Evaluation Strategy for General Surgery Education and Training Program (GSET) and Curriculum

Background

In 2022, General Surgery will introduce a revised surgical training program that is based upon a hybrid of competency-based principles and time-based training entitled the General Surgery Education and Training (GSET) program. The revised program will consist of:

- A new Curriculum with clear competencies and milestones across three levels of training.
- Defined milestones required to be met at key stages of training before being permitted to progress to the next stage.
- Assessment of those competencies through standardised Entrustable Professional Activities (EPAs) and Procedure Based Activities (PBAs).
- Clearly defined In-Training Assessment forms aligned to the Curriculum milestones
- Shifting of the Fellowship Examination from final year to fourth year.
- Milestones to be met prior to being approved to present for the Fellowship Examination.

The GSET program, indeed any educational program, fundamentally is about change – change in skills, knowledge, and behaviour.[1, 2] An evaluation strategy is therefore required to determine if the new program, Curriculum, and intervention has achieved the desired change and outcomes.[1-3]

Evaluation Purpose

The purpose of this evaluation strategy is to examine the effects of the GSET program and Curriculum in terms of:

- Ability to assess Trainees against the Curriculum milestones
- Ability of the assessments and milestones to guide feedback
- Ability to adequately progress Trainees
- Changes in behaviours based on clear competencies outlined in the Curriculum
- Increased learning and knowledge based on clearly outlined Curriculum competencies and milestones

The results of the evaluation will also be used to assist in making decisions pertaining to improvements to the GSET program and Curriculum.

Stakeholders

It is important to identify the stakeholders that will be affected by this evaluation strategy to ensure that the information provided is able to be utilised.[3] The scope of stakeholders has been broadened and the following model as described by Battista has been applied:

- Those who have the authority to make decisions about the program (Authority)
- Those responsible for delivery of the program (Delivery)
- Those who benefit from the program (Beneficiaries)
- Those potentially disadvantaged from the program (Disadvantaged)

A fifth group has been added to include those who have an interest in the outcomes of the evaluation but are not involved in either the decision making or the delivery. This group are those involved in surgical education and training who may have input into broadly developed frameworks or who have an interest in the methods utilised in the GSET program, Curriculum, and evaluation.

Authority	Delivery	Beneficiaries	Disadvantaged	Interest Groups
•Australian Board in General Surgery •RACS Board of Surgical Training •RACS Education Board •GSET Implementation Committee •Australian Medical Council •Curriculum Working Party	•Australian Training Committees •Supervisorss of Training •Trainers and Assessors •Hospitals with General Surgery Trainees •General Surgeons Australia	•GSET Trainees •Patients	•SET Trainees	•RACS Research and Education Department •RACS Rural Surgery Section •New Zealand Board in General Surgery •Other interested RACS groups •General Surgery subspecialty societies

Evaluation Model

A logic model has been utilised to construct the evaluation strategy as it is an outcomes-based approach that focusses on changes that have occurred.[1, 4] A key component of logic models is that it also allows for short, medium, and long-term outcomes to be evaluated. Hence this model is able to track outcomes, evaluate those, and if required make changes to both the Curriculum and GSET program based on collected information and in a timely manner. The short and medium term outcomes specifically allows for an understanding of which elements are functioning and which are not and therefore be able to make informed adjustments mid-way that enhance the probability of achieving the long-term outcomes.[2, 5]
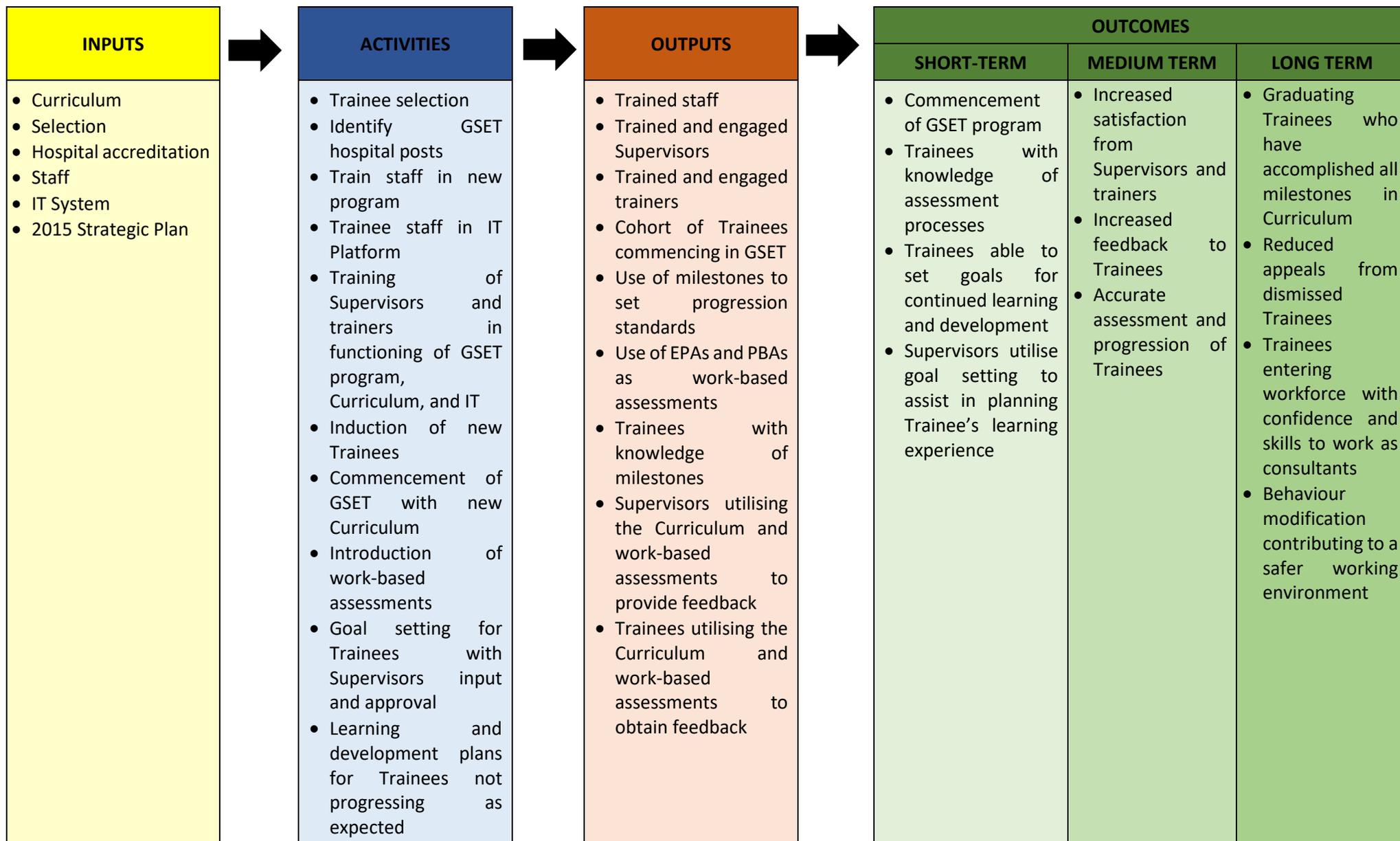
The Logic Model for Evaluation of the GSET Program and Curriculum

The logic model for the evaluation of the GSET Program and Curriculum includes the following aspects:[1, 2, 5-8]:

Component	Description
Inputs	The resources, materials, processes, or technology that are available to the training program.
Activities	Details the innovations, strategies, or changes planned. Generally, the activities are listed in order of implementation.
Outputs	The outputs indicate that the activity has been undertaken, that is the product of the activity.
Outcomes	The outcomes have been grouped into Short-term, Medium-term, and Long-term. The outcomes stipulate the intended changes as a result of the activities in the program. The long-term outcomes may also be viewed as impacts that have occurred due to the interventions included in the training program.

The inputs, activities, and outputs are referred to as the program structure and are under the control of the program managers and staff.[2] In the case of General Surgery this would be the Australian Board in General Surgery (AuBiGS) and the staff at General Surgeons Australia (GSA) who manage the GSET program and Curriculum. The outcomes identify the changes or benefits to those undertaking the program, the organisation, or others who may be affected.[2, 5] For this evaluation, this would be the Trainees, Supervisors, trainers, and the community.

INPUTS	ACTIVITIES	OUTPUTS	OUTCOMES		
			SHORT-TERM	**MEDIUM TERM**	**LONG TERM**
• Curriculum • Selection • Hospital accreditation • Staff • IT System • 2015 Strategic Plan	• Trainee selection • Identify GSET hospital posts • Train staff in new program • Trainee staff in IT Platform • Training of Supervisors and trainers in functioning of GSET program, Curriculum, and IT • Induction of new Trainees • Commencement of GSET with new Curriculum • Introduction of work-based assessments • Goal setting for Trainees with Supervisors input and approval • Learning and development plans for Trainees not progressing as expected	• Trained staff • Trained and engaged Supervisors • Trained and engaged trainers • Cohort of Trainees commencing in GSET • Use of milestones to set progression standards • Use of EPAs and PBAs as work-based assessments • Trainees with knowledge of milestones • Supervisors utilising the Curriculum and work-based assessments to provide feedback • Trainees utilising the Curriculum and work-based assessments to obtain feedback	• Commencement of GSET program • Trainees with knowledge of assessment processes • Trainees able to set goals for continued learning and development • Supervisors utilise goal setting to assist in planning Trainee's learning experience	• Increased satisfaction from Supervisors and trainers • Increased feedback to Trainees • Accurate assessment and progression of Trainees	• Graduating Trainees who have accomplished all milestones in Curriculum • Reduced appeals from dismissed Trainees • Trainees entering workforce with confidence and skills to work as consultants • Behaviour modification contributing to a safer working environment

Context

- The GSET program occurs in the work environment of accredited training hospitals.
- The GSET program is accredited by the Australian Medical Council (AMC) through the Royal Australasian College of Surgeons (RACS). The RACS and therefore GSET Program is obliged to comply with the AMC conditions and recommendations to ensure continuing accreditation.
- The changing nature of the working environment in terms of the need for greater work-life balance and flexibility.

Explanatory Notes

- The short-term goals, whilst on face value seem simplistic, have been included because the GSET program has been delayed by one year due to COVID19 therefore it is a vital goal.
- Many of the short and medium term goals will continue long term. They have been placed in the short and medium term as that is when the intended outcomes would expected.
- There will be a rolled-out approach to training of Supervisors. As only Trainees who commence in 2022 will be enrolled in the GSET Program, not all hospitals will have a GSET Trainee in their accredited posts.
- Training material and online learning modules will be developed to train the trainers who are more than 1000 across Australia.

Assumptions

The following assumptions have been made:

- 100% of hospital Supervisors who have a GSET Trainee at their hospital in 2022 will attend the training sessions that will introduce them to the work-based assessments, Curriculum, and GSET Program in the year prior to the introduction of GSET (2021).
- That remaining Supervisors will attend a training session during 2022.
- That all Trainees will attend the induction program prior to commencing training.

External factors

The following external factors will need to be monitored:

- The ongoing effects of the COVID19 pandemic on the ability to convene training sessions for Supervisors in 2021 and induction for Trainees
- Delays in the development of the IT platform
- Changes to the AMC conditions and recommendations
- Political agenda from external parties
- Social issues that arise and may need to be addressed

Evaluation Questions, Indicators and Methods

A logic model assists in identifying specific evaluation questions that should be investigated and the most appropriate methods to use taking into account both resources and the stakeholders.[7] The below table includes:

- various questions that will be required to be answered as part of the evaluation linked to the overall outcomes identified in the logic model
- anticipated outcome in more detail
- indicator as a measurement of success

It is important to note that an indicator is different to the outcome. Whilst the outcome describes the change, the indicator is a measurement or marker that demonstrates if the outcome has occurred.[8] The following methods will be used to collect data:

- Data analysis
- Questionnaires
- Snowball reviews

Focus groups have purposely not been included as it is extremely difficult to gather the required number of participants to run a meaningful focus group. Furthermore, the resources required to undertake this would not be available to the AuBiGS. Following the results of the questionnaire, there may be the possibility of convening targeted focus groups for example with Trainees who have not been able to progress through the program, or with Supervisors who have managed Trainees in difficulty. This may be a better use of the resources in terms of targeting a specific area. However, this has not been included in the evaluation at this stage. Once the GSET program launches and evaluation commences, the logic model and evaluation questions will expand, becoming a dynamic document. This is also acknowledging the concept of the evaluation results being able to provide insights, opportunities for further evaluation, and revision of the program.

Notes:

- *The indicators and methods for the short and medium term goals will continue to be undertaken long term. That is the indicators and methods will be repeated either each term or yearly as required to ensure that continuous data is collected across many cohorts of Trainees.*
- *The Curriculum is reviewed every three years to ensure accuracy of procedures and skills. The evaluation results would therefore feed into this review.*
- *Indicators for logbook numbers have not been included due to the ongoing work of the Logbook Working Party*

Outcome from Logic Model	Specific Evaluation Question	Anticipated Outcome	Indicator	Method
Short Term Outcomes				
• Commencement of GSET program	• Did the GSET program roll out in 2022?	• Selection of cohort of Trainees to commence on GSET program in 2022	• 70-80 candidates commence training in GSET in 2022	• There is no specific method to evaluate this, as this will be evident from the selection process
	• Was the IT system available for training?	• IT (soft) system available in October 2021 to commence training of Supervisors, new cohort of Trainees, and staff	• Training workshops have been able to be undertaken with the IT (soft) system available	• There is no specific method to evaluate this, as this will be evident from the material available in the lesson plan for the workshops
	• Was the IT system ready and available for 2022?	• IT System available from February 2022 (commencement of Term 1 – 2022)	• 100% of Trainees able to log into the system in February to commence entering goals	• Data analysis from IT system looking at number of Trainees who logged into the system and entered goals for first term on program
			• 100% of Supervisors and trainers able to log into the system and undertake assessments	• Data analysis from IT system looking at number of assessments completed per term
			• Staff with access to IT Platform	• N/A
• Trainees with knowledge of assessment processes	• Do Trainees have the necessary information, understanding, and tools to commence completing EPAs and PBAs?	• Trainees have confidence to commence undertaking EPAs and PBAs during first term in training	• 100% of Trainees complete a minimum of 5 EPAs and 5 PBAs by end of first term of training	• Data analysis from IT System
			• 95% of Trainees complete the minimum number of EPAs and PBAs by end of second term of training	• Data analysis from IT System and End of Term Assessment ratings
				• Preliminary review of EPAs and PBA completed at the end of the first six months of GSET

Outcome from Logic Model	Specific Evaluation Question	Anticipated Outcome	Indicator	Method
			• Trainees have a clear understanding of EPAs and PBAs and the process required to complete them	• Questionnaire at the end of the first year of the program looking specifically at knowledge of EPA and PBA process, IT system functionality, the intent of the EPAs and PBAs, and understanding of what aspects the EPAs and PBAs are assessing (including linking to the Curriculum milestones)
• Trainees able to set goals for continued learning and development	• Have Trainees been able to appropriately set goals for their continued learning?	• Trainees able to set specific goals related to the Curriculum, EPAs, and PBAs each term	• 100% of Trainees have completed the online goal setting within three weeks of the commencement of term	• Data analysis from IT System
	• Has the setting of goals assisted Trainees in having conversations with Supervisors about learning expectations for the term?	• Trainees able to have conversation with Supervisors	• Trainees report satisfaction with goal setting process and increased tailoring of experience	• Data analysis from IT System • Trainee Questionnaire to be able to collect qualitative information on experience in goal setting and ability to tailor experience in rotation based on those goals
• Supervisors utilise goal setting to assist in planning Trainee's learning experience	• Has the setting of goals assisted Supervisors in adapting the learning experience for Trainees?	• Supervisors increased understanding of goals Trainees wish to accomplish during the term and ability to provide those opportunities.	• 100% of Supervisors acknowledge goals set by Trainees • Supervisors report increased understanding of Trainees needs and ability to tailor experience in rotation	• Data analysis from IT System • Supervisor Questionnaire to be able to collect qualitative information on experience in goal setting and ability to tailor experience in rotation based on those goals
Medium Term Outcomes				

Outcome from Logic Model	Specific Evaluation Question	Anticipated Outcome	Indicator	Method
• Increased satisfaction from Supervisors and trainers	• Did the descriptions of the EPAs, PBAs, and Curriculum milestones provide direction for learning?	• Trainees and Supervisors have a clearer understanding of the standards required for each GSET level	• Reduced requests for reconsideration of assessment outcomes	• Data analysis of number of reconsiderations requested compared to those in the current SET program
			• Trainees and Supervisors indicate increase usage and reference of the Curriculum milestones (that is the milestones are being integrated in everyday practice)	• Questionnaire at end of third and fifth year of the program to both Trainees and Supervisors and Trainers
	• Has the satisfaction level of the Trainees and Supervisors in terms of the GSET program experience (learning, feedback, assessment, processes, IT system) increased from that of Trainees and Supervisors on the SET program?	• Trainees and Supervisors report increased satisfaction level with the GSET program	• Reduction in problems with the program identified	• Questionnaire at end of third and fifth year of the program to both Trainees and Supervisors.
• Increased feedback to Trainees	• Did the use of EPAs, PBAs, and Curriculum milestones allow for the provision of meaningful feedback to Trainees?	• Increased feedback comments on assessments	• Reduced requests for reconsideration of assessments based on lack of feedback provided.	• Data analysis of number of reconsiderations requested compared to those in the current SET program
			• 90% of all assessments include feedback comments from Supervisors	• Data analysis of IT System
		• Supervisors report ability to provide targeted feedback	• Supervisors report increased ability to provide feedback	• Data analysis of IT System • Questionnaire at end of first, third and fifth year of the program to Supervisors and Trainers
			• Each Trainee has a minimum of three Feedback Notes provided by Supervisors per term.	

Outcome from Logic Model	Specific Evaluation Question	Anticipated Outcome	Indicator	Method
	• Did the EPAs, PBAs, and Curriculum milestones provide a basis for Trainees to seek feedback?	• Trainees increase ability to request specific feedback	• Trainees report increased ability and comfort in seeking feedback	• Questionnaire at end of third and fifth year of the program to Trainees.
• Accurate assessment and progression of Trainees	• Did the PBAs assist in determining readiness for the Fellowship Exam both in terms of numbers required and feedback given?	• Trainees who are not ready to present for the Examination are not approved	• Decrease in number of failed attempts	• Data analysis of results
			• Increase in number of Trainees not attempting the Examination at the first available opportunity	• Data analysis of registrations • Questionnaire of Trainees exploring reasons why they did not present for the Examination
		• Supervisors report increased confidence in not approving Trainees to present for the Examination	• Increase in Supervisors confidence and ability to not approve Trainees who are not ready	• Questionnaire of Supervisors in fourth year of program
	• Were the milestones appropriate for the various GSET levels? • Did the Curriculum milestones allow for understanding of what was expected at the various GSET level?	• Milestones have been appropriately identified for the various GSET levels	• 90% of milestones are deemed at the appropriate level	• Questionnaire and Snowball Review with Supervisors and Trainees undertaken separately at end of first, third, and fifth year.
	• Did the EPAs, PBAs and Curriculum milestones provide greater clarity and ability to identify performance issues?	• Performance issues are identified early, and Learning and Development Plans (LDP) instigated to remedy issues	• Reduction in number of Trainees who are only identified as having performance issues 4-5 years into the program	• Data analysis comparing to results in current SET program
			• 99% of Trainees on LDPs successfully obtain Fellowship	
		• Trainees can better identify their own performance issues/areas for improvement	• Increased comments on performance from Trainees	• Data analysis of ratings and comments by both Trainees and Supervisors

Outcome from Logic Model	Specific Evaluation Question	Anticipated Outcome	Indicator	Method
		• Supervisors are more comfortable in articulating performance issues/areas for improvement	• Increased comments on performance from Supervisors	• Questionnaire and/or Snowball Review
	• Did the EPAs, PBAs, and Curriculum milestones identify Trainees who required extended learning?	• Trainees who require more time are identified earlier (by GSET3)	• Reduction in number of Trainees who require extended learning at the end of training	• Data analysis comparing number of Trainees in current SET program who required extended learning at the end of training
	• Did the EPAs, PBAs, and Curriculum Milestones assist in determining the assessment outcome for the rotation?	• Better informed assessment decisions	• Reduced requests for reconsideration of assessments based on lack of feedback provided	• Data analysis of number of reconsiderations requested compared to those in the current SET program • Questionnaire and/or Snowball Review
Long Term Outcomes				
• Graduating Trainees who have accomplished all milestones in Curriculum	• Have Trainees been able to accomplish all the milestones in the Curriculum?	• Trainees graduate with competencies identified as being necessary to be a General Surgeon	• 99% of Trainees who commence on the GSET program are awarded Fellowship	• Data analysis of Trainees commencing and graduating
	• Have Trainees been able to complete all courses and educational activities in the program?	• Trainees complete all mandatory courses by GSET3	• No Trainee is rated as Not progressing due to not completing courses	• Ongoing data analysis of course completion, Trainee's Day attendance, and Research Requirement point acquisition
		• Trainees attend minimum four Trainee's Day by GSET5	• 100% Trainees have completed requirement by GSET5	

	• Have Trainees been able to complete their Research Requirement	• Trainees complete Research Requirement by GSET5		• Analysis of how Research Requirement is being attained by:
				  o Gender
				  o Type

Outcome from Logic Model	Specific Evaluation Question	Anticipated Outcome	Indicator	Method
• Reduced appeals from dismissed Trainees	• Have the number of appeals due to issues with natural justice and procedural fairness reduced as a result of utilising work-based assessments and Curriculum milestones?	• Appeals made by Trainees are not upheld (that is the appeals process determines there were no issues with natural justice and procedural fairness)	• 100% of appeals are not upheld	• Data analysis comparing appeals in current SET program
• Trainees entering workforce with confidence and skills to work as consultants	• Did the PBAs enable Trainees to determine their own level of competency in a procedure both in terms of description and number required? • Did the EPAs enable Trainees to determine their own level of competency in the skill both in terms of description and number required?	• Trainees report increase in skill level over time	• 95% of Trainees meet the minimum requirements for EPAs and PBAs (the 95% is considering the possibility of Trainees not having access and the rate of Trainees in difficulty which is currently 10% per term)	• Data analysis
			• All Trainees report an increase in skills and knowledge from GSET1 to GSET3 to GSET5	• Undertake pre-knowledge questionnaire at the beginning of the training program in GSET1 pertaining to the Trainee's level of knowledge and confidence in performing procedure and EPA tasks. Repeat the questionnaire at GSET3 and at GSET5 to ascertain increase and changes in knowledge and confidence.
			• Trainee report increased confidence in skills and knowledge as consultants	• Follow-up questionnaire 3-5 years post Fellowship. *(Note: It would also be valuable to survey Trainees in the current SET program and be able to compare results)*

Outcome from Logic Model	Specific Evaluation Question	Anticipated Outcome	Indicator	Method
			• No Trainees are reported as regressing in skills and knowledge	• Data analysis
	• Did the EPAs, PBAs, and Curriculum milestones enable Trainees to undertake deeper reflection into their skill level and learning trajectory?	• Trainees ability to undertake deeper reflection increases progressively over training	• 90% of EPAs, PBAs, and In-Training Assessments include comments from Trainees.	• Data analysis of comments
		• Trainees continue to undertake deep reflection in their independent practice following Fellowship	• Trainees report deep reflection as standard component of their continuing professional development	• Follow-up questionnaire 3-5 years post Fellowship. *(Note: It would also be valuable to survey Trainees in the current SET program and be able to compare results)*
• Behaviour modification contributing to a safer working environment	• Did the Curriculum milestones for the professionalism skills allow for understanding of what was expected in the workplace?	• Greater understanding of professionalism skills and their impact on the working environment	• Trainees report greater understanding of how behaviours and attitudes affect the working environment	• Undertake pre-training questionnaire in GSET1 pertaining to the Trainees attitudes and behaviours focusing on the Professionalism Skills. Questions should be structured around the behaviours outlined in the Curriculum. Repeat the questionnaire at GSET3 and at GSET5 to ascertain increase and changes attitudes and behaviours.
	• What behaviours or attitudes changed?		• Trainees report changes in previously held attitudes and behaviours	

Limitations

A limitation of logic models is that they can often only search for evidence that the program is working as intended and overlook unintentional outcomes.[1, 4] Furthermore, it is difficult to predict all the factors that may influence not only the introduction of the GSET program but intended outcomes. It may also be the case that the outcomes are not because of the activities undertaken but of other interventions occurring outside of the GSET program and Curriculum. A logic model is also resource intensive however if the GSET program and Curriculum are to be successful, it is important that the necessary resources are provided both in terms of staff and budget.

Knowing these limitations, the purpose of investigating the outcomes will be fulfilled by using the logic model. Focus must occur in the development of the measurement tools to ensure the correct questions are being asked that elicit information from participants about the GSET Program and curriculum itself, whilst exploring if there were other factors that may have contributed.

Reporting and Dissemination

It is important to ensure that the results of the evaluation are published and presented appropriately. The results of an evaluation strategy should be used to stimulate improvements and change.[9] A logic model assists in reporting to the stakeholders the issues the revised program focussed on and how these were addressed.[2] The results must therefore be published and presented widely and particularly to the decision makers.

It is equally as important to disseminate the results to the Trainees and Supervisors. The results should be presented to both groups in an open forum to allow for discussion and understanding of not only the results but the evaluation process that was undertaken. Indeed, it should be noted that the evaluation strategy should be communicated widely to these groups prior to the commencement of the GSET program. This will allow both groups to be aware that an evaluation strategy is in place, that their contribution is vital, and that they will be provided with the results. It will also assist in instilling confidence that the GSET Program and Curriculum is being thoroughly evaluated.

The reporting of the evaluation must include not only the objectives of the evaluation, together with the methods used and results, but suggested areas for improvement and further development. The evaluation strategy should be able to provide recommendations to the AuBiGS on whether the outcomes have been met or if further developments are to be explored. Furthermore, the reporting should also include additional areas for evaluation.

Monica Carrarini

GSA – Director, Education and Training

References

1. Frye, A.W. and P.A. Hemmer, *Program evaluation models and related theories: AMEE Guide No. 67.* Medical Teacher, 2012. **34**(5): p. e288-e299.

2. McLaughlin, J.A. and G.B. Jordan, *Using Logic Models*, in *Handbook of Practical Program Evaluation*. 2015. p. 62-87.

3. Battista, A., et al., *Demystifying Program Evaluation in Surgical Education*, in *Advancing Surgical Education - Theory, Evidence and Practice*. 2019, Springer.

4. Rogers, P.J., *Logic Model*, in *Encyclopedia of Evaluation*, S. Mathison, Editor. 2005: Thousand Oaks, California.

5. Kneale, D., J. Thomas, and K. Harris, *Developing and Optimising the Use of Logic Models in Systematic Reviews: Exploring Practice and Good Practice in the Use of Programme Theory in Reviews.* PloS one, 2015. **10**(11): p. e0142187.

6. Lovato, C. and D. Wall, *Programme evaluation: Improving practice, influencing policy and decision-making*, in *Understanding Medical Education: Evidence, Theory and Practice*. 2014, Oxford Wiley Blackwell. p. 385-400.

7. Lawton, B., et al., *Logic Models: A Tool for Designing and Monitoring Program Evaluations*. 2014, Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Pacific.

8. Smart, J., *Planning an evaluation: STEP BY STEP*. 2020: https://aifs.gov.au/cfca/expert-panel-project/program-planning-evaluation-guide/plan-your-program-or-service/how-develop-program-logic-planning-and-evaluation.

9. Kern, D.E., P.A. Thomas, and M.T. Hughes, *Curriculum development for medical education : a six-step approach*. 2009: Baltimore, Md. : Johns Hopkins University Press, c2009. 2nd ed.