



**HAL**  
open science

## Role of advanced theory of mind in teenagers' evaluation of source information

Yann Dyoniziak, Anna Potocki, Jean-François Rouet

### ► To cite this version:

Yann Dyoniziak, Anna Potocki, Jean-François Rouet. Role of advanced theory of mind in teenagers' evaluation of source information. *Discourse Processes*, 2023, 60 (4-5), pp.363-377. 10.1080/0163853X.2023.2197691 . hal-04137782

**HAL Id: hal-04137782**

**<https://hal.science/hal-04137782>**

Submitted on 24 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Yann Dyoniziak<sup>1</sup>, Anna Potocki<sup>2</sup>, Jean-François Rouet<sup>1</sup>**

<sup>1</sup>Centre de Recherches sur la Cognition et l'Apprentissage (CeRCA), UMR 7295, CNRS,

Université de Tours, Université de Poitiers, CeRCA – CNRS, France

<sup>2</sup>Laboratoire de Recherches sur les Apprentissages en Contexte (LaRAC), Université

Grenoble Alpes, France

**\* Correspondence:**

Yann Dyoniziak

[yann.dyoniziak@univ-poitiers.fr](mailto:yann.dyoniziak@univ-poitiers.fr)

Dyoniziak, Y., Potocki, A., & Rouet, J.-F. (2023). Role of advanced theory of mind in teenagers' evaluation of source information. *Discourse Processes*, 1-15.

<https://doi.org/10.1080/0163853X.2023.2197691>

## **Accepted manuscript**

### **Abstract**

With the development of the Internet as a main source of information, teenagers are increasingly faced with multiple documents, which may contain contradictory statements, and whose reliability must be assessed. One way to assess information reliability is to evaluate the source of the information (e.g., author expertise, intention). However, teenagers rarely engage in such a sourcing process. The present study aims to explore the role of a potential explanatory factor of teenagers' sourcing abilities: Advanced Theory of Mind (AToM). We hypothesized that AToM would be significantly related to teenagers' evaluation skills when reading multiple documents, and in particular to their attribution of sources' intentions and benevolence. We also hypothesized that this contribution will occur over and above teenagers' word reading and textual inferencing skills. Seventy-two students in Grade 8 read a set of online documents about a fictitious socio-scientific controversy and answered comprehension and evaluation questions. AToM was a significant predictor of comprehension and evaluation performance. The role of AToM was especially important for source's evaluation and intentions questions. This study thus contributes to a better understanding of the possible factors of teenagers' developing sourcing skills.

### **Keywords:**

Sourcing; Theory of Mind; Multiple document comprehension; Inferences; Adolescents

## **Role of advanced theory of mind in teenagers' evaluation of source information**

### **Introduction**

Around 51 % of the global population and 80 % of the population in developed countries now have access and are using the Internet (ITU, 2018). The Internet offers a vast amount of information scattered over multiple documents. Cognitive products (e.g., hypothesis, ideas, opinions, etc.) circulating into this 'collective intellectual space' (Bronner, 2013) are almost completely unregulated. While readers increasingly turn to the Internet to gather information to help them make important decisions (Kiili et al., 2021), it has become less common for them to rely on traditional publications – most of which go through some editing process – to know more on an unfamiliar topic (Mason et al., 2022). Consequently, one's responsibility to evaluate the quality of information is being transferred from professional publishers' gatekeepers to laypeople (Strømsø & Kammerer, 2016). Individuals' ability to critically evaluate the contents they read on the Internet therefore appears more than ever as a core literacy skill (e.g., Bråten et al., 2017; Britt et al., 2019; Greene et al., 2016) that most teenagers do not fully master (e.g., Potocki et al., 2020; Wineburg & McGrew, 2016). The present study therefore seeks to explore the factors that could determine teenagers' source evaluation skills, and especially the role of an understudied factor: Advanced Theory of Mind (AToM). Such a study appears necessary to better understand teenagers' sourcing difficulties and could contribute to the future development of effective sourcing interventions for this population.

### ***From single to multiple text comprehension: The role of evaluation and sourcing processes***

Reading comprehension skills have been traditionally defined as the product of word identification and language comprehension (e.g., Gough & Tunmer, 1986). Within this framework, it has been argued that effective reading comprehension highly depends on the quality of lexical representations (i.e., the ability to easily and simultaneously retrieve the phonological, orthographic, and semantic representations of the read words; Perfetti & Hart, 2002). To understand a text, the reader also needs to process text information at a higher level in order to create a ‘situation model’ (i.e., a cognitive representation of the text content; van Dijk & Kintsch, 1983). Central to comprehension is the production of inferences that lead ‘to link up ideas and fill in details that are not explicitly mentioned’ in the text (Cain & Oakhill, 1999, p. 489). Inferences are necessary to achieve a coherent representation, by connecting information across sentences or by injecting background knowledge to fill in the gaps in the text (Cain & Oakhill, 1999). Graesser al. (1994) identified no less than 13 different types of inferences. Interestingly, Graesser et al. singled out ‘author’s intent’ as a specific class of inferences, suggesting that they may rely on specific mechanisms, although noting that ‘the status of author intent or attitude (...) is uncharted at this point of inference research’ (p. 382).

However, most of these seminal studies on reading comprehension have solely focused on the comprehension of single texts often drawn from the narrative genre. Yet, whether in school settings or while browsing the Internet for their own purposes, teenagers are generally faced with more diverse text types and, increasingly often, with multiple documents. Readers’ comprehension of multiple documents extends beyond the simple sum of the respective situation models of each single text. Perfetti et al. (1999) proposed that the comprehension of multiple texts involves the construction of a so-called ‘document model’, in which the reader connects an integrated situation model (i.e., the mental representation of each document’s content and of possible overlaps or links between texts contents) and an

intertext model, which includes connections between texts and between each text and their respective source. Some reading situations are likely to promote the creation of a full document model (Perfetti et al., 1999). For instance, according to the Discrepancy-Induced Source Comprehension effect (D-ISC), readers are likely to link contents to their respective sources if they notice contradictions across texts dealing with the same topic (Braasch et al., 2012). One way to deal with discrepant texts is to link the discrepant claims to their respective sources (i.e., to encode not just *what is said* but also *who says it*). The reader may then weigh the claims by assessing the source's reliability, and decide whom to trust and what to believe (see Content-Source Integration model; Stadtler & Bromme, 2014).

The construction of a document model thus requires the reader to identify and assess any available information regarding the sources of the texts, a process defined as *sourcing* (Wineburg, 1991). The construct of sourcing however is multifaceted and the term is 'used as an umbrella concept encompassing various mental activities focused on source information' (Scharrer & Salmerón, 2016, p. 1540). A source corresponds to the parameters surrounding the production of discourse (author, editor, date, audience, purpose, and so forth; Perfetti et al., 1999). It refers to both the origin of the document, its author, or the author's rationale (Britt & Rouet, 2020). Any of these parameters may be used to gauge the source's reliability. Major dimensions of reliability include the author's knowledgeability, their benevolence and the level of editorial control involved in the publishing process (e.g., Pérez et al., 2018). Sourcing, and henceforth the comprehension of multiple documents, relies on the assessment of any of these dimensions. However, research to date has found that teenagers rarely engage in sourcing processes while reading, at least in the absence of explicit prompts (e.g., Britt & Angliskas, 2002; Wineburg & McGrew, 2016). A study by Kiili et al. (2020) conducted with 6<sup>th</sup> graders found that participants **mostly drew on the most reliable sources when writing from a set of web pages**. However, as pointed out by the authors, most of the materials were

reliable, which could have biased participants' selections. Students also seldom cited any sources, which further suggests limited consideration for source information. Likewise, in a study involving 5<sup>th</sup>, 7<sup>th</sup>, 9<sup>th</sup> graders and college undergraduates, Potocki et al. (2020) found that students' ratings of competent and less competent sources only differed significantly at or above grade 7. Students even had greater difficulty assessing a source's intention (e.g., benevolence, conflict of interest). Students' justifications for reliability assessment focused more on text content than on source dimensions, especially at the earlier graders. Ineffective sourcing may hinder students' comprehension of multiple documents and increase students' vulnerability to misinformation.

### ***Cognitive development and teenagers' sourcing difficulties: The role of Advanced Theory of Mind***

To our knowledge, only a handful of studies have examined factors that may underlie teenagers' sourcing skills. Anmarkrud et al. (2021) suggested that a vast range of individual factors could play a role in sourcing abilities, like, for example, word reading skills (for younger readers), prior knowledge, working memory capacity, or epistemic beliefs. Surprisingly, social cognition-related constructs are missing from the literature exploring the determinants of adequate sourcing skills. Yet, evaluating an information source requires one to apprehend a perspective distinct from one's own and to infer what the source knows and what the source wants. One's ability to understand a third party's perspective (independent from one's own knowledge, belief or values) may condition the production of inferences regarding information sources' knowledgeability and intentions.

Theory of Mind, or ToM, is generally defined as a set of cognitive skills enabling a person to think about their own and others' mental states (e.g., Baron-Cohen et al., 1985; Beaudoin et al., 2020). ToM in its simple form may be defined as reasoning about a third party's beliefs about the world and is usually mastered by the age of 5 (Miller, 2009).

Traditionally, ToM is measured with false beliefs tasks, like the *Sally and Anne* task (Baron-Cohen et al., 1985) or the *faux-pas recognition test* (Baron-Cohen et al., 1999) in which children have to figure out what certain characters know or don't know about a specific situation or to detect social gaffes. Although basic ToM is typically acquired during the early years, research has shown that ToM develops further into more advanced forms throughout childhood and adolescence (Advanced Theory of Mind or AToM, e.g., Miller, 2009; Wang et al., 2016). The late development of more sophisticated forms of ToM has been linked with brain maturation and executive functions refinement (e.g., Blakemore & Choudhury, 2006; Burnett & Blakemore, 2009). AToM enables complex psychological inferences, for instance, to understand that a person can hold a false belief about another person's belief (e.g., Osterhaus & Koerber, 2021). Additionally, Barzilai and Weinstock (2020) suggested that sophisticated forms of ToM (also named Interpretative ToM) are precursors of multiple perspective understanding: An individual who understands that different perspectives arise not only from different objective knowledge but also from different subjective perspectives can handle multiple points of view adequately. Others have posited that such second-order mental state understanding (*What does A believe that B believes?*) is required to understand irony or lie (Sodian & Kristen, 2016), which is primordial to unravel malicious intent (e.g., conflict of interest).

In summary, AToM is a dimension of social cognition that develops throughout late childhood and adolescence. AToM supports individuals' inferences regarding third parties' beliefs and intentions. As such, ATOM could be a prerequisite for one's evaluation of key dimensions of source credibility, including an author's expertise (*what the author knows*) and intention (*what is the author's communicative purpose*). However, AToM is less likely to be involved in other dimensions of sourcing, for instance, the level of editorial control, which may be more related to one's knowledge about the world of media (i.e., knowledge-based



inferences). Likewise, AToM (and therefore, inferences about the source's knowledge or intent) may be less related to other types of inferences (e.g., causal or logical inferences).

Thus far, only a few studies have examined the role of theory of mind abilities in reading comprehension. Dore et al. (2018) suggested that ToM, as an ability to attribute mental states to others, could contribute to children's narrative text comprehension via its contribution to understanding characters' mental states. In Atkinson et al.'s (2017) longitudinal study, ToM measured at the age of three significantly predicted reading comprehension skills at age 6 after controlling for executive functions, visuospatial skills, decoding, and linguistic skills. Finally, in Kim's (2020) direct and indirect effects model of reading (DIER), ToM is considered an important skill contributing to higher-order comprehension processes. In a study conducted with 2<sup>nd</sup> and 4<sup>th</sup> graders, Kim found that ToM and knowledge-based causal inferences were moderately correlated with each other and that ToM was still independently related to reading comprehension, indicating that ToM does capture a 'unique' inferential ability compared to knowledge-based inferences.

To our knowledge, the study by Florit et al. (2020) is so far the only published source of evidence regarding the role of ToM in multiple text comprehension. They investigated the role of AToM (defined as 'the ability to reason about what a person believes or thinks about mental states held by a second person', p. 4) on 4<sup>th</sup> and 5<sup>th</sup> graders' comprehension of sets of three texts providing various perspectives on age-appropriate topics (e.g., pros and cons of eating chocolate). Florit and her colleagues measured AToM with the *Strange Stories task* (White et al., 2009), which required participants to infer mental states by interpreting nonliteral statements (e.g., characters fooling other characters, misunderstanding of a same situation, persuasion). This task differs from the Sally and Anne task in that it requires the understanding how one character can use another character's anticipated false beliefs to fool them. The results showed that AToM significantly predicted children's comprehension of

multiple documents after controlling for their word reading skills and prior knowledge. It should be noted, however, that the Strange stories task requires participants to read and comprehend written texts, which may partly explain the correlation with the criterion task. The present study sought to overcome this limitation by administering an AToM task based on pictorial scenarios. In addition, Florit et al.'s study examined the role of AToM in multiple documents comprehension in general, but not specifically on sourcing abilities, which was the main goal of the present study.

### ***Objectives and hypotheses of the present study***

The present study sought to examine the role of teenagers' advanced theory of mind (AToM) abilities in sourcing while reading multiple documents. Several studies have demonstrated that teenagers experience difficulties when evaluating information reliability through its source (e.g., Potocki et al., 2020; Wineburg & McGrew, 2016). But the underlying factors of such difficulties remain scarcely explored. In this study, we hypothesized that teenagers' AToM would especially be recruited when having to evaluate a source (by grasping adequately what it knows and wants – i.e., source's expertise and benevolence) or to directly infer a source's intention (Hypothesis 1). To provide students with a situation likely to enhance a sourcing process and the construction of a full document model (Perfetti et al., 1999), we made use of a multiple document reading task including contradictory statements (Braasch et al., 2012). In addition, the task was based on a fictitious socio-scientific controversy to reduce the role of prior knowledge and encourage students to use evaluation by sourcing processes. We further hypothesized that AToM would contribute to multiple document comprehension and source evaluation over and above more general linguistic skills such as inferencing or word reading (Hypothesis 2).

## Method

### *Participants*

Participants were 90 French 8<sup>th</sup> grade students ( $M$  age = 13.46;  $SD$  = .28; 50 female) from four different classes and two French urban and semi-urban middle schools (two classes per school). All students were fluent French speakers. Eighteen students were excluded from the analyses because they did not complete the procedure and/or had some diagnosed condition (e.g., ADHD, dyslexia). The final sample included 72 students ( $M$  age = 13.46;  $SD$  = .27; 40 female).

### *Materials*

All the tasks were presented on a computer screen through a website designed by the research team as part of a larger project (URL blinded for review). The website presented each task one by one and enabled the automated collection of participants' responses (i.e., multiple choice or short answer open questions) and corresponding time on task.

### *Multiple documents reading task: assessment of comprehension, sourcing, and intention inferential abilities*

The Web platform featured a set of four online documents (e.g., pseudo-website pages, forums, commercial, social network and scientific website) that dealt with a controversy regarding a fictitious social-technological issue (impact of a quantic belt supposed to increase positive mood). The details of the characteristics and sources presented in each website are presented in Table 1 below.

[Table 1 near here]

The tasks were scripted and contextualized in line with Sabatini et al. (2014). An introductory message explained the context of the task and gave guidelines to the student: 'You hear a lot about the Kotalia lately. It's a belt that is supposed to improve the mood of its

*wearer by sending waves into the body. Your friend Nathan has taken one, and he is in great shape. You are thinking about getting one, but you want to know more about the Kotalia before you see your friend Nathan again and discuss it with him. To do this, you go online and look at a set of sites that are open in tabs. We will ask you questions about the different sites. Since you have the question and the text in front of you, you don't have to read all the texts, you can look for the answer to the question directly in the text. Don't forget to use the scroll bar on the right to go down the site. Sometimes you will be asked what the characters want, what you think their intention might be. Make guesses, there are no right or wrong answers!*

The developed websites were presented together via the use of tabs (see Figure 1). On the left side of the screen, the students were presented comprehension, evaluation, and intention questions one by one in a lockstep fashion. These questions had either a multiple-choice or an open-ended format.

[Figure 1 near here]

Internal consistency was checked for each type of question. Analyses were run using Jamovi (version 2.3.18 - The Jamovi project, 2022) and we used McDonald's omega (Béland et al., 2018; Dunn et al., 2014). Comprehension questions ( $n = 4$ ) assessed participants' ability to identify simple information written in the text (e.g., 'How many volunteers were recruited for the experiment?'). All of them were multiple choice questions. Comprehension questions had a limited but acceptable internal consistency ( $\omega = .60$ ).

Evaluation questions ( $n = 6$ ) required participants to assess the reliability of the websites (four items), and of a selection of users within forums or social networks (two items). Regarding websites, participants were asked which website was the most reliable (MCQ, 1 point), which website they would recommend to a friend and why (open-ended, 1 point), which was the least reliable (MCQ, 1 point), and why (open-ended, 1 point).

Regarding characters, participants were provided with two pairs of characters who had expressed opposite claims and asked to tell which character was the most reliable (open-ended, 1/2 point) and why (open-ended, 1/2 point). Responses to why-questions were credited if participants mentioned relevant source dimensions (e.g., expertise). One of the coauthors and a research assistant scored participants' responses to the evaluation open-ended questions. Discrepancies were discussed and the scoring grid was iteratively revised until the two scorers reached a sufficient inter-rater agreement for each question (Mean  $\kappa = .796$ ; min  $\kappa = .717$ ; max  $\kappa = .935$ ). Evaluation questions reached a good internal consistency ( $\omega = .767$ ).

Finally, intention questions ( $n = 8$ ) assessed participants' ability to infer the sources' intentions with only open-ended questions (e.g., 'What does Ouroug want and why?'). Each question concerned only one source. Each of them required to infer the source's intention. For example, as Ouroug published a sponsorship link on a forum, s/he probably wants the reader to use this link to buy the product. Participants received one point if they mentioned at least one plausible intent. One of the coauthors and a research assistant scored participants' responses to the intention open-ended questions. Discrepancies were discussed and the scoring grid was iteratively revised until the two scorers reached a sufficient inter-rater agreement for each question (Mean  $\kappa = .762$ ; min  $\kappa = .623$ ; max  $\kappa = .907$ ). Intention questions reached a good internal consistency ( $\omega = .831$ ).

All materials and questions are available on the OSF:

[https://osf.io/4hkfp/?view\\_only=f7a34fe8b32a4e1687a9dbb8fd0b9215](https://osf.io/4hkfp/?view_only=f7a34fe8b32a4e1687a9dbb8fd0b9215)

#### *Advanced ToM assessment*

To measure participants' AToM abilities, we used the 'TOM15' task (Desgranges et al., 2012). This task is composed of a set of 7 short picture-based scenarios. For each scenario, participants were asked a second-order false belief question (e.g., 'What does X think that Y

believes?') and a control comprehension question (e.g., 'Who ate the cheese slice?'). One point was granted if participants correctly responded to the two questions (see Figure 2 for example).

[Figure 2 near here]

A full description of the task is available in Desgranges et al. (2012). [Since the original publication did not mention any reliability analysis, we assessed the task's internal consistency in our sample and found that it was barely within acceptable range \( \$\omega = .584\$ \). Nevertheless, our mean and standard deviation were quite similar to those reported by Desgranges et al. with their adult population, which suggests at least some robustness. We then opted for the author's original scoring method.](#)

#### *Word reading/lexical quality assessment*

A computerized word comparison task based on Auphan et al. (2019) was designed to assess participants' word reading fluency and lexical quality. The task involved three series of 18 items, in which participants had to decide whether pairs of words or pseudo-words were identical in terms of orthography, pronunciation (pseudo-words), or were semantically related. Correct responses and the response time were recorded. The comparison rule to be used was displayed at the beginning of each subtask. A speed-accuracy indicator was then calculated for each series (Auphan et al., 2019). The dependent measure used in the present study was a ratio of time spent per correct item, collapsed across the three subtasks. Internal consistency was good ( $\omega = .807$ ).

#### *Inferential abilities assessment*

To assess participants' inferential abilities, we used six items from the task '*Gestion de l'implicite*' [*Management of implicit content*] by Duchêne May-Carle (2000; see also Diemunsch et al., 2022). Each item was made of a short text (three sentences) followed by

three comprehension questions. Each question involved three response options: Yes/No/One cannot know (see Table 2 for an example).

[Table 2 near here]

We arranged two sets of three items (i.e., nine questions per set). Half the students received set A and the other half received set B. In both sets, three questions had to be removed to reach a sufficient internal consistency ( $N = 6$ ;  $\omega = .750$ ; and  $N = 6$ ;  $\omega = .733$  for sets A and B, respectively).

### ***Procedure***

The procedure, which was mostly made of computerized reading comprehension tasks, was approved as educationally relevant by the local school authorities. Students participated in two separate sessions of about 1h30 each. [In the first session, they completed the word reading task, the AToM task, and the inference task. The second session was dedicated to the multiple document reading task only.](#) The sessions were scheduled on [students'](#) usual class time. Two researchers met the whole class (between 18 and 28 students) in the computer room of their school and in the presence of their usual teacher. The procedure was framed as an instructional session dealing with study skills. Students were informed that they were going to have a special practice session and that they would be assigned to a series of new tasks involving reading and answering questions. They were informed that their data would be communicated to the researchers anonymously and for study purposes only. Students were sat at a computer and connected to the experiment website. In addition to the instructions provided by the website for each task, researchers also explained each task orally. Tasks were assigned one by one, and the next task only started after all students had completed the previous one.

## Results

### *Descriptive statistics*

Descriptive statistics were first computed on the scores obtained on each task to detect potential distribution issues. The two sets of inferential questions were of unequal difficulty (set A:  $M = 4.95$ ;  $SD = 1.4$ ; set B:  $M = 3.38$ ;  $SD = 1.84$ ). Therefore, raw scores on this measure were transformed into  $Z$  scores.

Table 3 shows the means, standard deviations, skewness, and kurtosis for the different variables of interest in this study: AToM, word reading, comprehension, evaluation, intentions, and inferences.

[Table 3 near here]

Mean scores showed that difficulty varied across tasks, with evaluation and intention questions being rather difficult, although not impossible to achieve. We then computed correlations to examine the interplay between our variables.

[Table 4 near here]

As expected, positive correlations were observed between word reading and all measures of multiple document comprehension. Within the multiple document task, measures of comprehension, evaluation, and intention were also significantly related. Furthermore, AToM was significantly related to all types of multiple document questions, and, consistent with Hypothesis 1, most strongly correlated with intention questions ( $r = .39$ ,  $p < .001$ ).



### ***Specific contribution of AToM to multiple document comprehension and sourcing***

To explore the specific contributions of AToM, word reading and inferential abilities in multiple document comprehension and sourcing, linear mixed model analyses were run with AToM, inferences, and word reading as fixed effects and class (intercept) as a random factor. The score on each type of multiple documents questions (i.e., comprehension, evaluation by sourcing, and intention inferences) was successively considered as the dependent variable (see Table 5 for a synthesis of the results).

For the comprehension questions of the multiple documents task, the tested model ( $R^2(m) = .17$ ) revealed **only** a significant effect of AToM ( $F(1, 66.3) = 4.73, p = .03$ ). The effects of inferential abilities ( $F(1, 65.5) = 3.26, p = .07$ ) and word reading ( $F(1, 66.1) = 1.72, p = .19$ ) were not significant.

As regards evaluation questions the tested model ( $R^2(m) = .06$ ) demonstrated a significant effect of AToM scores only ( $F(1, 66.3) = 4.36, p = .04$ ) while inferences and word reading scores were non-significant (respectively  $F(1, 65.7) = .39, p = .54$ ; and  $F(1, 66.2) = .45, p = .50$ ).

Finally, for intention questions the tested model ( $R^2(m) = .20$ ) revealed a significant effect of AToM scores ( $F(1, 66.3) = 7.09, p = .01$ ) and word reading scores ( $F(1, 66.1) = 6.29, p = .02$ ), while inferential abilities were not significant ( $F(1, 65.6) = .23, p = .63$ ). The latter finding was consistent with Hypothesis 2.

[Table 5 near here]

## Discussion

Research has shown that although teenagers increasingly turn to the internet to gather important information (Kiili et al., 2021), they display limited sourcing skills (e.g., Potocki et al., 2020; Wineburg & McGrew, 2016) at least when assigned tasks in a school context (Paul et al., 2017). One's ability to critically evaluate sources' competence, intention and reliability may depend on a combination of educational and developmental factors. However, data linking adolescents' sourcing skills to dimensions of their cognitive development are still scarce. The present study attempted to fill this gap by investigating the role of an understudied social cognitive factor, namely Advanced Theory of Mind (AToM; White et al., 2009) as a predictor of multiple document comprehension, controlling for traditional factors such as decoding fluency and inferencing ability. We hypothesized that AToM would be particularly involved in the understanding of what a source character knows and wants, and that it would specifically predict teenagers' abilities to evaluate information reliability when reading multiple documents. To prompt the creation of a full document model, in which readers connect sources to contents (Perfetti et al., 1999), we made use of a set of documents featuring contradictory information (Braasch et al., 2012). Moreover, the documents dealt with a fictitious socio-scientific controversy for which students could not have any prior knowledge. Finally, trying to overcome past studies' limitations, we used a mostly pictorial AToM task (Desgranges et al., 2012) to decrease the potential confound between AToM and reading comprehension.

The study yielded several interesting findings. First, and in line with previous studies (Potocki et al., 2020), source evaluation turned out to be challenging for most participants. Second, and consistent with our main prediction, AToM was a significant predictor of students' multiple document reading skills (Hypothesis 1) over and above basic reading skills and textual inferencing skills (Hypothesis 2). These results confirm previous studies of

younger children showing that ToM contributes to text comprehension (Atkinson et al., 2017; Kim, 2020). They also extend previous studies by examining the role of more advanced forms of theory of mind in more complex comprehension processes and in an older population. In addition, our measure of AToM appeared as the only significant predictor of comprehension and evaluation questions of the multiple document reading task. Intention questions were explained by both AToM and basic reading skills. We also observed that the variance explained by AToM was stronger for the latter type of questions, meaning that AToM could be particularly recruited to know *what the source wants*. The latter finding is consistent with Barzilai and Weinstock's (2020) theoretical suggestion that interpretative ToM could be a precursor of multiple perspectives understanding. Such a finding is also of importance given that assessing the benevolence of a source seems particularly challenging for adolescents (e.g., Potocki et al., 2020). AToM could therefore represent an interesting factor to work on in order to foster source's benevolence discrimination in teenagers.

Interestingly, AToM and inferential abilities were slightly correlated with each other, which is to be expected since both tasks require the production of inferences (see also Kim, 2020). However, inferential abilities **did** not significantly predicted **comprehension, evaluation and intention questions**. Inferences relying on AToM seem to rely on processes distinct from those relying on causality or logic. This is consistent with recent findings by Hagá et al. (2014): social inferences often require to be updated as additional information comes in, which is a demanding process that is mastered only later in development. AToM specifically captures inferences regarding mental states, which are likely to be involved in evaluation (sourcing) and intentions questions. This confirms that fluent decoding and inferential abilities are not enough to achieve a coherent understanding in this evaluative context (e.g., Potocki et al., 2020). **Surprisingly, inferential abilities were never significantly**

related to the dependent measures. This requires additional investigations in future study on links between multiple documents comprehension and ‘classical’ textual inferences.

However, it is worth noting that although being significant predictors, the factors examined in the present study explained only a relatively small share of the variance of multiple document comprehension and evaluation performance (maximum 20%). This suggests that other factors than those included in the present study might be considered to fully explain teenagers’ evaluation skills. For instance, some studies have reported that difficulties in multiple text comprehension – including sourcing – could be related to inadequate epistemic beliefs (e.g., Barzilai & Stromso, 2018). Thus, Barzilai and Zohar (2012) noticed that 6<sup>th</sup> graders having an *evaluatist* epistemic perspective (*‘knowledge and truth is complex, tentative, unreachable, but can be approximated on the basis of criteria we choose’*) outperformed those having an *absolutist* perspective (*‘truth exists, knowledge is right or wrong, the right answer is somewhere’*) on source evaluation tasks. However, knowing that the most elaborated epistemic beliefs (i.e., *evaluatism*; Kuhn et al., 2000) require the integration of both objectivity and subjectivity, it may be hypothesized that the ability to apprehend another’s point of view (ToM) could be a precursor of more elaborated epistemic perspectives (Kuhn & Udell, 2007). Likewise, Potocki et al. (2023) found that teenagers’ multiple document comprehension depends in part on the development of their executive functions. As Wang et al. (2016) reported possible links between ToM and executive functions, the inclusion of the latter factor in a developmental model of sourcing capabilities could also be of interest. Such hypotheses and the links between AToM, epistemic beliefs, executive functions and sourcing constitute a promising research horizon.

### *Limitations and perspectives*

Our study entails a number of limitations that must be acknowledged. First, our choice of a mostly pictorial AToM test (TOM15; Desgranges et al., 2012) failed to clearly disentangle the assessment of AToM from participants' reading skills, which was a limitation of prior work (i.e., Florit et al., 2020). Despite the presence of supportive pictures our AToM test also involved the understanding of written sentences. Therefore, underlying language abilities may explain in part the pattern of findings. Future studies should seek to include ToM tasks that do not overlap (or do so to a lesser extent) with verbal skills, and to improve their control for the influence of language comprehension abilities.

As a second limitation, the questions we used to assess teenagers' sourcing skills only provided a global measure of sourcing ability, even though sourcing has been described as involving several components (e.g., Pérez et al., 2018). As a result, we were not clearly able to examine whether AToM abilities were more involved in some dimensions of source reliability evaluation (e.g., benevolence) than others (e.g., editorial control). Future studies should address this limitation by administering sourcing assessment tasks that distinguish these different dimensions (see for example Potocki et al., 2020).

As a third limitation, the internal reliability of several of our measures was relatively low. Our comprehension task may have confounded distinct levels of processing (e.g., literal vs. inferential comprehension). Likewise, the reliability of our AToM measure was barely within the acceptable range, even though the distribution of performance was similar to those of the original study conducted with adults (Desgranges et al., 2012). This points to the need to further develop precise and reliable instruments to assess teenagers' ability to reason about third parties' mental states.

Finally, because our main tasks involved the sustained reading of several documents, students' interest, motivation and engagement (or lack thereof) may have also partly

determined their performance (Guthrie & Klauda, 2014). This range of factors would be worth investigating in the future.

## **Conclusion**

In sum, teenagers' Advanced Theory of Mind (AToM) abilities appear to help them evaluate sources as they read multiple texts about a controversial topic. More specifically, AToM seems to be involved in inferring source's intentions and may thus contribute to the evaluation of its benevolence. This study adds to the current understanding of the cognitive determinants of teenagers' source evaluation and suggests that teenagers' limited ability to reason about the interplay of third parties' mental states could partly their difficulties when assessing the reliability of sources.

In this way, this study also suggests interesting educational perspectives. Teenagers seem able to engage in this kind of reflexive process and to respond to sourcing-targeted interventions (e.g., Brante & Strømsø, 2018; Pérez et al., 2018). In addition, secondary education represents an optimal window of opportunity to help them acquire the strategies that may serve their purposes in and out of school (Kohnen et al., 2020). Thus, specific exposure and training in the psychological inferences related to knowledge and intent of information sources (including AToM) could provide individuals with the grounding necessary to develop their intellectual skills and may help students better understand the diverse and rather chaotic information world they live in.

## **Acknowledgments**

This study was supported by the French National Research Agency [grant ANR-17-CE28-0016-02]. We thank the teachers, staff, and students of the schools for their kind cooperation. We thank our colleagues Jean Ecalle, Émilie Dujardin and Annie Magnan for designing the

word reading task. Special thanks to Marjorie Guillou and Agathe Phillip for their help with data collection and data coding.

## References

- Anmarkrud, Ø., Bråten, I., Florit, E., & Mason, L. (2022). The Role of Individual Differences in Sourcing : A Systematic Review. *Educational Psychology Review*, 34(2), 749-792. <https://doi.org/10.1007/s10648-021-09640-7>
- Atkinson, L., Slade, L., Powell, D., & Levy, J. P. (2017). Theory of mind in emerging reading comprehension: A longitudinal study of early indirect and direct effects. *Journal of Experimental Child Psychology*, 164, 225-238. <https://doi.org/10.1016/j.jecp.2017.04.007>
- Auphan, P., Ecalle, J., & Magnan, A. (2019). Computer-based assessment of reading ability and subtypes of readers with reading comprehension difficulties : A study in French children from G2 to G9. *European Journal of Psychology of Education*, 34(3), 641-663. <https://doi.org/10.1007/s10212-018-0396-7>
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, 21(1), 37-46. [https://doi.org/10.1016/0010-0277\(85\)90022-8](https://doi.org/10.1016/0010-0277(85)90022-8)
- Baron-Cohen, S., O’Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of Faux Pas by Normally Developing Children and Children with Asperger Syndrome or High-Functioning Autism. *Journal of Autism and Developmental Disorders*, 29(5), 407-418. <https://doi.org/10.1023/A:1023035012436>
- Barzilai, S., & Strømsø, H. I. (2018). Individual Differences In Multiple Document Comprehension. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.), *Handbook of Multiple Source Use* (1<sup>st</sup> ed., p. 99-116). Routledge. <https://doi.org/10.4324/9781315627496-6>
- Barzilai, S., & Weinstock, M. (2020). Beyond trustworthiness: Comprehending multiple source perspectives. In *Handbook of learning from multiple representations and perspectives* (pp. 123-140). Routledge.

- Barzilai, S., & Zohar, A. (2012). Epistemic Thinking in Action: Evaluating and Integrating Online Sources. *Cognition and Instruction*, 30(1), 39-85.  
<https://doi.org/10.1080/07370008.2011.636495>
- Beaudoin, C., Leblanc, É., Gagner, C., & Beauchamp, M. H. (2020). Systematic Review and Inventory of Theory of Mind Measures for Young Children. *Frontiers in Psychology*, 10, 2905. <https://doi.org/10.3389/fpsyg.2019.02905>
- Béland, S., Cousineau, D., & Loye, N. (2018). Utiliser le coefficient omega de McDonald à la place de l'alpha de Cronbach. [Using the McDonald's Omega Coefficient Instead of Cronbach's Alpha]. *McGill Journal of Education*, 52(3), 791-804.  
<https://doi.org/10.7202/1050915ar>
- Blakemore, S.-J., & Choudhury, S. (2006). Development of the adolescent brain: Implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry*, 47(3-4), 296-312. <https://doi.org/10.1111/j.1469-7610.2006.01611.x>
- Braasch, J. L. G., Rouet, J.-F., Vibert, N., & Britt, M. A. (2012). Readers' use of source information in text comprehension. *Memory & Cognition*, 40(3), 450-465.  
<https://doi.org/10.3758/s13421-011-0160-6>
- Brante, E. W., & Strømsø, H. I. (2018). Sourcing in text comprehension: A review of interventions targeting sourcing skills. *Educational Psychology Review*, 30(3), 773-799.  
<https://doi.org/10.1007/s10648-017-9421-7>
- Bråten, I., Stadtler, M., & Salmerón, L. (2017). The Role of Sourcing in Discourse Comprehension. In M. F. Schober, D. N. Rapp, & M. A. Britt (Eds.), *The Routledge Handbook of Discourse Processes* (2<sup>e</sup> ed., p. 141-166). Routledge.  
<https://doi.org/10.4324/9781315687384-10>
- Britt, M. A., & Aglinskas, C. (2002). Improving Students' Ability to Identify and Use Source Information. *Cognition and Instruction*, 20(4), 485-522.  
[https://doi.org/10.1207/S1532690XCI2004\\_2](https://doi.org/10.1207/S1532690XCI2004_2)
- Britt, M. A. & Rouet, J.-F. (2020). Multiple Document Comprehension. In L. Zhang (Ed.) *Oxford Research Encyclopedia of Education* (pp 1-23). Oxford University Press.  
doi:10.1093/acrefore/9780190264093.013.ORE\_EDU-00867.R2



- Britt, M. A., Rouet, J.-F., Blaum, D., & Millis, K. (2019). A Reasoned Approach to Dealing with Fake News. *Policy Insights from the Behavioral and Brain Sciences*, 6(1), 94-101. <https://doi.org/10.1177/2372732218814855>
- Bronner, G. (2013). *La démocratie des crédules* [The democracy of the Gullible] (1<sup>st</sup>. ed). Presses Universitaires de France [University Press of France].
- Burnett, S., & Blakemore, S.-J. (2009). The Development of Adolescent Social Cognition. *Annals of the New York Academy of Sciences*, 1167(1), 51-56. <https://doi.org/10.1111/j.1749-6632.2009.04509.x>
- Cain, K., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and writing*, 11(5-6), 489-503.
- Desgranges, B., Laisney, M., Bon, L., Duval, C., Mondou, A., Bejanin, A., Fliss, R., Beaunieux, H., Eustache, F., & Muckle, G. (2012). TOM-15 : Une épreuve de fausses croyances pour évaluer la théorie de l'esprit cognitive. [TOM 15: a false-belief task to assess cognitive theory of mind]. *Revue de Neuropsychologie [Journal of Neuropsychology]*, 4(3), 216. <https://doi.org/10.3917/rne.043.0216>
- Diemunsch, C., Boiseau, M., & May-Carle, A. D. (2022). Étalonnage de « La Gestion de l'implicite » pour l'évaluation de la compréhension des inférences auprès de patients adultes et étude de l'impact des variables socio-démographiques sur les résultats au test. [Calibration of "Managing Implicit" for the assessment of inference comprehension with adult patients and study of the impact of socio-demographic variables on test scores]. *Glossa*, 28-43.
- Dore, R. A., Amend, S. J., Golinkoff, R. M., & Hirsh-Pasek, K. (2018). Theory of Mind: A Hidden Factor in Reading Comprehension? *Educational Psychology Review*, 30(3), 1067-1089. <https://doi.org/10.1007/s10648-018-9443-9>
- Duchêne May-Carle, A. (2000). *La gestion de l'implicite* [The Management of the implicit.]. Isbergues: Ortho Edition.
- Dunn, T. J., Baguley, T., & Brunson, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), 399-412. <https://doi.org/10.1111/bjop.12046>

- Florit, E., De Carli, P., Giunti, G., & Mason, L. (2020). Advanced theory of mind uniquely contributes to children's multiple-text comprehension. *Journal of Experimental Child Psychology*, 189, 104708. <https://doi.org/10.1016/j.jecp.2019.104708>
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and special education*, 7(1), 6-10.
- Guthrie, J. T., & Klauda, S. L. (2014). Effects of classroom practices on reading comprehension, engagement, and motivations for adolescents. *Reading research quarterly*, 49(4), 387-416.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101(3), 371-395. <https://doi.org/10.1037/0033-295X.101.3.371>
- Greene, J. A., Sandoval, W. A., & Bråten, I. (2016). *Handbook of Epistemic Cognition*. Routledge. <https://doi.org/10.4324/9781315795225>
- International Telecommunication Union. (2018). *ITU releases 2018 global and regional ICT estimates*. <https://www.itu.int/en/mediacentre/Pages/2018-PR40.aspx>
- Hagá, S., Garcia-Marques, L., & Olson, K. R. (2014). Too young to correct: A developmental test of the three-stage model of social inference. *Journal of Personality and Social Psychology*, 107(6), 994–1012. <https://doi.org/10.1037/pspa0000012>
- Kiili, C., Bråten, I., Kullberg, N., & Leppänen, P. H. T. (2020). Investigating elementary school students' text-based argumentation with multiple online information resources. *Computers & Education*, 147, 103785. <https://doi.org/10.1016/j.compedu.2019.103785>
- Kiili, C., Forzani, E., Brante, E. W., Rääkkönen, E., & Marttunen, M. (2021). Sourcing on the Internet Examining the Relations Among Different Phases of Online Inquiry. *Computers and Education Open*, 100037. <https://doi.org/10.1016/j.caeo.2021.100037>
- Kiili, C., Laurinen, L., & Marttunen, M. (2008). Students Evaluating Internet Sources: From Versatile Evaluators to Uncritical Readers. *Journal of Educational Computing Research*, 39(1), 75-95. <https://doi.org/10.2190/EC.39.1.e>
- Kim, Y.-S. G. (2020). Theory of mind mediates the relations of language and domain-general cognitions to discourse comprehension. *Journal of Experimental Child Psychology*, 194, 104813. <https://doi.org/10.1016/j.jecp.2020.104813>

- Kohnen, A. M., Mertens, G. E., & Boehm, S. M. (2020). Can middle schoolers learn to read the web like experts? Possibilities and limits of a strategy-based intervention. *Journal of Media Literacy Education, 12*(2), 64-79. <https://doi.org/10.23860/JMLE-2020-12-2-6>
- Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive development, 15*(3), 309-328.
- Kuhn, D., & Udell, W. (2007). Coordinating own and other perspectives in argument. *Thinking & Reasoning, 13*(2), 90-104. <https://doi.org/10.1080/13546780600625447>
- Macedo-Rouet, M., Potocki, A., Scharrer, L., Ros, C., Stadtler, M., Salmerón, L., & Rouet, J.-F. (2019). How Good Is This Page? Benefits and Limits of Prompting on Adolescents' Evaluation of Web Information Quality. *Reading Research Quarterly*. <https://doi.org/10.1002/rrq.241>
- Mason, L., Moè, A., Tornatora, M. C., & Ronconi, A. (2022). Promoting Web-Source Evaluation and Comprehension of Conflicting Online Documents: Effects of Classroom Interventions. In P. Limone, R. Di Fuccio, & G. A. Toto (Eds.), *Psychology, Learning, Technology* (p. 3-21). Springer International Publishing. [https://doi.org/10.1007/978-3-031-15845-2\\_1](https://doi.org/10.1007/978-3-031-15845-2_1)
- Miller, S. A. (2009). Children's understanding of second-order mental states. *Psychological Bulletin, 135*(5), 749-773. <https://doi.org/10.1037/a0016854>
- Osterhaus, C., & Koerber, S. (2021). The Development of Advanced Theory of Mind in Middle Childhood: A Longitudinal Study From Age 5 to 10 Years. *Child Development, 92*(5), 1872-1888. <https://doi.org/10.1111/cdev.13627>
- Paul, J., Macedo-Rouet, M., Rouet, J.-F., & Stadtler, M. (2017). Why attend to source information when reading online? The perspective of ninth grade students from two different countries. *Computers & Education, 113*, 339-354. <https://doi.org/10.1016/j.compedu.2017.05.020>
- Pérez, A., Potocki, A., Stadtler, M., Macedo-Rouet, M., Paul, J., Salmerón, L., & Rouet, J.-F. (2018). Fostering teenagers' assessment of information reliability: Effects of a classroom intervention focused on critical source dimensions. *Learning and Instruction, 58*, 53-64. <https://doi.org/10.1016/j.learninstruc.2018.04.006>
- Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. *Precursors of functional literacy, 11*, 67-86.

- Perfetti, C.A., Rouet, J.-F., & Britt, M.A. (1999). Towards a theory of documents representation. in H. van Oostendorp & S.R. Goldman (Eds.) *The Construction of Mental Representations During Reading* (pp. 99-122). Lawrence Erlbaum Associates.
- Potocki, A., Châtelier, M., Dyoniziak, Y., Pylouster, J., & Rouet, J. (2023). Disponibilité des textes et performances en compréhension écrite à l'adolescence : Quelle implication des fonctions exécutives ? *L'année psychologique*. [Role of text availability on reading performance and strategies in adolescence: What is the involvement of executive functions?].
- Potocki, A., de Pereyra, G., Ros, C., Macedo-Rouet, M., Stadtler, M., Salmerón, L., & Rouet, J.-F. (2020). The development of source evaluation skills during adolescence : Exploring different levels of source processing and their relationships (El desarrollo de las habilidades de evaluación de las fuentes durante la adolescencia: una exploración de los distintos niveles de procesamiento de las fuentes y sus relaciones). *Journal for the Study of Education and Development*, 43(1), 19-59. <https://doi.org/10.1080/02103702.2019.1690848>
- Sabatini, J. P., O'reilly, T., Halderman, L., & Bruce, K. (2014). Broadening the scope of reading comprehension using scenario-based assessments: Preliminary findings and challenges. *L'Année Psychologique*, 114(04), 693-723. <https://doi.org/10.4074/S0003503314004059>
- Sodian, B., & Kristen, S. (2016). Theory of mind. In J. Greene, W. Sandoval, & I. Bråten, *Handbook of Epistemic Cognition*. Routledge.
- Stadtler, M., Babel, S., Rouet, J. F., & Bromme, R. (2014). Ninth-grade students possess good sourcing skills, but do not apply them spontaneously while reading. *Annual Meeting of the American Educational Research Association, Philadelphia, USA*.
- Stadtler, M., & Bromme, R. (2014). The content-source integration model: A taxonomic description of how readers comprehend conflicting scientific information. In *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (p. 379-402). MIT Press.
- Strømsø, H. I., & Kammerer, Y. (2016). Epistemic cognition and reading for understanding in the internet age. In *Handbook of Epistemic Cognition*. Routledge.
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*.

- Wang, Z., Devine, R. T., Wong, K. K., & Hughes, C. (2016). Theory of mind and executive function during middle childhood across cultures. *Journal of Experimental Child Psychology*, 149, 6-22. <https://doi.org/10.1016/j.jecp.2015.09.028>
- White, S., Hill, E., Happé, F., & Frith, U. (2009). Revisiting the strange stories: Revealing mentalizing impairments in autism. *Child development*, 80(4), 1097-1117.
- Wineburg, S.S. (1991). Historical problem solving: a study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, 83, 73-87. <https://doi.org/10.1037/0022-0663.83.1.73>
- Wineburg, S., & McGrew, S. (2016). *Evaluating information: The cornerstone of civic online reasoning*.  
<https://stacks.stanford.edu/file/druid:fv751yt5934/SHEG%20Evaluating%20Information%20Online.pdf>