Journal of Numerical Cognition

PsychOpen GOLD

Theoretical Contributions

Thirty Years of the Give-N Task: A Systematic Review, Reflections, and Recommendations

Theresa Elise Wege[1], Taeko Bourque[2], Rebecca Merkley[2], Pierina Cheung[3]

[1] *Centre for Mathematical Cognition, Loughborough University, Loughborough, United Kingdom.* **[2]** *Department of Cognitive Science, Carleton University, Ottawa, Canada.* **[3]** *National Institute of Education, Nanyang Technological University, Singapore (City), Singapore.*

Journal of Numerical Cognition, 2025, Vol. 11, Article e10577, https://doi.org/10.5964/jnc.10577

Received: 2022-11-03 • **Accepted:** 2025-02-01 • **Published (VoR):** 2025-06-06

Handling Editor: Andreas Obersteiner, Technical University of Munich, Munich, Germany

Corresponding Author: Pierina Cheung, Centre for Research in Child Development, National Institute of Education, Nanyang Technological University, 1 Nanyang Walk, Singapore 637616. E-mail: cheung.pierina@gmail.com

Supplementary Materials: Data, Preregistration [see Index of Supplementary Materials]

Abstract

The Give-N (give-a-number) task has become a popular assessment of children's number words and counting knowledge since Wynn's (1990, 1992) seminal work over 30 years ago. Using the Give-N task, numerous studies have shown that children learn the first few number words slowly, before they understand how counting represents number. This learning trajectory and children's associated behaviors on the Give-N task are represented by "knower-levels" and form the basis for a large body of research assessing children's number learning. Recent research has started to critically analyze the theoretical conceptualisation and reliability of knower-levels. We added to this work by conducting a systematic review of studies using the Give-N task. This review provides an overview of methodological practices and variations in the task's administration and scoring of knower-levels which have theoretical and methodological implications. We argue that advancing methodology and theory for research in children's number learning requires (1) consideration of Give-N task administration and scoring in study design and reporting and (2) reflection on the assumptions and limitations of classifying children's performance on the Give-N task in the knower-level framework.

Keywords

number words, counting, cardinal principle, give-a-number, Give-N, knower-levels

Highlights

- The Give-N task has become a standard assessment for measuring children's number knowledge.
- We began the study with the realization that knower-levels from the Give-N task are not coded in the same way across studies, and that information was lacking on what are common methodological practices.
- We conducted a systematic review of studies that used the Give-N task.
- Our analysis highlights how the core assumptions of the knower-level framework have become intertwined with Give-N task methodology and identifies gaps in our understanding of early number development.

The Give-N task is a set production task commonly used to measure the development of children's knowledge of the cardinal meaning of number words and knowledge of counting (e.g., Almoammer et al., 2013; Barner, Libenson, et al., 2009; Le Corre et al., 2016; Sarnecka et al., 2007; Schaeffer et al., 1974; Sella et al., 2017; Wynn, 1990, 1992). In this task,

children are asked to produce a set of objects corresponding to a number word ("Can you give me four pebbles?"). Their success or failure at doing so is assessed for different number words. The premise is that knowledge of the cardinal meaning of number words and the cardinal principle of counting allows children to accurately produce sets of the exact number asked for on the Give-N task.

The Give-N task was first used to assess children's number word and counting knowledge by Schaeffer and colleagues (1974), who found that children who correctly gave larger numbers can also correctly give all the small numbers. Wynn (1990) later refined the task and showed that some children act as "counters" and some as "grabbers" (Wynn, 1990, 1992; see also Frye et al., 1989). Counters responded with similar accuracy to requests for small and large number words and used counting to give N objects. In contrast, grabbers only responded correctly for requests up to "one", "two" or "three" and, instead of using counting for larger requests, they grabbed an arbitrary number of objects. When asked how many objects they gave, these children sometimes responded with counting, e.g., "one, two, three, four, five" but were often unable to act on that information when prompted, e.g., "but I wanted four, can you fix it so that I have four?". Counters were mostly children older than 3.5 years and younger children were mostly grabbers. By observing these patterns of performances on the Give-N task longitudinally, Wynn (1992) concluded that children begin to learn the cardinal meaning of the small number words and only later learn the cardinal principle of counting. Specifically, once children know the cardinal principle of counting, they understand that sets corresponding to a given number word can be produced by applying the counting procedure. Therefore, children who are cardinal principle knowers can generate any set – up to the highest number they can count to – upon request on the Give-N task. Since Wynn, numerous studies across different languages and cultures have replicated the finding that children go through different 'knower-levels' (e.g., Barner, Chow, & Yang, 2009; Barner, Libenson, et al., 2009; Ceylan & Aslan, 2018; Chang & Sandhofer, 2019; Knudsen et al., 2015; Koike & Yamagata, 2014; Lê & Noël, 2020; Le Corre et al., 2016; Marchand & Barner, 2021; Marušič et al., 2016; Piantadosi et al., 2014; Sarnecka et al., 2007; Sorariutta et al., 2017; Stojanov, 2011).

The Knower-Level Framework and Give-N Task Are Intertwined

The knower-level framework, and correspondingly the Give-N task, have been used extensively in the last 30 years of research into children's number word learning (see Barner, 2017; Carey & Barner, 2019; Sarnecka, 2015; Sella et al., 2021, for reviews). Supported by empirical data drawn primarily from the Give-N task, current theories of early number learning are based on the observations that (1) children learn to recite the count list "one, two, three, four, five…" as early as age 2 but have not yet ascribed numerical meaning to the number words and the counting procedures; (2) between usually 2 and 4 years of age children learn the meanings of the number words up to "three" or "four" sequentially; (3) after learning the meaning of the first number words, children understand the cardinal principle of counting. This stage-like development is correlated with other aspects of numeracy development. For example, children who can only produce some of the first few numbers on the Give-N task were found to lack knowledge of cardinality in a counting puppet task compared to those who can accurately produce larger numbers on the Give-N task (Le Corre et al., 2006). Other studies found a correlation between non-symbolic numerical comparisons and cardinal principle knowledge on the Give-N task (Abreu-Mendoza et al., 2013; Cheung & Le Corre, 2018; Negen & Sarnecka, 2015). Geary et al. (2018) also found that those who learned the cardinal principle of counting early are better at later mathematics (e.g., arithmetic, number line estimation). These findings support the stage-like development of children's learning of number words and counting that is inherent in the knower-level framework.

The knower-level framework is based on three core assumptions, which are in turn underpinned by children's performance on the Give-N task.

These core assumptions are:

1. knowledge of exact number word meanings is represented as all-or-none
2. knowledge of a subset of small number word meanings up to around "four" is acquired one at a time in sequence → (1) together with (2) leads children who are just learning the small number words to perform accurately on the Give-N task up to a certain number (i.e., their knower-level) and to fail the Give-N task for any larger number

3. knowledge of counting changes when children learn the cardinal principle
   → which leads children who have learned the cardinal principle to count to solve the Give-N task, especially for larger numbers.

In the knower-level framework children who only know the small number words are called 'subset-knowers', and those who understand the cardinal principle of counting are called 'cardinal principle knowers' (CP-knowers). Differences in performance on the Give-N task between subset-knowers and CP-knowers provides one of the foundations for current theories that propose a developmental discontinuity between knowledge of a subset of the small number words (knowing numbers "one" through "three" or "four") and knowledge of the cardinal principle of counting (see Barner, 2017; Carey & Barner, 2019; Sarnecka, 2015; Sella et al., 2021, for review). Classifying children as subset- and CP-knowers has facilitated insights into number learning through the investigation of group differences, targeted interventions, and developmental trajectories (e.g., Geary et al., 2018; Huang et al., 2010; O'Rear & McNeil, 2019; Sella & Lucangeli, 2020). However, recent studies have presented nuances of early number learning that pose some challenge to the core assumptions of the knower-level framework.

Subset-knowers may not have all-or-none knowledge of just certain small numbers, but may have partial knowledge of higher numbers. Some studies have argued that, for example, a child might have partial knowledge of the number "two", if they reliably give 2 objects when asked to give "two", but also sometimes make overextension errors and give 2 objects when asked for other numbers (Barner & Bachrach, 2010; O'Rear & McNeil, 2019).

The boundary between subset knowledge and cardinal principle knowledge may not be fixed at "three" or "four". Krajcsi and Fintor (2023) found a group of children who succeed on some large numbers such as "five" but fail to give larger ones such as "six" or "seven" and termed these children large-number subset-knowers (see also Rousselle & Vossius, 2021). To some, these results suggest that what constitutes the subset-knower stages does not necessarily stop at "three" or "four" as previously assumed, which have theoretical implications for what cognitive representations may underlie subset-knowers' knowledge of number words (Carey, 2009; Cheung & Le Corre, 2018; Le Corre & Carey, 2007; Schneider et al., 2022; vanMarle et al., 2018).

Recent studies also show that tasks that are thought to be interchangeable with the Give-N task may, in fact, tap into different aspects of cardinal principle knowledge (e.g., Baroody & Lai, 2022; Baroody et al., 2023; O'Rear et al., 2024; Mou et al., 2021). They show that creating a set that matches a number word (as assessed by the Give-N task) is likely distinct from identifying or labelling the exact number of objects in a given set (as assessed by How Many Task, What's-on-this-Card Task).

Give-N Task Methodology and Issues of Measurement Reliability

Some of the above-mentioned nuances of early number learning may relate to variations in how the Give-N task is administered and scored, which affect our ability to compare across studies (Krajcsi, 2021; Marchand & Barner, 2021; see also Le Corre et al., 2006). While keeping all other aspects of the Give-N methodology constant, these studies show that (1) administering the trials of the Give-N task as a fixed list or in a titration method leads to similarly reliable knower-level *categories* (e.g., non-knowers vs. subset-knowers vs. CP-knowers), but individual knower-levels were not always reliable (Marchand & Barner, 2021), (2) asking for numbers larger than "six" leads to differences in knower-level classification for children who may otherwise be equally classified as CP-knowers (Krajcsi & Fintor, 2023), (3) not asking follow-up questions leads to lower knower-level classifications, especially for potential CP-knowers (Krajcsi, 2021; Le Corre et al., 2006) and (4) using different types of objects affects knower-level classifications (Petersen & McNeil, 2013). Overall, even seemingly minor, isolated variations in the administration of the Give-N task affect knower-level classifications.

Furthermore, studies vary in how children's performance on the Give-N task is scored and classified within the knower-level framework. For example, some studies classify children as CP-knowers if they can give sets correctly up to "six" (e.g., Le Corre et al., 2006), other studies use "four" as the cut-off (e.g., Geary et al., 2017), or apply stricter rules and require children to give sets correctly up to "ten" to be classified as CP-knowers (e.g., Kimura et al., 2013). In addition, children can be classified as an n-knower or an n+1 knower depending on whether one scores overextension errors (e.g., Barner & Bachrach, 2010; O'Rear et al., 2020). When applying this scoring criterion, a child who gives two

for numbers other than "two" would be considered a "one-knower", even if they also correctly gave two when asked for "two". Children's knower-level classification can thus vary depending on whether overextension errors are taken into account when coding knower-levels. While some theoretical and modelling work has considered the effect of scoring (Lee & Sarnecka 2011; Sella et al., 2021), empirical investigations into effects of scoring variations for the measurement reliability of the Give-N task are still sparse.

The Present Study

To inform future debates on methodological issues surrounding the Give-N task and theoretical ones about how children learn the meaning of number words and counting, we conducted a systematic review of the Give-N literature. We investigated what aspects of the Give-N task are commonly found across studies and which aspects vary from study to study as an indicator of common and diverging assumptions across researchers. The systematic review minimizes biases by including all relevant studies that meet the search criteria (see details below). We selected variables on the administration and scoring of the Give-N task that are key to the core assumptions of the knower-level framework.

The aims of our review are (1) to provide an overview of the Give-N task literature to date; (2) to examine the scope of methodological similarities and differences in the administration and scoring of the Give-N task, specifically in reference to the core assumptions of the knower-level framework; and (3) to discuss potential theoretical and methodological implications for current research on number word learning that draws heavily on the Give-N task.

Method

Protocol

We followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) guidelines (Page et al., 2021) in developing this systematic review protocol.

Information Sources

The literature search was done in May 2021. The search terms were (("Give-N task") OR ("Give-N task") OR ("give-me-x") OR ("give-x") OR ("give-me-n") OR ("make a set of") OR ("give-n") OR ("give-n") OR ("Give-Number task") OR ("knower-level") OR ("knower level") OR ("cardinality") OR ("cardinal principle")) AND (("preschool") OR ("children") OR ("kindergarten") OR ("nursery")). We searched PsycINFO, Web of Science, ERIC Electronic databases, and SCOPUS for primary research articles. We also searched for all primary research articles that cited Wynn (1990) or Wynn (1992), given those citations are most commonly used for the Give-N task and the knower-level framework. Grey literature sources such as dissertations, theses and preprints were included. Published conference proceedings were included; oral or poster conference presentations were not included.

Eligibility Criteria

All published empirical studies that reported the use of a Give-N task were eligible to be included. A Give-N task was defined as a set production task with manipulable objects in response to verbal requests for multiple numbers across several trials. Articles that reported a Give-N task were not included if participants were non-human animals, adults, or children over the age of 6 years old. Articles were further excluded if they reported secondary data (in this case only the study reporting the primary data was included) or if the full text was not available in English.

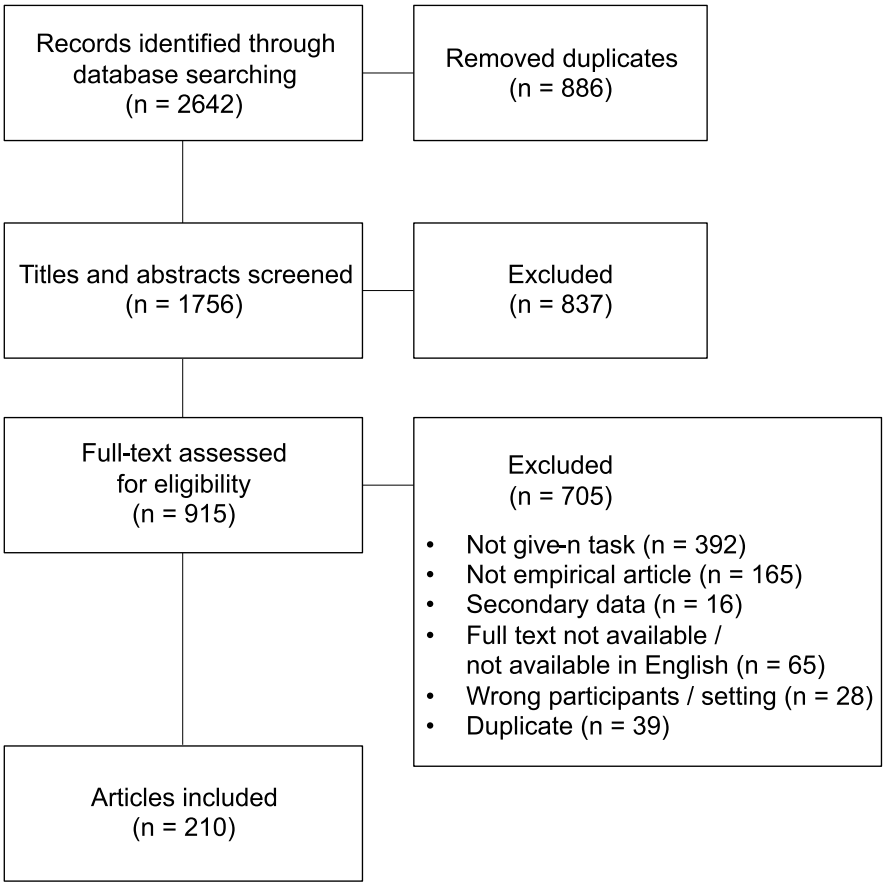
Article Screening and Selection

After conducting the initial search, we stored the search results with a commercial software for managing systematic reviews, Covidence (n.d., www.covidence.org). After removing duplicates, search results were then screened for eligibility. In a first step, three of the authors screened titles and abstracts and excluded articles that obviously did not meet our eligibility criteria. Each article was assessed independently by two randomly assigned coders. Coders were in moderate

agreement about inclusion or exclusion at this stage of the screening process (agreement: 78%, Cohen's kappa = .58). Disagreement was largely due to one coder making more conservative judgements and being almost twice as likely to exclude an article at this stage. All disagreements between two coders were resolved by the third coder. In a second step, the full text of the remaining articles was assessed for eligibility. Articles were included or excluded based on the consensus of two coders randomly selected from all four authors. Coders were in almost perfect agreement at this stage (agreement: 94%, Cohen's Kappa = .84). Where coders did not agree, the entire research team discussed whether to include or exclude an article. Each step in the selection process was recorded with a PRISMA flow chart (Page et al., 2021) (Figure 1).

Figure 1

PRISMA Flowchart of Article Screening Process and Selection



Coding Scheme and Data Extraction

We used Covidence to extract data from all 210 included articles based on a coding scheme that was developed before we started the review process but was further fine-tuned during the planning and training stages. We coded information for three broad categories: setting (e.g., study location, participant information, details of data collection), Give-N administration (e.g., details of materials and procedure used in the Give-N task), and Give-N scoring (e.g., scoring framework and knower-level scoring criteria, number of children assigned to each knower-level and their age distribution). We selected elements on administration and scoring that are key to interpreting the assumptions of the knower-level framework, and these elements may affect knower-level classifications.

To refine the coding scheme and ensure coding consistency, the entire research team jointly coded a small selection of articles. To further ensure the coding scheme is used consistently across coders, 10 randomly selected articles were coded by two coders and the extracted data was compared. Coders agreed on over 90% of the extracted data and disagreements were again resolved by discussion.

The remaining articles were coded by one of the two independent coders, the two first authors. Only articles that posed a challenge to our coding scheme were coded by both coders and discussed with the entire research team. For each article, we coded only the first study that included a Give-N task. Consequently, for articles that used the Give-N task in multiple studies or at multiple timepoints in a longitudinal study we only created one data entry. This decision was made to increase the independence of data entries from each other.

A codebook with detailed explanations of all variables and all extracted data, including data not reported in this review, are publicly available on the OSF (Wege et al., 2022S). Data may be reused and adapted under a CC-BY-NC-SA license.

Results

We present the results of our systematic review in two parts: (1) overview of the Give-N literature including a brief analysis of Give-N methodology in studies that did not apply the knower-level framework and (2) an in-depth analysis of methodological practices and variations in Give-N methodology within the knower-level framework.

Overview of the Give-N Literature

Our search identified 210 unique articles that described an empirical study using a Give-N task, with about 50% of them published since 2014 (see the Appendix for a visual summary). Taken together these studies report data on the Give-N task from over 15,000 children between mainly 2-5-years of age. Almost 90% of all articles using the Give-N task reported studies conducted in North America and Europe with most studies specifically conducted in the USA, Canada, and the UK. This is in line with the broader field of cognitive development, where most research samples English speaking children based in North America or the United Kingdom (Draper et al., 2023; Nielsen et al., 2017; Singh et al., 2021).

Give-N tasks that meet our definition of a set production task with a verbal request and manipulable objects are referred to by a variety of names (e.g., “give-a-number”, “give-me-x”, “cardinality task”) and reported to measure a variety of constructs with a variety of methodologies. About half of the studies included in our review (n = 109) used the knower-level framework to score and interpret children’s performance on a Give-N task. In the remaining studies (n = 101), the Give-N task is reported to measure cardinality knowledge or number word knowledge, either on its own or in combination with other measures (e.g., scores on a “how many?” task); it is also often reported as one of several measures of general numeracy. The two most prevalent scoring frameworks for the Give-N task beyond the knower-level framework are proportion of correct responses and Grabbers vs. Counters (but note that these scoring frameworks are also used alongside the knower-level framework in some studies). We provide a brief overview of the Give-N methodology of these two scoring frameworks before reporting in-depth analyses of variations in Give-N methodology within the knower-level framework.

Proportion Correct

Rather than classifying children into knower-levels, some studies (24%) report children’s performance as the proportion of trials in which they give n correctly, regardless of which particular number they give correctly (e.g., Cahoon et al., 2021; Kirk et al., 2017; Nanu et al., 2018; Nicoladis et al., 2010; Salsa & Martí, 2015; Steele et al., 2012; Verdine et al., 2014). Unlike studies that adopt the knower-level framework, these studies frequently include trials that request “ten” or larger numbers (e.g., Bakker et al., 2019; Cankaya et al., 2014; Purpura & Simms, 2018). Scoring the Give-N task as proportion of correct responses allows researchers to derive a continuous variable, which has been interpreted as a measure of counting accuracy, a proxy of general numerical skills or as a continuous measure of knowledge of the cardinal principle.

Grabbers vs. Counters

The distinction between Grabbers and Counters based on performance on the Give-N task was first introduced by Wynn (1990) and has since been used in a few more studies (e.g., Bruce & Threlfall, 2004; Dowker, 2005; Fluck et al., 2005; Munn, 1994). Scoring criteria in the Grabbers vs. Counters framework often consider children's performance holistically across all trials and are generally more flexibly defined. Wynn (1990) for instance scored children as Counters, even if they counted incorrectly, as long as they counted in response to the Give-N requests. Children who do not spontaneously count aloud may also be scored as Counters if they give items deliberately one by one or point at items one by one (Rousselle & Vossius, 2021). It may not always be obvious whether children are engaged in counting behavior and reliable scoring can require laborious double coding of children's behaviour (e.g., Jara-Ettinger et al., 2016; Posid & Cordes, 2018). Thus, although the knower-level framework predicts a shift in the use of strategies in solving the Give-N task, few studies have analyzed children's counting behavior.

Methodological Practices and Variations Within the Knower-Level Framework

We identified 109 articles that report the use of a Give-N task within the knower-level framework, and we refer exclusively to these articles in this section.

One of the most frequently used protocols of the Give-N task is a narrative set-up that asks children to give *n* as "Can you make four ducks/fish jump in the pond?" and provides 10 to 15 toy ducks/fish and a bowl or a blue piece of paper. Although approximately two-thirds of the studies we reviewed used a narrative set-up, studies vary in details of administration and scoring. In what follows, we review four aspects of the administration of the Give-N task because they bear on the core assumptions of the knower-level framework and might affect knower-level classifications.

1. Administration

The administration of the Give-N task varies across both the materials and the procedure used (Figure 2). As shown in Figure 2, few aspects of the Give-N task apply to more than 50% of the studies (see dark grey bars), and many studies lack sufficient reporting detail (see dotted bars).

1.1. Trial Structure: Was a Titration or Fixed Trial List Used? — One of the variations in the procedure of the Give-N task that has been examined empirically is the trial structure of administration (see Marchand & Barner, 2021): number words are either requested according to the titration method (e.g., Gibson et al., 2020; Sarnecka & Carey, 2008; Wynn, 1992) or a fixed list of trials (e.g., Brueggemann & Gable, 2018; Negen & Sarnecka, 2015; Slusser et al., 2019; Wagner et al., 2019). The use of the titration method assumes the sequential learning pattern in the knower-level framework, since children are asked the minimum number of trials needed for a knower-level classification. The titration method is more commonly used (69% of reviewed studies) than the fixed trial list method (27%); likely because the titration method was reported in Wynn (1990, 1992) and is on average quicker to administer since it often includes fewer trials (but note that the administration is also overall more difficult and more prone to experimenter error). The strong preference for the titration method could also indicate the assumptions many researchers make about the knower-level framework, namely the sequential learning pattern of small numbers and the all-or-none nature of number knowledge. In the fixed trial list, all children received a fixed number of requests for a fixed set of number words, including both small and large number words.

1.2. What Numbers Are Requested and How Often? — The range of numbers requested is closely related to the trial structure and assumptions underlying the knower-level framework. Under the knower-level framework, small number words from "one" to "four" are acquired in sequence. These are the numbers that are asked in most studies so children can be classified into one-knowers, two-knowers, etc. based on accuracy on these trials. Additionally, larger numbers are asked to test whether children are CP-knowers, and this often includes "five" and "six" (e.g., Condry & Spelke, 2008; Gunderson et al., 2015; Sarnecka & Carey, 2008), with few asking up to "ten" (e.g., Almoammer et al., 2013; Kersey et al., 2018; Liang et al., 2021; Shusterman et al., 2016). Numbers larger than "six" are not always asked. This links to the core assumption that once children know the cardinal principle of counting, they can correctly generate sets

for all numbers within their count list. Thus their performance on "six" is not expected to differ from "eight", "nine", or "ten" (but see Krajcsi & Fintor, 2023).

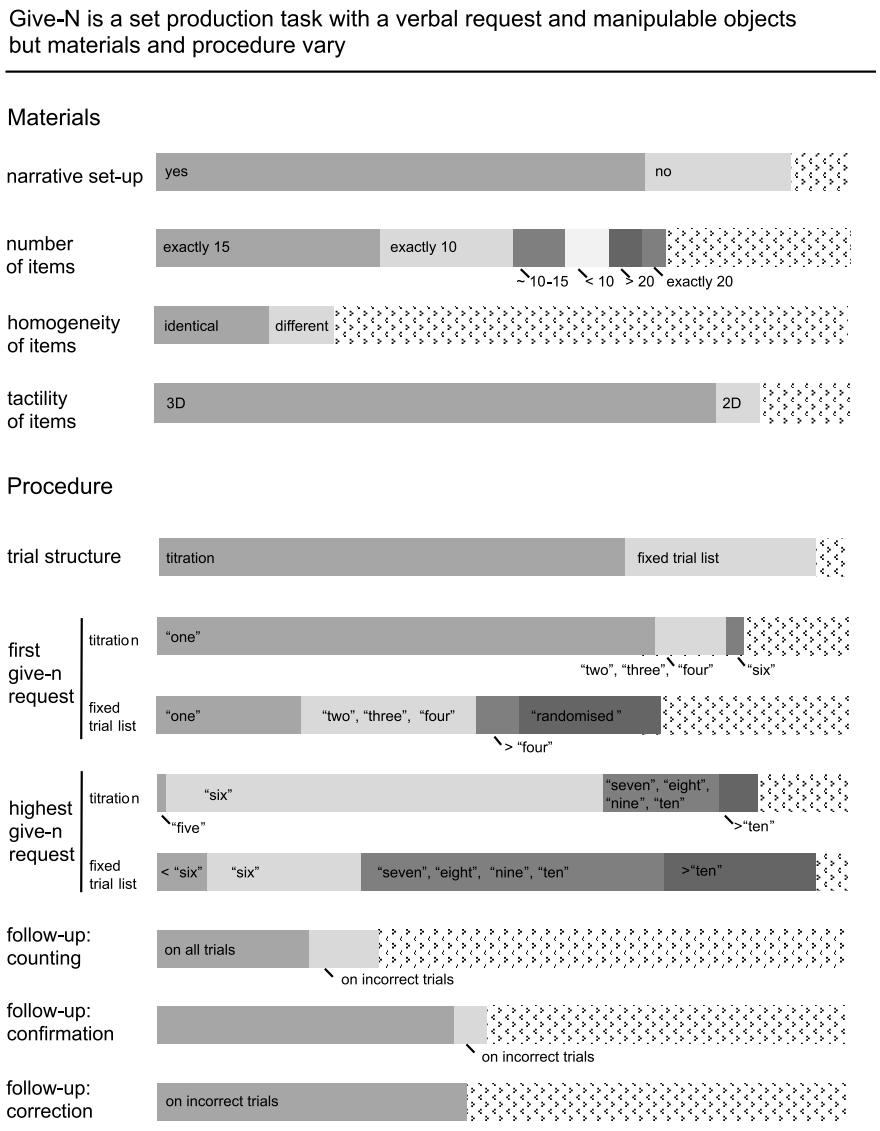
The range of numbers requested also differs as a function of trial structure. The titration method optimizes the administration of the Give-N task for the knower-level framework, since it only requires the administration of as many trials as necessary to find the highest number a child can accurately give. It can thus include as few as four trials (e.g., two trials each of "one" and "two" if a child is a non-knower) and rarely includes requests for numbers larger than "six". In contrast, the fixed trial list method always requests the numbers "one", "two", "three", "four", "six" and commonly also "five", "eight" and/or "ten" (e.g., Marušič et al., 2016; Sella et al., 2017; Wagner et al., 2015; see Figure 2, highest Give-N request), with each number commonly requested 2 to 3 times. Numbers larger than "six" are more often requested under the fixed trial list, which is likely more suitable for testing the core assumptions of the knower-level framework than the titration method.

1.3. What Stimuli Are Used? — The type and size of stimuli used could affect children's counting and grabbing behavior of sets, which might affect response accuracy. Most studies use three-dimensional manipulatives (e.g., toy animals or counters; Desoete et al., 2013; Jamalian, 2014; Merritt & Brannon, 2013; Pinhas et al., 2014; Scalise & Ramani, 2021) but flat disks or cards have also been used (e.g., Mateu & Hyams, 2016; Meyer et al., 2020; Rousselle et al., 2004; Sella et al., 2017). The size of an object is rarely reported, despite large objects being harder to grab than smaller ones. Recently, a digital set-up was introduced with stimuli on a screen (de Vries et al., 2021). Although children could show signs of grabbing behavior with digital stimuli, systematic comparisons between the physical and digital versions are lacking. In addition, we note an important feature of stimuli that was insufficiently reported, namely, whether the stimuli were identical-looking objects (e.g., 15 identical looking green counters) or objects of the same kind but had different physical attributes such as color (e.g., 10 counters consisting of 5 green and 5 yellow one, or 10 different fish). While varying object features might make the task more engaging for young children, having different numbers of copies of each item could bias children's responses (e.g., if children prefer a particular color).

1.4. Are Follow-up Questions Asked or Not? — Previous studies have shown that whether children are asked to count could affect their knower-level classification, especially those who may be CP-knowers (Krajcsi & Fintor, 2023; Le Corre et al., 2006). We analyzed the types of follow-up questions that are typically asked *after* the child has provided an initial response to the request to give a number. There are three types of follow-up questions: confirmation, counting, and correction. The most frequently used follow-up question is to ask children after a trial to confirm that the number of items they provided corresponds to the number requested, e.g., "Is this three fish?". The confirmation question is included to ensure that children are done putting objects on a plate; thus, if this follow-up question is included, it is typically asked on each trial (e.g., Chernyak et al., 2016; Lindskog & Simms, 2021; Savelkouls et al., 2020). Another type of follow-up question is to ask children to count the items they provided. Studies differ with respect to the consistency of this request, some asking it on every trial (e.g., Davidson et al., 2012; Moore et al., 2016) and others only if children gave an incorrect number of items (e.g., Barner, Chow, & Yang, 2009; Spaepen et al., 2018). Finally, confirmation and counting requests may result in correction opportunities where children are prompted to change the number of items they provided and make it n items (e.g., Gibson et al., 2020; Savelkouls et al., 2020; Szkudlarek & Brannon, 2018). Surprisingly, about 50% of the articles we reviewed (out of 109) do not report the use of follow-up questions, but it is possible that follow-up questions were asked but were not explicitly reported. The inclusion of follow-up questions, specifically the counting requests, provides a separate test for the assumption that only children who have knowledge of the cardinal principle use counting to solve the Give-N task. This is because counting requests that result in correction opportunities provide another view on how children solve the Give-N task by testing whether children can use counting to fix an already given set.

Figure 2

Overview of Variations in Key Aspects of the Administration of the Give-N Task



Note. Each bar represents the total of 109 reviewed articles. Dotted bar segments represent the number of articles that did not report details on a particular aspect of administration.

2. Scoring

As expected, most studies (80%) included a distinction between subset- and CP-knowers in the scoring. Of the studies that did not classify children as subset- or CP-knowers, most scored knower-levels for all numbers that were asked for in the task, resulting in six-knowers, seven-knowers, etc. (e.g., Dresen et al., 2022; Gordon et al., 2021; Mussolin et al., 2014). In these studies, children's knower-level was often used as a continuous variable for analysis. Some studies (8%) only classified particular knower-levels for the purpose of sample inclusion and hence did not report information beyond those particular knower-levels (e.g., only three- and four-knowers were scored in Huang et al., 2010, for a number word training study). Despite most studies classifying children into subset vs. CP-knowers, studies vary in the

criteria for classifying children. In the following subsections, we analyzed the criteria used to classify (1) children as n-knowers within the subset knower range and (2) children as CP-knowers.

2.1. What Makes an n-Knower? — A core assumption of the knower-level framework is that knowledge of small number words is all-or-none, such that children respond correctly on the Give-N task when asked for known numbers and incorrectly for unknown numbers. In about 75% of studies, the criterion for knowing a number is a "two out of three" rule: A child who gives n correctly twice can be classified as an n-knower even if they also give n incorrectly once. Conversely, the largest known n is commonly identified by children's failure to give the next larger number $n + 1$ "two out of three" times. Although most studies use a "two out of three" rule, studies vary on what responses qualify as giving n incorrectly. To score how often children gave n incorrectly, about 50% of studies also scored whether children gave n when asked for any other number. For example, in those studies, giving three items when asked for "four" would be scored as incorrectly giving both "three" and "four" (e.g., Batchelor et al., 2015; Condry & Spelke, 2008; Slusser et al., 2013).

In the titration method, an n-knower must have also correctly given numbers $< n$ at least once, since this is the prerequisite for even reaching trials requesting n. This is not the case for the fixed trial list. In the fixed trial list, performance on numbers $< n$ is considered when classifying children as n-knowers using the Bayesian scoring model (e.g., Sella et al., 2017; Sella & Lucangeli, 2020). The Bayesian scoring model calculates the likelihood of a child being each of the different knower-levels based on theoretical assumptions about the response patterns of different knower-levels. This includes the assumption that an n-knower successfully gives all numbers $< n$ (Lee & Sarnecka, 2010, 2011; Negen et al., 2012). Studies have also applied the "two out of three" rule and not allowed for numbers to be skipped in classifying n-knowers under the fixed trial list method.

As noted, some studies included follow-up questions that allowed for corrections ("Can you make this n?") after incorrect trials. We did not find any explicit reporting on how these corrections are incorporated into the scoring criteria. We assume that the correction replaced the erroneous response. However, the correction could also be counted as an additional response, and this would affect scoring rules (e.g., if a child incorrectly gives four objects when asked for "three" and then incorrectly fixed to two, does this count against knowing "four" *and* "two"?)

2.2. What Makes a Cardinal Principle Knower? — We found two main criteria for performance on the Give-N task that result in children being classified as CP-knowers: (1) minimum knower-level and (2) overall success. The minimum knower-level criterion requires children to succeed at or up to at least a certain number without a requirement to also succeed at trials requesting numbers larger than the minimum knower-level. In contrast, the overall success criterion requires children to succeed at all requested numbers. Note that for certain administrations of the Give-N task, these criteria can be identical, e.g., scoring six-knowers as CP-knowers in a titration method that only requests numbers up to "six" also satisfies "overall success". We coded and summarized such instances as the minimum knower-level criterion.

Out of the studies that scored CP-knowers, most studies applied a minimum knower-level criterion (82%). For instance, Wagner and colleagues (2015) report: "If n [the knower-level] was five or higher, the child was classified as a CP-knower." While we did find a handful of studies in which four-knowers were considered cardinal principal knowers (6%) (e.g., Villarroel et al., 2011; see also Wynn, 1992, who included four-knowers with the CP-knowers), most frequently the minimum knower-level was six-knower (56%) or five-knower (38%). Note that even if the minimum knower-level to be classified as a cardinal principle knower was six-knower, five-knowers are frequently not included in the subset knower range. This is because studies may not request "five" in the trial list or may recode five-knowers as CP-knowers for theoretical considerations about the subset knower range being limited at four (for discussions of five-knowers see e.g., Krajcsi & Fintor, 2023; Marchand et al., 2022; Posid & Cordes, 2018; Sarnecka et al., 2015).

The "overall success" criterion was applied in 10% of studies that scored CP-knowers. Given that studies vary in how "success" is defined, and which numbers are requested (see above), the actual performance on the Give-N task that indicates overall success and classifies children as CP-knowers can vary. Overall success criteria, for instance, include correctly giving numbers "five", "six", "seven" and "nine" once (Schneider et al., 2021), giving all numbers between "one" and "ten" correctly in "two out of three" trials (Kimura et al., 2013) and correctly giving all requested numbers to the extent where it likely indicates cardinal principle knowledge in the Bayesian model (Sella et al., 2017).

We found only a few studies (8%) after Wynn's introduction of the knower-levels that explicitly considered children's overt counting behavior (e.g., whether they follow the stable-order or word-object correspondence principle) when scoring cardinal principle knowledge (e.g., Fluck & Henderson, 1996; Le Corre & Carey, 2007). Hence, although the cardinal principle is directly related to how counting maps onto number and some studies explicitly include counting in the Give-N task ("Can you count and make sure?"), overt counting behavior is not often considered when scoring Give-N performance.

Discussion

The Give-N task has become an increasingly popular task to assess children's number word and counting knowledge. Unsurprisingly, with increasing popularity of the task comes variability in how it is run. In our review, half of the published papers applied the knower-level framework. The other half of Give-N studies mainly used overall proportion of correct responses as a continuous measure of early numeracy. We began the study with the realization that knower-levels from the Give-N task are not coded in the same way across studies, and that information was lacking on what are common methodological practices. Documenting methodological practices – with a focus on both similarities and differences – of the Give-N task provides data on whether there is a standard version of the Give-N task that most researchers use, or whether efforts to optimize reliability of the task are needed. Although our study cannot address whether methodological variations of the kind we documented here influence knower-level classifications (see Krajcsi, 2021; Marchand & Barner, 2021; for this endeavour), our descriptive analysis highlights key assumptions researchers have about the knower-level framework and gaps in our current knowledge of early number development.

On one hand, we found that some aspects of the knower-level framework have become entrenched across the literature. For example, most studies used the titration method to classify children into knower-levels, and always included small numbers up to "four". Studies also overwhelmingly focused on children's accuracy on giving a number and rarely analyzed counting behavior. These findings suggest the general acceptance of the core assumptions of the knower-level framework, specifically that children's number word knowledge is all-or-none, that small number words are learned in sequence, that all larger number words (at least those up to "ten") are learned in a similar fashion, and that children's counting knowledge changes when they become CP-knowers.

On the other hand, we found that studies vary on aspects of the task that may influence knower-level classifications. There is variation in whether researchers score overextension errors (i.e., scoring as incorrect for *n* when children gave *n* objects for other numbers), the type of follow-up questions asked upon set production, and the coding of children's performance on numbers larger than "four". We also found that these aspects of the task are frequently not reported in sufficient detail to replicate the Give-N methodology of a given study. These findings indicate that there may not be a "standard" Give-N task, which could make it challenging to compare knower-level classifications and associated findings between studies.

Moving forward with theory and methodology for research in children's learning of number words and counting requires (1) reflecting on the assumptions and limitations of classifying children's performance on the Give-N task in the knower-level framework and (2) consistent reporting of Give-N methodology.

Current Gaps and Implications for the Knower-Level Framework

Our review found that almost all studies relied on accuracy on giving a number in scoring children's performance on the task. The Grabbers vs. Counters distinction was not commonly used, and when used, was flexibly defined. Few studies considered children's overt counting behavior in classifying whether they know the cardinal principle, or whether they use counting to generate a set and to fix a set. This is important, because it is argued that children's knowledge of counting changes when they become CP-knowers (one of the core assumptions of the knower-level framework) yet it is often not directly assessed. Focusing on accuracy alone may provide an incomplete picture of children's knowledge of number words and counting (e.g., Barclay et al., 2017; Cheung et al., 2022). Errors on the Give-N Task could reflect performance or distraction errors, or a genuine lack of knowledge of numbers (Jara-Ettinger et al., 2016; Sarnecka & Lee, 2009), which can be difficult to disentangle for larger numbers. In addition, recent studies argue that there may be

large-number-subset-knowers (i.e., children who succeed on some large numbers such as "five" but fail on even larger ones) and small-number-subset-knowers (i.e., Wynn's *n*-knowers), but children may fail to give *n* correctly for different reasons (see Barclay et al., 2017; Fuson, 1988). Information on *how* children generate a set of objects on the Give-N task remains a gap that may help resolve the current debate on the nature of subset-knowers' knowledge of numbers (Krajcsi & Fintor, 2023).

Second, we found that certain aspects of the Give-N task were under-reported but were nevertheless important for knower-level classifications. Specifically, follow-up questions were reported less than half of the time. Not asking follow-up questions can underestimate children's accuracy on giving a number, because children who may otherwise be CP-knowers are not provided with an opportunity to fix an incorrect response (Krajcsi, 2021; Le Corre et al., 2006).

Third, we found that studies differ on what counts as evidence that a child *knows* the meaning of a certain number word. Most studies that adopted the knower-level framework used a "two out of three" rule, classifying children as knowing *n* if they give n correctly two out of three times, and not giving *n* + 1 two out of three times. But only half of the studies scored overextension errors and categorized children as *not* knowing *n* if they gave *n* for other numbers. Without this criterion, children with the same performance on the Give-N task could thus be coded as *n*-knowers in one study but *n* + 1 knowers in another study. In addition, it is often not clear how the scoring of overextension errors is applied. For example, do "other numbers" include all numbers in the study or only numbers larger than *n*? How does the number of total trials affect the application of this rule?

Fourth, we found that the boundary for what counts as a cardinal principle knower is not always clear. Some studies classified children as CP-knowers if they can give up to "five", and some require children to be correct on multiple numbers up to "six", or also numbers larger than "six". Children's success at giving larger numbers has been taken as evidence that learning the cardinal principle involves acquiring a conceptual insight into how counting maps number words to set sizes (Le Corre & Carey, 2007). But our data show that what counts as a "large" number is not always the same across studies, which has implications for the subset- vs. CP-knower distinction.

After 30 years of the Give-N task, a challenge for future research into children's early number learning will be to disentangle how the knower-level framework has been shaped by the methodology of the Give-N task and vice versa. Throughout the paper, we note how intertwined core assumptions of the knower-level framework are with methodological decisions of the Give-N task. We see several issues with this.

First, the core assumptions of the knower-level framework should stand the test from diverging methods. On this front, however, existing studies that compare the Give-N task with another task that assesses number word knowledge such as the Whats-on-this card task or the Point-to-X task show that one- and two-knowers, along with CP-knowers may be more reliable than other knower-levels (Le Corre et al., 2006; Le Corre et al., 2016; Marchand & Barner, 2021). More work is needed to validate the number word learning trajectory, beyond using the Give-N task itself, and beyond the use of primarily cross-sectional designs. Future studies can address the following questions: are the small numbers always learned one at a time and in sequence, or could a child learn more than one number word at a time (e.g., Palmer & Baroody, 2011)? Can we reliably differentiate subset-knowers from CP-knowers based on accuracy on the Give-N task alone? And is there a fixed boundary (e.g., "four") that differentiates between the two groups of children? Converging evidence from diverging methods will help advance research on early number learning.

Second, a focus on accuracy on the Give-N task as *the* measurement of number word and counting knowledge may limit our current conceptualization of cardinal principle knowledge (e.g., Baroody et al., 2023; Gelman, 1993; Gelman & Gallistel, 1986). In particular, the operationalization of cardinal principle knowledge taps into whether children can produce a set of 5, 6, 7, or 8 objects upon request. Accuracy on giving sets correctly may reflect children's understanding about when to stop counting (e.g., you ask for *N*, it means once I count to *N*, I can stop; see also Baroody & Lai, 2022; Baroody et al., 2023). But, do CP-knowers understand the meaning of number as an exact representation of number – two sets that are in one-to-one correspondence have the same cardinality (Frege, 1960; Izard et al., 2008; Sarnecka & Wright, 2013; Wiese, 2003)? And do they understand the relationship between the last word of the count and the counting procedures? Existing studies suggest that CP-knowers know two sets that have the same number can be placed in one-to-one correspondence, but they have limited knowledge about when two sets do *not* have the same number (Le et al., 2024; Soto-Alba & Le Corre, 2019). CP-knowers also seem to accept that the last word of a count denotes the numerosity even in the face of violations of counting procedures (e.g., they accepted a set has 16 fish if a puppet

double-counted each of the eight fish; Cheung et al., 2020). Analyzing children's accuracy on giving sets of 5 to 8 correctly may thus provide an incomplete picture of cardinal principle knowledge. Future studies should explore other aspects of learning the cardinal principle knowledge and how they vary across tasks.

Third, equally important as replicable empirical data is theory development (see van Rooij & Baggio, 2021, for a recent discussion on this issue). Explaining why and how children learn small numbers differently from large ones and what it means to have cardinal principle knowledge can guide empirical investigation of number learning (e.g., Barner, 2017; Baroody et al., 2023; Carey, 2009; Carey & Barner, 2019; Mix et al., 2002; Simon et al., 2023). The Give-N task may be a highly optimized methodology to pinpoint children's number word and counting knowledge, but validating the knower-level framework requires us to go beyond binning children into knower-levels based on the titrated Give-N task. Theorising and directly testing the cognitive processing that underlies the various patterns of performance we ascribe to subset-knowers and CP- knowers will allow us to make great strides in understanding the learning of number words and counting.

Methodological Considerations for Researchers Using the Give-N Task

Our review suggests that clarifying methodological practices and detailed reporting would be beneficial for research on children's early number learning. In the following section we discuss avenues to improve on methodological practices in the Give-N literature. We developed a short checklist to aid researchers in the comprehensive selecting and reporting of their Give-N methodology (see Figure 3 for the checklist). The checklist draws on the coding scheme we used in this review and our observations about aspects of the task that commonly were not reported. We hope this checklist may (1) facilitate future studies that aim to explore the impact of different methodological practices, (2) provide an overview of design considerations for the Give-N task, and (3) be used as a guideline for researchers to use as a starting point when reporting their study.

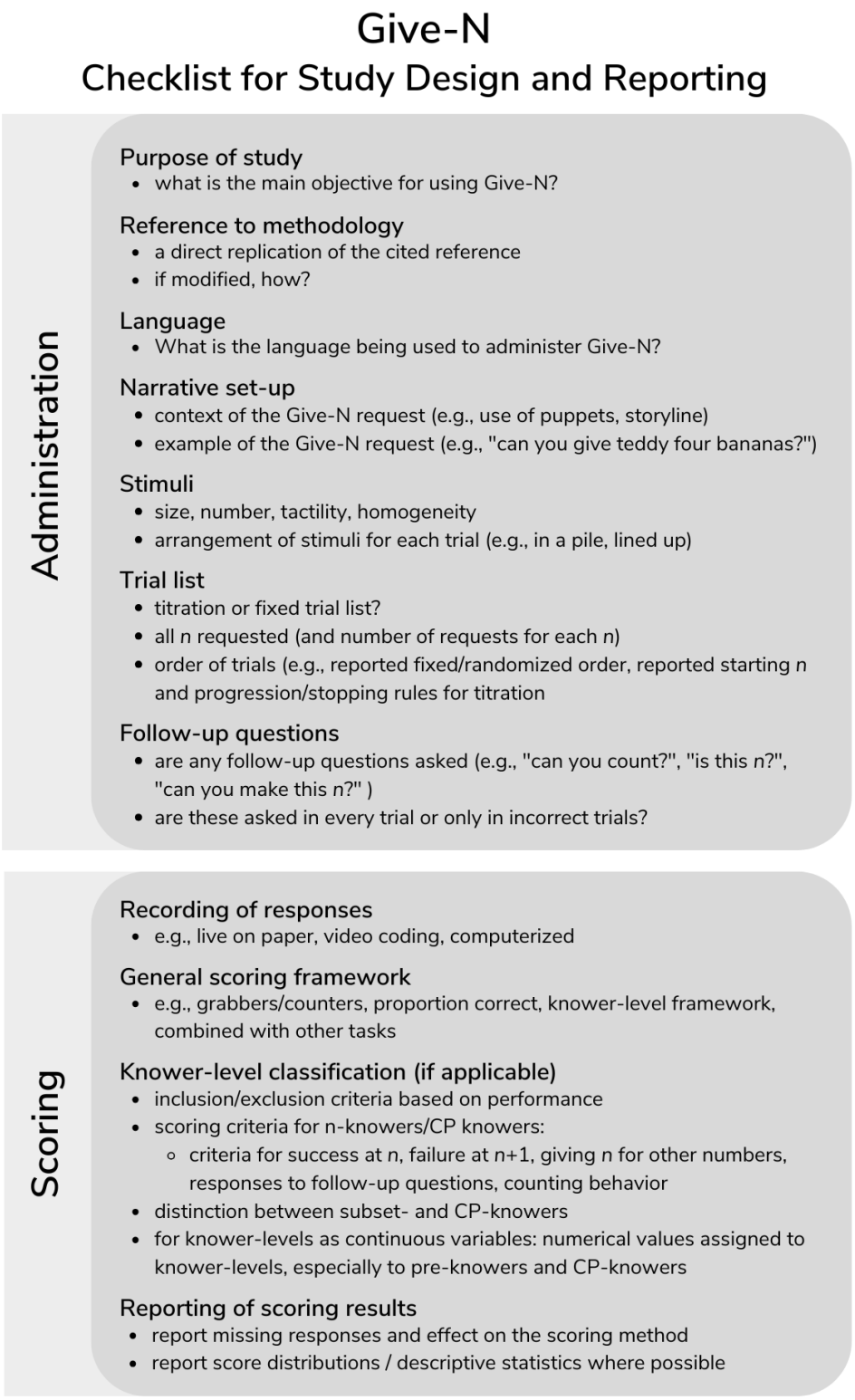
To aid researchers in designing a Give-N task for their study, we have the following recommendations. If the goal is to classify children into knower-levels, the titration method is more efficient (see also Marchand & Barner, 2021, for discussion). If, however, the goal is to examine the relationship between cardinal number knowledge and other aspects of numerical understanding such as estimation, comparison, or ordering, then the fixed trial list allows for a more balanced analysis as it provides the same amount of data on a set of numbers. We recommend using a simple narrative set-up and 3D stimuli that are easily graspable but do not roll around. Following studies by Krajcsi and Fintor (2023), if the goal of the study includes identifying CP-knowers, children should be allowed to fix their responses so as not to underestimate their performance. As suggested by a reviewer, the initial confirmation question ("Is that *N*?") to ensure that a child is done placing objects on the plate can be simplified to "Are you done?" as the original request might be perceived as a challenge to children's initial response. In short, methodological decisions reflect researchers' assumptions and goals of the research and should be accompanied by justification or explanation.

Although the current paper focuses on the Give-N task as a standard assessment of number word knowledge and counting, expanding on and adapting the Give-N methodology has proven a versatile research tool. The Give-N task has been expanded with clever variations to measure other aspects of numerical knowledge. For instance, Give-N using number words with different types of nouns or noun phrases ("one person, two animals, and three pieces of furniture"; Sorariutta et al., 2017), ordinal number ("give me the *third* toy"; Colomé & Noël, 2012; Meyer et al., 2018), the base-10 number system ("give me twenty-three" with tens and units; Guerrero et al., 2020), and quantifiers ("give me *some* toys"; Barner, Chow, & Yang, 2009). Researchers have also adapted the task to suit their research purpose (e.g., Baroody et al., 2023; Huang et al., 2010). We expect to see more variations and adaptations beyond its original purpose and design features, including in educational settings for pedagogical purposes (McDonald et al., 2021).

Finally, large scale research collaboration promises to be key in evaluating Give-N methodology and theories of early number learning. Different protocols for the Give-N task may be collected, peer-reviewed and disseminated via open methods databases (Krajcsi & Reynvoet, 2019, http://numericalcognitionmethods.org; UCSD Language & Development Lab, 2022, https://Give-N.ucsd.edu) and large representative datasets with consistent Give-N methodology can be generated in Many-Labs collaborations (such as the Many-Numbers research consortium, www.manynumbers.org/home).

Figure 3

A Checklist to Aid Researchers in the Comprehensive Selecting and Reporting of Their Give-N Task Methodology



Conclusion

Since Wynn's seminal work 30 years ago, the Give-N task has been instrumental for our understanding of how children learn number words and counting. Current theoretical challenges, such as investigating the mechanisms of number word learning or what it truly means to grasp the cardinal principle of counting, will require further interrogation into what the Give-N task does (and does not) measure. Our review contributes to these efforts by providing an overview of past research, highlighting how the core assumptions of the knower-level framework have become intertwined with Give-N task methodology, and discussing Give-N methodology in relation to gaps in our current understanding of early number learning. We have shared the data from this review and developed a checklist of suggestions for detailed reporting that we hope can lay the groundwork for future studies. More reflection on the use of the Give-N task and large-scale collaboration on testing theories related to the knower-level framework can lead to greater strides in understanding how children learn number words.

Funding: This research was partly supported by an award from Carleton University Faculty of Arts and Social Sciences' Covid-19 Research Assistantship Fund.

Acknowledgments: The authors have no additional (i.e., non-financial) support to report.

Competing Interests: The authors have declared that no competing interests exist.

Data Availability: The data that support the findings of this study are openly available in the Open Science Framework (OSF) (see Wege et al., 2022S).

Supplementary Materials

The Supplementary Materials contain the following items:

- The preregistration for the study (Wege et al., 2021S)
- A codebook with detailed explanations of all variables (Wege et al., 2022S)
- All extracted data, including data not reported in this review (Wege et al., 2022S)

Index of Supplementary Materials

Wege, T. E., Merkley, R., Bourque, T., & Cheung, P. (2021S). *Give-N review* [Preregistration]. OSF Registries. https://doi.org/10.17605/OSF.IO/TWSE8

Wege, T. E., Bourque, T., Merkley, R., & Cheung, P. (2022S). *Thirty years of knower-levels: A systematic review of the Give-N task. Data & codebook* [Research data and codebook]. OSF. https://osf.io/ymunr

References

Abreu-Mendoza, R. A., Soto-Alba, E. E., & Arias-Trejo, N. (2013). Area vs. density: Influence of visual variables and cardinality knowledge in early number comparison. *Frontiers in Psychology, 4*, Article 805. https://doi.org/10.3389/fpsyg.2013.00805

Almoammer, A., Sullivan, J., Donlan, C., Marušič, F., Žaucer, R., O'Donnell, T., & Barner, D. (2013). Grammatical morphology as a source of early number word meanings. *Proceedings of the National Academy of Sciences of the United States of America, 110*(46), 18448–18453. https://doi.org/10.1073/pnas.1313652110

Bakker, M., Torbeyns, J., Wijns, N., Verschaffel, L., & De Smedt, B. (2019). Gender equality in 4- to 5-year-old preschoolers' early numerical competencies. *Developmental Science, 22*(1), Article e12718. https://doi.org/10.1111/desc.12718

Barclay, M., Cheung, P., & Shusterman, A. (2017). *Limitations in children's induction of the cardinality principle: Evidence from the Give-a-Number task with higher quantities* [Poster presentation]. Budapest CEU Conference on Cognitive Development, 2017.

Barner, D. (2017). Language, procedures, and the non-perceptual origin of number word meanings. *Journal of Child Language, 44*(3), 553–590. https://doi.org/10.1017/S0305000917000058

Barner, D., & Bachrach, A. (2010). Inference and exact numerical representation in early language development. *Cognitive Psychology, 60*(1), 40–62. https://doi.org/10.1016/j.cogpsych.2009.06.002

Barner, D., Chow, K., & Yang, S. J. (2009). Finding one's meaning: A test of the relation between quantifiers and integers in language development. *Cognitive Psychology, 58*(2), 195–219. https://doi.org/10.1016/j.cogpsych.2008.07.001

Barner, D., Libenson, A., Cheung, P., & Takasaki, M. (2009). Cross-linguistic relations between quantifiers and numerals in language acquisition: Evidence from Japanese. *Journal of Experimental Child Psychology, 103*(4), 421–440. https://doi.org/10.1016/j.jecp.2008.12.001

Baroody, A. J., & Lai, M. (2022). The development and assessment of counting-based cardinal number concepts. *Educational Studies in Mathematics, 111*(2), 185–205. https://doi.org/10.1007/s10649-022-10153-5

Baroody, A. J., Mix, K. S., Kartal, G., & Lai, M.-l. (2023). The development and assessment of early cardinal-number concepts. *Journal of Numerical Cognition, 9*(1), 182–195. https://doi.org/10.5964/jnc.10035

Batchelor, S., Keeble, S., & Gilmore, C. (2015). Magnitude representations and counting skills in preschool children. *Mathematical Thinking and Learning, 17*(2-3), 116–135. https://doi.org/10.1080/10986065.2015.1016811

Bruce, R. A., & Threlfall, J. (2004). One, two, three and counting. *Educational Studies in Mathematics, 55*(1-3), 3–26. https://doi.org/10.1023/B:EDUC.0000017676.79430.dc

Brueggemann, A., & Gable, S. (2018). Preschoolers' selective sustained attention and numeracy skills and knowledge. *Journal of Experimental Child Psychology, 171*, 138–147. https://doi.org/10.1016/j.jecp.2018.02.001

Cahoon, A., Gilmore, C., & Simms, V. (2021). Developmental pathways of early numerical skills during the preschool to school transition. *Learning and Instruction, 75*, Article 101484. https://doi.org/10.1016/j.learninstruc.2021.101484

Cankaya, O., LeFevre, J.-A., & Dunbar, K. (2014). The role of number naming systems and numeracy experiences in children's rote counting: Evidence from Turkish and Canadian children. *Learning and Individual Differences, 32*, 238–245. https://doi.org/10.1016/j.lindif.2014.03.016

Carey, S. (2009). *The origin of concepts.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195367638.001.0001

Carey, S., & Barner, D. (2019). Ontogenetic origins of human integer representations. *Trends in Cognitive Sciences, 23*(10), 823–835. https://doi.org/10.1016/j.tics.2019.07.004

Ceylan, M., & Aslan, D. (2018). Cardinal number acquisition of Turkish children. *Journal of Education and e-learning Research, 5*(4), 217–224. https://doi.org/10.20448/journal.509.2018.54.217.224

Chang, A., & Sandhofer, C. M. (2019). How bilingual parents talk to children about number in Mandarin and English. *Frontiers in Psychology, 10*, Article 1090. https://doi.org/10.3389/fpsyg.2019.01090

Chernyak, N., Sandham, B., Harris, P. L., & Cordes, S. (2016). Numerical cognition explains age-related changes in third-party fairness. *Developmental Psychology, 52*(10), 1555–1562. https://doi.org/10.1037/dev0000196

Cheung, P., & Le Corre, M. (2018). Parallel individuation supports numerical comparisons in preschoolers. *Journal of Numerical Cognition, 4*(2), 380–409. https://doi.org/10.5964/jnc.v4i2.110

Cheung, P., Sohail, S., Barclay, M., Ellman, J., May, N., & Shusterman, A. (2020). *Do full counters understand cardinality?* [Conference talk]. 3rd Mathematical Cognition and Learning Conference, Dublin, Ireland.

Cheung, P., Toomey, M., Jiang, Y. H., Stoop, T. B., & Shusterman, A. (2022). Acquisition of the counting principles during the subset-knower stages: Insights from children's errors. *Developmental Science, 25*(4), Article e13219. https://doi.org/10.1111/desc.13219

Colomé, À., & Noël, M.-P. (2012). One first? Acquisition of the cardinal and ordinal uses of numbers in preschoolers. *Journal of Experimental Child Psychology, 113*(2), 233–247. https://doi.org/10.1016/j.jecp.2012.03.005

Condry, K. F., & Spelke, E. S. (2008). The development of language and abstract concepts: The case of natural number. *Journal of Experimental Psychology: General, 137*(1), 22–38. https://doi.org/10.1037/0096-3445.137.1.22

Covidence. (n.d.). [Systematic review software]. Veritas Health Innovation, Melbourne, Australia. Available at www.covidence.org

Davidson, K., Eng, K., & Barner, D. (2012). Does learning to count involve a semantic induction? *Cognition, 123*(1), 162–173. https://doi.org/10.1016/j.cognition.2011.12.013

Desoete, A., Praet, M., Titeca, D., & Ceulemans, A. (2013). Cognitive phenotype of mathematical learning disabilities: What can we learn from siblings? *Research in Developmental Disabilities, 34*(1), 404–412. https://doi.org/10.1016/j.ridd.2012.08.022

de Vries, H., Meyer, C., & Peeters-Podgaevskaja, A. (2021). Learning strategies in Russian ordinal acquisition. *First Language, 41*(1), 90–108. https://doi.org/10.1177/0142723720934261

Dowker, A. (2005). Early identification and intervention for students with mathematics difficulties. *Journal of Learning Disabilities, 38*(4), 324–332. https://doi.org/10.1177/00222194050380040801

Draper, C. E., Barnett, L. M., Cook, C. J., Cuartas, J. A., Howard, S. J., McCoy, D. C., Merkley, R., Molano, A., Maldonado-Carreño, C., Obradoviæ, J., Scerif, G., Valentini, N. C., Venetsanou, F., & Yousafzai, A. K. (2023). Publishing child development research from around the world: An unfair playing field resulting in most of the world's child population under-represented in research. *Infant and Child Development, 32*(6), Article e2375. https://doi.org/10.1002/icd.2375

Dresen, V., Moeller, K., & Pixner, S. (2022). Association between language and early numerical development – The case of quantifiers. *European Journal of Developmental Psychology, 19*(4), 477–493. https://doi.org/10.1080/17405629.2021.1916463

Fluck, M., & Henderson, L. (1996). Counting and cardinality in English nursery pupils. *British Journal of Educational Psychology, 66*(4), 501–517. https://doi.org/10.1111/j.2044-8279.1996.tb01215.x

Fluck, M., Linnell, M., & Holgate, M. (2005). Does counting count for 3-to 4-year-olds? Parental assumptions about preschool children's understanding of counting and cardinality. *Social Development, 14*(3), 496–513. https://doi.org/10.1111/j.1467-9507.2005.00313.x

Frege, G. (1960). *The foundations of arithmetic: A logico-mathematical enquiry into the concept of number* [J. L. Austin, Trans.] (2nd ed.). Northwestern University Press.

Frye, D., Braisby, N., Lowe, J., Maroudas, C., & Nicholls, J. (1989). Young children's understanding of counting and cardinality. *Child Development, 60*(5), 1158–1171. https://doi.org/10.2307/1130790

Fuson, K. C. (1988). *Children's counting and concepts of number*. Springer.

Geary, D. C., Nicholas, A., Li, Y., & Sun, J. (2017). Developmental change in the influence of domain-general abilities and domain-specific knowledge on mathematics achievement: An eight-year longitudinal study. *Journal of Educational Psychology, 109*(5), 680–693. https://doi.org/10.1037/edu0000159

Geary, D. C., vanMarle, K., Chu, F. W., Rouder, J., Hoard, M. K., & Nugent, L. (2018). Early conceptual understanding of cardinality predicts superior school-entry number-system knowledge. *Psychological Science, 29*(2), 191–205. https://doi.org/10.1177/0956797617729817

Gelman, R. (1993). A rational-constructivist account of early learning about numbers and objects. In D. L. Medin (Ed.), *Psychology of learning and motivation: Advances in research and theory* (Vol. 30, pp. 61-96). Academic Press. https://doi.org/10.1016/S0079-7421(08)60294-7

Gelman, R., & Gallistel, C. R. (1986). *The child's understanding of number*. Harvard University Press. https://doi.org/10.4159/9780674037533

Gibson, D. J., Gunderson, E. A., & Levine, S. C. (2020). Causal effects of parent number talk on preschoolers' number knowledge. *Child Development, 91*(6), e1162–e1177. https://doi.org/10.1111/cdev.13423

Gordon, R., Scalise, N. R., & Ramani, G. B. (2021). Give yourself a hand: The role of gesture and working memory in preschoolers' numerical knowledge. *Journal of Experimental Child Psychology, 208*, Article 105145. https://doi.org/10.1016/j.jecp.2021.105145

Guerrero, D., Hwang, J., Boutin, B., Roeper, T., & Park, J. (2020). Is thirty-two three tens and two ones? The embedded structure of cardinal numbers. *Cognition, 203*, Article 104331. https://doi.org/10.1016/j.cognition.2020.104331

Gunderson, E. A., Spaepen, E., Gibson, D., Goldin-Meadow, S., & Levine, S. C. (2015). Gesture as a window onto children's number knowledge. *Cognition, 144*, 14–28. https://doi.org/10.1016/j.cognition.2015.07.008

Huang, Y. T., Spelke, E., & Snedeker, J. (2010). When is four far more than three? Children's generalization of newly acquired number words. *Psychological Science, 21*(4), 600–606. https://doi.org/10.1177/0956797610363552

Izard, V., Pica, P., Spelke, E. S., & Dehaene, S. (2008). Exact equality and successor function: Two key concepts on the path towards understanding exact numbers. *Philosophical Psychology, 21*(4), 491–505. https://doi.org/10.1080/09515080802285354

Jamalian, A. (2014). *Grouping gestures promote children's effective counting strategies by adding a layer of meaning through action* [Doctoral thesis, Columbia University]. ProQuest Dissertations & Theses. https://www.proquest.com/dissertations-theses/grouping-gestures-promote-childrens-effective/docview/1548332297/se-2

Jara-Ettinger, J., Gibson, E., Kidd, C., & Piantadosi, S. (2016). Native Amazonian children forego egalitarianism in merit-based tasks when they learn to count. *Developmental Science, 19*(6), 1104–1110. https://doi.org/10.1111/desc.12351

Kersey, A. J., Braham, E. J., Csumitta, K. D., Libertus, M. E., & Cantlon, J. F. (2018). No intrinsic gender differences in children's earliest numerical abilities. *npj Science of Learning, 3*(1), Article 12. https://doi.org/10.1038/s41539-018-0028-7

Kimura, K., Wagner, K., & Barner, D. (2013). Two for one? Transfer of conceptual content in bilingual number word learning. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 35, pp. 2734-2739). https://escholarship.org/uc/item/6k33v8nz

Kirk, H., Gray, K., Ellis, K., Taffe, J., & Cornish, K. (2017). Impact of attention training on academic achievement, executive functioning, and behavior: A randomized controlled trial. *American Journal on Intellectual and Developmental Disabilities, 122*(2), 97–117. https://doi.org/10.1352/1944-7558-122.2.97

Knudsen, B., Fischer, M. H., & Aschersleben, G. (2015). Development of spatial preferences for counting and picture naming. *Psychological Research, 79*(6), 939–949. https://doi.org/10.1007/s00426-014-0623-z

Koike, W., & Yamagata, K. (2014). Development of numerical notations: Relation between number reading and number concepts. In *Proceedings of the 16th European conference on Developmental Psychology* (pp. 109-111). Medimond International Publishers.

Krajcsi, A. (2021). Follow-up questions influence the measured number knowledge in the Give-a-Number task. *Cognitive Development, 57*, Article 100968. https://doi.org/10.1016/j.cogdev.2020.100968

Krajcsi, A., & Fintor, E. (2023). A refined description of initial symbolic number acquisition. *Cognitive Development, 65*, Article 101288. https://doi.org/10.1016/j.cogdev.2022.101288

Krajcsi, A., & Reynvoet, B. (Eds.). (2019). *Numerical Cognition Methods Directory*. https://numericalcognitionmethods.org

Le, K. N., Kwon, C., Wu, M., & Barner, D. (2024). 'Five' is the number of bunnies and hats: Children's understanding of cardinal extension and exact number. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 46, pp. 1630-1637). https://escholarship.org/uc/item/8293g6hd

Lê, M.-L. T., & Noël, M.-P. (2020). Transparent number-naming system gives only limited advantage for preschooler's numerical development: Comparisons of Vietnamese and French-speaking children. *PLoS One, 15*(12), Article e0243472. https://doi.org/10.1371/journal.pone.0243472

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition, 105*(2), 395–438. https://doi.org/10.1016/j.cognition.2006.10.005

Le Corre, M., Li, P., Huang, B. H., Jia, G., & Carey, S. (2016). Numerical morphology supports early number word learning: Evidence from a comparison of young Mandarin and English learners. *Cognitive Psychology, 88*, 162–186. https://doi.org/10.1016/j.cogpsych.2016.06.003

Le Corre, M., Van de Walle, G., Brannon, E. M., & Carey, S. (2006). Re-visiting the competence/performance debate in the acquisition of the counting principles. *Cognitive Psychology, 52*(2), 130–169. https://doi.org/10.1016/j.cogpsych.2005.07.002

Lee, M. D., & Sarnecka, B. W. (2010). A model of knower-level behavior in number concept development. *Cognitive Science, 34*(1), 51–67. https://doi.org/10.1111/j.1551-6709.2009.01063.x

Lee, M. D., & Sarnecka, B. W. (2011). Number-knower levels in young children: Insights from Bayesian modeling. *Cognition, 120*(3), 391–402. https://doi.org/10.1016/j.cognition.2010.10.003

Liang, Y., Zhang, L., Wang, C., & Liu, Y. (2021). Performance patterns and strategy use in number line estimation among preschool children with different spontaneous focusing on numerosity tendencies. *Infant and Child Development, 30*(1), Article e2203. https://doi.org/10.1002/icd.2203

Lindskog, M., & Simms, V. (2021). *Measuring toddlers non-symbolic number processing*. PsyArXiv. https://doi.org/10.31234/osf.io/v835g

Marchand, E., & Barner, D. (2021). *The acquisition of French un.* PsyArXiv. https://doi.org/10.31234/osf.io/6hcxy

Marchand, E., Lovelett, J. T., Kendro, K., & Barner, D. (2022). Assessing the knower-level framework: How reliable is the Give-a-Number task? *Cognition, 222*, Article 104998. https://doi.org/10.1016/j.cognition.2021.104998

Marušič, F., Žaucer, R., Plesnièar, V., Razboršek, T., Sullivan, J., & Barner, D. (2016). Does grammatical structure accelerate number word learning? Evidence from learners of dual and non-dual dialects of Slovenian. *PLoS One, 11*(8), Article e0159208. https://doi.org/10.1371/journal.pone.0159208

Mateu, V. E., & Hyams, N. (2016). One is the loneliest number: The acquisition of Spanish indefinite 'un'. In *Proceedings of the 6th conference on generative approaches to language acquisition North America (GALANA 2015)* (pp. 73-80). https://www.lingref.com/cpp/galana/6/paper3207.pdf

McDonald, J. A., Merkley, R., Mickle, J., Collimore, L. M., Hawes, Z., & Ansari, D. (2021). Exploring the implementation of early math assessments in kindergarten classrooms: A research-practice collaboration. *Mind, Brain and Education, 15*(4), 311–321. https://doi.org/10.1111/mbe.12293

Merritt, D. J., & Brannon, E. M. (2013). Nothing to it: Precursors to a zero concept in preschoolers. *Behavioural Processes, 93*, 91–97. https://doi.org/10.1016/j.beproc.2012.11.001

Meyer, C., Barbiers, S., & Weerman, F. (2018). Ordinals are not as easy as one, two, three: The acquisition of cardinals and ordinals in Dutch. *Language Acquisition, 25*(4), 392–417. https://doi.org/10.1080/10489223.2017.1391266

Meyer, C., Barbiers, L. C. J., & Weerman, F. (2020). Many systems, one strategy: Acquiring ordinals in Dutch and English. *Glossa: A Journal of General Linguistics, 5*(1), Article 100. https://doi.org/10.5334/gjgl.595

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). Multiple cues for quantification in infancy: Is number one of them? *Psychological Bulletin, 128*(2), 278–294. https://doi.org/10.1037/0033-2909.128.2.278

Moore, A. M., vanMarle, K., & Geary, D. C. (2016). Kindergartners' fluent processing of symbolic numerical magnitude is predicted by their cardinal knowledge and implicit understanding of arithmetic 2 years earlier. *Journal of Experimental Child Psychology, 150*, 31–47. https://doi.org/10.1016/j.jecp.2016.05.003

Mou, Y., Zhang, B., Piazza, M., & Hyde, D. C. (2021). Comparing set-to-number and number-to-set measures of cardinal number knowledge in preschool children using latent variable modeling. *Early Childhood Research Quarterly, 54*, 125–135. https://doi.org/10.1016/j.ecresq.2020.05.016

Munn, P. (1994). The early development of literacy and numeracy skills. *European Early Childhood Education Research Journal, 2*(1), 5–18. https://doi.org/10.1080/13502939485207491

Mussolin, C., Nys, J., Content, A., & Leybaert, J. (2014). Symbolic number abilities predict later approximate number system acuity in preschool children. *PLoS One, 9*(3), Article e91839. https://doi.org/10.1371/journal.pone.0091839

Nanu, C. E., McMullen, J., Munck, P., Hannula-Sormunen, M. M., & the Pipari Study Group. (2018). Spontaneous focusing on numerosity in preschool as a predictor of mathematical skills and knowledge in the fifth grade. *Journal of Experimental Child Psychology, 169*, 42–58. https://doi.org/10.1016/j.jecp.2017.12.011

Negen, J., & Sarnecka, B. W. (2015). Is there really a link between exact-number knowledge and approximate number system acuity in young children? *British Journal of Developmental Psychology, 33*(1), 92–105. https://doi.org/10.1111/bjdp.12071

Negen, J., Sarnecka, B. W., & Lee, M. D. (2012). An Excel sheet for inferring children's number-knower levels from Give-N data. *Behavior Research Methods, 44*(1), 57–66. https://doi.org/10.3758/s13428-011-0134-4

Nicoladis, E., Pika, S., & Marentette, P. (2010). Are number gestures easier than number words for preschoolers? *Cognitive Development, 25*(3), 247–261. https://doi.org/10.1016/j.cogdev.2010.04.001

Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology, 162*, 31–38. https://doi.org/10.1016/j.jecp.2017.04.017

O'Rear, C. D., & McNeil, N. M. (2019). Improved set-size labeling mediates the effect of a counting intervention on children's understanding of cardinality. *Developmental Science, 22*(6), Article e12819. https://doi.org/10.1111/desc.12819

O'Rear, C. D., Kirkland, P. K., & Purpura, D. J. (2024). The how many and Give-N tasks: Conceptually distinct measures of the cardinality principle. *Early Childhood Research Quarterly, 66*, 61–74. https://doi.org/10.1016/j.ecresq.2023.08.010

O'Rear, C. D., McNeil, N. M., & Kirkland, P. K. (2020). Partial knowledge in the development of number word understanding. *Developmental Science, 23*(5), Article e12944. https://doi.org/10.1111/desc.12944

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., & Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Systematic Reviews, 10*, Article 89. https://doi.org/10.1186/s13643-021-01626-4

Palmer, A., & Baroody, A. J. (2011). Blake's development of the number words "one," "two," and "three". *Cognition and Instruction, 29*(3), 265–296. https://doi.org/10.1080/07370008.2011.583370

Petersen, L. A., & McNeil, N. M. (2013). Effects of perceptually rich manipulatives on preschoolers' counting performance: Established knowledge counts. *Child Development, 84*(3), 1020–1033. https://doi.org/10.1111/cdev.12028

Piantadosi, S. T., Jara-Ettinger, J., & Gibson, E. (2014). Children's learning of number words in an indigenous farming-foraging group. *Developmental Science, 17*(4), 553–563. https://doi.org/10.1111/desc.12078

Pinhas, M., Donohue, S. E., Woldorff, M. G., & Brannon, E. M. (2014). Electrophysiological evidence for the involvement of the approximate number system in preschoolers' processing of spoken number words. *Journal of Cognitive Neuroscience, 26*(9), 1891–1904. https://doi.org/10.1162/jocn_a_00631

Posid, T., & Cordes, S. (2018). How high can you count? Probing the limits of children's counting. *Developmental Psychology, 54*(5), 875–889. https://doi.org/10.1037/dev0000469

Purpura, D. J., & Simms, V. (2018). Approximate number system development in preschool: What factors predict change? *Cognitive Development, 45*, 31–39. https://doi.org/10.1016/j.cogdev.2017.11.001

Rousselle, L., Palmers, E., & Noël, M. P. (2004). Magnitude comparison in preschoolers: What counts? Influence of perceptual variables. *Journal of Experimental Child Psychology, 87*(1), 57–84. https://doi.org/10.1016/j.jecp.2003.10.005

Rousselle, L., & Vossius, L. (2021). Acquiring the cardinal knowledge of number words: A conceptual replication. *Journal of Numerical Cognition, 7*(3), 411–434. https://doi.org/10.5964/jnc.7029

Salsa, A. M., & Martí, E. (2015). Objects, pictures and words: Effects of representational format on four-year-olds' quantity knowledge/ Objetos, imágenes y palabras. Efectos del formato representacional en el conocimiento de la cantidad a los cuatro años. *Estudios de Psicología, 36*(1), 71–91. https://doi.org/10.1080/02109395.2014.1000031

Sarnecka, B. W. (2015). Learning to represent exact numbers. *Synthese, 198*(S5, Suppl.), S1001–S1018. https://doi.org/10.1007/s11229-015-0854-6

Sarnecka, B. W., & Carey, S. (2008). How counting represents number: What children must learn and when they learn it. *Cognition, 108*(3), 662–674. https://doi.org/10.1016/j.cognition.2008.05.007

Sarnecka, B. W., Goldman, M. C., & Slusser, E. B. (2015) How counting leads to children's first representations of exact, large numbers. In R. Cohen Kadosh & A. Dowker (Eds.), *Oxford handbook of numerical cognition* (pp. 291-309). Oxford University Press.

Sarnecka, B. W., Kamenskaya, V. G., Yamana, Y., Ogura, T., & Yudovina, Y. B. (2007). From grammatical number to exact numbers: Early meanings of 'one', 'two', and 'three' in English, Russian, and Japanese. *Cognitive Psychology, 55*(2), 136–168. https://doi.org/10.1016/j.cogpsych.2006.09.001

Sarnecka, B. W., & Lee, M. D. (2009). Levels of number knowledge during early childhood. *Journal of Experimental Child Psychology, 103*(3), 325–337. https://doi.org/10.1016/j.jecp.2009.02.007

Sarnecka, B. W., & Wright, C. E. (2013). The idea of an exact number: Children's understanding of cardinality and equinumerosity. *Cognitive Science, 37*(8), 1493–1506. https://doi.org/10.1111/cogs.12043

Savelkouls, S., Hurst, M. A., & Cordes, S. (2020). Preschoolers' number knowledge relates to spontaneous focusing on number for small, but not large, sets. *Developmental Psychology, 56*(10), 1879–1893. https://doi.org/10.1037/dev0001099

Scalise, N. R., & Ramani, G. B. (2021). Symbolic magnitude understanding predicts preschoolers' later addition skills. *Journal of Cognition and Development, 22*(2), 185–202. https://doi.org/10.1080/15248372.2021.1888732

Schaeffer, B., Eggleston, V. H., & Scott, J. L. (1974). Number development in young children. *Cognitive Psychology, 6*(3), 357–379. https://doi.org/10.1016/0010-0285(74)90017-6

Schneider, R. M., Brockbank, E., Feiman, R., & Barner, D. (2022). Counting and the ontogenetic origins of exact equality. *Cognition, 218*, Article 104952. https://doi.org/10.1016/j.cognition.2021.104952

Schneider, R. M., Sullivan, J., Guo, K., & Barner, D. (2021). What counts? Sources of knowledge in children's acquisition of the successor function. *Child Development, 92*(4), e476–e492. https://doi.org/10.1111/cdev.13524

Sella, F., Berteletti, I., Lucangeli, D., & Zorzi, M. (2017). Preschool children use space, rather than counting, to infer the numerical magnitude of digits: Evidence for a spatial mapping principle. *Cognition, 158*, 56–67. https://doi.org/10.1016/j.cognition.2016.10.010

Sella, F., & Lucangeli, D. (2020). The knowledge of the preceding number reveals a mature understanding of the number sequence. *Cognition, 194*, Article 104104. https://doi.org/10.1016/j.cognition.2019.104104

Sella, F., Slusser, E., Odic, D., & Krajcsi, A. (2021). The emergence of children's natural number concepts: Current theoretical challenges. *Child Development Perspectives, 15*(4), 265–273. https://doi.org/10.1111/cdep.12428

Shusterman, A., Slusser, E., Halberda, J., & Odic, D. (2016). Acquisition of the cardinal principle coincides with improvement in approximate number system acuity in preschoolers. *PLoS One, 11*(4), Article e0153072. https://doi.org/10.1371/journal.pone.0153072

Simon, M. A., Della Volpe, D., & Velamur, A. (2023). Construction of the cardinality principle through counting: Critique and conjecture. *Mathematical Thinking and Learning, 25*(4), 465–480. https://doi.org/10.1080/10986065.2021.2012737

Singh, L., Cristia, A., Karasik, L. B., & Oakes, L. (2021). *Why are we still so WEIRD? Barriers and bridges towards a diversified science of early development.* PsyArXiv. https://doi.org/10.31234/osf.io/hgukc

Slusser, E., Ditta, A., & Sarnecka, B. (2013). Connecting numbers to discrete quantification: A step in the child's construction of integer concepts. *Cognition, 129*(1), 31–41. https://doi.org/10.1016/j.cognition.2013.05.011

Slusser, E., Ribner, A., & Shusterman, A. (2019). Language counts: Early language mediates the relationship between parent education and children's math ability. *Developmental Science, 22*(3), Article e12773. https://doi.org/10.1111/desc.12773

Sorariutta, A., Hannula-Sormunen, M. M., & Silvén, M. (2017). Maternal sensitivity in responding during play and children's pre-mathematical skills: A longitudinal study from infancy to preschool age. *European Journal of Developmental Psychology, 14*(1), 1–15. https://doi.org/10.1080/17405629.2016.1140641

Soto-Alba, E. E., & Le Corre, M. (2019). *Do CP-knowers know that all numerically equal collections are in one-to-one correspondence?* [Conference talk]. 78th Biennial Meeting of the Society for Research in Child Development. Baltimore, MD, USA.

Spaepen, E., Gunderson, E. A., Gibson, D., Goldin-Meadow, S., & Levine, S. C. (2018). Meaning before order: Cardinal principle knowledge predicts improvement in understanding the successor principle and exact ordering. *Cognition, 180*, 59–81. https://doi.org/10.1016/j.cognition.2018.06.012

Steele, A., Karmiloff-Smith, A., Cornish, K., & Scerif, G. (2012). The multiple subfunctions of attention: Differential developmental gateways to literacy and numeracy. *Child Development, 83*(6), 2028–2041. https://doi.org/10.1111/j.1467-8624.2012.01809.x

Stojanov, A. (2011). Children's understanding of quantitative and non-quantitative transformations of a set. *Review of Psychology, 18*(2), 55–62. https://hrcak.srce.hr/81458

Szkudlarek, E., & Brannon, E. M. (2018). Approximate arithmetic training improves informal math performance in low achieving preschoolers. *Frontiers in Psychology, 9*, Article 606. https://doi.org/10.3389/fpsyg.2018.00606

UCSD Language & Development Lab. (2022). *Give-N app*. https://Give-N.ucsd.edu

vanMarle, K., Chu, F. W., Mou, Y., Seok, J. H., Rouder, J., & Geary, D. C. (2018). Attaching meaning to the number words: Contributions of the object tracking and approximate number systems. *Developmental Science, 21*(1), Article e12495. https://doi.org/10.1111/desc.12495

van Rooij, I., & Baggio, G. (2021). Theory before the test: How to build high-verisimilitude explanatory theories in psychological science. *Perspectives on Psychological Science, 16*(4), 682–697. https://doi.org/10.1177/1745691620970604

Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., Newcombe, N. S., Filipowicz, A. T., & Chang, A. (2014). Deconstructing building blocks: Preschoolers' spatial assembly performance relates to early mathematical skills. *Child Development, 85*(3), 1062–1076. https://doi.org/10.1111/cdev.12165

Villarroel, J. D., Miñón, M., & Nuño, T. (2011). The origin of counting: A study of the early meaning of 'one', 'two' and 'three' among Basque-and Spanish-speaking children. *Educational Studies in Mathematics, 76*(3), 345–361. https://doi.org/10.1007/s10649-010-9291-0

Wagner, K., Chu, J., & Barner, D. (2019). Do children's number words begin noisy? *Developmental Science, 22*(1), Article e12752. https://doi.org/10.1111/desc.12752

Wagner, K., Kimura, K., Cheung, P., & Barner, D. (2015). Why is number word learning hard? Evidence from bilingual learners. *Cognitive Psychology, 83*, 1–21. https://doi.org/10.1016/j.cogpsych.2015.08.006

Wiese, H. (2003). *Numbers, language, and the human mind*. Cambridge University Press.

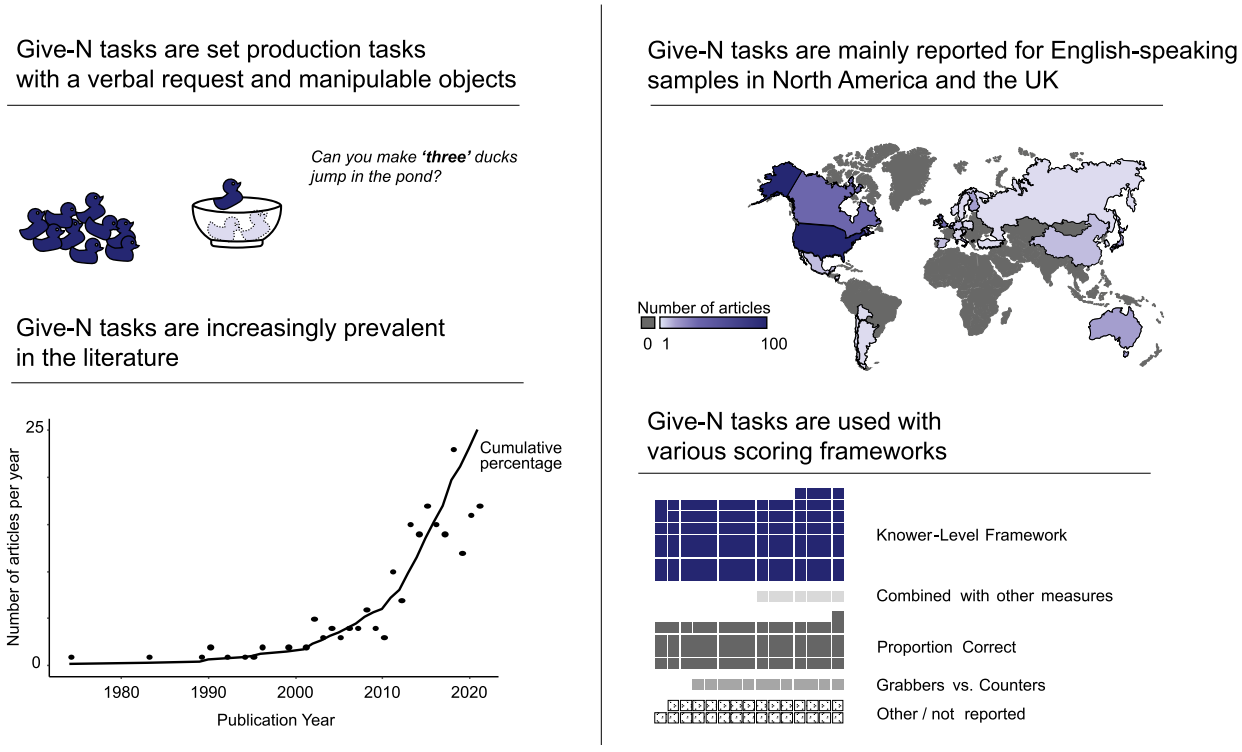
Wynn, K. (1990). Children's understanding of counting. *Cognition, 36*(2), 155–193. https://doi.org/10.1016/0010-0277(90)90003-3

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology, 24*(2), 220–251. https://doi.org/10.1016/0010-0285(92)90008-P

Appendix

Figure A1

Visual Summary and Overview of the Give-N Literature Based on the 210 Articles Included in This Review, 109 of Which Used the Give-N Task With the Knower-Level Framework



Note. A few articles reported multiple scoring frameworks, such that not every datapoint in the figure represents a unique article.

Journal of Numerical Cognition (JNC) is an official journal of the Mathematical Cognition and Learning Society (MCLS).

leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology (ZPID), Germany.