7** shows the relationship between age and gender to APACS-Heb performance in our sample, and the graphic representation is represented in **Figure 1**.

Table 7 shows that both age and gender had effects on most tasks and composite scores. In *Interview*, there were no significant variables, and therefore the performance was consistent across all participants. In *Description*, as age increases the performance increases as well. In addition, females outperformed males in the current task. In *Narratives*, there was a significant positive effect of age. In *Figurative Language 1* and *Humor*, no variable was significant. For *Figurative Language 2* and *Pragmatic Production* age had a positive significant linear effect, and gender had a significant effect as well, in which females outperformed males. For *Pragmatic Comprehension* age had a positive linear effect. Finally, for the *APACS-Heb Total Score* age had a positive significant effect, and gender had a significant effect as well, as females outperformed males.

3.6 Effect of Screening Tests on APACS-Heb Tasks and Composite Score

The correlations of all screening tests and APACS-Heb tasks and composite scores were assessed with Pearson correlation coefficients, as presented in **Table 8**.

As can be seen in **Table 8**, *vocabulary* did not have a significant correlation with any of the tasks or composite scores. All tasks and scores excluding *Description*, *Figurative Language 1* and *Humor* had moderate to good correlations with *Theory of Mind*. Interestingly, significant negative correlations were found for *Lack of Social Responsiveness* and *Interview*, *Narratives*, *Pragmatic Production*, *Pragmatic Comprehension* and *APACS Total*.

In order to assess the contribution of the screening tests (vocabulary, ToM, Social Responsiveness) to the explained variance of the APACS-Heb tasks and composite scores, a series of multiple regressions was performed with all APACS-Heb tasks and composite scores as dependent variables. *Vocabulary*, *Theory of Mind* and *Social Responsiveness* were included as predictors in the regression model. Each analysis

TABLE 7 | Effect of demographic variables on APACS-Heb tasks and composite scores.

Task or composite score	Term	Estimate	Standard error	t-value	p-value	Model R ²
Interview	No significant variables					
Description	(Constant)	46.71	0.25	185.90	<0.001*	0.24
	Age	0.40	0.17	2.30*	0.03*	
	Gender ¹	0.93	0.35	2.69	0.01*	
Narratives	(Constant)	49.99	1.06	47.28	<0.001*	0.15
	Age	1.70	0.73	2.32	0.03*	
Figurative Language 1	No significant variables					
Humor	No significant variables					
Figurative Language 2	(Constant)	21.61	0.97	22.26	<0.001*	0.36
	Age	2.65	0.67	3.92	<0.001*	
	Gender ¹	3.51	1.34	2.62	0.01	
Pragmatic Production	(Constant)	0.98	0.01	198.37	<0.001*	0.20
	Age	0.01	0.00	2.17	0.04	
	Gender ¹	0.01	0.01	2.27	0.03	
Pragmatic Comprehension	(Constant)	0.87	0.01	61.88	<0.001*	0.30
	Age	0.04	0.01	3.68	<0.001*	
APACS Total	(Constant)	0.93	0.01	108.46	<0.001*	0.31
	Age	0.02	0.01	3.68	<0.001*	
	Gender ¹	0.02	0.01	2.07	0.04	

The table reports the task or composite score name; the term in the regression model; coefficient estimate; standard error; t-value associated with the term; p-value with stars *** denoting significant terms; adjusted R². The "age" variable was mean centered. ¹Gender: 0 = Males, 1 = Females.

initially included all predictors, and was followed by backward elimination, in which non-significant items were removed. The data in **Table 9** shows the relationship between all screening tests to APACS-Heb performance in our sample.

Only *Theory of Mind* had effects on most tasks and composite scores, excluding *Description*, *Figurative Language 1* and *Humor*. A positive linear effect was found for all mentioned tasks and composite scores, suggesting that as performance increases the ability to understand other's intention increases. In addition, in order to rebut multicollinearity, Pearson correlation was calculated for scores of all screening tests. The following correlations were weak and not significant: Vocabulary and Theory of Mind ($r(38) = -0.02, p = 0.92$) and Vocabulary and Lack of Social Responsiveness ($r(38) = 0.05, p = 0.75$). While the correlation between Theory of Mind and Lack of Social Responsiveness is significant, it is still considered weak ($r(38) = -0.33, p = 0.04$).

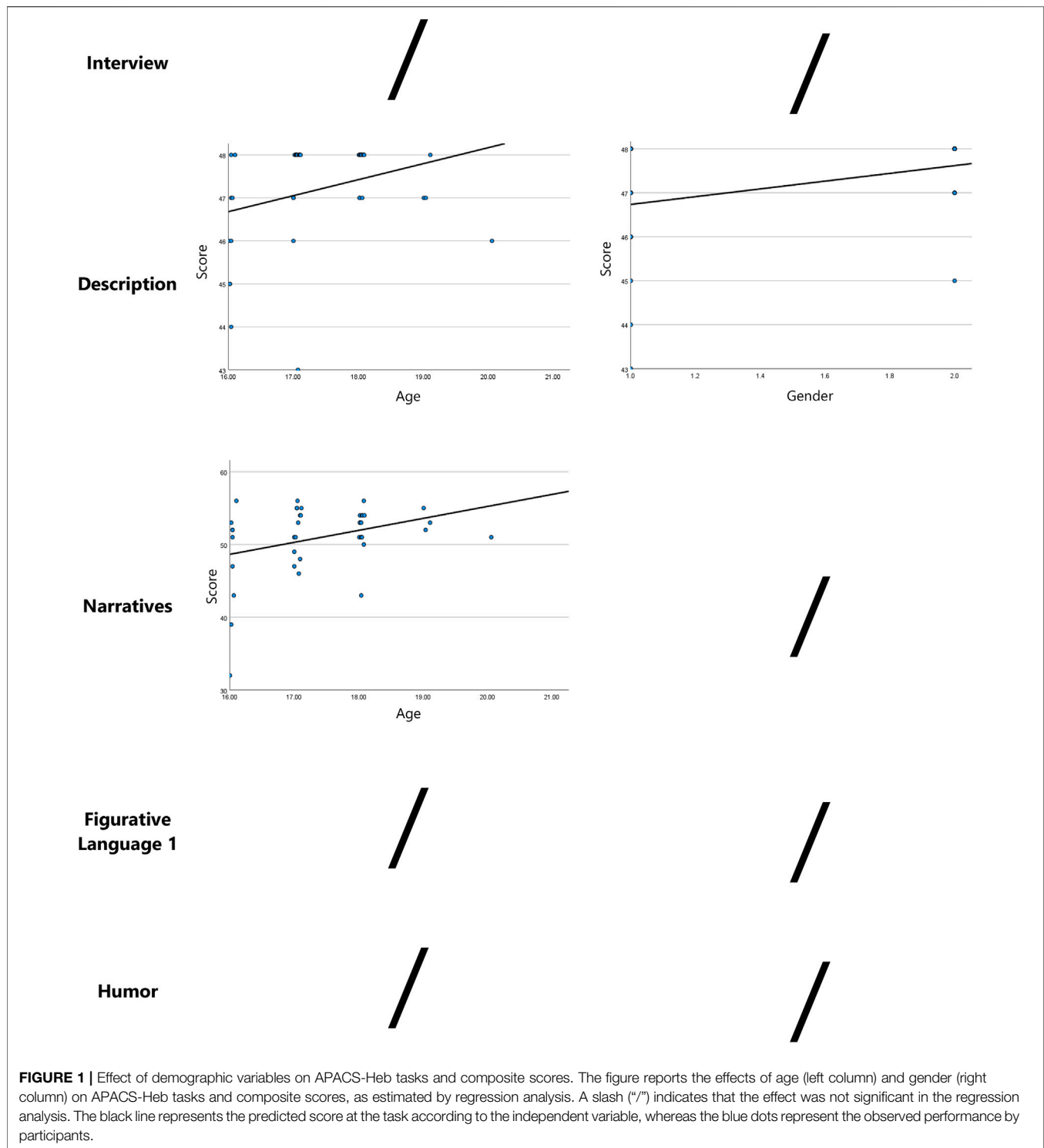
4 DISCUSSION

The aim of the present study was to provide normative data for the Hebrew version of Assessment of Pragmatic Abilities and Cognitive Substrates (APACS-Heb) and to assess internal consistency, validity, and test-retest reliability, among Hebrew speaking adolescents. In particular, the content validity of APACS-Heb confirmed the relevancy and

representativity of the items in the battery to the evaluation of pragmatic abilities.

Consistent with the study's hypotheses APACS-Heb shows a satisfactory reliability, as measured by internal consistency and Test-Retest reliability. Firstly, acceptable internal consistency (Cronbach's alpha >0.60) was found for most APACS-Heb tasks, except for the *Description* task ($r = 0.58$). This might be explained by a ceiling effect, as the mean score for the task is 47.20 and the maximum score is 48. Such ceiling effect might be explained by the bias in the sample, as most of the participants are with high socioeconomic status, as reported in the demographic data. Analyses of Test-Retest data revealed good to excellent stability in scores for most tasks, excluding *Humor* and *Figurative Language 1*, obtained from the subset following a 2-week interval, as previously reported for APACS (Arcara and Bambini, 2016).

As far as individual performance, individual performance of participants for all composite scores was performed (Arcara et al., 2020). Interestingly, more participants scored below mean for Pragmatic Production when compared to Pragmatic Comprehension. This finding might be explained by the high scores obtained for Pragmatic Production (mean = 0.99, maximum score = 1) in comparison to lower scores for Pragmatic Comprehension (mean = 0.89, maximum score = 1). Therefore, it is possible that it is more difficult to achieve a score over the mean \pm SD for Pragmatic Production when compared to Pragmatic Comprehension.



Consistent with APACS findings, a factor analysis on APACS-Heb scores revealed two factors, in which the first one accounted for all production and comprehension tasks excluding humor, and the second factor accounted mostly for *Humor* (Arcara and Bambini, 2016). Thus, the factor analysis supports the construct validity of APACS-Heb, as the battery addresses different

domains of pragmatics, that are possibly related to separate cognitive substrates. In regard to the first factor, accounting for all tasks excluding *Humor* and *Description*, it might be suggested that it is related to *Theory of Mind* abilities. As found in a regression analysis performed for all screening tests, understanding other's intentions predicted performance

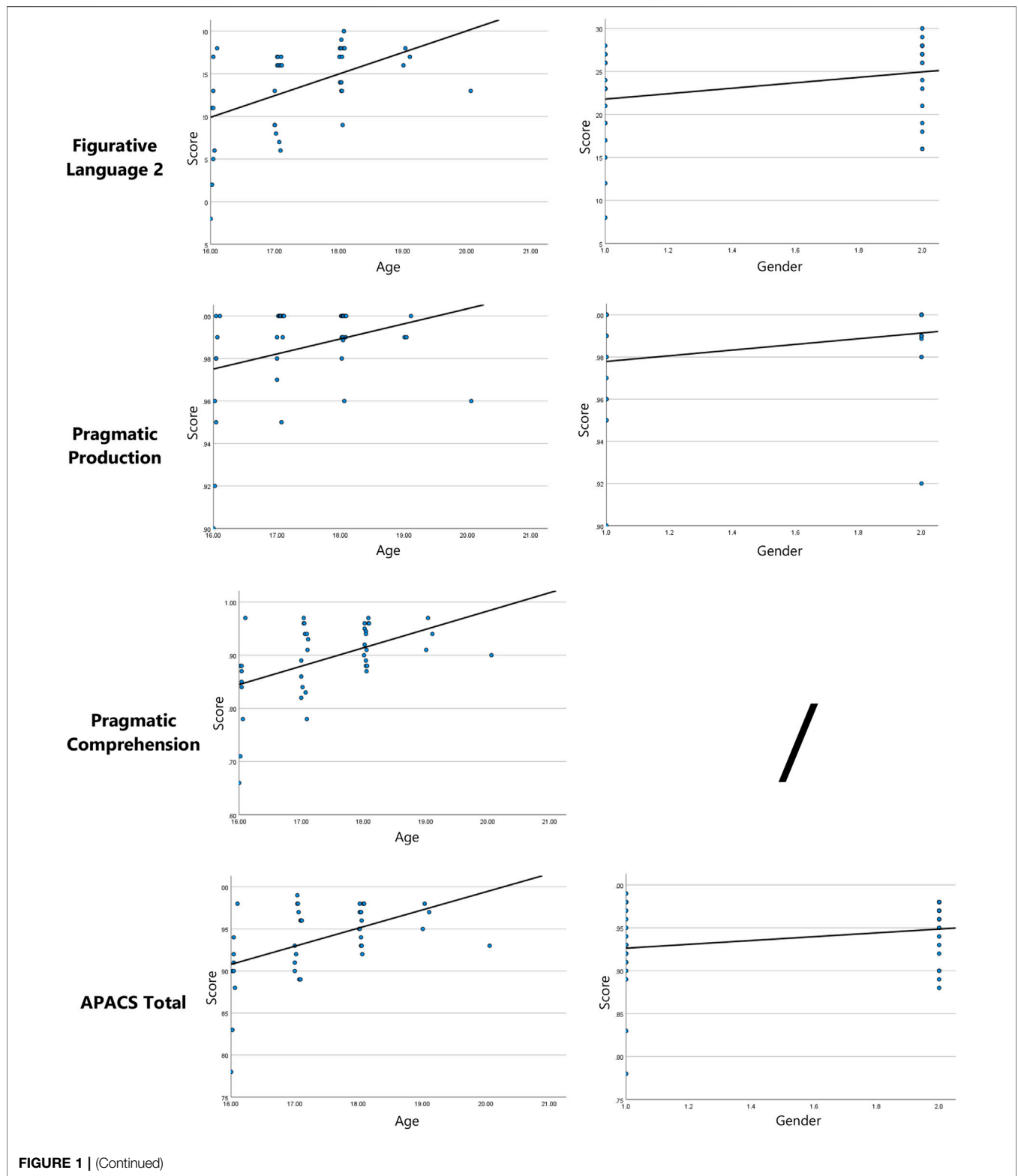


FIGURE 1 | (Continued)

of Interview, Narratives, Figurative Language 2, Pragmatic Production, Pragmatic Comprehension and APACS Total. The same tasks had significant correlations with Theory of Mind as revealed using Pearson correlations. These findings are consistent

with previous research, addressing the role of Theory of Mind in several pragmatic skills and in the comprehension of metaphorical expressions (Happé, 1993). Theory of Mind is required for narrative comprehension, as the listener must

TABLE 8 | Pearson correlation of screening tests and APACS-Heb tasks and composite scores.

Task or composite score		Vocabulary	Theory of mind	Lack of social responsiveness
Interview	Pearson Correlation	0.13	0.69 ^a	-0.34 ^b
	p-value	0.44	<0.001	0.04
Description	Pearson Correlation	0.04	0.31	-0.21
	p-value	0.80	0.05	0.19
Narratives	Pearson Correlation	-0.12	0.54 ^a	-0.39 ^b
	p-value	0.48	<0.001	0.01
Figurative Language 1	Pearson Correlation	-0.15	0.30	-0.23
	p-value	0.37	0.06	0.16
Humor	Pearson Correlation	-0.14	-0.06	-0.06
	p-value	0.4	0.71	0.72
Figurative Language 2	Pearson Correlation	-0.23	0.60 ^a	-0.24
	p-value	0.15	<0.001	0.142
Pragmatic Production	Pearson Correlation	0.12	0.63 ^a	0.34 ^b
	p-value	0.45	<0.001	0.03
Pragmatic Comprehension	Pearson Correlation	-0.23	0.56 ^a	-0.37 ^b
	p-value	0.45	<0.001	0.03
APACS Total	Pearson Correlation	-0.17	0.61 ^a	-0.33 ^b
	p-value	0.30	<0.001	0.04

^bCorrelation is significant at the 0.05 level (2-tailed).

^aCorrelation is significant at the 0.01 level (2-tailed).

TABLE 9 | Effect of screening tasks on APACS-Heb tasks and composite scores.

Task or composite score	Term	Estimate	Standard error	t-value	p-value	Model R ²
Interview	(Constant)	29.58	2.34	12.67	<0.001*	0.48
	Theory of Mind	0.71	0.12	5.91	<0.001*	
Description	No significant variables					
Narratives	(Constant)	12.30	9.83	1.25	0.22	0.29
	Theory of Mind	1.99	0.51	3.94	<0.001*	
Figurative Language 1	No significant variables					
Humor	No significant variables					
Figurative Language 2	(Constant)	-21.94	9.90	-2.22	0.03	0.36
	Theory of Mind	2.34	0.51	4.60	<0.001*	
Pragmatic Production	(Constant)	0.77	0.04	18.05	<0.001*	0.40
	Theory of Mind	0.01	0.00	5.00	<0.001*	
Pragmatic Comprehension	(Constant)	0.31	0.14	2.18	0.04*	0.31
	Theory of Mind	0.03	0.01	4.16	<0.001*	
APACS Total	(Constant)	0.54	0.08	6.46	<0.001*	0.37
	Theory of Mind	0.02	0.004	4.74	<0.001*	

The table reports the task or composite score name; the term in the regression model; coefficient estimate; standard error; t-value associated with the term; p-value with stars *** denoting significant terms; adjusted R².

interpret the intentions, goals, and actions of characters within a narrative (Mason and Just, 2009). These authors have characterized the relation of those two skills by neuroimaging,

in which it was found that narrative comprehension shares some of the neural substrate of Theory of Mind, evoking a specific neural network named “Protagonist Perspective Network”

(Mason and Just, 2009). However, when the effect of Theory of Mind on metaphor comprehension was measured for children (8–15 years), first-order Theory of Mind skills did not predict metaphor comprehension (Norbury, 2005). Thus, future studies should examine the effect of using different methods of assessing Theory of Mind on pragmatic tasks more thoroughly.

Neither vocabulary nor social responsiveness (using SRS) were found as predictors to the tasks comprising the first factor. However, significant Pearson correlations were found for social responsiveness with *Interview*, *Narratives*, *Pragmatic Production*, *Pragmatic Comprehension* and *APACS Total*. All mentioned correlations were negative attesting to the fact that higher scores in SRS questionnaire, assessing social responsiveness, indicate greater social impairment. Therefore, the negative trend indicates that better social responsiveness predicts better performance in the mentioned tasks and composite scores. The current findings are consistent with previous findings, suggesting that the narrative ability has an important role in human communication, since the ability to tell experiences and stories is a main component of social interactions in everyday life (Duinmeijer et al., 2012). The described finding may also explain the correlation between social responsiveness and the *Interview* task, as participants were required to describe several domains of their everyday life, similarly to experience telling.

Consistent with previous findings, the factor analysis suggested that humor comprehension is a separate factor, relying on other cognitive resources (Arcara and Bambini, 2016). Our finding obtained from the regression analysis, show that Theory of Mind did not predict better performance in the *Humor* task. This finding might be explained by previous research, addressing a specific effect of Theory of Mind on Humor (Bischetti et al., 2019). The specific effect was significant for mental jokes, but was not significant for phonological jokes. For the humor task in APACS-Heb, participant were required to choose a correct ending to a brief story. The incorrect ending was either straightforward non-funny ending, or an unrelated non-sequitur ending. Unlike mental jokes, correct endings in the current battery either play with literal and ambiguous meanings, or require to derive non-explicit and unexpected scenarios, and therefore this finding might be in line with our current results. As none of the screening tests, including Theory of Mind, social responsiveness and vocabulary, predicted performance in the *Humor* task, nor any of them had a significant correlation with *Humor*, it is interesting to think of other underlying processes that may be related to humor comprehension. For instance, it was previously found for ALS patients that humor comprehension accuracy was predicted by pragmatic skills (Bambini et al., 2020). Therefore, it is possible that humor comprehension impairment might be part of a larger cognitive impairment, being linked to pragmatic impairment. Another possible explanation, as mentioned in the introduction, the detection and resolution of incongruities is necessary for the comprehension of verbal jokes (Chan et al., 2013). The two processes are separate and activate two different brain regions, however, both are required in order for the listener to understand

a joke. The neural substrates of the mentioned processes might be the underlying mechanism for the second factor, distinguishing humor comprehension from the other tasks. Another possible factor that might affect humor comprehension is executive functions. When presented with a joke, the listener encounters a two-stage process (Suls, 1972). When listening to a joke and encountering incongruity, usually in the punchline, the participant is required to suppress the mismatch literal interpretation that was first given to the joke. Next, he is required to reevaluate the entire joke and shift to a novel and humorous interpretation. While some studies addressed the effect of executive functions on humor comprehension, the relative contribution of each specific mechanism remains unclear. While it was found that cognitive flexibility, abstract reasoning and short-term memory affect humor altogether, individual effects were not found (Mak and Carpenter, 2007). Therefore, future research should examine the contributing factors to the comprehension of humor, and in particular the specific executive functions taking part in the process.

As for the demographic variables, effects of age and gender were evident in the current study (see **Table 6, Figure 1**). Three of the six tasks and all composite scores identified either age or gender as predictors of performance. The pattern shown for age had been consistent across tasks (excluding *Interview*, *Humor* and *Figurative Language 1*) with a positive effect for age. Interestingly, while both tasks had the same type of stimuli: metaphors, idioms and proverbs, an age effect was only found for *Figurative Language 2* and not for *Figurative Language 1*. Therefore, it is interesting to think whether age effects the specific task required from the participants, rather than the comprehension of the stimuli. That is, while *Figurative Language 1* uses multiple-choice questions, for *Figurative Language 2* the participant is required to give a verbal explanation to the expression he is presented with. It is possible that the effect found for *Figurative Language 2* points on the ability to give a verbal interpretation that develops as age increases, rather than a development of figurative language comprehension. The age effect found for several APACS-Heb tasks, supports previous research which suggested that the ability of comprehending figurative language improves during adolescence (Nippold et al., 1997). When assessed for the APACS battery, age had a negative effect on most tasks and composite scores (Arcara and Bambini, 2016). However, it is important to note that all APACS participants were 19 years old or older. Previous studies have found positive correlation between idiom comprehension and age, for TD children and adolescents (Saban-Bezael and Mashal, 2019). In addition, it was found that the described correlation was mediated by executive functions among the TD participants. Therefore, it is possible to assume that while pragmatic skills improve during adolescence, possible due to the development of executive functions, they decrease during adulthood. As for the gender effects, unlike APACS in which gender did not have a significant effect on performance, the current findings show that gender had a significant effect on *Description*, *Figurative Language 2*, *Pragmatic Production* and *APACS-Heb Total Score* performance for APACS-Heb (Arcara and Bambini, 2016). In all mentioned tasks and composite scores, females scored higher

than males. The current findings suggest for the first time that during adolescence females are better at some pragmatic skills than males.

A limitation of the study is that while our normative data is assumed to be representative of the general population, many of the participants were recruited from youth movements and most of them are with high socioeconomic status, and therefore might not represent a heterogenic sample of participants. Thus, although the APACS-Heb battery is suitable for assessing pragmatics in neurodevelopmental disorders, the data obtained in the current study would not be reliable to diagnose pragmatic impairments in children and/or adolescents and adults with such disorders. It is important that future research will assess young adults with low socioeconomic status as well. However, it should be noted that while some participants in the APACS battery were with low socioeconomic status, their scores were high as well (Arcara and Bambini, 2016). Future studies should use additional screening tests, including executive functions. Furthermore, more ecological research settings should be used in order to assess pragmatic skills in a natural environment.

Despite the described limitation, in creating norms for adolescents, the study will enable the use of APACS-Heb in assessing pragmatic skills and detecting specific impairments among individuals with special needs. Pragmatic impairment is prevalent among several populations, such as ASD, schizophrenia, specific learning disorders and intellectual disabilities (Bruce et al., 2006; Vulchanova et al., 2015; Cappelli et al., 2018). APACS was previously used to assess the pragmatic abilities of individuals with dyslexia, Parkinson's disease, multiple sclerosis, and traumatic brain injuries (Cappelli et al., 2018; Montemurro et al., 2019; Carotenuto et al., 2018; Arcara et al., 2020). As APACS-Heb is relatively short (35–40 min) and does not require effortful training for the clinician administering it, it is suitable for individuals with neurodevelopmental or psychiatric disorders as mentioned above. We believe the findings reported here will foster future research aimed at investigating pragmatic abilities among the

mentioned populations and within a broader age group, in order to detect the specific impaired domains for a suitable intervention to be suggested.

In conclusion, normative data was collected for the APACS-Heb battery, and it was found that it is a valid and reliable measure of pragmatic skills. As it addresses several domains of both pragmatic production and pragmatic comprehension, the APACS-Heb battery can evaluate pragmatic skills among Hebrew speaking adolescents.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Israeli Ministry of Education. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

SF collected the data, analyzed the data and wrote the paper; NM initiated the study, conceptualized the study, supervised the study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcomm.2021.758384/full#supplementary-material>

REFERENCES

- Angeleri, R., Bosco, F. M., Gabbatore, I., Bara, B. G., and Sacco, K. (2012). Assessment Battery for Communication (ABaCo): Normative Data. *Behav. Res.* 44 (3), 845–861. doi:10.3758/s13428-011-0174-9
- Arcara, G., and Bambini, V. (2016). A Test for the Assessment of Pragmatic Abilities and Cognitive Substrates (APACS): Normative Data and Psychometric Properties. *Front. Psychol.* 7, 70. doi:10.3389/fpsyg.2016.00070
- Arcara, G., Tonini, E., Muriago, G., Mondin, E., Sgarabottolo, E., Bertagnoni, G., et al. (2020). Pragmatics and Figurative Language in Individuals with Traumatic Brain Injury: fine-grained Assessment and Relevance-Theoretic Considerations. *Aphasiology* 34 (8), 1070–1100. doi:10.1080/02687038.2019.1615033
- Bambini, V., Bischetti, L., Bonomi, C. G., Arcara, G., Lecce, S., and Ceroni, M. (2020). Beyond the Motor Account of Amyotrophic Lateral Sclerosis: Verbal Humour and its Relationship with the Cognitive and Pragmatic Profile. *Int. J. Lang. Commun. Disord.* 55 (5), 751–764. doi:10.1111/1460-6984.12561
- Bauminger-Zviely, N., Karin, E., Kimhi, Y., and Agam-Ben-Artzi, G. (2014). Spontaneous Peer Conversation in Preschoolers with High-functioning Autism Spectrum Disorder versus Typical Development. *J. Child Psychol. Psychiatry* 55 (4), 363–373. doi:10.1111/jcpp.12158
- Beaty, R. E., and Silvia, P. J. (2012). Why Do Ideas Get More Creative across Time? an Executive Interpretation of the Serial Order Effect in Divergent Thinking Tasks. *Psychol. aesthetics, creativity, arts*, 6(4), 309, 319. doi:10.1037/a0029171
- Bischetti, L., Ceccato, I., Lecce, S., Cavallini, E., and Bambini, V. (2019). Pragmatics and Theory of Mind in Older Adults' Humor Comprehension. *Curr. Psychol.* 1–17. doi:10.1007/S12144-019-00295-W
- Bishop, D. V. M. (1998). Development of the Children's Communication Checklist (CCC): A Method for Assessing Qualitative Aspects of Communicative Impairment in Children. *J. Child. Psychol. Psychiat.* 39 (6), 879–891. doi:10.1017/s0021963098002832
- Botting, N. (2002). Narrative as a Tool for the Assessment of Linguistic and Pragmatic Impairments. *Child Language Teaching and Therapy* 18 (1), 1–21. doi:10.1191/0265659002ct224oa
- Bruce, B., Thernlund, G., and Nettelbladt, U. (2006). ADHD and Language Impairment. *Europ. Child Adolesc. Psych* 15 (1), 52–60. doi:10.1007/s00787-006-0508-9
- Cappelli, G., Noccetti, S., Arcara, G., and Bambini, V. (2018). Pragmatic Competence and its Relationship with the Linguistic and Cognitive Profile of Young Adults with Dyslexia. *Dyslexia* 24 (3), 294–306. doi:10.1002/dys.1588
- Carotenuto, A., Coccozza, S., Quarantelli, M., Arcara, G., Lanzillo, R., Brescia Morra, V., et al. (2018). Pragmatic Abilities in Multiple Sclerosis: The Contribution of the Temporo-Parietal Junction. *Brain Lang.* 185, 47–53. doi:10.1016/j.bandl.2018.08.003

- Chan, Y.-C., Chou, T.-L., Chen, H.-C., Yeh, Y.-C., Lavallee, J. P., Liang, K.-C., et al. (2013). Towards a Neural Circuit Model of Verbal Humor Processing: An fMRI Study of the Neural Substrates of Incongruity Detection and Resolution. *Neuroimage* 66, 169–176. doi:10.1016/j.neuroimage.2012.10.019
- Chiappe, D. L., and Chiappe, P. (2007). The Role of Working Memory in Metaphor Production and Comprehension. *J. Mem. Lang.* 56 (2), 172–188. doi:10.1016/j.jml.2006.11.006
- Constantino, J. N., and Gruber, C. P. (2012). *Social Responsiveness Scale: SRS-2*. Torrance, CA: Western Psychological Services.
- Corcoran, R., Mercer, G., and Frith, C. D. (1995). Schizophrenia, Symptomatology and Social Inference: Investigating “Theory of Mind” in People with Schizophrenia. *Schizophrenia Res.* 17 (1), 5–13. doi:10.1016/0920-9964(95)00024-g
- Del Sette, P., Bambini, V., Bischetti, L., and Lecce, S. (2020). Longitudinal Associations between Theory of Mind and Metaphor Understanding during Middle Childhood. *Cogn. Develop.* 56, 100958. doi:10.1016/j.cogdev.2020.100958
- Duinmeijer, I., de Jong, J., and Scheper, A. (2012). Narrative Abilities, Memory and Attention in Children with a Specific Language Impairment. *Int. J. Lang. Commun. Disord.* 47 (5), 542–555. doi:10.1111/j.1460-6984.2012.00164.x
- Dunn, D. M. (2019). *Peabody Picture Vocabulary Test*. in *Measurement Instrument*. 5th ed. (Bloomington, MN: NCS Pearson).
- Giora, R. (1997). Understanding Figurative and Literal Language: The Graded Salience Hypothesis. *Cogn. Linguistics* 8, 183–206. doi:10.1515/cogl.1997.8.3.183
- Glucksberg, S., McGlone, M. S., Grodzinsky, Y., and Amunts, K. (2001). *Understanding Figurative Language: From Metaphor to Idioms*. Oxford, New York: Oxford University Press. No. 36.
- Greig, T. C., Nicholls, S. S., Wexler, B. E., and Bell, M. D. (2004). Test–retest Stability of Neuropsychological Testing and Individual Differences in Variability in Schizophrenia Outpatients. *Psychiatry Res.* 129 (3), 241–247. doi:10.1016/j.psychres.2004.09.006
- Happé, F. G. (1993). Communicative Competence and Theory of Mind in Autism: A Test of Relevance Theory. *Cognition* 48 (2), 101–119.
- Kasirer, A., and Mashal, N. (2016). Comprehension and Generation of Metaphors by Children with Autism Spectrum Disorder. *Res. Autism Spectr. Disord.*, 32, 53–63. doi:10.1016/j.rasd.2016.08.003
- Kerbel, D., and Grunwell, P. (1997). Idioms in the Classroom: An Investigation of Language Unit and Mainstream Teachers’ Use of Idioms. *Child Language Teaching and Therapy*, 13 (2), 113–123. doi:10.1177/026565909701300201
- Landa, R., Piven, J., Wzorek, M. M., Gayle, J. O., Chase, G. A., and Folstein, S. E. (1992). Social Language Use in Parents of Autistic Individuals. *Psychol. Med.* 22 (1), 245–254. doi:10.1017/s0033291700032918
- Lecce, S., Ronchi, L., Del Sette, P., Bischetti, L., and Bambini, V. (2019). Interpreting Physical and Mental Metaphors: Is Theory of Mind Associated with Pragmatics in Middle Childhood. *J. Child. Lang.* 46 (2), 393–407. doi:10.1017/s030500091800048x
- Mak, W., and Carpenter, B. D. (2007). Humor Comprehension in Older Adults. *J. Int. Neuropsychol. Soc.* 13 (4), 606–614. doi:10.1017/S1355617707070750
- Marjoram, D., Gardner, C., Burns, J., Miller, P., Lawrie, S., and Johnstone, E. (2005). Symptomatology and Social Inference: a Theory of Mind Study of Schizophrenia and Psychotic Affective Disorder. *Cogn. neuropsychiatry* 10 (5), 347–359. doi:10.1080/13546800440000092
- Martin, I., and McDonald, S. (2003). Weak Coherence, No Theory of Mind, or Executive Dysfunction? Solving the Puzzle of Pragmatic Language Disorders. *Brain Lang.* 85 (3), 451–466. doi:10.1016/s0093-934x(03)00070-1
- Mashal, N., and Kasirer, A. (2011). Thinking Maps Enhance Metaphoric Competence in Children with Autism and Learning Disabilities. *Res. Develop. Disabilities* 32 (6), 2045–2054. doi:10.1016/j.ridd.2011.08.012
- Mashal, N., Vishne, T., Laor, N., and Titone, D. (2013). Enhanced Left Frontal Involvement during Novel Metaphor Comprehension in Schizophrenia: Evidence from Functional Neuroimaging. *Brain Lang.* 124 (1), 66–74. doi:10.1016/j.bandl.2012.11.012
- Mason, R. A., and Just, M. A. (2009). The Role of the Theory-Of-Mind Cortical Network in the Comprehension of Narratives. *Lang. Linguistics Compass* 3 (1), 157–174. doi:10.1111/j.1749-818x.2008.00122.x
- Menashe, S., Leshem, R., Heruti, V., Kasirer, A., Yair, T., and Mashal, N. (2020). Elucidating the Role of Selective Attention, Divergent Thinking, Language Abilities, and Executive Functions in Metaphor Generation. *Neuropsychologia* 142, 107458. doi:10.1016/j.neuropsychologia.2020.107458
- Montemurro, S., Mondini, S., Signorini, M., Marchetto, A., Bambini, V., and Arcara, G. (2019). Pragmatic Language Disorder in Parkinson’s Disease and the Potential Effect of Cognitive Reserve. *Front. Psychol.*, 10, 1220. doi:10.3389/fpsyg.2019.01220
- Nippold, M. A., Uhden, L. D., and Schwarz, I. E. (1997). Proverb Explanation through the Lifespan. *J. Speech Lang. Hear. Res.* 40 (2), 245–253. doi:10.1044/jslhr.4002.245
- Norbury, C. F. (2005). The Relationship between Theory of Mind and Metaphor: Evidence from Children with Language Impairment and Autistic Spectrum Disorder. *Br. J. Dev. Psychol.* 23 (3), 383–399. doi:10.1348/026151005x26732
- Rapp, A. M., and Wild, B. (2011). Nonliteral Language in Alzheimer Dementia: a Review. *J. Int. Neuropsychol. Soc.* 17 (2), 207–218. doi:10.1017/s1355617710001682
- Reyes, A., Rosso, P., and Buscaldi, D. (2012). From Humor Recognition to Irony Detection: The Figurative Language of Social Media. *Data & Knowledge Engineering* 74, 1–12. doi:10.1016/j.data.2012.02.005
- Saban-Bezalel, R., and Mashal, N. (2019). Different Factors Predict Idiom Comprehension in Children and Adolescents with ASD and Typical Development. *J. Autism Dev. Disord.* 49 (12), 4740–4750. doi:10.1007/s10803-019-04193-9
- Sacco K., K., Bara B.G., R., Colle L., F. M., Mate D., L., Angeleri R., D., and Bosco F.M., B. G. (2008). Assessment Battery for Communication - ABACo: A New Instrument for the Evaluation of Pragmatic Abilities. *J. Cogn. Sci.* 9 (2), 111–157. doi:10.17791/jcs.2008.9.2.111
- Skarakis-Doyle, E., and Dempsey, L. (2008). Assessing story Comprehension in Preschool Children. *Top. Lang. Disord.* 28 (2), 131–148. doi:10.1097/01.tld.0000318934.54548.7f
- Suls, J. M. (1972). “A Two-Stage Model for the Appreciation of Jokes and Cartoons: An Information-Processing Analysis,” in *Psychology of Humor*. Editors J. H. Goldstein and P. E. McGhee (New York: Academic Press), 81–100. doi:10.1016/b978-0-12-288950-9.50010-9
- Tantam, D. (2000). “Adolescence and Adulthood of Individuals with Asperger Syndrome,” in *Asperger Syndrome*. Editors A. Klin and F. R. Volkmar (New York: Guilford Press), 367–399.
- Tierney, C. D., Kurtz, M., Panchik, A., and Pitterle, K. (2014). ‘Look at Me when I Am Talking to You’. *Curr. Opin. Pediatr.*, 26(2), 259–264. doi:10.1097/mop.0000000000000075
- Tompkins, V., Guo, Y., and Justice, L. M. (2013). Inference Generation, story Comprehension, and Language Skills in the Preschool Years. *Read. Writ* 26 (3), 403–429. doi:10.1007/s11145-012-9374-7
- Vicente, A., and Falkum, I. L. (2021). *Accounting for the Preference for Literal Meanings in Autism Spectrum Conditions*. Mind & Language.
- Vrticka, P., Black, J. M., and Reiss, A. L. (2013). The Neural Basis of Humour Processing. *Nat. Rev. Neurosci.* 14 (12), 860–868. doi:10.1038/nrn3566
- Vulchanova, M., Saldaña, D., Chahboun, S., and Vulchanov, V. (2015). Figurative Language Processing in Atypical Populations: the ASD Perspective. *Front. Hum. Neurosci.*, 9, 24. doi:10.3389/fnhum.2015.00024
- Vulchanova, M., Vulchanov, V., and Stankova, M. (2011). Idiom Comprehension in the First Language: a Developmental Study. *Vigo Int. J. Appl. Linguistics* 8, 206–234.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Fussman and Mashal. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.