2019 ■ Vol. 15 ■ no. 3

Computerized Item Modeling Practices using Computer Adaptive Formative Assessment Automatic Item Generation System: A Tutorial

Jaehwa Choi [a,✉] and Xinxin Zhang [b]

[a]Graduate School of Education and Human Development, the George Washington University
[b]Centre For Research in Applied Measurement and Evaluation University of Alberta

Abstract ■ This tutorial provides a user-friendly technical document to assist those interested in developing item models using Computer Adaptive Formative Assessment (CAFA) Automatic Item Generation (AIG), which is a specialized web-based system for AIG. The first section of the tutorial introduces AIG theory and background, clarifying its definition, listing its potential advantages, and explaining the underlying philosophy/theory. CAFA AIG system, a web-based system specialized in implementing AIG, is introduced in the second section. The tutorial concludes by providing three computerized item modeling examples to enable readers to understand how AIG can be applied in their own research.

Keywords ■ Automatic Item Generation; Assessment Engineering; Evidence-centered Design; Digital Assessments; Computer Adaptive Formative Assessments.

Acting Editor ■ Denis Cousineau (Université d'Ottawa)

Reviewers
■ One anonymous reviewer

✉ jaechoi@gwu.edu

JC: 0000-0003-4225-9582; *XZ*: 0000-0003-4926-7980

10.20982/tqmp.15.3.p214

Introduction

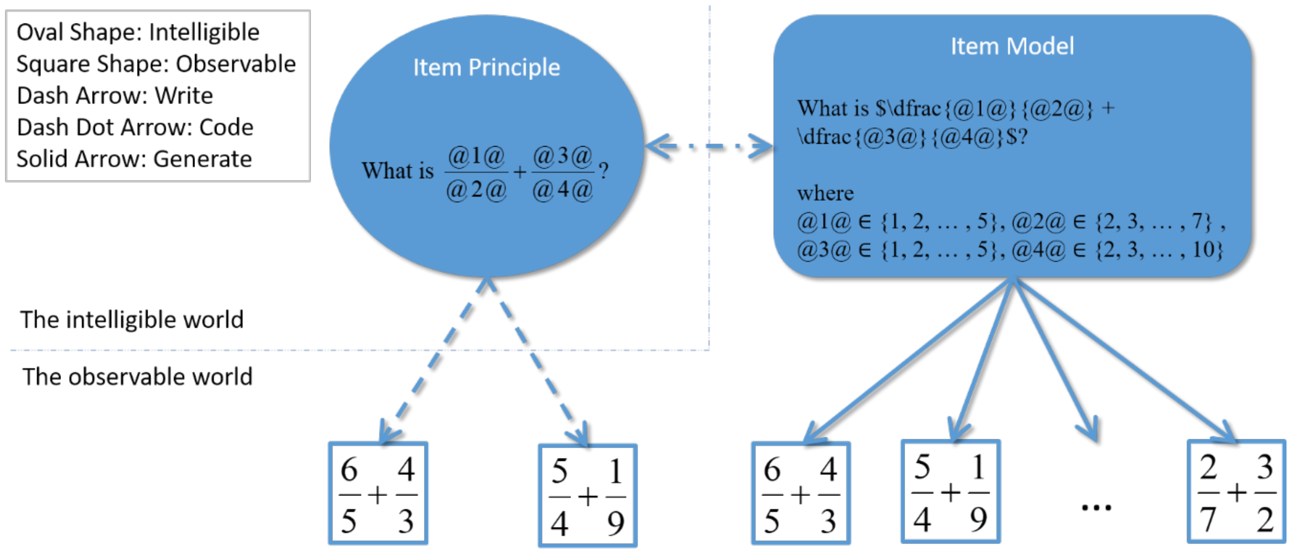
Automatic Item Generation (AIG) is an innovative development and management strategy that integrates cognitive and psychometric theories for futuristic assessment services in digital framework (Bejar, 1993; Choi, Yoon, & Kim, 2017; Embretson & Yang, 2007; Gierl & Haladyna, 2013; Irvine & Kyllonen, 2002). Within an AIG framework, a computerized item model composed of specialized computer codes/modules is utilized for machine to systematically generate massive, high-quality digital items/assessments in digital or Information and Communication Technology (ICT) environments. AIG provides several innovative digital solutions to overcome the shortcomings of the traditional assessment development/management process in which human item writers manually create, review, analyze and modify individual items, which is a slow, expensive and often subjective procedure (Choi, 2017).

Conceptual Perspectives for AIG

To help readers gain a deeper understanding of the theoretical and/or conceptual foundations of AIG, Choi (2017) illustrates the differences between an AIG-based approach and the traditional approach for developing items from a conceptual perspective. The traditional approach —*human writing*— requires an item writer to manually create individual items based on an item principle/design held in their mind's eye. Specifically, it requires the item writer to represent, clarify, and verify the principle and then to transform this principle into item instances. Suppose that an item writer is creating multiple math items based on the principle of adding fractions: he or she needs to repeatedly represent, clarify, and verify this principle and then transform it into item instances, which requires time and a great deal of effort to achieve.

The AIG-based approach —*machine generating*— requires an item writer to code an item model using a computer and then use a computer system to generate item in-

Figure 1 ■ A conceptual framework of the item principle, item model, and item instances using a fraction addition task as an example



stances by automatically varying the various elements involved. Suppose that an item writer is creating multiple math items based on the principle of adding fractions using machine generation, the example shown in Figure 1. The first step is to code an item model "What is @1@/@2@ + @3@/@4@?", where @1@, @2@, @3@ and @4@ are four elements (or item parameters) to be varied. The item writer then sets values for each element (in this case, @1@ has 5 values from 1 to 5; @2@ has 6 values from 2 to 7; @3@ has 5 values from 1 to 6; and @4@ has 9 values from 2 to 10). Once the model is coded, items can be directly generated. For the item model in this example, a total of 1,350 ($5 \times 6 \times 5 \times 9$) items can be generated if there are no conditions or constraints imposed on the elements.

However, in practice, one or more constraints may be required to generate specific item instances. For example, to generate only simplified and different denominator fraction items, an item developer could specify necessary conditions such as @2@ being the mutual prime of @1@, and @4@ being the mutual prime of @3@ while being unequal to @2@. By doing so, this item model would generate a total of 558 items with simplified and different denominators.

Beyond massive generation, AIG researchers have also emphasized the validity implications of AIG. For example, an item model is a higher order representation of actual items (Bejar, 2002), and is a manifested form or idea based on item principle or design. These item models can therefore help us understand, specify, infer, and validate what we are measuring. This perspective of AIG, which focuses on such validity, plays a very important role in guiding us to understand the theoretical foundations of AIG (Choi, 2017). The next section of this paper will cover potential advantages/utilities of AIG over different perspectives.

Advantages and Utilities of AIG over Three Different Perspectives

The AIG approach has a number of implications for assessment theory and practice. In particular, AIG is rapidly expanding the scope and impact of psychometrics related practices and policies. It is thus necessary to examine and understand the potential implications and advantages of utilizing AIG from multiple points of view. In order to understand, evaluate, and successfully implement AIG appropriately, Choi (2017) proposes adopting three different perspectives for AIG, namely theoretical, practical, and technological. In this section of this paper we will particularly focus on the practical and technological perspectives.

A major use of AIG is as a solution for producing massive high-quality digital assessments by systematically generating large numbers of items. For example, Lai, Alves, and Gierl (2009) generated 64,260 item instances from 34 parent items – the pre-made items on which item models are based– for assessing student performance over a wide range of different subjects (in their example, math, literature, science, and social studies assessments for grades 3, 6 and 9). Choi, Kim, and Yoon (2016, 2017) showed how more than more than two million math items can be digi-

Table 1 ■ Psychometric Features of CAFA AIG System

Other AIG Tools	CAFA AIG
Quantitative domain	Supports various content domains such as cognitive skills beyond quantitative knowledge (e.g., Math or Statistics)
Generating items	Capable of assembling and generating assessments (e.g., parallel forms or testlets) beyond item-level generations
Stem parameterization	In multiple choice (MC) format, choices can be parametrized to support distractor modeling
Static or no-feedback	Feedback can also be parametrized (so-called, automatic and adaptive feedback generator) for maximizing utilities such as diagnostic/formative assessments and learning analytics
Fixed item format	Supports various item formats (constructed responses, MC3, MC4, MC5, matching, true/-false)

tally generated from 350 item models to assess students in grades 6 and 7.

AIG is also known for offering a variety of potential practical advantages beyond the mass production of assessment items, ranging from examining theoretical/psychometric excellence to practical aspects (Choi, 2017; Choi & Lee, 2016). In particular, there is evidence to suggest that AIG can offer a variety of solutions to help improve assessment practices and address policy related concerns (e.g., exam leakages, teaching to test, financial burden associate with assessments, etc.; Choi, Kang, Kim, Dardick, & Zhang, 2015; Choi, Kim, & Pak, 2018; Dardick & Choi, 2016).

Choi (2017) specifically highlights AIG's role as a useful tool for digitization (or digital transformation) of assessments, one of the top priorities of nearly every government, organization, and enterprise in recent years. An AIG approach can integrate various types of interactive and intelligent components in a way that has not previously been possible utilizing traditional items with static elements because computerized item models can be seamlessly incorporated with digital standards such as HTML5, the standard for structuring and presenting content on the World Wide Web, to sustainably handle multiple types of dynamic and interactive multimedia such as graphics, audio and video clips as well as external digital technologies such as text-to-speech or speech-to-text modules. AIG has recently emerged as an integral part of digital assessments through its capacity to support and integrate a variety of other psychometric techniques and components, including automatic scoring, adaptive algorithms, and automated test assembly (Choi, 2017).

Notwithstanding the many potential advantages of AIG listed above, AIG has not been massively adapted in the field for three reasons. First, practitioners are still not fully aware of how this new technology can be used in assessment practices; second, as yet there are very few user-friendly and accessible AIG tools that practitioners can utilize; and third, not enough resources and materials have been developed to assist practitioners seeking to learn how to implement AIG. In the next section of this paper, we will provide several practical examples to illustrate how AIG technology can be utilized in real-world situations to enable practitioners to fully appreciate the benefits of using a specialized AIG tool, in this case, CAFA AIG system.

CAFA AIG System

Features

Computer Adaptive Formative Assessment (CAFA) AIG system (**cyk17**) is an innovative assessment tool developed by CAFA Lab, Inc., which is an incorporated research institution founded in 2012 to conduct ICT-based assessment, testing, and measurement research and development projects. CAFA AIG integrates both cognitive and psychometric theories into assessment development practices to generate high-quality digital assessment instruments using state-of-the-art digital technologies. CAFA Lab currently provides two different types of CAFA AIG system accounts, Non-commercial Account and Institutional Account. Visit the system website for more information about the accounts. As a futuristic AIG tool in ICT environments, CAFA AIG incorporates a number of psychometric and technological features, as summarized in Tables 1 and 2 below.

One thing to emphasize here is that the system supports CAFA AIG WordPress plug-in. WordPress, which is an online, open source website creation tool written in PHP, is one of the easiest and most popular blogging and website content management systems (CMS) on the Internet. CAFA AIG WordPress plugin enables site admins to seamlessly add AIG features to their assessment service websites without knowing a single line of code. This CAFA AIG feature represents a major breakthrough for efforts to spread and disseminate AIG technology and related services to assessment organizations, services, and researchers at every

Table 2 ■ ICT Features of CAFA AIG System

Other AIG Tools	CAFA AIG
Paper-pencil delivery oriented	Web/Mobile publication oriented, incorporating Web Standards such as Hypertext Markup Language (HTML) 5 and ePub 3
Numbers or word parameters	Only Supports various technological/multimedia parameters, including dynamic/interactive figures/diagrams Non-compatible databases Compatible with Instructional Management Systems (IMS) Question & Test Interoperability (QTI; QTILite Specification Version 1.2)
Closed-application	Capable of integrating external smart/intelligent digital solutions (e.g., Text-to-Speech or Speech-to-Text Artificial Intelligences)
Local application	Client-server Assessment Platform via CAFA AIG WordPress Plugin

level. Note that the features of CAFA AIG introduced in this paper are focused on the current version, i.e., developing item models in order to generate innovative digital item instances. CAFA AIG will evolve into a full-scale item modeling tool including: 1) assembling assessments with generated items, 2) deploying the assessments to collect response data in both on- and off-line modes, and 3) conducting advanced psychometric analyses with the collected data.

Examples of Computerized Item Models and Item Instances

Five types of item instances generated from item models developed using CAFA AIG system are presented in Figures 2, 3, and 4. These examples demonstrate how CAFA AIG incorporates various digital innovations (such as, 3D diagrams, interactive multimedia, text-to-speech Artificial Intelligence (AI), speech-to-text AI, or non-verbal cognitive ability measure figures) into computerized item modeling practices.

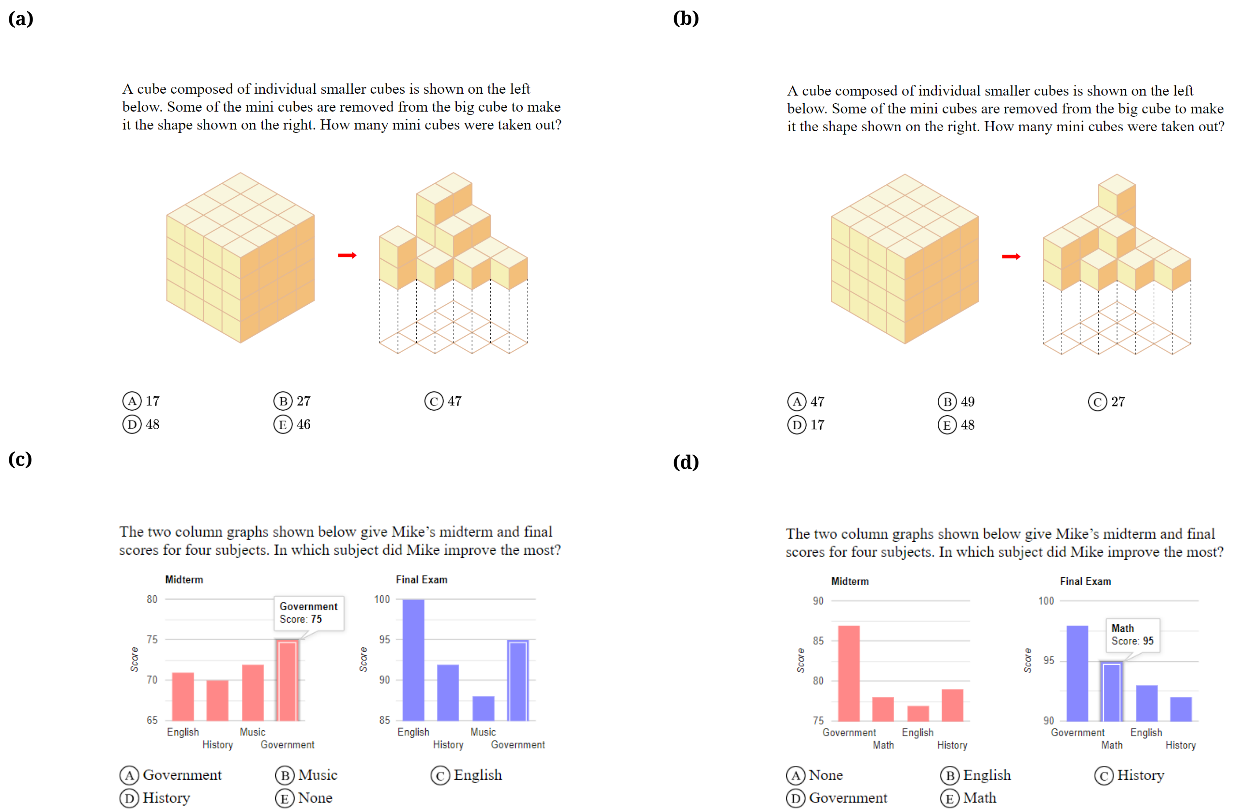
The 3D diagram item shown in Panel 1, Figure 2 is designed to improve students' spatial imagination, particularly their ability to mentally recreate (or alter) 3D objects. This provides a useful aid to encourage students to approach a problem from more than one angle. 3D diagrams such as the one shown can be parameterized within the computerized item model to produce massive numbers of variations on the same theme. Similarly, the interactive multimedia items shown in Panel 2, Figure 2 offer assesses interactive functionalities such as exploring bar graphs, zooming graphs, and scaling coordinates. Manipulating these interactive components can increase assesses' engagement and help them focus on the assessment exercise. The text-to-speech AI items shown in Panel 1, Figure 3 fulfill three main functions. First, this type of item model shows how external digital solutions can be integrated into an AIG framework to produce innovative items that read dynamically-generated passages for respondents without the need for human voice actors. Second, this type of item model allows us to investigate the impact of providing additional and/or optional information (i.e., a voice reading a passage out loud) on assessment results. Rather than simply reading the stem and the various answer choices as in conventional items, respondents could also listen to the stem and each option presented audibly. Third, this type of item model illustrates a potential utility of AIG as a test accommodation tool, dramatically improving the accessibility and the assessment accommodation options.

The items in Panel 2, Figure 3 also integrate external speech recognition, but here different artificial intelligence plays a role to receive/grade responses as voice. When respondents respond verbally when answering the randomized questions, the system automatically and instantaneously scores the sounds by comparing them with the correct response; this option could also be helpful for learning foreign languages by correcting the pronunciation offered by the student. The non-verbal cognitive reasoning items shown in Figure 4 can be used to measure various cognitive abilities, such as inductive, logical, abstract, and spatial reasoning. Using AIG, we can produce massive amounts of digitally mastered multimedia items adjusted appropriately for different cognitive components and various cognitive demand levels.

The items generated using CAFA AIG have proven to be useful in measuring different types of ability, including knowledge, comprehension, application, and analysis, at many levels. Incorporating such technology-enhanced features in assessments can improve the learning motivation of students and the efficiency of teaching (Choi et al., 2018). In addition, thanks to the mass production capability of AIG, these examples can be clear evidence of the claim that AIG is a sustainable and efficient tool for the transition from paper pencils to genuine digital assessments (Choi, 2017).

The next section provides several computerized item modeling examples, focusing particularly on how to develop item models and generate items from the models using the AIG system.

Figure 2 ■ Item instance pairs generated from five computerized item models using CAFA AIG system. Panel 1: 3D diagram items; Panel 2: Interactive multimedia items.



Computerized Item Modeling Examples

Before demonstrating how item models can be developed using CAFA AIG, it is necessary to present an overview of the interface for CAFA AIG system and CAFA AIG's syntax structure. Readers can access CAFA AIG system (WordPress plugin version) via the following link: http://aig.cafalab.com/.

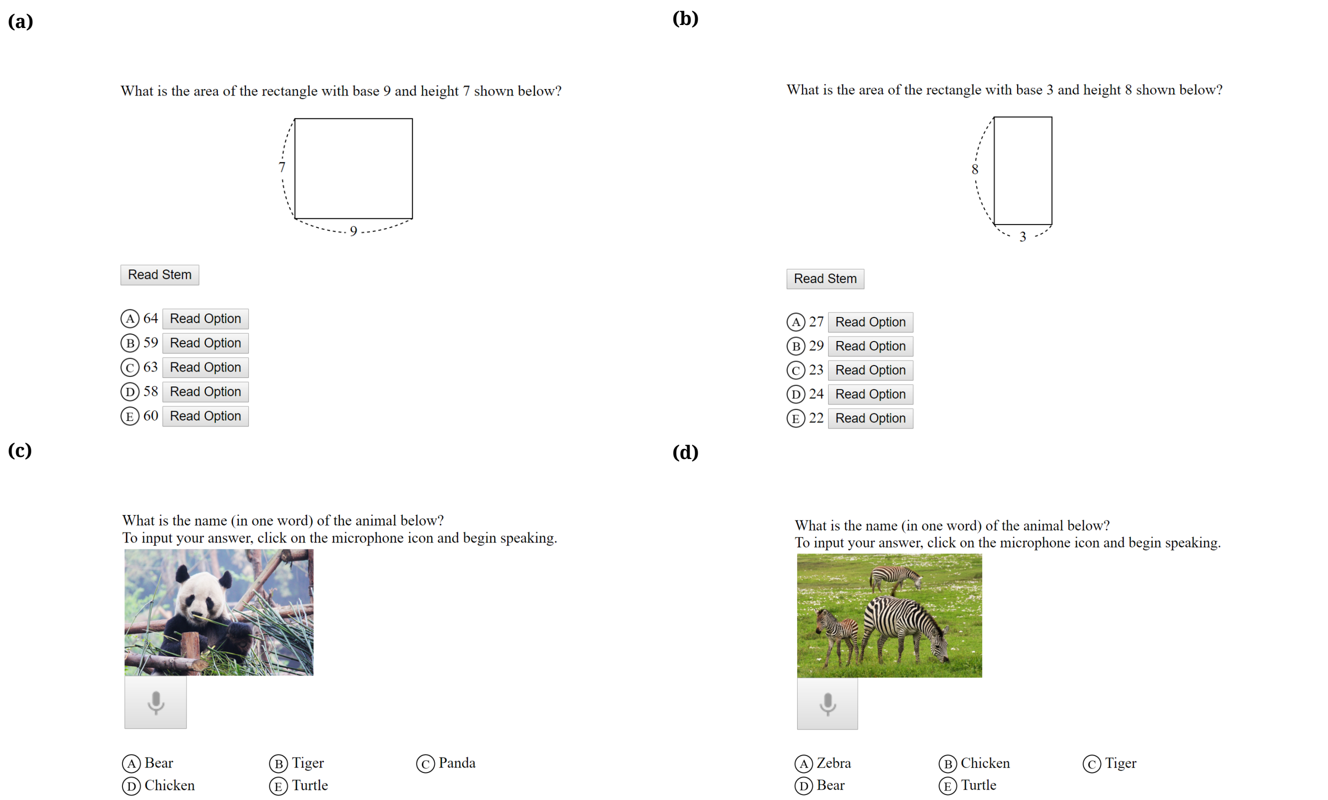
CAFA AIG Item Model

Once logged in to CAFA AIG system, users begin to develop an item model by clicking **CAFA** => **Create Item Model** in the CAFA AIG WordPress plugin located at the lower-left side panel. Then, a new dialogue box (Figure 5) will open, in which users can create an item model, or AIG template. Although the default is a selected-response format, CAFA AIG also supports a constructed-response format. As it is more widely used, in the following sections we focus on the default format.

The user interface of CAFA AIG item model includes several sections. Users can assign an **ID** number for each item model through ID section. For example, 100101 could be used to designate the 10th-grade/1st domain/1st cluster item model. **Standard** specifies measurement domain such as standard or construct, while **Description** provides additional descriptive information of an item model. **Stem** is one of the core components of an item model and includes statements, illustrations, and/or question; users can include item model parameter(s) in this stem section. **Difficulty** helps users specify the difficulty level (easy, normal or difficult) for the item model[1], and **Ability** identifies the cognitive ability level (knowledge, comprehension, application, and analysis) required to solve the item. **Number of Choices in a Row** determines how many options will be displayed in a row; the default value is three. **Parameters** is one of the most important components of item model as this is where users to identify and specify the elements of an item model. Users can set multiple parameters for an item model. For each one, the token "@number@" is used

[1] Note that this difficulty level is determined by item modelers or subject matter experts as an auxiliary item model information. Thus, it is different from the item difficulty index which is typically estimated with empirical response data within Classical Test Theory and/or Item Response Theory frameworks.

Figure 3 ■ Item instance pairs generated from five computerized item models using CAFA AIG system. Panel 1: Items with text-to-speech AI; Panel 2: Items with speech-to-text AI .



to name the parameter. The current version of CAFA AIG supports up to 24 parameters for each item model, with @1@ ... @24@ being used to name the 1st to 24th parameter. **Condition** sets the value boundaries for each parameter using various functions, and **Choice Options** and **Answer Key** specifies the choice options, including both the incorrect and correct choices for multiple choice format item models.

CAFA AIG Syntax Components

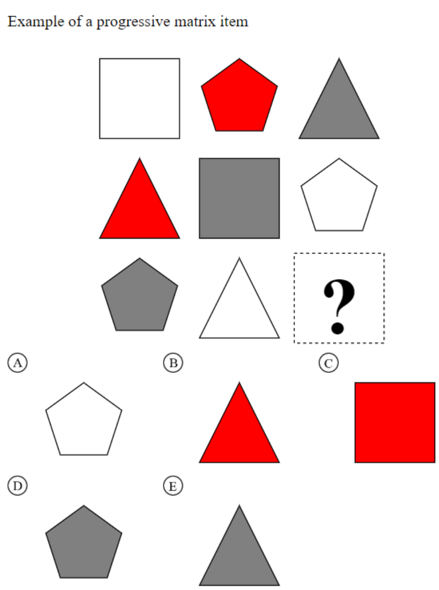
CAFA AIG syntax is a straightforward programming language that is implemented within CAFA AIG system to specify an item model. There are six categories of syntax, namely: 1) Normal text; 2) LaTex syntax for math expressions; 3) Parameters and system defined symbols; 4) HTML code; 5) Figure codes; and 6) Evaluated expressions. In this tutorial, we will focus on the parameters, system defined symbols, and figure code because these categories are the most widely used in practice. Visit the system website for more detailed references about the CAFA AIG.

Parameters. Parameters play an important role in generating new items because more new items are created when parameters are varied. There are three types of parameters: **Number**, **Enumeration**, and **Array**. Number type parameter is instantiated as a number and is defined by a range (minimum and maximum) and a precision. For example, a natural number ranging from 0 to 9 with a precision of 1 represents a one-digit number type parameter. Enumeration type parameter is used when non-numeric values are required. To define an enumeration type parameter, a list of possible values separated by | (a vertical bar character) can be used. For instance, when one enumeration parameter is instantiated as being one of "car," "ship," or "plane," its values can be presented as car|ship|plane. The third type of parameter is Array, which is used to instantiate a list of values. To define an array type parameter, in addition to its range and precision, a size — the number of values — must also be specified. For instance, in order to create an array parameter that represents 100 values uniformly distributed across the range [0, 1], we can present this as a list of numbers ranging from 0 to 1 with a precision of 0.001 (or smaller), and a size of 100.

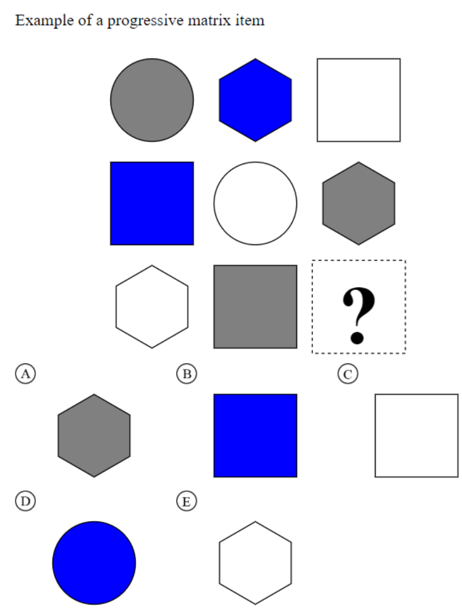
Conditions. Conditions ensure each parameter takes a reasonable value and guarantees the quality of the generated items. The conditions can be independent or they may depend on prior parameters. These are usually presented

Figure 4 ■ Item instance pairs generated from five computerized item models using CAFA AIG system; non-verbal cognitive ability measurement items.

(a)



(b)



as an expression with various operators. Suppose @1@ + @2@ is created as a stem and @1@ and @2@ are single-digit numbers. In order to ensure the answer is also a single-digit number, we can use the following parameter setting for @1@, @2@:

```
@1@: Min = 1; Max = 8; Precision = 1;
@2@: Min = 1; Max = 8; Precision = 1;
   Condition: < 10 - @1@
```

Figure commands. As a part of the CAFA AIG library, the figure commands are used to add dynamically varying figures or graphs into an item model. A figure command starts with two hash marks (##) and ends with another pair of hash marks. The specific explanation and application of the figure code will be described later in this tutorial. Visit the system website for more detailed references about the item model parameters, conditions, and figure commands.

Item Model Example 1

Figure 6 presents the CAFA AIG interface for generating items for the example shown earlier in Figure 1. These items meet the standard: "Add and subtract fractions with unlike denominators (including mixed numbers) by replacing given fractions with equivalent fractions in such a way as to produce an equivalent sum or difference of fractions with like denominators."

In *Stem*, the statement "What is $\dfrac@1@@2@ + \dfrac@3@@4@$?" uses LaTeX to display a math equation that contains four parameters: @1@, @2@, @3@ and @4@. In this example, parameters 1 and 2 are the numerator and denominator of the first fraction, respectively, while parameters 3 and 4 are numerator and denominator of the second fraction. These four parameters, along with two additional parameters (parameters 5 and 6) are defined in Parameters. Parameter 1 is an integer from 1 to 5; parameter 2 is an integer from 2 to 7. To create a simplified fraction, *mprime*, which stands for "mutual prime", is used to make sure that parameter 2 is a mutual prime with @1@. Parameter 3 is an integer from 1 to 5. As its condition "<>@1@" suggests, parameter 3 is not equal to the parameter 1. Parameter 4 is an integer from 2 to 10, is a mutual prime with parameter 3, and is not equal to parameter 2. To present the correct final result in simplified form, parameter 5 is set as the greatest common factor of the numerator (@1@ × @4@ + @2@ × @3@) and the denominator (@2@ × @4@) of the answer. In a similar way, parameter 6 is set as the greatest common factor of (@1@ + @3@) and (@2@ + @4@), which is used to present the wrong answers in simplified form. Five options, including one correct choice (option A), are identified in **Options** and **Answer**. Typical examples of the items generated in response to this procedure are presented in Figure 7.

Figure 5 ■ CAFA AIG item model specification interface

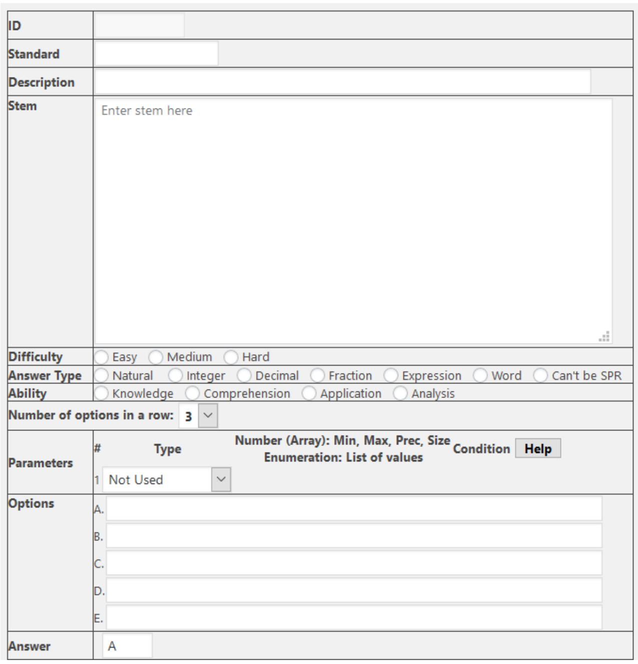
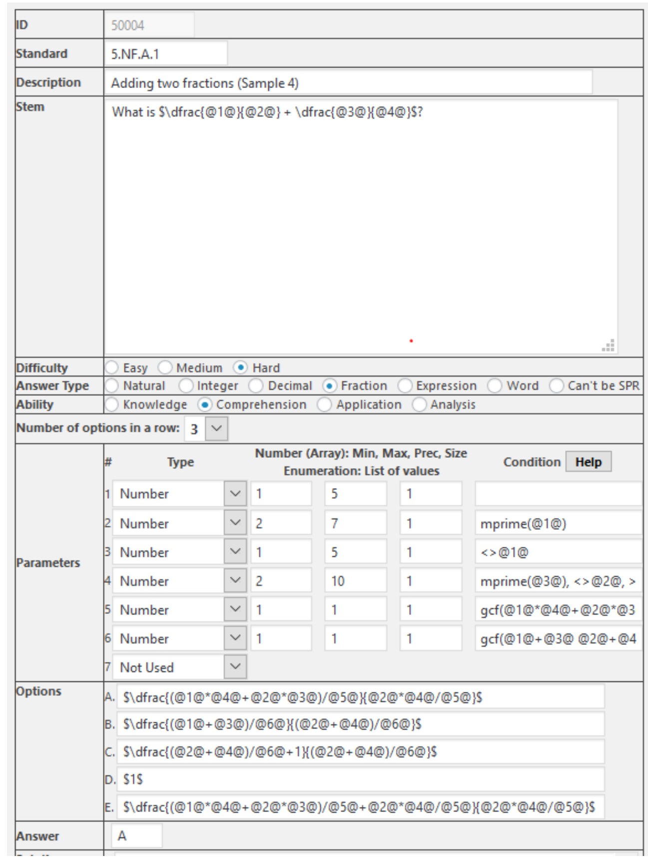


Figure 6 ■ The CAFA AIG syntax for the item model for Example 1



By clicking **Add Item button**, users can save an item model into the CAFA AIG server. Users also check the developed item models by clicking **CAFA => Create Item Models** in the plugin menu list. In this menu, there are 4 different action buttons for generating item instances from each item model a user developed: 1) firstly, by clicking Gen Insts button, users can generate item instances and save them into the CAFA AIG server; 2) by clicking **Copy to Local button**, users can copy the generated item instances into the local database where the plugin is installed and running; 3) by clicking **Export** button, users can directly download the item instances in a CSV (tab separated) format file; and 4) users also directly check the first 20 instances with a QTI format by clicking **QTIs** button. These items are all designed to comply with the digital standard such as HTML5, so they are optimized for use in a variety of digital/mobile devices.

Example 2

Figure 8 presents CAFA AIG interface and an item model for generating biology items. Four lines are written in Stem. The first three lines present a statement containing three parameters — @1@, @2@, and @3@. The fourth line "## 300 200 CP 0 1 0 1 1 1 0 0 0 IM @5@ 0 1 1 1 ##" functions as a figure code to display a picture. The first two numbers ("300 200") define the size of the drawing area in pixels and "CP 0 1 0 1 1 1 0 0 0" is a coordinate plane command with 9 arguments (numbers) that defines the coordinate plane. The first four arguments are, respectively, the minimum coordinate of x, the maximum coordinate of x, the minimum coordinate of y, and the maximum coordinate of y. In this example, the coordinate plane is [0, 1] x [0, 1]. The fifth and the sixth arguments define the gaps between the x and y coordinates. In this example, the gaps are 1 and 1, respectively. The last three arguments are "on" (1) or "off" (0) for the axis, grid, and labels. In this example, none of the axis, grid, and labels are displayed, as the values for the axis, grid, and labels are all set to 0. The "IM @5@ 0 1 1 1" image command is used to display an image specified by an URL. The IM command has five arguments. The first argument is the URL, which in this example is presented as a parameter. This means that this template provides different URLs for displaying different images. The second and third arguments are the coordinates of the left-top corner of the image; in this example, the coordinates are [0,1]. The fourth and the fifth arguments are the width and height of the image, which in this example are both set to 1.

Five parameters are identified in Parameters. The corresponding values for parameters 1, 2, 3 and 5, all of which appear in the stem, and parameter 4, which appears in the options, are filled with appropriate values. Here, the

Figure 7 ■ Item instances generated from the item model in Example 1.

(a)

Question: What is $\frac{5}{4} + \frac{1}{9}$?

A. $\frac{49}{36}$

B. $\frac{6}{13}$

C. $\frac{14}{13}$

D. 1

E. $\frac{85}{36}$

(b)

Question: What is $\frac{2}{7} + \frac{3}{2}$?

A. $\frac{25}{14}$

B. $\frac{5}{9}$

C. $\frac{10}{9}$

D. 1

E. $\frac{39}{14}$

value list for parameter 1 contains white-tailed jackrabbit, white-tailed ptarmigan, white-tailed ptarmigan, long-tailed weasel, long-tailed weasel, and arctic fox; the value list for parameter 2 contains prairies, tundra, mountains, prairies, mountains, and tundra; the value list for parameter 3 contains fur, feather, feather, fur, fur, and fur; the value list for parameter 4 contains community and population; and the value list for parameter 5 contains six different URLs. To ensure the generated items are accurate, index (1) is set as a condition to assemble elements with the same index as parameter 1 for parameters 2, 3 and 5. Five options, including one correct choice (adaptation) and four incorrect choices (ecosystem, niche, species, and either community or population, as set by parameter 4) are identified in Options and Answer. Examples of the items generated are presented in Figure 9.

Example 3

Figure 8 presents the CAFA AIG interface for generating mathematical items that require students to identify the unit rate of a linear function. Four lines are written in Stem. The first line presents the question “What is the unit rate of the following graph?”, the second line “

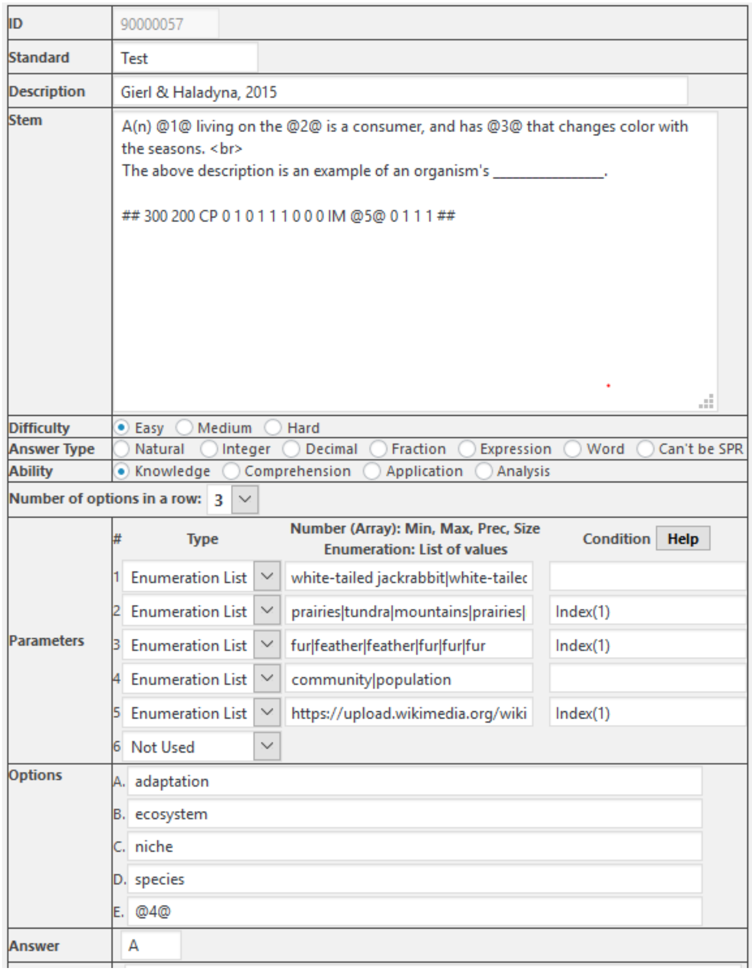
” works as a line breaker, and the third and fourth lines form a figure code “## 200 300 CP (-5) 8 (-8) 22 1 1 1 0 0 FL @2@ 0 DT @1@ @3@ FS 17 TX (@1@+1.5) (@3@-1) "(@1@, @3@)" ##” to display a linear function. There are three parameters in this figure code, namely @1@, @2@, and @3@. The first two numbers “200 300” define the size of the drawing area in pixels. “CP (-5) 8 (-8) 22 1 1 1 0 0” is a coordinate plane command that specifies that the coordinate plane is [-5, 8] x [-8, 22]; the gaps between the x and y coordinates are both 1; and the axis is displayed but the grids and labels are not shown. The command "FL @2@ 0" draws a linear function, y = @2@x + 0, which specifies a linear function that passes through the origin (0, 0) and whose slope is set by parameter @2@. "DT @1@ @3@ " orders the system to draw a dot at (@1@, @3@). FS 17, which is a font size command, sets the font size of the text as 17; the TX command “TX (@1@+1.5) (@3@-1) "(@1@, @3@)” then displays the text "(@1@, @3@)” at the location (@1@+1.5) (@3@-1).

As mentioned above, three parameters appear in the stem. These are identified in Parameters. The corresponding values for parameters 1, 2, and 3 range from 2 to 4, 2 to 5, and 1 to 20, respectively, and all have a precision of 1. Parameter 2 is conditioned not to be equal to parameter 1 and parameter 3 is conditioned to be equal to the product of parameters 1 and 2. Five options, including one correct choice and four incorrect choices, are identified in Options and Answer. As Figure 8 shows, the correct option is parameter 2; the incorrect options are dependent on parameters 1 and 3 and their operation. For example, option D is equal to the quotient of dividing 1 by parameter 2. The generated items are presented in Figure 9.

Conclusion

The rapid advances in science and technology have had widespread impacts throughout society, and psychometric or educational assessment is no exception. AIG, a state-of-the-art measurement technology, is rapidly developing into a very useful tool and is thus drawing a great deal of interest from assessment practitioners and researchers. However, despite its theoretical excellence and many potential advantages, it is also true that it remains quite difficult to use AIG in practice. This tutorial is specifically designed to help practitioners understand and use AIG tech-

Figure 8 ■ The CAFA AIG syntax for the item model for Example 2. (Note: some of the parameter value lists extend beyond the items shown in the window for that parameter.)



nology for real-life assessment practice and research using CAFA AIG system. In the first section of this tutorial, we introduced AIG theory, explaining its purpose, describing its advantages, and presenting its underlying philosophy. CAFA AIG system was introduced in the second section, after which we demonstrated how computerized item modeling is performed using CAFA AIG system with three practical examples. We hope this tutorial inspires readers to try this new approach for themselves, thus facilitating the transfer and dissemination of AIG technology, raising its profile and boosting its utilities for empirical research and practice.

References

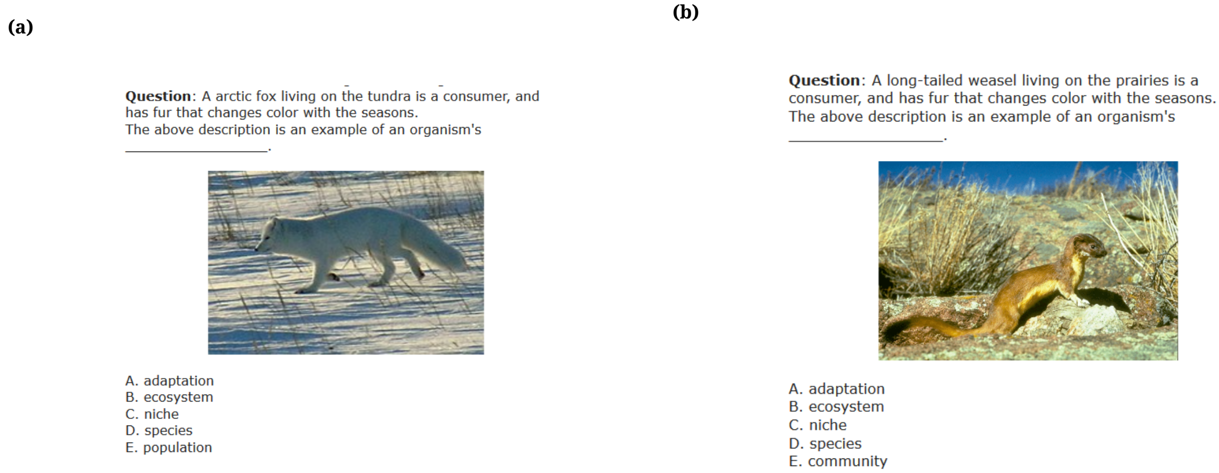
Bejar, I. I. (1993). A generative approach to psychological and educational measurement. In R. J. M. Frederickson & I. I. Bejar (Eds.), *N* (pp. 323–359). Test Theory for a New Generation of Tests . Mahwah, NJ: Lawrence Erlbaum Associates.

Bejar, I. I. (2002). Generative testing: From conception to implementation. In S. H. Irvine & P. C. Kyllonen (Eds.), *Item generation for test development* (pp. 199–217). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Choi, J. (2017). *Next generation item and test development: A practical introduction to automatic item generation*. Assessment, Testing and Measurement Series. Washington, D.C.: the George Washington University.

Choi, J., Kang, M., Kim, N., Dardick, W., & Zhang, X. (2015). A smart way of coping with common core challenges - introduction to cafa smartworkbook. *Journal of Educational Issues*, *1*(2), 70–89.

Figure 9 ■ Item instances generated from the item model in Example 2.



Choi, J., Kim, H., & Pak, S. (2018). Evaluation of automatic item generation utilities in formative assessment application for korean high school students. *Journal of Educational Issues*, *4*(1), 68–89.

Choi, J., Kim, S., & Yoon, K. (2016). *K-math workbook grade 6*. Clarksville, MD: CAFA Lab, Inc.

Choi, J., Kim, S., & Yoon, K. (2017). *K-math workbook grade 7*. Clarksville, MD: CAFA Lab, Inc.

Choi, J., & Lee, L. (2016). Automatic item generation: Beyond mass production, Washington, DC: Paper presented at the annual meeting of the Korean-American Educational Researchers Association.

Choi, J., Yoon, K., & Kim, S. (2017). *Computerized item modeling practice with CAFA Automatic Item Generation*. Clarksville, MD: CAFA Lab, Inc.

Dardick, W., & Choi, J. (2016). Teacher empowered assessment system: Assessment for the 21st century. *Journal of Applied Educational and Policy Research*, *2*, 1–11.

Embretson, S. E., & Yang, X. (2007). Automatic item generation and cognitive psychology. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics: Vol* (pp. 747–768). 26. Psychometrics North Holland, UK: Elsevier.

Gierl, M. J., & Haladyna, T. M. (2013). *Automatic item generation: Theory and practice*. New York, NY: Routledge.

Irvine, S. H., & Kyllonen, P. C. (2002). *Item generation for test development*. Washington, D.C.: Air Force Human Resources Laboratory & Educational Testing Service.

Lai, H., Alves, C., & Gierl, M. (2009). Applying item model taxonomy for automatic item generation for cat. In 1. a. Retrieved April (Ed.), *Proceedings of the 2009 gmac conference on computerized adaptive testing* (pp. 1–11). Washington, D.C.: 2011.

Citation

Choi, J., & Zhang, X. (2019). Computerized item modeling practices using computer adaptive formative assessment automatic item generation system: A tutorial. *The Quantitative Methods for Psychology*, *15*(3), 214–225. doi:10.20982/tqmp.15.3.p214

Received: 18/12/2018 ~ Accepted: 20/09/2019

Figures 10 and 11 follows.

Figure 10 ■ The CAFA AIG syntax for the item model in Example 3

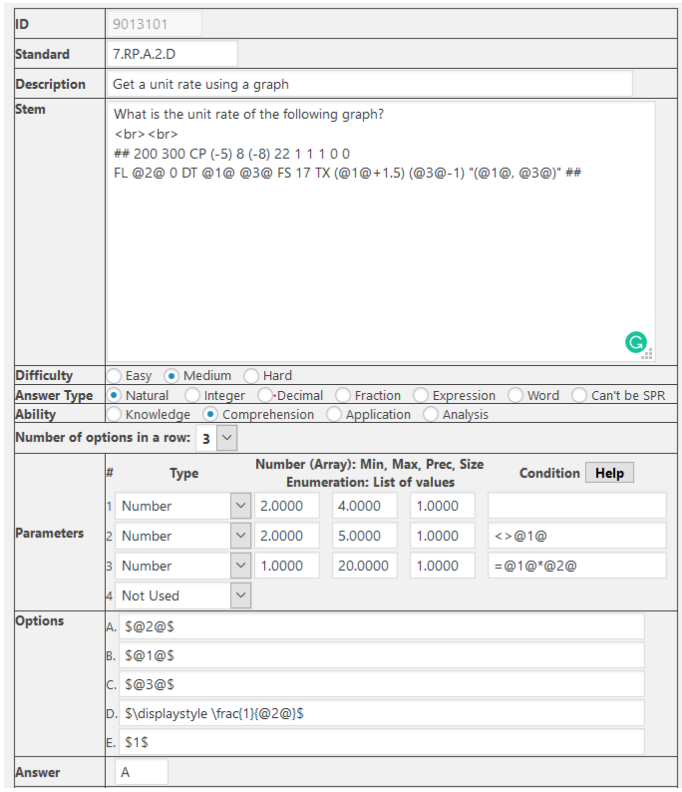
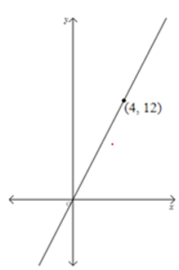


Figure 11 ■ Item instances generated from the item model in Example 3.

(a)

Stem: What is the unit rate of the following graph?



Options:

A. 4

B. 1

C. 3

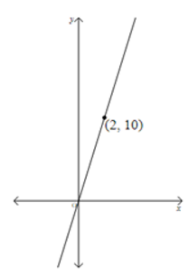
D. $\frac{1}{3}$

E. 12

Answer: C

(b)

Stem: What is the unit rate of the following graph?



Options:

A. 2

B. 5

C. 10

D. 1

E. $\frac{1}{5}$

Answer: B