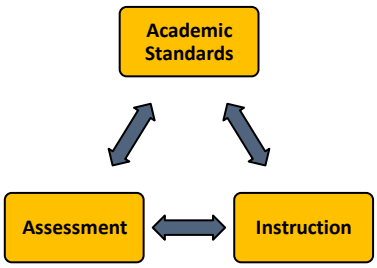
Virginia Alternate Assessment Program (VAAP)
Item Writing, Review, and Revision Style Guide
Updated April 2023

Style Guide Purpose & Updates: The purpose of the VAAP Style Guide is to update guiding principles for developing—writing, reviewing, & editing/revising—VAAP test items. There are three guiding principles for developing high-quality test items for students with significant cognitive disabilities (SWSCD) that we term (1) *alignment* (2) *appropriateness/relevance*, and (3) *accessibility* that are defined, below. Following these principles, a series of item writing and review pointers are described. Updates (compared to the previous versions of the style guide) are included based on stakeholder/user feedback, internal review/observations by VDOE, and adjustments to test development since the first year of the new VAAP (2021-22).

Test Security Reminder: VAAP test items are secure, please do not download, share them with anyone outside of VDOE/BRT, or discuss test item content through email or other insecure means. Test item writing/review templates should be shared via Pearson's SFTP portal (https://vaptsftp.pearson.com/) or some other agreed upon and secure means of test information transfer.

THREE GUIDING PRINCIPLES FOR ITEM WRITING AND REVIEW

1. ***Alignment*** is the degree to which the academic content and performance demands of a test item matches what is described in a corresponding standard (VESOL + complexity continuum). When comparing a test item and VESOL, there should be a close match (or alignment) between the content and performance expectations that allows students to demonstrate what they know (i.e., an item should be "spot on", with no improvements needed to improve the strength of alignment).



   - Collectively, test items should be a representative sample of the academic content expressed in VESOLs and offer an opportunity for students to demonstrate that they have learned that content — classroom instruction/assessment help make sure they have learned it.
   - Assessments, academic standards, & instruction should therefore be ***strongly aligned*** — working together and reinforcing one another — to make valid assessment-guided decisions across grade-level performance demands. Thus, close,

strong alignment between test items and VESOL is crucial as a prerequisite to other aspects of high-quality test items.

Key questions to consider as you write and review test items that are *strongly aligned with VESOL:*

- Does the test item include prerequisite / fundamental skills?
  Test items, especially those written to align with the lower end of the complexity continuum of the VESOL, may target prerequisite / fundamental skills. Test items that target prerequisite / fundamental skills are aligned with the VESOL if the content & performance demands build toward demonstrating proficiency described in the lower end of the complexity continuum of the VESOL. For example:
  - Science — identifying/distinguishing plants and animals from other "everyday" objects is a prerequisite to identifying parts of plants and animals and their function/purpose;
  - Math — counting unit squares is a prerequisite to using multiplication to solve more complex surface area problems using the area formula (A = l x w); and
  - Reading — identifying a word in a sentence/paragraph is a prerequisite to being able to identify and understand the meaning of grade-level vocabulary.

2. ***Appropriateness/Relevance*** is the degree to which the academic content and performance demand included in the test item and VESOL are connected to the learning needs and everyday lives of SWSCD. For example, *radioactive decay* is an interesting topic. However, the scientific concept itself, understanding radioactivity and quantum mechanics, likely holds little relevance for the vast majority of SWSCD. Alternatively, understanding the connection between *form and function* and applying that to understand how key parts of animals, plants, and humans help them grow and survive is a universal scientific concept that is relevant, including for individuals with significant disabilities. For example, teaching SWSCD that plant roots take in water and nutrients, or that eyes, ears, nose (etc) help animals and people make sense of the world (sometimes supported by assistive technologies), is both educationally germane and appropriate/relevant for helping build towards post-secondary success and employment. Addressing content that is relevant to living in the Virginia is also important for some VESOL (e.g., learning about watersheds and their importance).

   Key questions to consider as you write and review test items that are *appropriate/relevant*:

   - Does the test item require academic knowledge and skills that are important for success in school? (e.g., academically, socially)

- Is the test item connected to knowledge and skills that are important for success in post-secondary contexts? (e.g., life skills, employment opportunities)

3. ***Accessibility*** is the degree to which the academic content and performance demands in the test item are understandable, reachable, and demonstrable for SWSCD. Test items should clearly present targeted academic content in ways that allow students to access and demonstrate their knowledge and skills while lessening the chance that construct irrelevant variance affects the test results. For example, SWSCD might have multiple exceptionalities, including visual impairment or blindness. Test items should be constructed in ways that allow students to access and demonstrate

   Key questions to consider as you write and review test items that are *accessible*:
   - Will SWSCD understand the language in the test item?
     e.g., clear vocabulary and terminology
   - Will SWSCD understand the intent/performance demand of the test item?
     e.g., academic knowledge and skills targeted
   - Are there sensory impairments that may make this test item inaccessible for the SWSCD population or a particular segment of the SWSCD population?
     i.e., even if administered in Large Print, Braille, or another appropriate communication modality that is aided by available supports/accommodations
   - Is the test item free of superfluous details—in text, graphics, or concepts?
     e.g., details unrelated to demonstrating proficiency around the targeted content/performance demand, free of jargon
   - Is the test item free of errors that inhibit meeting performance demands?
     - Typos / mistakes in spelling or grammar
     - Incorrect or missing graphics
     - Incorrect answer
   - Is the test item consistent with other test items?
     - Broadly consider the uniformity of item design and format—item development is both a science and an art
     - Adhere to UDA principles (see, for example, Downing's (2006) *Twelve Steps for Effective Test Development* and Haladyna's (2004) *Developing and Validating Multiple-choice Test Items*)
     - Avoid items based solely on visual information/being able to see graphics, where providing text would give away the correct answer and/or graphics are not accessible to visually-impaired students

- E.g., ***More*** VI-biased — Here are three natural areas. [Point to options.] Which is an ocean: A, B, or C? (pics only for answer options offer less access to test content for VI students)
- E.g., ***Less*** VI-biased — Here are three natural areas. [Point to options.] Which is a large body of water: ocean, desert, or forest? (pics **and** words for answer options offer greater access to test content for VI students)

- **Does the test and test items adhere to Universal Design for Assessment (UDA) Principles Important to Developing Quality Multiple-choice VAAP Test Items?**
  - Clear/brief directions
  - Simplified and consistent vocabulary and phrasing
  - Parallel grammar structure
  - Brief stem/prompt and answer options
  - Sufficient and succinct background/framing information
  - Work seamlessly with embedded and available supports/accommodations
  - Randomly presented [correct] answer options
  - Free of bias (specific disability/exceptionality related; demographic characteristics (e.g., gender identity, race/ethnicity, religion)

OTHER CONSIDERATIONS FOR ITEM WRITING AND REVIEW

1. ***Four General Item Types & Examples —***

   1) **Text-only items** (click for example *practice* items)
   2) **Text w/ prompt graphic** (click for example *practice* items)
   3) **Text w/ answer option graphics** (click for example *practice* items
   4) **Text w/ prompt and answer option graphics** (click for example *practice* items)

2. ***'Writing New Items w/ Item Template Sheets —***

   Template Dictionary:
   item_id = internal item ID code (TBD)
   group = grade
   standards = VESOL ID code
   reporting_cat = VAAP score reporting category
   vesol = revised VESOL text
   lmh = Low, Medium, High parameters associated with VESOL and that guide item writing

complexity_cont = teacher/parent friendly version of VESOL LMH parameters
lmh_target = LMH parameter that item should target (options = L, M, or H)

item_information = includes all text-based parts of item —
verbal preamble + hidden/examiner text [in brackets w/ period at end of sentence] for paper pencil + prompt/stem + three answer options, colon after stem, question mark at end of sentence)
e.g., Here are 10 cars. [Point to the cars.] Mike will divide these 10 cars into two groups. How many cars will be in each group: 2, 3, or 5?

Prompt = item prompt or stem — what student will see on screen/paper
e.g., Mike will divide these 10 cars into two groups. [prgrph] How many cars will be in each group?

Note paragraph html code to begin a new line of text in prompt. i.e., [prgrph]
Typically, this code is inserted before main question prompt, including: after passages/text in Reading items, before math equations in Math items, etc.

option_a = answer option A (text only – *1 to 4 words, consistent length across A, B, & C options*) — e.g., 2
option_b = answer option B (text only – *1 to 4 words, consistent length across A, B, & C options*) — e.g., 3
option_c = answer option C (text only – *1 to 4 words, consistent length across A, B, & C options*) — e.g., 5

correct = correct answer option (options = A, B, or C) — e.g., C

Graphic Notes = succinct details that specifically describe any graphics associated with prompt and/or answer options (see Pearson specs, below, too) —
e.g., Prompt: Two rows with 5 identical cars in each row (2 by 5 array). Put the division problem 10 ÷ 2 = ____ under array. No graphics for answer options.

feedback = item feedback/revision instructions, to possibly be returned to writer(s) for study and to implement in other items

3. ***Pearson Item (Graphics) Specs to Consider —***

- All graphics are line art

- Adobe Illustrator templates (for BRT graphics artist)
  - ***VAAP_template_response_v2.ait*** = template to be used for all answer option art
  - ***VAAP_template_stem_1_para_text_v2*** = template to be used for stem art if the stem also contains <u>one</u> line of text in a paragraph tag
  - ***VAAP_template_stem_2_para_text.ait*** = a template to be used for stem art if the stem also contains <u>two</u> lines of text in a paragraph tag
  - ***VAAP_template_stem_3_para_text.eps*** = a template to be sued for stem art if the stem also contains <u>three</u> lines of text in a paragraph tag (hardly ever used)
- All files have white background to make art accessible with the color contrast tool in TestNav
- Ensure all stroke widths are at least 1pt
- Ensure everything in the color black is in true, 100% black and the black is not being created by a combination of colors
- Ensure all art is on one layer
- Font = Arial MT Pro (for any labeling/text in graphics)

<u>VDOE/BRT REVIEW PHASES & ASSOCIATED STEPS</u>

Review Phase 1 — Internal pre-delivery to Pearson — Initial item writing/harvesting/reviewing/editing

- Leverages item templates
- Item writers (VDOE + BRT)
- Item reviewers (VDOE + BRT)
- Item editors (VDOE + BRT)
- Selecting/writing/reviewing/revising/reconciliation
- Informal, tho specific feedback
- Goal = develop “semi-clean” versions of items to be uploaded to DIR for more formal review

Review Phase 2 — Internal pre-delivery to Pearson — Follow-up item editing/refinement, more formal [DIR-based] reviews by VDOE, BRT, and/or Virginia educators/specialists

- Leverages DIR (likely at least two formal reviews)
- Item reviewers (VDOE + BRT … VA special educators/education specialists, when appropriate)
- Item editors (BRT)
- Formal feedback based on co-developed DIR questions

- Goal = develop “clean” versions of items to be included in practice and operational test forms

Review Phase 3 — Internal post-delivery to Pearson — Item editing/refinement via ABBI and TestNav8

- VDOE review in ABBI (Pearson proprietary platform for reviewing items)
    - Items in test form order
    - Single item view
    - Ensure rendering of text and graphics is correct
    - Preview mode [student view]
    - Check text to speech
    - Student paper copy
    - Scoring correct
    - Standard assigned correctly
- VDOE review in Testnav8 (software that administers VAAP to students online)
- Examiner's copies and student materials created post-Testnav8 review

Examples of text-only items — Reading and Math, respectively

Spoken: Here are three words. (Point to answer choices.) I will read a sentence to you, follow along. (Point to and read sentence to student.) What is another word for father: mom, dad, or brother?

Ty likes to do things with his father.

What is another word for father?

mom	dad	brother

Spoken: Here is a subtraction problem: 19 minus 13. (Point to student materials.) If you have 19 and take away 13, how many are left: 2, 3, or 6?

19 - 13 = ___?

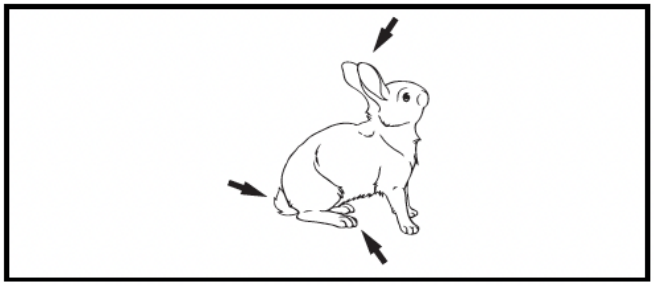
If you have 19 and take away 13, how many are left?

2	3	6

Examples of text w/ prompt graphic items — Science and Math, respectively

Spoken: Here is a rabbit. (Point to student materials.) Which trait helps the rabbit hear predators when they come near: feet, tail, or ears?

Which trait helps the rabbit hear predators when they come near?



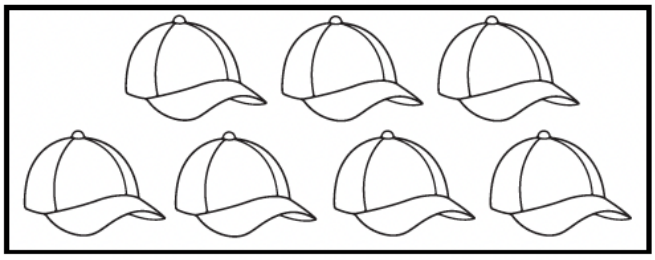
feet

tail

ears

Spoken: Here are baseball caps. (Point to student materials.) How many baseball caps are there: 5, 6, or 7?

How many baseball caps are there?



5

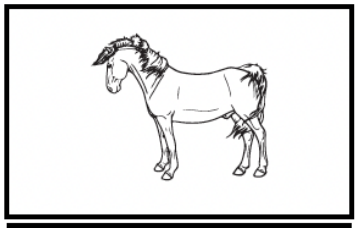
6

7

Example of text w/ answer option graphics items — Science, Reading, & Math, respectively

Spoken: Here are three living things. (Point to student materials.) Which is an animal: horse, mushroom, or tree?

Which is an animal?



horse

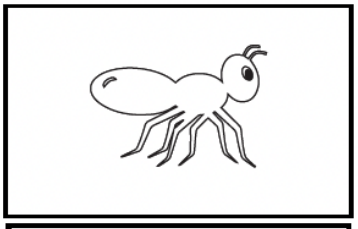
mushroom

tree

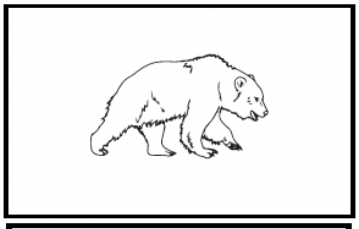
Spoken: Here are three pictures with words. (Point to answer choices.) I will read a sentence to you, follow along. (Point to and read sentence to student.) What likes to eat fish: ants, bears, or lions?

Bears like to eat fish.

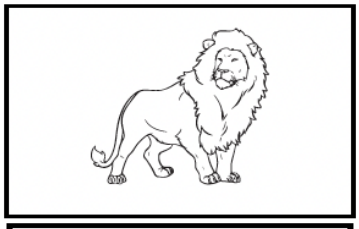
What likes to eat fish?



ants



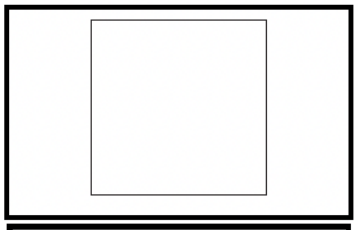
bears



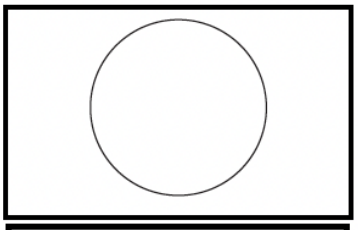
lions

Spoken: Here are three shapes. (Point to student materials.) Which shape has 4 sides and 4 angles: A, B, or C?

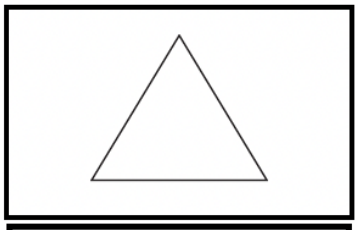
Which shape has 4 sides and 4 angles?



A



B



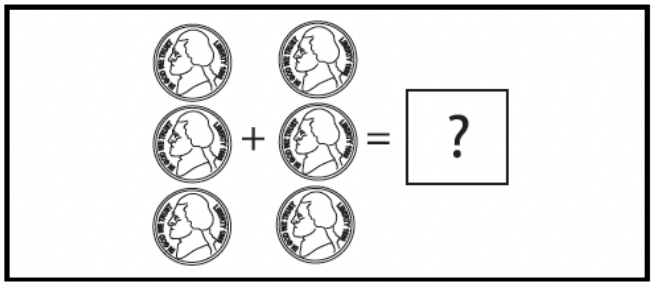
C

Example of text w/ prompt and answer option graphics — Math and Science, respectively

Spoken: Here is an addition problem with nickels. (Point to student materials.) A nickel is 5 cents. How much money is 3 nickels plus 3 nickels: 20 cents, 25 cents, or 30 cents?

n = 5 cents

How much money is 3 nickels plus 3 nickels? 3n + 3n



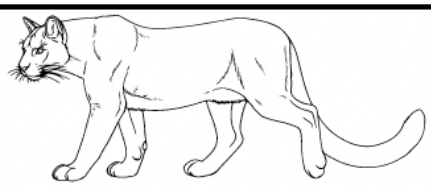
20 cents

25 cents

30 cents

Spoken: This is a cougar. (Point to student materials.) Which could compete with the cougar for food: grass, ant, or wolf?

Which could compete with the cougar for food?



grass

ant

wolf