


inflexion



Exploring the Practices for and Effects of Open Book Assessment

Final Report

June 2021

inflexion

360 E 10th Ave Suite 300

Eugene, OR 97401

1.541.246.2600

Prepared by

Jessica N. Jacovidis, PhD

Kristine L. Chadwick, PhD

Paul Beach, PhD

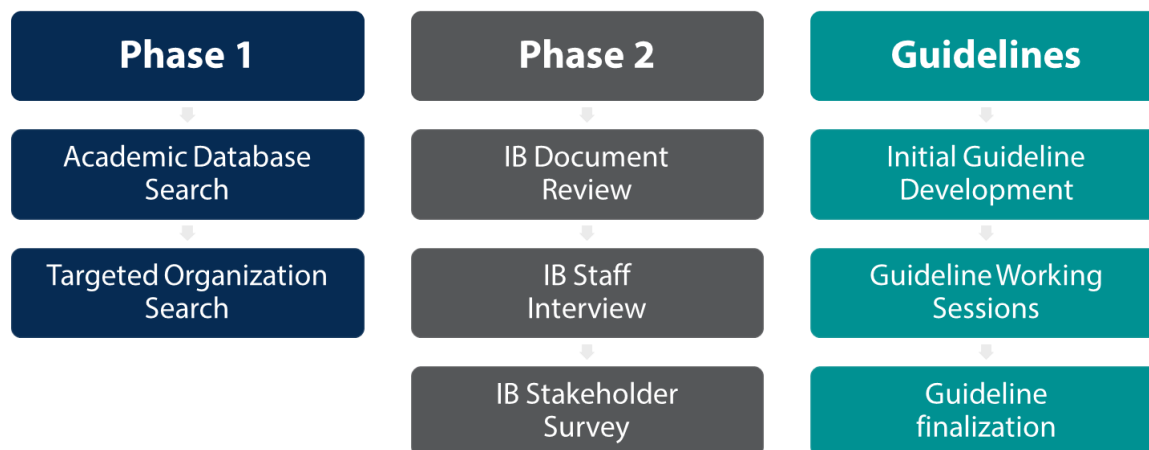
Executive Summary

In July 2020, the IB contracted with Inflexion to explore the effects and practices for open book examinations (OBEs). This study sought to inform IB’s understanding of the best practices in the use of resource aids from a practical perspective (e.g., students with specific educational and assessment needs, cultural differences, academic integrity, teacher assessment literacy and expertise, and school resources). Ultimately, the goal of the study was to develop guidelines specific to OBEs and use of reference aids on IB summative and classroom-based assessments anchored in the best available research evidence and contextualized to the IB. This study addressed the following five overarching research questions; some of the research questions had additional sub-questions:

1. What are the effects of the use of reference aids in high-stakes summative assessment, compared to assessment without these aids?
2. What are the effects of using reference aids in classroom assessment, compared to assessment without these aids?
3. What are good practices regarding assessments using reference aids for high-stakes summative assessment?
4. What are good practices regarding classroom assessment and examination preparation for assessments using reference aids and how can these be made inclusive for all students?
5. What are the essential teaching and assessment design skills, and school resources required to implement assessment using reference aids in the classroom?

Research Design and Methodology

Inflexion researchers employed a sequential, multiphase mixed methods design. In Phase 1, Inflexion conducted a comprehensive literature review by exploring academic databases and the websites of targeted organizations. In Phase 2, Inflexion performed a targeted review of IB documents, conducted interviews with key IB staff, and administered a survey of IB stakeholders. Taken together, the data from Phase 1 and Phase 2 were used to develop guidelines on OBEs and the use of reference aids within IB.



Phase 1 Findings

- **Limited Literature:** The literature base for OBEs is unsystematic, uncomprehensive, and splintered. Most of the research on OBEs is correlational, based on small-scale interventions in higher education, and lacks sufficient detail about the nature of the reference aids examined, with many studies simply noting students had access to reference materials.
- **Student Performance:** Results on student performance are mixed and highly dependent on context. OBEs and closed book exams (CBEs) are generally compared in terms of long-term retention; however, this type of comparison may not be appropriate, as it does not align with the purpose of OBEs.
- **Student Test Preparation:** Research is inconclusive on whether students prepare differently for OBEs and CBEs. In the absence of clear expectations or specific training, students will likely prepare less for OBEs than CBEs. Preparation may also vary by students' learning orientation.
- **Student Well-being:** Research on student anxiety and stress relies almost exclusively on student self-report responses to surveys or interviews. Generally, students report less anxiety and reduced stress from OBEs relative to CBEs; however, OBEs can increase anxiety and stress when students underestimate the difficulty of OBEs, are unfamiliar with the OBE format, or have little experience preparing and using reference aids.
- **Backwash Effects on Learning:** OBEs can provide powerful backwash effects on learning. For this to be true, a number of factors must be addressed, including designing assessments to target higher order thinking skills, providing clear expectations around the use of reference aids, and providing teachers and students training and practice with OBEs.
- **Assessment Objectives/Purpose:** Well-designed OBE questions require students to go beyond factual knowledge to demonstrate higher order thinking, problem solving, application, and analysis. Teaching, learning, and assessment format and item design must all be aligned to effectively target and measure higher order thinking on OBEs.
- **Exam and Item Design:** OBEs can be designed and administered in ways that do not sacrifice psychometric quality. However, particular attention should be given to balancing breadth and depth of content coverage, reliability, and the impact on test score validity.
- **Assessment literacy and design competency:** Very little research exists on the type of assessment design skills needed to effectively implement OBEs.
- **Academic Integrity:** Clear expectations, honor code agreements, and cheating detection software can be used to combat academic dishonesty on OBEs.
- **School resources:** Very little research exists on what type of school resources are needed to support and train teachers to effectively implement OBEs.

Phase 2 Findings

- **Receptivity:** Generally, receptiveness to OBEs and the use of reference aids is high. However, the purpose of the assessment is key. IB staff and stakeholders recognize that reference aids should only be used if it is the best way to assess the knowledge and skills that are being targeted by the assessment.

- **Alignment with IB Philosophies:** The use of reference aids aligns with existing IB policies and pedagogical practices. Further, IB is already using reference aids in some content areas and has some existing policies around specific reference aids.
- **Ease of Incorporation of OBEs:** Reference aids should be fairly easy to incorporate into IB's practice. However, there are too many unknowns at this point to fully understand what the implementation of OBEs would look like for IB.
- **Suitability of Reference Aids for Different Subjects:** Some reference aids may be better suited for certain subject areas. However, it is difficult to know *which* reference aids will be best suited for what subject areas.
- **Perceived Impact of OBEs:** IB staff and stakeholders generally felt introducing OBEs would have positive effects, including increasing assessment authenticity and the ability to assess higher order thinking instead of recall and memorization, and potentially, decreasing student stress.
- **Relevance of Recommendations from the Literature:** Even with the limited literature base, the recommendations from research seem to be relevant for IB's work.
- **Concerns Regarding OBEs:** There are some concerns about the effects of OBEs on academic honesty and equity. However, most other concerns from interviewees stemmed from a lack of clarity around what reference aids may be used and in what subject areas.
- **Assessment Skills for OBEs:** Assessment-related skills will be key for the success of OBEs. However, there is limited information on the teaching and assessment design skills that are necessary for teachers to use reference aids effectively.
- **Professional Learning:** Professional learning will be needed both internally and externally. There are too many unknowns regarding how IB might proceed with OBEs and the use of reference aids to know how this might affect schools and IB's work. Training sessions, templates, and examples will be invaluable.
- **Impact on IB Work:** Given there are still several decisions to be made, IB staff were unaware of how OBEs would impact their work.

Guideline Development and Working Session Conclusions

Inflexion researchers developed initial guidelines specific to OBEs and use of reference aids on IB summative and classroom-based assessments and held three working sessions with IB staff. The goal of the working sessions was to react to the initial guidelines, further articulate the suggested guidelines, and add relevant guidelines or identify areas of work for which guidelines still need to be developed. The development of actionable guidelines was dependent on the quality of information available in the literature, and on specific decision points for IB. Thus, the guidelines were not as specific and actionable as originally envisioned. The following guidelines present higher-level recommendations and considerations that will need to be further articulated as IB's understanding of OBEs and decisions around their use evolve. Further, Inflexion researchers provided a subjective categorization of each guideline based on the level of importance for successful implementation of OBEs. This coding was designed to facilitate the prioritization of the guidelines and assist IB staff in deciding where to focus their resources.

Importance	Guideline
Overarching and Policy-Related Guidelines	
Critical	1. Provide clear rules and guidelines and set clear expectations for OBEs and the use of reference aids.
Critical	2. Provide professional learning opportunities to all stakeholders.
Critical	3. Review and update policies around accessibility to cover the use of reference aids.
Curriculum Development Guidelines	
Critical	4. Allocate time to teach the skill of creating and/or using reference aids.
Moderate	5. Determine whether use of reference aids should be a skill targeted by the assessment.
Moderate	6. Provide guidance for teachers on preparing students for OBEs.
Assessment Design Guidelines	
Moderate	7. Use reference aids when they are the best choice for assessing the knowledge and skills of interest.
Critical	8. Leverage reference aids to assess higher order thinking skills.
Moderate	9. Ensure that time constraints are realistic given the number of questions and number of reference aids.
Assessment Implementation Guidelines	
Moderate	10. Provide explicit instructions for checking reference aids and exam invigilation.
Moderate	11. Review and update scoring models as needed to account for OBEs and the use of reference aids.
Critical	12. Ensure exams allowing reference aids still meet adequate assessment quality standards.
Professional Learning Guidelines	
Critical	13. Provide learning opportunities on how to write assessment questions that optimize the use of reference aids.
Moderate	14. Provide learning opportunities on how to create and/or use reference aids.
Moderate	15. Provide learning opportunities on checking reference aids and exam invigilation.

Areas for Future Research

Given the lack of information in the literature, there are a number of future research studies that could be conducted to better understand how OBEs and the use of reference aids would work in IB's context.

- If the IB is interested in incorporating OBEs, additional empirical evidence is needed directly comparing key outcomes of interest for open and closed book exams. This information would not only benefit the IB but would contribute significantly to the literature base on open book assessments.
- IB should consider engaging IB teachers as subject matter experts to explore the appropriateness of specific reference aids in content areas for which there is interest in implementing OBEs, and to suggest which existing assessments might benefit from an OBE approach. Once information is gathered from subject matter experts, theoretical findings can be tested empirically. Further, IB could partner with pilot schools to implement plan-do-act (PDSA) cycles to sufficiently trial the use of reference aids prior to scaling up.
- IB should consider a study with schools to examine teacher assessment design skills; access to, familiarity with, and use of reference aids; and school resources needed for successful implementation of OBEs. This study could include the guidance that schools and teachers would like to have regarding implementing OBEs in the classroom, and the format in which schools would prefer to receive the guidance and training.
- IB should consider an environmental scan to identify and learn from organizations that have already been successful at implementing OBEs and determine the best way to learn from their experiences. Further, though exemplars exist, IB will need to consider how to tailor these examples to its context.
- IB will need explore the impact of OBEs on its specific tests and in its specific contexts. IB may consider the process that is used to validate exams and/or conduct a series of pilot studies to explore the impact before scaling up the use of OBEs. IB may also want to consider, a priori, an acceptable psychometric benchmark that OBEs must meet.
- IB staff should consider what they would hope to learn about OBEs in paper-pencil and digital assessment, access and inclusion, and testing accommodations and consider how to best capture that knowledge.

Conclusion

Ultimately, this study sought to develop actionable guidelines that could be used as IB expanded OBEs and the use of reference aids within the organization. However, after extensive searching, it became clear that this was not possible. We are coming away with many more questions than answers. There is much more to learn about OBEs before IB can implement OBEs on a large scale. As such, general guidelines and considerations were provided to help IB staff frame their thinking about OBEs and to inform the next steps of the OBE journey.

