

# Integrated testlets in optics and photonics: an assessment tool suitable for textbook and online delivery

Rayf C. Shiell<sup>a</sup> and Iain R. McNab<sup>b</sup>

<sup>a</sup>Department of Physics and Astronomy, Trent University, Peterborough, Ontario, K9L 0G2, Canada<sup>1</sup>

<sup>b</sup>McNab Group, 152 Park town lotte APT 127-205. Naejeong-ro, Bundang-gu, Seongnam-si, 13599, South Korea<sup>2</sup>

## ABSTRACT

Integrated testlets are a means to assess a student's understanding of complex knowledge through a set of scaffolded questions within an answer-until-correct format, and with grades that can, if desired, be awarded according to the number of attempts made by the student. Integrated testlets have been delivered to students at several universities in Canada, in physics, chemistry, and biology. In comparison with traditional multiple-choice-based assessments that ask items that are wholly independent of each other, an integrated testlet purposefully poses a set of integrated and dependent items that build on each other. This allows for the assessment and formative learning of deeper and more interconnected aspects of the course material. This is important in all STEM disciplines, and especially beneficial for cumulative and interdisciplinary fields such as optics and photonics. In the past few years we have for the first time extended integrated testlet delivery to an online format using the WeBWoRK delivery system, and more recently included these within end-of-chapter questions throughout a textbook in optics as part of our significant update and revision of the classic, internationally-known, *Introduction to Optics* text by F. L. Pedrotti, L. S. Pedrotti, and L. M. Pedrotti. For students this provides the benefits of integrated testlets as originally conceived, while also connecting topics within the book. This has encouraged us as authors to use deliberate and mindful composition practices, and advanced our skills in conveying the breadth and depth of the many concepts within optics and photonics.

Keywords: Integrated testlets, answer-until-correct, multiple choice, formative assessment, immediate feedback

## 1. INTRODUCTION

In recent years one of us (RS) has co-created a new student assessment tool, called an integrated testlet,<sup>1,2</sup> which has been successfully delivered to students at several educational institutions across Canada, through paper-based and online means. An integrated testlet assesses students' understanding of complex knowledge through a set of scaffolded multiple-choice questions, each adopting an answer-until-correct format. Compared to typical multiple-choice-based tests that ask questions (called items) that are wholly independent of each other, an integrated testlet adopts an answer-until-correct format to then purposefully allow items to then become interdependent so they can build on one another. This scaffolding allows for the assessment and learning of deeper and more interconnected aspects of a course. Testing such interconnectivity of knowledge is particularly useful within STEM disciplines, and is especially beneficial in cumulative and interdisciplinary fields such as optics and photonics.

Since their inception, originally targeted to university students during tests and examinations using paper-based question sheets and immediate-feedback scratch cards<sup>3</sup>, the scope of integrated testlets for learners has broadened enormously. Students who can benefit from this tool include those in technical courses, those studying remotely, and also those in grades 9-12. Additional settings for deployment include being employed as "ice-breakers" before class (to then act as the genesis of the topic of the day), as small group activities midway through class (to refocus student engagement and promote peer-to-peer learning), and as low stakes end-of-module assessments (to consolidate a topic and its utility within a tangible narrative). In this paper we expand on the many benefits afforded by integrated testlets and describe various

<sup>1</sup> [rayfshiell@trentu.ca](mailto:rayfshiell@trentu.ca)

<sup>2</sup> [iain.mcnab@mcnabgroup.com](mailto:iain.mcnab@mcnabgroup.com)

modes of delivery, from the original incarnation that adopted physical scratch cards to more recent delivery, both remotely, and as a self-study tool within an upcoming upper-year textbook in optics.

## 2. THE “WHY” AND THE “WHAT” OF INTEGRATED TESTLETS

Figure 1 illustrates a topic within optics and photonics for which an instructor might want their students to glean and to demonstrate an understanding of several core concepts, and to be able to envision these within a real-world setting. Here, an aperture,  $S$ , acts as a light source, but one that is neither a point source, nor a monochromatic one. Of interest to an optical scientist, and in a way that has real geometric meaning, is a *volume of coherence*, centered at some point,  $P$ , downstream of the aperture, from within which one can sample the light field through interferometric means and a visible fringe pattern would result. While this volume could be directly asked of students as in, for example, the last stage of a traditional, multi-part, free-response question, this does bring many drawbacks. First and foremost, it provides numerous pitfalls, with learners erring midway through such a task due to a single case of incorrect recall, or incorrect computation, or lack of conceptual understanding. This can hinder a student’s progress and impact their confidence, and can lead to a (somewhat justified) sense of injustice from the student feeling trapped due to the double jeopardy at play.

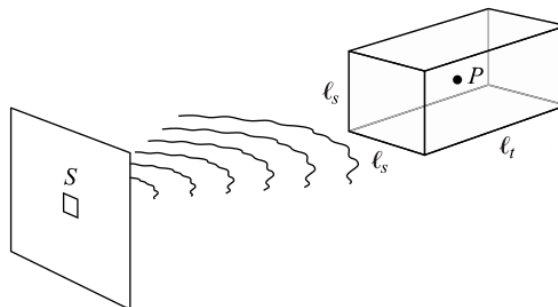


Figure 1: A depiction of concepts covered by the general topic of coherence: here, the spatial coherence width,  $l_s$ , the temporal coherence length,  $l_t$ , and the overall volume of coherence for a light field emerging from an aperture,  $S$ .

This wish to assess complex topics through more empathic assessments was one driver that led to the inception of integrated testlets. There was also, particularly at the introductory college level, a desire for assessments to evolve from asking only a barrage of unrelated, and often superficial, multiple-choice (MC) questions, and to instead probe higher rungs of Bloom’s revised taxonomy<sup>4</sup>, beyond simply recalling facts or implementing equations in an isolated manner. This goal is particularly important in STEM disciplines, where the intrinsic nature of the subject requires a knowledge and understanding that is cumulative and interconnected. Traditional, isolated MC questions, with their atomistic approach of small and disconnected items, lend themselves best to the assessment of small and disconnected parcels of knowledge.

Figure 2 shows an integrated testlet that can employ any of the answer-until-correct modalities we discuss below. This particular integrated testlet assesses student knowledge of the spatial and temporal coherence of a light field, with the pinnacle of understanding being that of determining the volume of coherence at some distance from an illuminated square aperture. Such complex assessments, adopting an answer-until-correct approach, serves many purposes beyond simply ranking test-takers. Rather, student learning is now a primary goal, with such a formative assessment, and students gleaning the scope and level of their knowledge, with positive reinforcement of each correct answer, rather than the negative reinforcement associated with incorrect answers with which they would otherwise leave a traditional assessment.<sup>5</sup>

Immediate feedback from multiple-choice items thus provides a real-time, two-way narrative between student and content expert, centered around a physical scenario that requires an extended period of focus, and which has the flexibility through various extensions (see below) to connect more strongly to the real world. The pedagogical messaging is particularly valuable as it encourages students to “press pause”, and to fully digest a scenario and its setting, with many items based around one scenario. Integrated testlets can be deployed at the beginning of a class to provide a springboard for the ensuing class discussion, with all students having been immersed in the same scenario, independent of the subject matter of any preceding class. Integrated testlets may also be deployed during regular class time to reinforce and connect concepts that have just been taught. Often taking the role of an unscored, formative, mid-class assessment, students can work in groups to jointly answer an integrated testlet. We have observed groups cheering with delight when they uncover the correct answer to an item and then enthusiastically progressing through the IT as they acquire and/or reinforce their knowledge. Group discussions in particular enhance peer-to-peer learning, as the very

process of explaining a concept aids in its crystallization.<sup>6,7</sup> In all the above, immediate feedback proves to be crucial for establishing such an environment of active learning.

An extended light source has the shape of a square aperture with sides 1 cm, and emits light with wavelength 500 nm and linewidth 0.035 nm. We are interested in the *volume of coherence* of the light field at a distance  $r$  away from this aperture, where  $r$  is much greater than the dimensions of the aperture.

(i) Which of the following statements is most correct about the *volume of coherence* at distance  $r$  away from the aperture?

A. The volume of coherence decreases as distance  $r$  increases, and here takes the shape of a *cylindrical* prism.  
B. The volume of coherence is independent of  $r$ , and here takes the shape of a *cylindrical* prism.  
C. The volume of coherence increases as distance  $r$  increases, and here takes the shape of a *rectangular* prism.  
D. The volume of coherence increases as the distance  $r$  increases, and here takes the shape of a *cylindrical* prism.  
E. The volume of coherence is independent of distance  $r$ , and here takes the shape of a *rectangular* prism.

(ii) Which of the following is the *length* of the volume of coherence along the direction towards/away from the source?

A. 1.2 mm      B. 1.2 cm      C. 7.1 mm      D. 7.1 cm      E. 7.1 m

(iii) Which of the following is the *width* of the volume of coherence at a distance 1 m from the source?

A. 50  $\mu\text{m}$       B. 61  $\mu\text{m}$       C. 100  $\mu\text{m}$       D. 1 mm      E. 122  $\mu\text{m}$

(iv) Which of the following is the *volume of coherence* at a location 1 m from the source?

A.  $3.5 \times 10^{-12} \text{ m}^3$       B.  $5.2 \times 10^{-12} \text{ m}^3$       C.  $12 \times 10^{-12} \text{ m}^3$       D.  $18 \times 10^{-12} \text{ m}^3$       E.  $26 \times 10^{-12} \text{ m}^3$

Figure 2: An integrated testlet comprising four items that, through one of the answer-until correct approaches described below, assesses students' understanding of spatial coherence width, temporal coherence length, and the volume of coherence.

When used within a test setting, the availability of partial credit for multiple attempts is perceived to be a fair approach by most students, who also appreciate that they leave the test with immediate knowledge of how well they did. The marking itself is also objectively fairer, as evidenced by a control-treatment study of first year physics exams with identical scenarios presented by both an integrated testlet and a free-response approach.<sup>8</sup> In our experience a student who has little idea how to begin a problem will often leave an entire page of their exam booklet blank, leaving the grader with no choice but to award a zero; yet it is likely the student knew something, and grades from the equivalent integrated testlet indeed confirmed this.

Finally, significant pedagogical benefit ensues when instructors become actively involved with either integrated testlet adaptation or composition, for much of the pedagogical power of integrated testlets derives from their extensible nature. Three kinds of items are commonly added, each with slightly different purposes. These are termed *intra-*, *ultra-*, and *extra-*testlet items. They each allow for uniquely-targeted measurement of knowledge by either expanding the measured domain or improving the operation of the testlet with modified cuing and scaffolding. Intra-testlet items provide additional scaffolding without expanding the breadth of the tested knowledge. This provides an additional step at a particular point to consolidate learning. Ultra-testlet addition incorporates an additional item with the goal of testing knowledge that goes beyond that of the original testlet, increasing the height of the pinnacle of the original question. This additional item now represents a more accessible rung in question difficulty than if it had appeared in the free-response analogue. Such additions allow instructors to reach beyond that which they would be comfortable testing for within a traditional free-response setting without immediate feedback. They add richness to a course and allow for low-stakes discrimination between the top students in a class. Lastly, extra-testlet addition tests knowledge that is either adjacent to, fundamental to, or complementary to the domain of the original integrated testlet. In particular, adding a conceptual item that underpins a given rung or even the entire knowledge domain can reinforce the foundational principles within a topic.

Such knowledge is often presumed, but rarely assessed. Thus, an extra-testlet addition is a non-integrated “one-off” item that can be placed anywhere within the testlet to increase the dimensionality of the sequence of items.

Composition of integrated testlets provides benefits to students through the training of instructors. Composing a testlet focuses the mind of the instructor on the purpose, the relevance, and the interconnected nature of the subject matter, and this activity is in practice both an art and a science. Basic composition requires identifying a scenario, arriving at distractors for each item, and considering distractors across all items. Refining an integrated testlet involves an awareness of several steps. First, this involves providing a narrative that speaks to the student interests (in optics and photonics, perhaps connecting to astronomy through the recently-launched James Webb Space Telescope, or connecting to life sciences through the differential depths of fields of herbivores with horizontal pupils compared to prey animals with vertical pupils). Composing alternate versions, and being mindful of rounding artefacts as students progress from one item to another, all connect the instructor to the material in a student-empathic way that the “chalk and talk” modes of delivery of times past did not.

### 3. APPROACHES FOR DEPLOYING INTEGRATED TESTLETS

#### 3.1 Scratch card deployment

Methods to provide the immediate feedback required of integrated testlets were for some time relatively difficult to implement. Early approaches used cumbersome mechanical boxes (Pressey, 1950). For the past two decades, scratch cards, known commercially as the *immediate feedback assessment technique* (IF-AT<sup>®</sup>), became available and these allow the adoption of the answer-until-correct model for classroom examinations. These contain boxes coated in a similar way to scratch-and-win lottery tickets, and conceal a star within the keyed-response option, with the distractor options being blank. Students answer each item until a star is revealed, and they then advance to the next item within the testlet with full knowledge of the answers to all previous items. The grade assigned for each item is dependent on the number of boxes that are scratched.

There are several aspects of good practice specific to the adoption of scratch cards for integrated testlet deployment, which are outlined in the literature.<sup>9</sup> Being mindful of the student experience includes delaying the distribution of the cards so this is a chance to pause, to think, and to avoid “scratch-fever”. In our experience, many students appreciate this opportunity to begin synthesizing answers in their supplied workbooks, as they would in traditional constructed-response exams. When finally given the cards, few of these students actually begin to scratch, while most instead continue to answer questions as if they were constructed-response, only taking the time later to transfer their answers to the IF-AT cards for validation. We provide different versions of proctored assessments to an exam room of students to discourage any temptation from looking at others’ cards, and because a scratch card-based exam is new to many students, we release the test instructions to the class in advance of the test. This allows the students to read and digest them in advance which saves time and reduces student anxiety.

Scratch cards require “keying” of all items’ options to a particular pattern, with the keyed responses placed at locations that are consistent with that of the card. This provides an obligation and an opportunity to compose a new set, or to re-order the existing set, of distractors in such a way that the flow of the entire integrated testlet is maintained.

#### 3.2 Online deployment

The recent pandemic and associated remote delivery of courses for a period of several months catalyzed the exploration of moving integrated testlet assessments to an online framework. The WeBWorK platform was chosen for this, as it allows flexible coding for each question and could be configured to allow up to four attempts for each item, with repeated (but initially randomized) variables for each student, and partial marks according to the number of attempts, with a scoring scheme that could be defined by the instructor.

Figure 3 shows a screenshot of the WeBWorK platform delivering the second item of the integrated testlet shown in Figure 2 in the Trent University PHYS 4240H *Modern Optics* course, although with unique variables in both stem and options for this particular student. The column on the bottom left shows, through numbered entries that are either

integers or end with a “.1”, the progress of the student through the integrated testlet. It can be seen here that they required two attempts at the first item, and one attempt for the other three items.

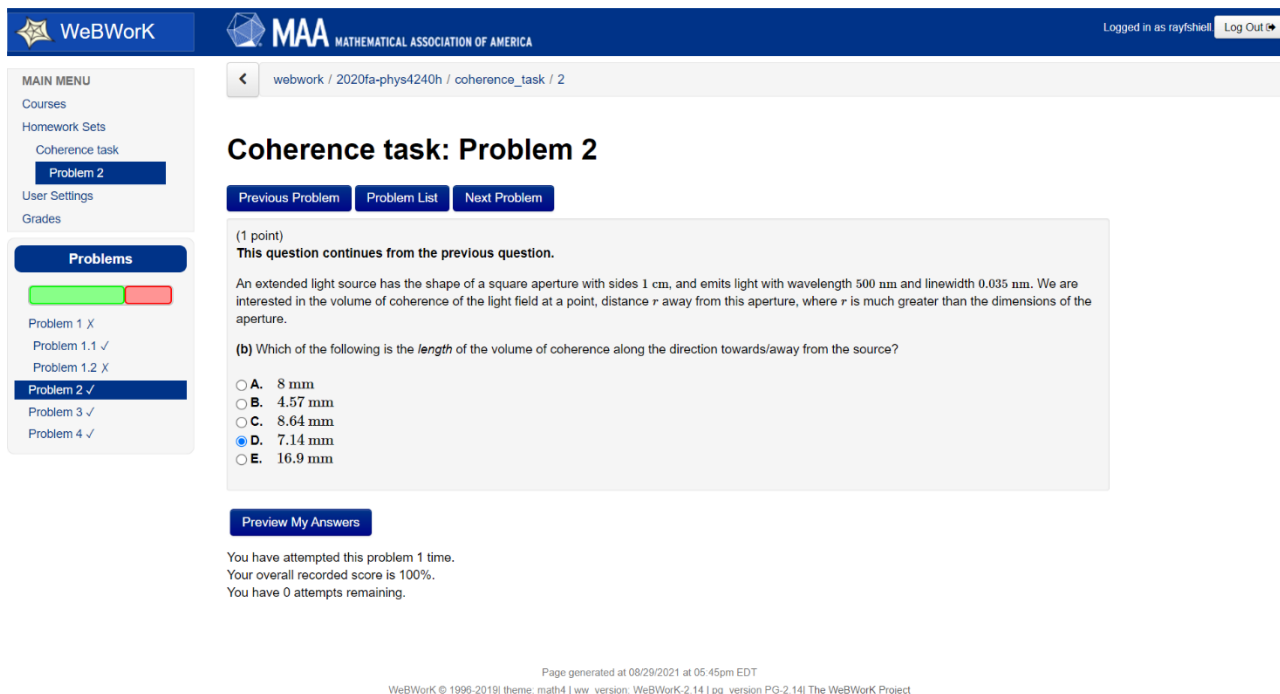


Figure 3: Online deployment of the second item of the item integrated testlet shown in Figure 2, although with variables and options unique to this student. The progress of the student through the four items is shown on the bottom left column.

Some advantages of online delivery through the WeBWorK platform in particular include:

- (a) the instructor is able to see how long, and in which order, each student answers each item within a question set. This enables the duration of a test or exam to be optimized, and facilitates both psychometric analysis and research on science pedagogy
- (b) marking is achieved entirely digitally and instantaneously, which avoids having to add up scratches on IFAT cards (and obviates the need for purchase of the same).
- (c) individualized, randomized questions can be asked of students such that the question’s input variables, distractors, and the option order together all constitute a question that is unique for every student. This helps the question become quite ‘cheat-resistant’, with the sharing or uploading of a problem to an external “homework help” website uniquely identifying the student.

### 3.3 Textbook deployment

The textbook delivery of integrated testlets will soon be implemented within each chapter of an update of *Pedrotti’s Introduction to Optics*,<sup>10</sup> which will be a 4<sup>th</sup> edition authored by both of us, with expected publication in 2024. The goal here is to provide for students a self-study tool that they can explore in their own time, and also to serve as a pre-class task given by the instructor from which the topic of the class can be developed. Further, these can be adapted, either by simply changing variables and redeploying the integrated testlet to the class, or by the addition of intra-, ultra-, and extra-items addition as described in Section 2. In the textbook deployment there is no penalty associated with the number of incorrect attempts that a student may make.

This form of deployment involves placing a code next to each option in each item, with a lookup table located in the back of the book that students can reference. The size of the lookup table, together with the random ordering of entries and the use of non-sequential numbering of codes in each item, is sufficient to discourage attempts to look up each

answer in advance of attempting the question. Figure 4 shows such a form of delivery, with an excerpt from the lookup table.

(iv) Which of the following is the <i>volume of coherence</i> at a distance 1 m from the source?		3297	YES	5788	NO
A.	$3.5 \times 10^{-12} \text{ m}^3$	3323	YES	5852	YES
B.	$5.2 \times 10^{-12} \text{ m}^3$	3327	NO	5879	NO
C.	$12 \times 10^{-12} \text{ m}^3$	3354	NO	5938	YES
D.	$18 \times 10^{-12} \text{ m}^3$	3451	NO	6025	NO
E.	$26 \times 10^{-12} \text{ m}^3$	3502	NO	6032	NO
		3519	NO	6095	YES
		3540	YES	6117	NO
		3641	NO	6168	YES
		3813	NO	6172	YES
		3822	YES	6174	NO
		3891	NO	6194	NO

Figure 4: (left) Excerpt of textbook deployment of the fourth item of the integrated testlet shown in Figure 2. (right) A small excerpt of the integrated testlet lookup table that is to be affixed to the back of the textbook.

### CONCLUSIONS

We have here introduced some motivation for, and a concrete example of, an integrated testlet, along with three distinct approaches for their deployment “in the wild”. Each approach provides the requisite answer-until-correct approach that is necessary for the integrated testlet to function as intended. The approach of choice will vary according to the instructor and their setting; we have extensive experience with the scratch card and online deployment, and expect to soon have within an optics textbook the third approach, with its goal of assisting students and instructors alike.

### ACKNOWLEDGMENTS

We are grateful to Matthew Romerein for masterfully blending the skills of art and science in producing Figure 1, and figures for the upcoming text. We acknowledge Aaron Slepkov at Trent University as the other co-creator of the integrated testlet, and with whom one of us (RS) has collaborated on much of the work discussed here for over a decade. We are grateful to Nicholas Gibbons, Stefanie Seaton and Tineke Bryson at Cambridge University Press and Assessment for providing seemingly endless encouragement as they patiently await the complete manuscript of the new edition of the upcoming textbook.

### REFERENCES

[1] Slepkov, A. D., "Integrated Testlets and the Immediate Feedback Assessment Technique", *American Journal of Physics*, 81, 782-791 (2013).  
 [2] Shiell, R. C., and Slepkov, A. D., "Integrated testlets: A new form of expert-student collaborative testing," *CELT*, **VIII**, 201-210 (2015).  
 [3] Epstein M. L., Lazarus A. D., Calvano T. B., Matthews K. A., Hendel R. A., Epstein B. B., Brosvic G. M., "Immediate feedback assessment technique promotes learning and corrects inaccurate first responses", *The Psychological Record*, 52(2), 187-201 (2002).  
 [4] Anderson, L.W. & Krathwohl, D.R. (Eds.). [A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives]. New York: Longman, (2001).  
 [5] Dihoff, R. E., G. M. Brosvic, M. L. Epstein, and M. J. Cook. "Provision of feedback during preparation for academic testing: Learning is enhanced by immediate but not delayed feedback". *The Psychological Record* 54(2), 207-231. (2004).  
 [6] de Carvalho Filho, M. K. "Assessing changes in performance and monitoring processes in individual and collaborative tests according to students' metacognitive skills," *European Journal of Cognitive Psychology*, 22(7), 1107-1136. (2010).  
 [7] Gilley, B. H. & Clarkston, B., "Collaborative testing: Evidence of learning in a controlled in-class study of undergraduate students." *Journal of College Science Teaching*, 43(3), 83-91, (2014).  
 [8] Slepkov, A. D., and Shiell, R. C., "Comparison of integrated testlet and constructed-response question formats," *Phys. Rev. ST Physics Ed. Research*, **10**, 020120, 1-15 (2014).  
 [9] DiBattista, D., Mitterer, J. O. & Gosse, L., "Acceptance by undergraduates of the immediate feedback assessment technique for multiple-choice testing", *Teaching in Higher Education* 9.1, 17-28 (2004).  
 [10] Pedrotti, S. J., Pedrotti, L. S., Pedrotti, L. M., [Introduction to Optics]. 3<sup>rd</sup> ed. Cambridge University Press (2007).