

1/31/2016

Inspire>Aspire Final Report

Planning Project and Process
Evaluation



John
Templeton
Foundation



MONTCLAIR STATE
UNIVERSITY

Research on Evaluation and
Developmental Systems Science Lab



Research on Evaluation and Developmental Systems
Science (REDSS) Lab

MONTCLAIR STATE UNIVERSITY

JENNIFER BROWN URBAN & MIRIAM LINVER (PRINCIPAL INVESTIGATORS)

Table of Contents

EXECUTIVE SUMMARY.....	p. 4
Program Description.....	p. 4
Big Questions and Study Aims.....	p. 5
Study Design and Participants.....	p. 6
Findings.....	p. 7
Aim 1: Assess variations in program implementation.....	p. 7
Aim 2: Pilot test quantitative measures of key character outcomes.....	p. 8
Aim 3: Validate the theory of change using qualitative interviews of youth.....	p. 8
Aim 4: Assess the relationship between variations in program implementation (e.g., differences in dosage) and poster quality.....	p. 9
Big Question 1: Can reflecting on and writing about the virtues espoused in the Laws of Life improve adolescent character development?.....	p. 9
Big Question 2: What is the overarching theory of change for Inspire>Aspire: Global Citizens in the Making?.....	p. 9
Big Question 3: How is the theory of change best assessed?.....	p. 10
Big Question 4: Does preliminary evidence support the theory of change?... 	p. 10
Recommendations.....	p. 10
Create a Foundation for Success.....	p. 10
Provide Best Practices to Teachers.....	p. 11
Cultivate More Character.....	p. 12
Expand Guidelines and Resources.....	p. 12
Create Additional Activities.....	p. 13
Extend Access to Younger Participants.....	p. 13
Revisit the Programmatic Theory of Change.....	p. 13
INTRODUCTION.....	p. 14
Program Activities.....	p. 15
GOALS OF PROJECT.....	p. 15
BIG QUESTION 2: THE THEORY OF CHANGE.....	p. 16
Evolutionary Evaluation and the Systems Evaluation Protocol.....	p. 16
Stakeholder Analysis.....	p. 18
Boundary Analysis.....	p. 20
Lifecycle Analysis.....	p. 21
Logic and Pathway Model.....	p. 23
Evaluation Scope.....	p. 27
Program-System Links.....	p. 28
Results of Articulating the Inspire>Aspire Theory of Change: ‘Aha!’ Moments.....	p. 29
‘Aha!’ Moment 1.....	p. 30
‘Aha!’ Moment 2.....	p. 30
‘Aha!’ Moment 3.....	p. 31
Finalizing the Evaluation Plan.....	p. 31
Reflection and Synthesis.....	p. 31

Evaluation Plan Creation.....	p. 31
BIG QUESTIONS 1, 3, & 4: THE PROCESS AND PILOT OUTCOME	
EVALUATION.....	p. 31
Design and Procedure.....	p. 32
Sample.....	p. 33
Teachers.....	p. 33
Pupils.....	p. 34
Measures.....	p. 34
Poster Quality.....	p. 34
Goal Setting/Selection.....	p. 35
Future Mindedness.....	p. 36
Sense of Purpose.....	p. 37
Positive Youth Development (PYD).....	p. 37
Preliminary Analyses.....	p. 39
SOC.....	p. 39
Sense of Purpose.....	p. 40
Aspiration Index.....	p. 41
Future Orientation.....	p. 41
PYD.....	p. 41
Results.....	p. 42
Variation in Program Implementation (Aim 1).....	p. 43
Qualitative Analysis.....	p. 45
Poster Quality (Aims 2 and 4).....	p. 54
Pupil Poster Quality Profile Analysis.....	p. 60
Pupil Outcomes: Comparing Wave 1 and Wave 2 Surveys (Aim 2).....	p. 62
Pupil Outcomes: Comparing Wave 1 and Wave 2 Interviews (Aim 3).....	p. 72
Goal Setting and Future Mindedness.....	p. 73
Goal Setting: Connection – Sense of Purpose.....	p. 76
The 5 Cs.....	p. 79
Self-knowledge and Self-care.....	p. 86
Self-Awareness.....	p. 87
Being Inspired.....	p. 89
Knowledge of Values, Virtues and Qualities.....	p. 90
Relationship between Teacher Characteristics and Poster Quality (Aim 4)..	p. 92
CONCLUSIONS.....	p. 100
Aim 1: Assess variations in program implementation.....	p. 101
Aim 2: Pilot test quantitative measures of key character outcomes.....	p. 101
Aim 3: Validate the theory of change using qualitative interviews of youth..	p. 102
Aim 4: Assess the relationship between variation in program	
implementation (e.g., differences in dosage) and poster quality.....	p. 102
Big Question 1: Can reflecting on and writing about the virtues espoused in	
the Laws of Life improve adolescent character development?.....	p. 102
Big Question 2: What is the overarching theory of change for	
Inspire>Aspire: Global Citizens in the Making?.....	p. 103
Big Question 3: How is the theory of change best assessed?.....	p. 103

Big Question 4: Does preliminary evidence support the theory of change?...	p. 103
Recommendations.....	p. 103
Create a Foundation for Success.....	p. 103
Provide Best Practices to Teachers.....	p. 104
Cultivate More Character.....	p. 105
Expand Guidelines and Resources.....	p. 105
Create Additional Activities.....	p. 106
Extend Access to Younger Participants.....	p. 106
Revisit the Programmatic Theory of Change.....	p. 107
REFERENCES.....	P. 108

EXECUTIVE SUMMARY

The picture we get of today's young people often seems bleak: they are disengaged from civic life, they engage in risky sexual behavior, they display confused moral reasoning, and have materialistic goals (Smith, Christoffersen, Davidson, & Herzog, 2011). Youth today are at a crossroads. Despite the challenges they face, they have the potential to be active, engaged, contributing members of civil society if they are provided with appropriate and much needed support from peers, families, and schools. Character education programs that specifically focus on helping youth to identify and reflect on positive personal values and transform those values into meaningful, sustained action hold tremendous promise for avoiding the pitfalls of adolescence and young adulthood and producing a generation of youth who are able to demonstrate moral fortitude and thrive in today's complex, global world. Inspire>Aspire: Global Citizens in the Making is one such promising character education program.

Program Description

Inspire>Aspire: Global Citizens in the Making was developed in the United Kingdom by Character Scotland. The program has been implemented in over **60** countries and has reached around **100,000** youth ages 10-18. Using a unique poster template and web-based resources, students engage in a process of self-discovery where they reflect on: their strengths and areas in need of improvement, who and what inspires them, and who they want to become and what they want to achieve in life. The poster process can be summarized as follows: 1) Critical self-reflection on one's own virtues and personal qualities which is hypothesized to lead to enhanced self-awareness and empathy – each person has a list of their strong points and what they need to work on; 2) Relating these virtues and qualities to an inspirational figure of real substance and how they helped this person to outstanding achievement – this translates theoretical concepts into

practical living; 3) Sourcing and commenting on inspirational quotations as maxims for life and living; 4) Choosing an inspirational story that has an important message for life; 5) Translating this inspiration into aspiration by focusing on key questions relating to what kind of person one wants to become and what one wants to achieve and contribute to others; and, 6) Soliciting feedback on what the young person has learned, which helps them reflect on the value of the whole process.

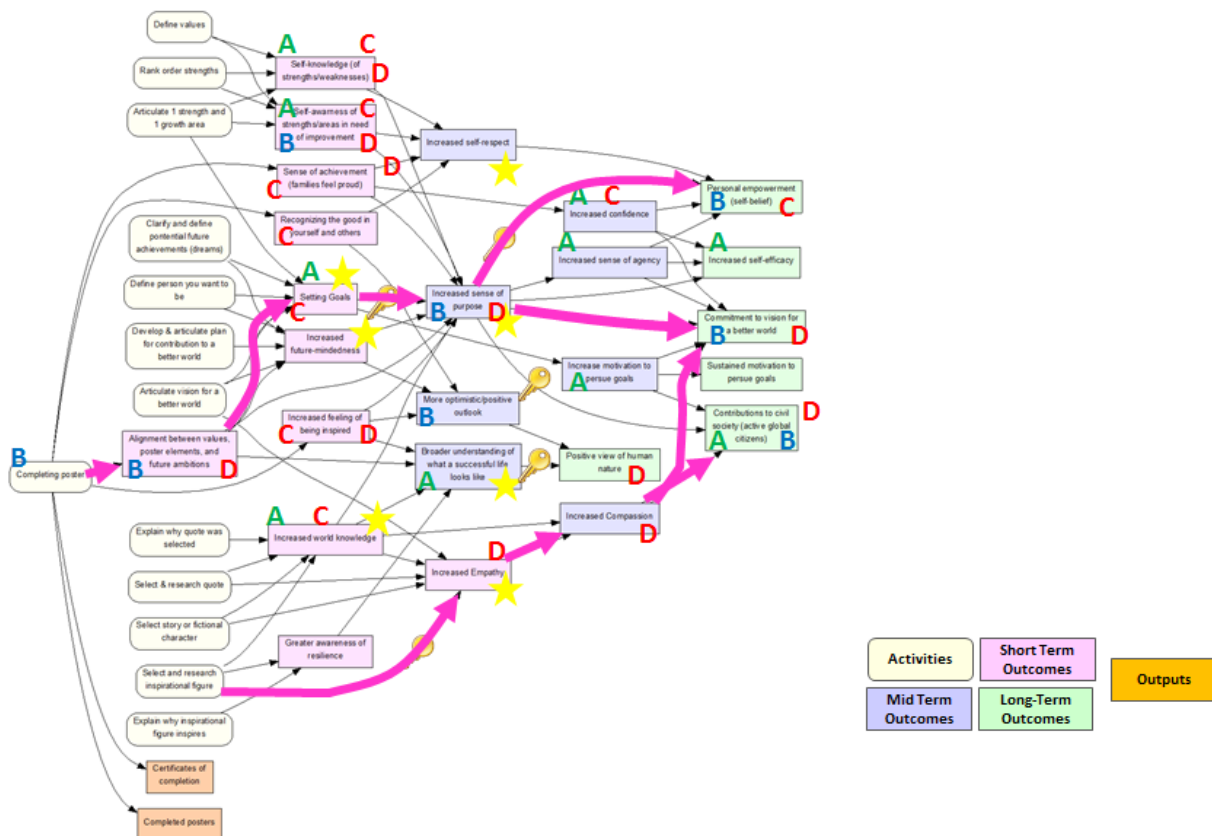
This project is a collaboration between Character Scotland, who developed and distributes Inspire>Aspire, and the Research on Evaluation and Developmental Systems Science (REDSS) Lab at Montclair State University, who conducted the planning, process, and pilot outcome evaluation of Inspire>Aspire. Inspire>Aspire and the research and evaluation presented in this report have been supported by grants from the John Templeton Foundation.

Big Questions and Study Aims

The primary **Big Question** this planning project and the subsequent process and pilot outcome evaluation aimed to address were: (1) Can reflecting on and writing about the virtues espoused in the Laws of Life improve adolescent character development? In order to begin to address the primary Big Question, this project investigated the following additional Big Questions: (2) What is the overarching theory of change for Inspire>Aspire: Global Citizens in the Making Values Poster program (INSPIRE>ASPIRE)?, (3) How is the theory of change best assessed?, and (4) Does preliminary evidence support the theory of change? To address the Big Questions, the project had the following **Aims**: (1) Assess variations in program implementation; (2) Pilot test quantitative measures of key character outcomes; (3) Validate the theory of change using qualitative interviews of youth; and, (4) Assess the relationship between variations in program implementation (e.g., differences in dosage) and poster quality.

Study Design and Participants

In order to assess the second Big Question (What is the overarching theory of change for Inspire>Aspire: Global Citizens in the Making Values Poster program), researchers from REDSS Lab worked with the program developers at Character Scotland to develop a theory of change. Using the Systems Evaluation Protocol (Trochim et al., 2012), we developed a detailed pathway model. The pathway model then guided subsequent analyses.



In order to assess variations in how teachers implement the Inspire>Aspire curriculum (**Aim 1**), trained researchers at Montclair State University conducted telephone interviews with 23 teachers who were implementing Inspire>Aspire. Quantitative measures of several key pupil outcomes including goal setting, future mindedness, sense of purpose, future aspirations, and a global measure of positive youth development were pilot tested using a matched pre-test (fall

2014)/post-test (spring 2015) design (**Aim 2**). We also examined the Inspire>Aspire theory of change using qualitative interviews with Inspire>Aspire students that were administered as a matched pre-test (fall 2014) and post-test (spring 2015) (**Aim 3**). In addition to assessing variations in how teachers implement Inspire>Aspire, we also assessed whether these implementation variations (e.g., dosage, teacher experience, use of supplemental activities) were related to the quality of the posters (the primary output of Inspire>Aspire) produced by the pupils (**Aim 4**).

The process and pilot outcome evaluation of Inspire>Aspire was conducted in Scotland and included Inspire>Aspire participating teachers and students in S2 grade level classes (12-14 year-olds) in Scotland, United Kingdom from September 2014-August 2015. A total of 32 teachers participated in online surveys of program implementation assessment, 23 of whom participated in telephone interviews and submitted their class' completed posters. A total of 123 pupils completed pre-tests and 108 did post-tests. A total of 26 pupils completed pre-program interviews and 24 provided post-program interviews.

Findings

The following section highlights the key findings from the research project as they relate to the Aims and Big Questions.

Aim 1: Assess variations in program implementation

There is a fair amount of variation in the ways in which teachers implement Inspire>Aspire. Teachers who have more experience with Inspire>Aspire tend to focus more on supplemental activities. Pupils whose teachers focused on poster activities and completed these activities almost entirely in class spent the most total number of minutes on Inspire>Aspire (807 minutes on average). Teachers who fully integrated Inspire>Aspire with the broader curriculum

tended to have pupils with higher quality posters. The language level of the template and instructions is challenging for many S2 pupils. While teachers lauded the program's flexibility, this often came from experienced teachers.

Teachers vary in what they consider to constitute the “program.” Some teachers view the poster itself as the program, whereas other teachers see the program (and accompanying activities) as the process leading up to completion of the poster. In the latter, the poster is viewed as a culminating activity that reinforces the previous lessons.

Aim 2: Pilot test quantitative measures of key character outcomes

Pupils did not show significant differences between pre- and post-test on most of the quantitative measures. Some significant results were found for future aspirations. This may be due to several reasons: 1) the poor psychometric properties of the measures; the level of sophistication of the measures for this population (i.e., the measures are designed for adolescents but the pupils in this study were at the younger end of the age range; not enough time had passed for the constructs to emerge, and/or even shorter-term outcomes need to be measured (e.g., precursors to caring).

Aim 3: Validate the theory of change using qualitative interviews of youth

The qualitative results provided some support for the theory of change. Pupils often demonstrated precursors to many of the constructs that were measured but many pupils did not manifest the construct itself (e.g., caring/compassion). This indicates that shorter-term outcomes need to be added to the theory of change to capture the immediate effects of participating in Inspire>Aspire.

Some students struggled with the word ‘values.’ It is important to consider how important it is for students to understand what the word ‘values’ itself means? Is it, instead, more important for students to understand what the meaning of the values themselves are?

Aim 4: Assess the relationship between variations in program implementation (e.g., differences in dosage) and poster quality

Teachers who fully integrated Inspire>Aspire with the broader curriculum (as opposed to using it as a standalone program) tended to have pupils with higher quality posters. The highest quality posters are produced by teachers who spend significant time on poster related activities in class but do not immediately have pupils work on the poster template. Preparing the students to work on the poster by using activities that lay the foundation for the poster elements is effective. Posters are of higher quality when the poster is completed in school rather than at home.

Big Question 1: Can reflecting on and writing about the virtues espoused in the Laws of Life improve adolescent character development?

Perhaps. Evidence from quantitative measures is inconclusive. The specific outcomes of interest need to be reconsidered (perhaps looking at shorter-term outcomes or other outcomes such as self-esteem, confidence and self-reflection which emerged from the qualitative data). Qualitative evidence suggests some indications of personal growth after participating in Inspire>Aspire but it is unclear if this is due to program participation or normative maturation of the youth.

Big Question 2: What is the overarching theory of change for Inspire>Aspire: Global Citizens in the Making?

This is represented in the pathway model. Significant strides were made in articulating the theory of change. The pathway model will be revisited and revised in light of the findings.

Big Question 3: How is the theory of change best assessed?

A mixed methods approach is critical. Teacher report measures should also be used. Some quantitative measures did not have good psychometric properties and will need to be refined and/or replaced. Cognitive interviewing would help with testing and refining measures.

Big Question 4: Does preliminary evidence support the theory of change?

Some elements of the theory of change show preliminary support. Based on these findings, it is important to revisit whether earlier markers of change should be included in the theory of change. Variations in teacher implementation style have an impact on pupil poster quality. Ceiling effects were also present. Therefore, it is important to recruit a more diverse group of pupils to participate in Inspire>Aspire in order to adequately test for program effects.

Recommendations

To advance the mission of Inspire>Aspire, the REDSS Lab leadership team synthesized the findings to create design and implementation recommendations.

Create a Foundation for Success. First, the program developers should *clarify the program goal and the role of the poster*. Character Scotland should determine whether youth character development is meant to occur by simply completing the poster or whether it occurs through classroom activities that provide foundational lessons that prepare pupils to complete the poster. **Is the poster a tool that reflects student character knowledge gained from preceding classroom lessons or is poster completion itself the vehicle for character education development?** Currently, some teachers seem to understand the program goal as poster completion, with little need for intensive classroom instruction on character elements. However, the current findings raise the question of whether the program goal should be learning what's necessary *in order to* complete the poster. With this approach, the poster would be a culminating

project reflecting student learning from program participation. For teachers who spend a lot of time discussing the poster components that foster character development (including values and inspirational figures), the poster serves as a tool to complement and reinforce lessons. For teachers who simply give the posters to their students with little or no direction, simply completing the poster *is* the program. As such, there are currently two different implicit and competing program goals co-occurring. If the goal of the program is NOT only to complete the poster, the program guidelines and website should make it clear that a thoughtful approach to the preceding character content is the crux of character development. To create a foundation for programmatic success, it is critical that teachers have a shared understanding of the program goal.

Provide Best Practices to Teachers. Teachers are instrumental to the success of Inspire>Aspire, so providing them with guidance on activities and best practices is critical. The findings suggest that Character Scotland should *recruit and encourage teachers who are able to integrate Inspire>Aspire with their broader curriculum*. Findings indicate that students of teachers that used classroom time as a central implementation space to hold discussions of values, inspirational figures, and complete the poster with other course/curriculum materials had, on average, the highest percentage of posters classified as High Aspirations. Thus, utilizing class time to integrate the program is important to the program impact.

Similarly, Character Scotland should *strongly encourage teachers to not assign Inspire>Aspire activities as homework*. Findings indicate that pupils who had teachers that primarily used homework time for Inspire>Aspire work tended to have the least favorable poster quality outcomes.

When establishing teacher expectations regarding program delivery time, Character Scotland should *strongly encourage teachers to spend about 800 minutes on implementing Inspire>Aspire*. Pupils whose teachers focused on poster activities and completed these activities almost entirely in class for about 800 minutes had the best outcomes. Character Scotland should provide guidance on how that time is best used to optimize pupil experiences.

Cultivate More Character. Once teachers are informed of the ideal structure for implementing Inspire>Aspire, the focus should turn next to content. The data suggests Character Scotland should *advise teachers to include (more) discussions* on character elements, ensuring pupils understand and reflect on specific character values to enhance their comprehension. Discussion can also be used to enhance other poster elements. To help facilitate these discussions, Inspire>Aspire should *add discussion prompts to the Website*, ensuring that all teachers, regardless of experience level, are equipped to implement successful lessons that deliver the intended objectives.

Expand Guidelines and Resources. To ensure that all teachers deliver the program content with adequate depth and pupil engagement, Character Scotland should *provide more specific guidelines about teaching the program*. Techniques for creating fun, engaging lessons should be shared, along with specific tips and guidelines about meeting learning objectives for each aspect of the program. To do so, *create approach strategies for each section of the poster*. Teachers who have completed the poster program and have classroom-tested strategies may be excellent resources for compiling new resources. Moreover, gathering great lessons and strategies from prior implementers may serve as an opportunity for Character Scotland to recognize stellar teachers and foster program buy-in. Character Scotland can also *offer guidelines on making appropriate adjustments to lessons and materials by offering*

suggestions for abridging or expanding them. Doing so helps maintain the program quality and adherence to the program goals while meeting real world demands.

Create Additional Activities. In addition to expanding classroom activities, Character Scotland should *develop additional supplemental activities that help pupils create high quality posters.* Character Scotland should *invite teachers who have customized and adapted existing materials or created their own supplemental materials to contribute to a growing database of resources.* As previously mentioned, teachers experienced in Inspire>Aspire delivery may have resources to share, which highlights their work and encourages others to use those resources in their implementation. Resources should adequately focus on all poster elements, so Character Scotland should *provide resources for each section of the poster (not just Inspirational Figures and Quotes) on the Website.*

Extend Access to Younger Participants. As the program is being delivered to younger pupils, the materials must be age- and ability-appropriate. Character Scotland should *revise the template and Website materials to be more accessible to a younger age group* by revisiting the vocabulary, reading level and specificity of directions. The newly developed and collected resources and teaching materials recommended above should also be age-appropriate to reach target pupils.

Revisit the Programmatic Theory of Change. The findings of the current study produced new insights and these should be incorporated into a revised theory of change (pathway model). To integrate these findings, the team should: *Revise the pathway model to reflect shorter-term outcomes or precursors and measure these pre-cursors.* The data suggest a number of shorter-term outcomes or precursors may be present in the program pathway, so

adding them to the model and ensuring they are sufficiently measured will allow understanding of how the program works and better serve new generations of pupil participants.

INTRODUCTION

Inspire>Aspire: Global Citizens in the Making Values Poster Program is a character education program that utilizes the strategic opportunity of the Glasgow 2014 Commonwealth Games and builds on the Learning for Life and Inspire>Aspire projects across over 60 countries. Inspire>Aspire is focused on creating educational experiences for youth that foster the virtues of generosity, reliability, future mindedness, joy, purpose, curiosity, and humility; these virtues directly correspond with the qualities of character that Sir John Templeton described in the charter of the John Templeton Foundation. The program theory also reflects Sir John's belief that disciplining one's mind through qualities of character (what many researchers refer to as intentional self-regulation) is the fundamental mechanism by which one can successfully and positively affect one's own positive developmental trajectory as well as the wider world. The Glasgow Games have a special emphasis on developing global citizenship; the values poster that students complete as part of Inspire>Aspire highlights these Commonwealth values as well as national educational outcomes. The fundamental purpose of Inspire>Aspire is the translation of inspiration into transformative and enduring aspiration. The current report shares preliminary findings from the planning, process and pilot outcome evaluation conducted by the Research on Evaluation and Developmental Systems Science (REDSS) Lab co-directed by Dr. Jennifer Brown Urban and Dr. Miriam Linver.

Program Activities

Teachers are provided access to the Inspire>Aspire website and materials, which include a Session Plan document that provides suggestions for launching the program, the Poster Values Activity where pupils are asked to define and rank each of the values, and the Inspirational Figures power point. Throughout the teacher-led classroom-based Inspire>Aspire activities, youth are encouraged to reflect on their life experiences, research an inspirational figure and consider how they will enact their own vision for a better world. The results of these activities are shared as part of a culminating poster project. The poster template's eight sections include personal *values and qualities* as *active, confident individuals; connected, responsible citizens; sustainable, effective contributors*; and *flourishing successful learners*. Pupils are also asked to select a *favorite story or fictional character* that means something to them and explain why it is their favorite and what they learned about life, select an *inspirational figure* to research and write about, choose *inspiring quotations* and comment on them, and answer key questions about *the kind of person they want to be, what they want to achieve in their life, their vision for a better world, and how they are going to contribute to that vision*. In addition, youth are asked to reflect on *what they have learned* from the exercises, and there are opportunities for both peer evaluation and teacher assessment. At each school, a subset of posters are awarded bronze, silver, or gold level awards, all of which entitle winners to attend an awards ceremony.

GOALS OF PROJECT

The primary **Big Question** this planning project and the subsequent process and pilot outcome evaluation aims to address is: (1) Can reflecting on and writing about the virtues espoused in the Laws of Life improve adolescent character development? In order to begin to address the primary Big Question, this project investigated the following additional Big

Questions: (2) What is the overarching theory of change for Inspire>Aspire: Global Citizens in the Making Values Poster program (Inspire>Aspire)?, (3) How is the theory of change best assessed?, and (4) Does preliminary evidence support the theory of change?

To address the Big Questions, Drs. Urban and Linver worked with staff at Character Scotland to develop a detailed theory of change and conducted a process and pilot outcome evaluation with the following **Aims**: (1) Assess variations in program implementation; (2) Pilot test quantitative measures of key character outcomes; (3) Validate the theory of change using qualitative interviews of youth; and, (4) Assess the relationship between variations in program implementation (e.g., differences in dosage) and poster quality. Each Big Question, as well as the aim and approach used to answer it through the current evaluation, is explained below. To answer the primary Big Question, we first determined the program theory of change, part of Big Question 2.

BIG QUESTION 2: The Theory of Change

The goal of Big Question 2 was to develop a detailed theory of change of Inspire>Aspire. A theory of change refers to the hypothesized process through which the program is expected to produce change in participants. The theory of change for Inspire>Aspire was derived using a systems evaluation approach to program and evaluation planning, specifically Evolutionary Evaluation (Urban, Hargraves, & Trochim, 2014).

Evolutionary Evaluation and the Systems Evaluation Protocol

Evolutionary Evaluation is a cutting-edge approach that considers the complex factors inherent in the larger systems within which a program is embedded (Trochim et al., 2012; Urban, Hargraves, Hebbard, Burgermaster, & Trochim, 2011; Urban et al., 2014; Urban & Trochim, 2009). Evolutionary Evaluation integrates principles from theories of evolution, systems science,

epistemology, and developmental science in order to ensure that such principles are incorporated when developing program logic and pathway models, identifying key pathways and nodes (outputs and outcomes), determining the boundary conditions for program models, assessing program lifecycles, and selecting evaluation designs that are appropriate to program evolution. This approach emphasizes: the need for constructing a causal diagram of how the program is believed to work; identifying and consulting both internal and external stakeholders about their perspectives on and priorities for the program; recognizing how the program is related globally to other programs, in part by identifying current scientific research on similar or related outcomes which can help link the program to more universal long-term goals; and continually assessing and revising the evaluation plans to improve the program and collect evidence on the program's success (Urban & Trochim, 2009).

Evolutionary Evaluation can be operationalized using the Systems Evaluation Protocol (SEP) which is a step-by-step guide to engaging in evaluation and program planning. The Systems Evaluation Protocol includes three stages of evaluation planning: (1) Preparation, (2) Modeling, and (3) Evaluation Plan Development (Figure 1). When followed, the SEP results in key outputs including a stakeholder map, logic model, and pathway model.

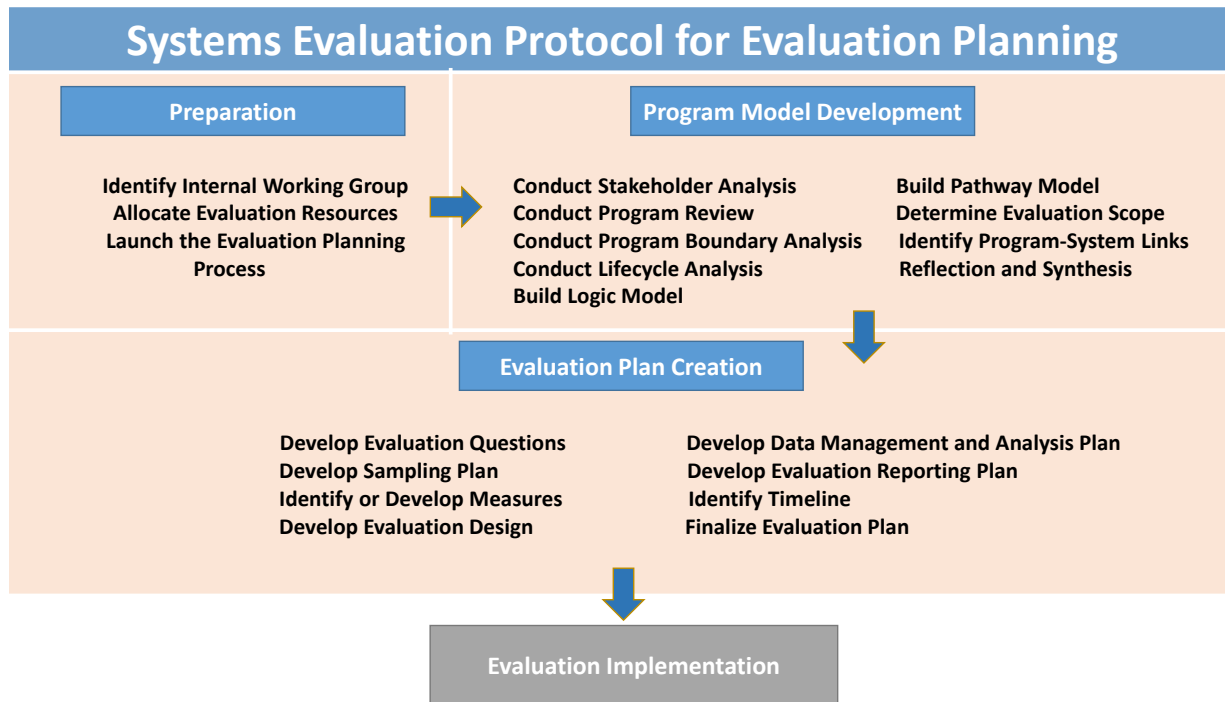


Figure 1. Steps in the Systems Evaluation Protocol (SEP)

Drs. Urban and Linver traveled to Scotland to meet with staff at Character Scotland and conduct in-person facilitation of the SEP with the goal of developing the theory of change for Inspire>Aspire (Figure 2). The Netway (a supporting cyber-infrastructure) was also used. This led to the development of the pathway model, the stakeholder map and some important ‘Aha!’ Moments all of which guided the subsequent process and pilot outcome evaluation.

Stakeholder Analysis. The Program Model Development stage began with a stakeholder analysis and the creation of a stakeholder map (Figures 3 and 4). The goal of this step was to identify all of the potential people and/or groups that have a stake in Inspire>Aspire and its evaluation.

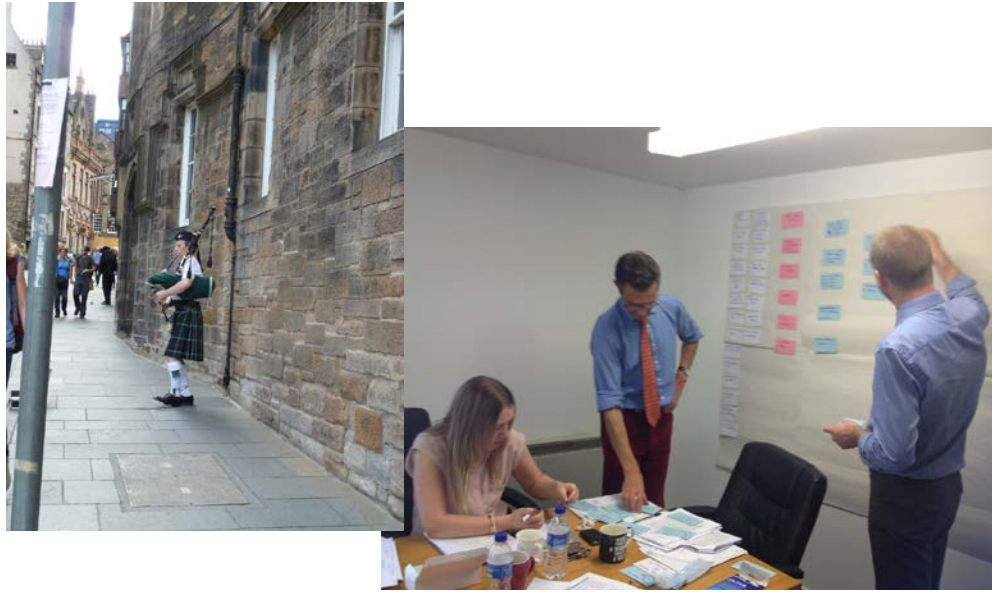


Figure 2. Systems Evaluation Protocol in Action. Pictured on the right is staff at Character Scotland – Julie Thompson, David Lorimer, & Gary Walsh.

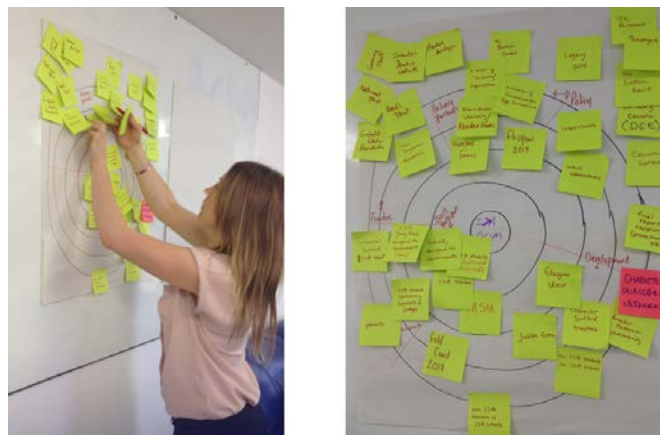


Figure 3. Julie Thompson engaged in the Stakeholder Mapping Process

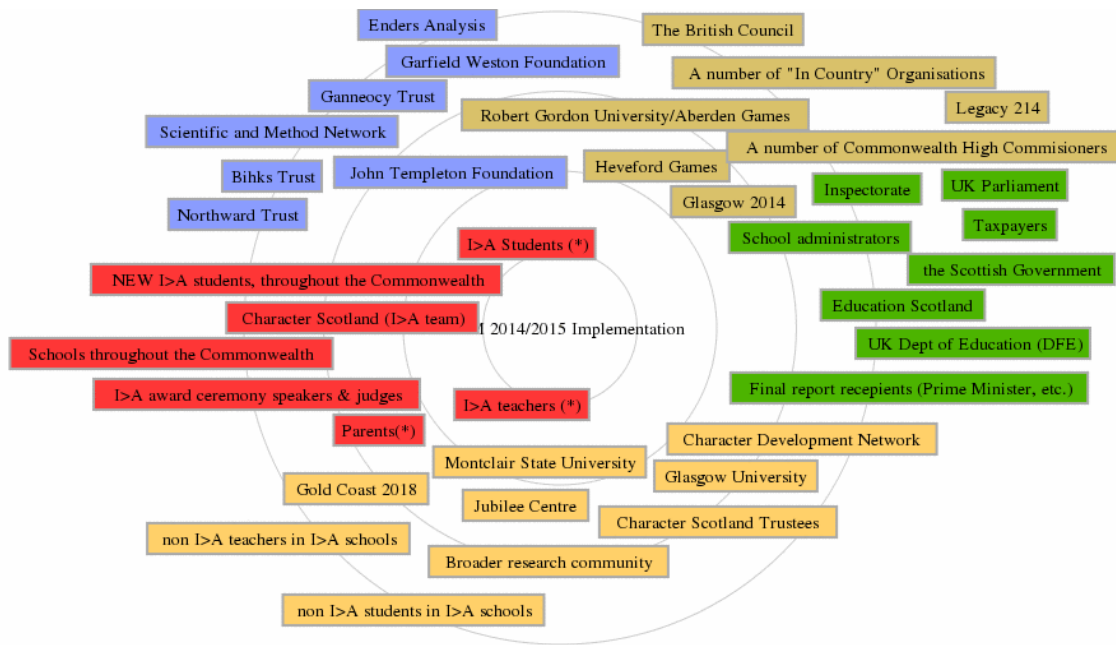


Figure 4. Stakeholder Map as Entered in the Netway. Blue boxes are funders; Red boxes are program recipients and direct supporters; Tan boxes are the research community & others who have a tangential interest in the program and outcomes of the evaluation; Green boxes are policy makers/government; and, Brown boxes are the “hook” (e.g., Glasgow 2014).

Boundary Analysis. This involved deciding, very precisely, which components are “in” the program and which ones are “out”. Programs are interconnected with the environment around them, and they are preceded, followed and accompanied by a complex array of other activities, events, programs, and effects. There is no single “right” program boundary (Figure 5).

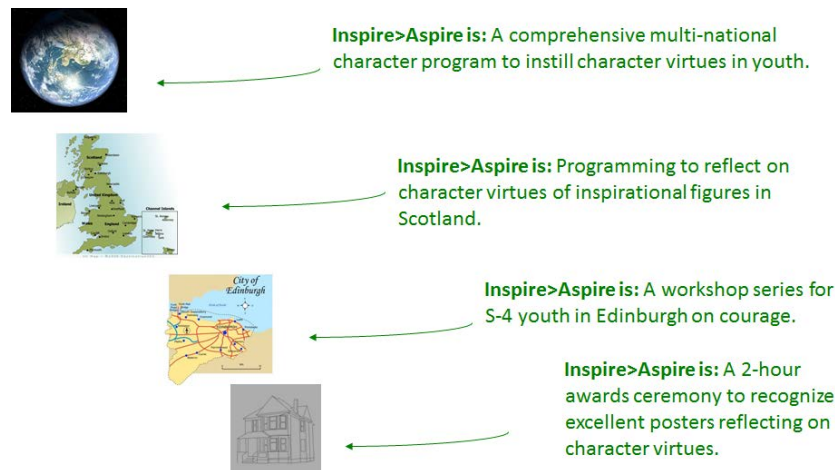


Figure 5. Boundary Analysis

Lifecycle Analysis. A lifecycle analysis is critical to understanding how a program and its corresponding evaluations change over time. Systems evaluation characterizes program development and evaluation as evolutionary processes with inherent lifecycle qualities (Trochim et al., 2012; Urban et al., 2014). Program lifecycles are divided into four phases: Phase 1 - Initiation, Phase 2 - Development, Phase 3 - Stability, and Phase 4 - Dissemination. For any given program lifecycle stage, there is an appropriate type of evaluation work to be done – that is, a corresponding evaluation lifecycle stage. “Appropriate” in this case means an evaluation that provides the kind of information that will be most useful and relevant to the decisions that arise for the program in its current lifecycle phase. Evaluation lifecycles are divided into four phases: Phase 1 - Process and Response, Phase 2 - Change, Phase 3 - Comparison and Control, and Phase 4 - Generalizability. An essential step in the SEP is to assess what lifecycle stage the program is in currently, and the lifecycle stage of prior evaluation efforts. If the lifecycles are in alignment, then the evaluation planning focuses on what kind of information will be needed in order to most effectively move the program along its lifecycle path. If the two are not aligned, then evaluation planning should address that gap and either focus on “filling in” information that has not been properly established yet (reining in the evaluation), or pushing for a higher level of evidence about the program (see Figure 6; Urban et al., 2014).

Figure 6 visually depicts Lifecycle Analysis. If the program and evaluation phases are perfectly aligned, the program would fall somewhere along the diagonal red line (e.g., Program A). Programs above the red line (e.g., Program C) are doing evaluations that are more advanced than their program lifecycle phase calls for. Program C is in the “Initiation” program lifecycle phase but it is being evaluated using a “Comparison & Control” evaluation such as an RCT

design. Programs below the red line (e.g., Program B) have evaluations that are “lagging behind” the program lifecycle phases. Program B is in the “Stability” phase of its program lifecycle, but is doing “Process & Response” evaluation. We determined that Inspire>Aspire is similar to Program B. It has a fairly advanced program lifecycle. The program has been conducted for many years and the program components are stable (there is little change to the program across implementation rounds). However, Inspire>Aspire had not yet been systematically evaluated. Most prior evaluation work focused on examining descriptive statistics based on poster data and anecdotes from teachers and pupils who had participated in the program. We determined that moving toward alignment of program and evaluation phases – and promoting the healthy evolution of the program was essential. It is important to recognize that for a program whose program and evaluation phases are not currently aligned, the move toward alignment does not necessarily occur within one evaluation cycle. Rather, the focus is on building evidence over successive evaluation cycles while simultaneously striving for phase alignment. This was the goal of the process and pilot outcome evaluation.

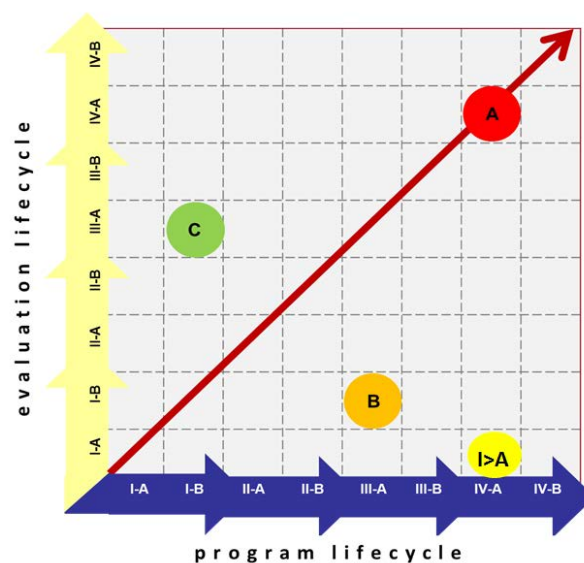


Figure 6. Evaluation and Program Lifecycle Phase Alignment. This figure depicts the relationship between program phases (on the x axis) and evaluation phases (on the y axis).

Logic and Pathway Model. A logic model captures the program activities, outputs, and short-, medium-, and long-term outcomes. The components of a logic model and their corresponding definitions are presented in Table 1 and the logic model for Inspire>Aspire is presented in Figure 7. Although logic models can be an effective tool for communicating the underlying program logic, they have several shortfalls. Columnar logic models typically do not provide enough detail to gain a real understanding of the program. The notion of causality is critically important in systems evaluation. In a program logic model there is a general idea of causality in the background – activities are expected to lead to outputs which are in turn expected to produce short-, medium-, and long-term outcomes and ultimately impacts – but logic models do not spell out the specific cause-effect relationships. They also fail to show the hypothesized key pathways or throughlines that tell the story of the program. In addition, logic models do not highlight key outcomes.

Table 1. Components of a logic model

Components of a logic model	
Activities	Actions and events that directly reach people who participate or are targeted.
Outputs	By-products of activities that serve as evidence that the activity took place, and on what scale (# reports produced or projects completed, records of workshops completed, etc.).
Outcomes	Effects on participants. Short-term outcomes often include changes in participant knowledge, attitudes, and skills; medium-term outcomes are a more consistent change in behavior; and long-term outcomes are those that affect the larger social context.

Logic Model

Input:	Activities:	Outputs:	Short-Term Outcomes:	Mid-Term Outcomes:	Long-Term Outcomes:
	Articulate 1 strength and 1 growth area	Certificates of completion	Alignment between values, poster elements, and future ambitions	Broader understanding of what a successful life looks like	Commitment to vision for a better world
	Articulate vision for a better world	Completed posters	Greater awareness of resilience	Increase motivation to pursue goals	Contributions to civil society (active global citizens)
	Clarify and define potential future achievements (dreams)		Increased Empathy	Increased Compassion	Increased self-efficacy
	Completing poster		Increased feeling of being inspired	Increased confidence	Personal empowerment (self-belief)
	Define person you want to be		Increased future-mindedness	Increased self-respect	Positive view of human nature
	Define values		Increased world knowledge	Increased sense of agency	Sustained motivation to pursue goals
	Develop & articulate plan for contribution to a better world		Recognizing the good in yourself and others	Increased sense of purpose	
	Explain reason for selecting story or fictional character		Self-awareness of strengths/areas in need of improvement	More optimistic/positive outlook	
	Explain why inspirational figure inspires		Self-knowledge (of strengths/weaknesses)		
	Explain why quote was selected		Sense of achievement (families feel proud)		
	Rank order strengths		Setting Goals		
	Reflect on what was learned				
	Select & research quote				
	Select and research inspirational figure				
	Select story or fictional character				

Figure 7. Inspire>Aspire Logic Model

Pathway models overcome the shortfalls listed above. A pathway model is a visual program model closely related to the columnar logic model. A pathway model shows causation, using the elements of the logic model and incorporating the logical connections that lead from an

activity to one or more short-term outcomes, and from there to medium-term outcomes, and ultimately to long-term outcomes. It tells the story of how a program works. Most importantly, the pathway model provides an invaluable framework for focusing the evaluation. Figure 8 presents the pathway model that corresponds with the logic model provided in Figure 7.

The process of moving from a logic model to a pathway model can be complicated. The knowledge or insight expressed by the pathway model is often held subconsciously by people involved with the program, and the opportunity to bring these insights out, articulate them, and combine the insights of a number of key people is actually quite rare. It is in the shared building of the model that many of the gains are achieved – this is where people manage to get on the same page in terms of their understanding of what the program is about (Figure 9).

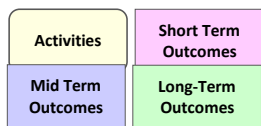
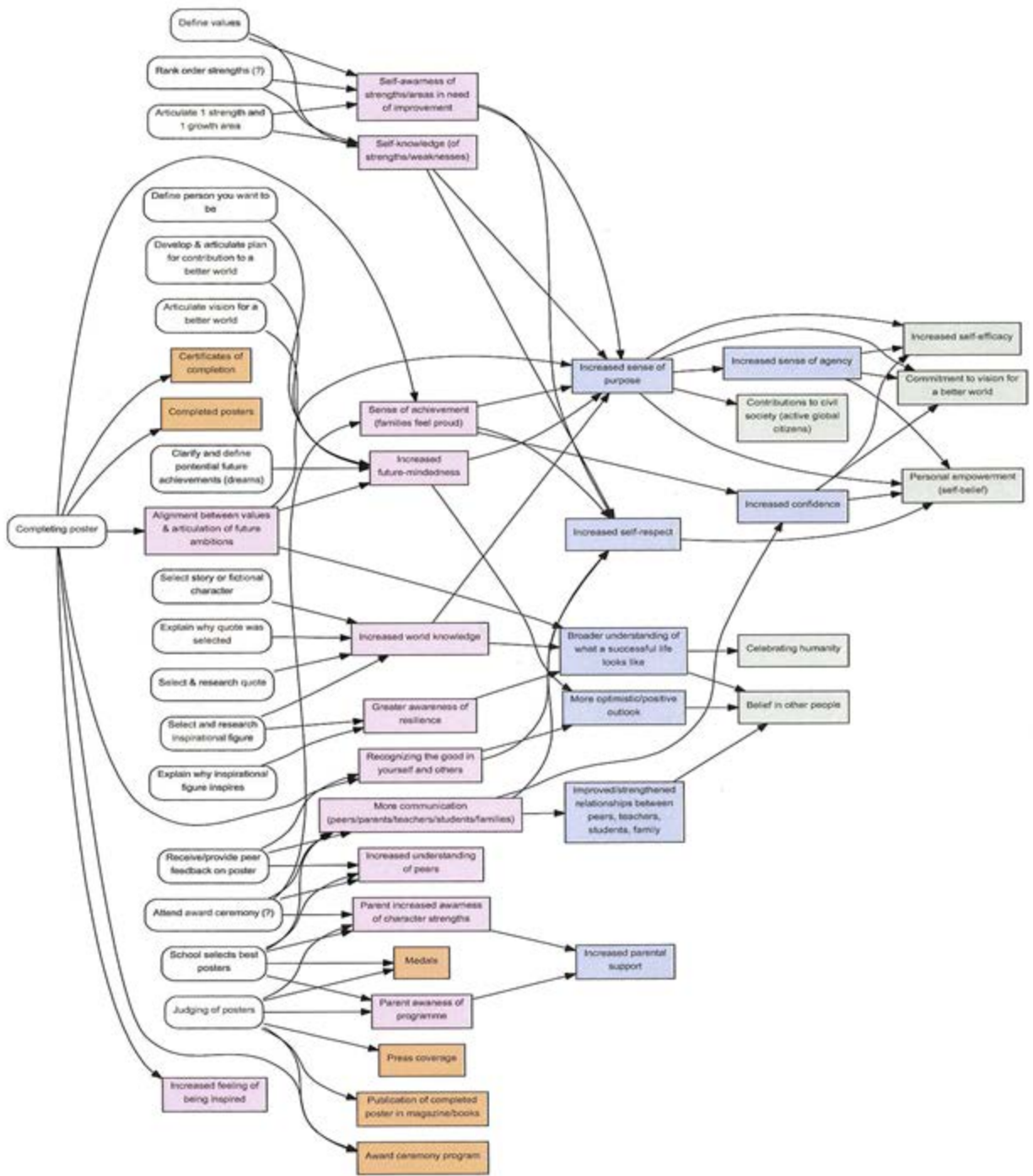


Figure 8. Inspire>Aspire Pathway Model. This figure was the result of a collaboration between Character Scotland program developers and MSU REDSS Lab facilitators.



Figure 9. Process of Creating a Pathway Model. Pictured: Julie Thompson, David Lorimer, Gary Walsh, and Ronnie Davidson of Character Scotland.

Evaluation Scope. Determining the evaluation scope involves making decisions about where to focus the upcoming evaluation. The completed pathway model becomes the framework on which to hang information that will help narrow the scope of the evaluation for one evaluation cycle. Evaluation Scope refers to the range of model elements that could reasonably be within the purview of the upcoming evaluation. Scope does not go as far as identifying specific evaluation questions but serves as an intermediate step supporting careful decision-making.

In order to determine the evaluation scope, we used the pathway model as a tool for engaging in the “Mining the Model” activity. As part of this activity, we asked several questions of the model. These included: 1) Are there any activities that are not connected to any outcomes?; 2) Are there any outcomes that are not connected to any activities?; 3) Are there any pathways that dead-end at short- or medium-term outcomes?; 4) Are there any big leaps in logic (i.e., an arrow from an activity to a medium- or long-term outcome)?; 5) Are there any boundary issues?; 6) Is there anything that might be confusing to an outsider; and, 7) Are there themes or

common threads among outcomes? Next, we examined the outcomes to see if we could identify “Prime Destinations” (outcomes with a lot of arrows going IN to them), “Gateways” (outcomes with a lot of arrows going FROM them), and “Hubs” (outcomes with a lot of arrows going IN AND OUT of them). These are labeled with stars in Figure 10. We then examined the connections between nodes and identified the most important linkages (labeled with keys in Figure 10). Next, we returned to the Stakeholder Map and identified the three primary external stakeholders of the program (A = Education Scotland (national government organization), B = John Templeton Foundation (funders), C = Inspire>Aspire teachers). We then identified the outcomes that we thought would be of greatest interest to these stakeholders and labeled them with letters in Figure 10. We also identified the outcomes that CES would most like to be able to say something about and labeled these with the letter D in Figure 10. Finally, we noted the throughlines that go from an activity all the way to a long-term outcome. Keeping in mind what we learned from the previous “Mine the Model” steps, we determined the main causal story of the program (highlighted in pink in Figure 10). The first highlighted pathway became the focus for the pilot outcome evaluation.

Program-System Links. In order to assist with determining the evaluation scope, it is important to identify program-system links. This step involves examining other related programs as well as the research literature primarily in an effort to identify potential validated measures as well as research evidence that could help support the logic conveyed by the links in the pathway model. Articulating clear connections between the program logic and the research evidence helps to build the case for anticipated long-term outcomes. In addition, the research evidence base provides opportunities for cost and time savings by applying previously developed tools to the

current evaluation questions. This step was conducted by staff at REDSS Lab and informed the measures selected for the process and pilot outcome evaluation.

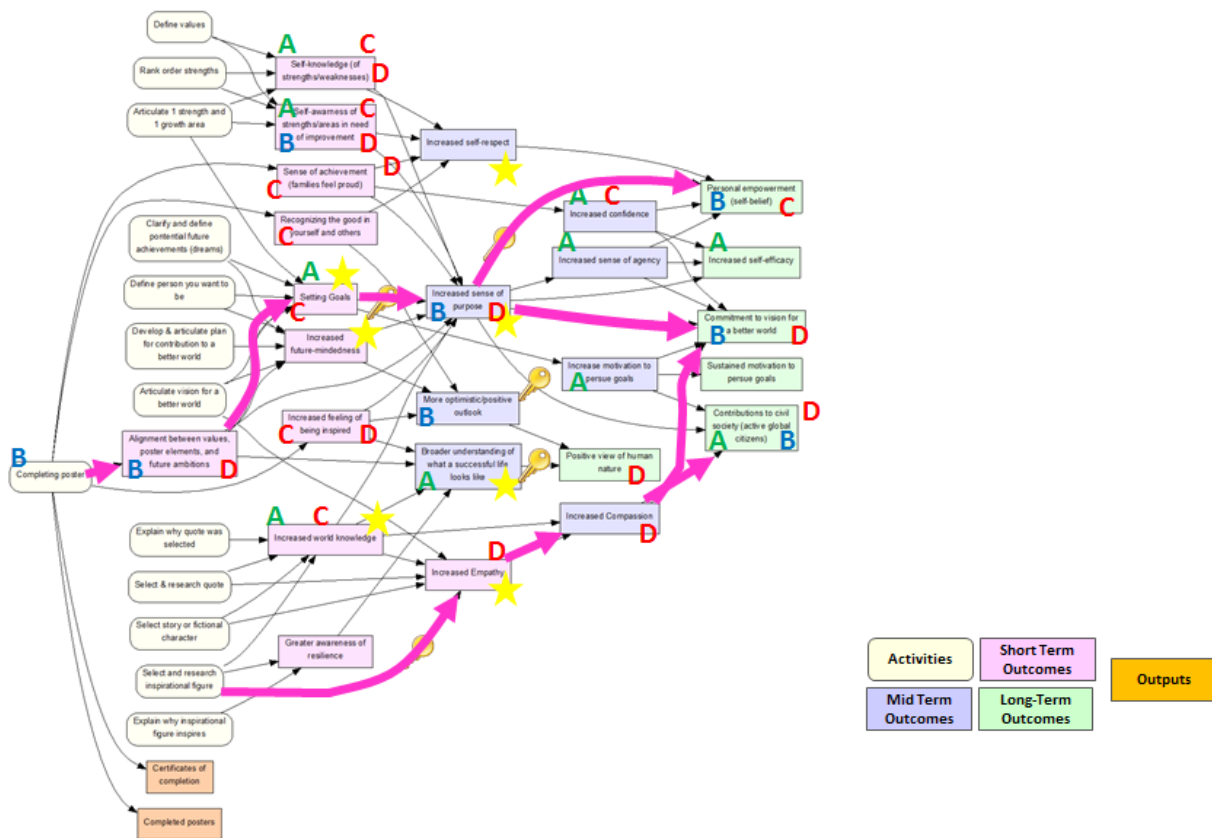


Figure 10. Mining the Model

Results of Articulating the Inspire>Aspire Theory of Change: ‘Aha!’ Moments

By ‘Aha!’ Moments, we mean the experience of an epiphany or the insight that often occurs when people are not even aware they are thinking of the problem. ‘Aha!’ Moments are important because they lead to evaluative thinking or critical thinking in the context of evaluation motivated by an attitude of inquisitiveness that leads to deeper understanding. ‘Aha!’ Moments provide evidence that evaluative thinking has occurred. The REDSS Lab facilitators and Character Scotland program developers experienced several ‘Aha!’ Moments during the evaluation planning process.

Two of these ‘Aha!’ Moments occurred during the Evaluation Scope phase of the SEP and the third ‘Aha!’ Moment occurred during the Program-System Links phase of the SEP.

‘Aha!’ Moment 1. Through discussions, we learned that the program developers had been emphasizing the importance of a year-end awards ceremony attended by selected participants representing a small fraction of the total number of youth who participated in Inspire>Aspire. In the original iteration of the pathway model, one section of the model focused on the awards ceremony. The program developers had emphasized the awards ceremony because they committed a lot of time and resources to organizing and hosting the ceremony. When they examined the model closely, they identified several “dead-end” outcomes: A short-term outcome that did not connect to any medium- or long-term outcomes and two medium-term outcomes that did not connect to any long-term outcomes. Through the modeling exercise, they realized that the awards ceremony was not actually at the core of the program; this realization resulted in a shift in thinking regarding the distribution of resources and focus (e.g., more resources dedicated to developing teacher resources rather than the awards ceremony). As a result of ‘Aha!’ Moment #1, several activities, outputs, as well as short- and medium-term outcomes were cut from the model.

‘Aha!’ Moment 2. Representatives of several stakeholder groups were asked to review the preliminary model and provide feedback. The outcome “setting goals” was not present in the initial model but was added to the final model when suggested by a youth stakeholder. The SEP deliberately incorporates multiple stakeholder perspectives and prioritizes youth voice which is an integral component of positive youth development (Urban, 2008).

‘Aha!’ Moment 3. The SEP facilitates research-practice integration particularly during the process of linking the pathway model with the research evidence-base. In the case of Inspire>Aspire, when the model was linked to the research evidence-base, we determined that goal setting is a critical short-term outcome.

Finalizing the Evaluation Plan

The last two steps of the Systems Evaluation Protocol focus on using the information learned during the prior steps to articulate a clear evaluation plan.

Reflection and Synthesis. It is important to pause for reflection and synthesis before creating the evaluation plan. The purpose of this step is to stop and reflect in order to summarize the current thinking about the program and its evaluation. All of the materials that have been produced from the previous steps of the protocol were reviewed and revised as needed.

Evaluation Plan Creation. This stage includes several sub-phases which ultimately led to the development of an evaluation plan for the process and pilot outcome evaluation (see Big Questions 1, 3, and 4 below). One of the key aspects of systems evaluation that distinguishes it from traditional evaluation planning is the concept of lifecycles as described earlier. The lifecycle phases of Inspire>Aspire were carefully considered when making decisions about the *Evaluation Questions, Sampling Plan, Measures, Design, Analysis Plan, and Timeline*. In addition to developing an evaluation plan for the process and pilot outcome evaluation we also developed an outcome evaluation plan which was submitted to the John Templeton Foundation and was successfully funded (August 2015-July 2018).

BIG QUESTIONS 1, 3, & 4: The Process and Pilot Outcome Evaluation

In order to address Big Questions 1, 3 and 4, the process and pilot outcome evaluation aimed to: (1) Assess variations in program implementation; (2) Pilot test quantitative measures

of key character outcomes; (3) Validate the theory of change using qualitative interviews of youth; and, (4) Assess the relationship between variations in program implementation (e.g., differences in dosage) and poster quality.

Design and Procedure

A single-group post-test design was used to address aims 1 and 4 and a single-group matched pre-/post-test design was used to address aims 2 and 3. In order to assess variations in how teachers implement the Inspire>Aspire curriculum (**Aim 1**), trained researchers at Montclair State University conducted telephone interviews with 23 teachers who were implementing Inspire>Aspire. These interviews took place during the 2014/2015 academic school year and were recorded and transcribed for analysis. Prior to completing the telephone interviews, teachers were asked to complete a short online survey that included questions with closed-ended responses (e.g., how many years have you been teaching Inspire>Aspire).

Quantitative measures of several key outcomes including goal setting, future mindedness, sense of purpose, future aspirations, and a global measure of positive youth development were pilot tested using a matched pre-test (fall 2014)/post-test (spring 2015) design (**Aim 2**). Staff at REDSS Lab prepared survey packets including detailed instructions regarding informed consent/assent and survey administration. Character Scotland distributed and collected completed surveys and consent forms and returned them to REDSS Lab for analysis. The Inspire>Aspire teachers administered the surveys to their pupils during their Inspire>Aspire class sessions.

In addition to pilot testing quantitative measures, we also examined the Inspire>Aspire theory of change using qualitative interviews with Inspire>Aspire students that were

administered as a matched pre-test (fall 2014) and post-test (spring 2015) (**Aim 3**). We expected that the quantitative measures of key outcomes could possibly not be sensitive enough to detect changes in youth behaviors. Therefore, the youth interviews allowed us to both triangulate the quantitative measures and to assess early markers of character development. The telephone interviews were conducted by trained researchers from REDSS Lab and the interviews were recorded and transcribed for analysis.

In addition to assessing variations in how teachers implement Inspire>Aspire, we also assessed whether these implementation variations (e.g., dosage, teacher experience, use of supplemental activities) were related to the quality of the posters (the primary output of Inspire>Aspire) produced by the pupils (**Aim 4**). Character Scotland collected completed pupil posters from the teachers at the end of the Inspire>Aspire and sent them to the REDSS Lab for analysis.

Sample

The process and pilot outcome evaluation of Inspire>Aspire was conducted in Scotland and included Inspire>Aspire participating teachers and students in S2 grade level classes (12-14 year-olds) in Scotland, United Kingdom from September 2014-August 2015.

Teachers. The goal was to sample for breadth in order to obtain a sample of teachers with substantial variability in terms of implementation strategies. We used a purposive stratified sampling approach and identified teachers that varied in terms of their experience using Inspire>Aspire and their perceived effectiveness (based on anecdotal evidence) and stratified the sampling frame into four groups based on experience and effectiveness: 1) high experience/high effectiveness; 2) high experience/low effectiveness; 3) low experience/high effectiveness; and 4) low experience/low effectiveness. A total of 32 teachers participated in online surveys of

program implementation assessment, 23 of whom participated in telephone interviews and submitted their class' completed posters. Of the 574 posters received, 198 have been coded to date and are included in the current analyses. Six of these teachers also had pupils in their classrooms participate in the evaluation.

Pupils. A total of 280 pupils were recruited from 10 classes in 6 schools in the Northeast and Central regions of Scotland to participate in pre- and post-test surveys. Of the 280 pupils recruited, 123 (44%) completed pre-tests and 108 did post-tests (88% retention from Wave 1 to Wave 2). A subset of pupils who completed the quantitative surveys were invited to participate in 15 minute phone interviews. A total of 26 pupils from the participating classes completed pre-program interviews and 24 provided post-program interviews.

Measures

Poster Quality. A poster coding rubric and dictionary was developed as a tool for the systematic scoring of posters. The rubric focuses on: a) degree of *completion*, b) degree of *development* in responses (or robust thoughtfulness), and c) degree of response *alignment* (or thematic consistency) across poster sections. The poster coding rubric uses an *a priori* coding system whereby data are coded according to pre-existing categories (Creswell, 2007).

During training, coders individually scored sets of 3 posters and then discussed their individual scores until group consensus was reached. Character Scotland reviewed the proposed rubric and their feedback was incorporated into the final rubric. A set of 15 posters were assigned to each of five coders; coders entered their final scores into an electronic platform containing the rubric, and responses were analyzed for percent agreement, Pearson's *r*, and intraclass correlations (ICC). The final ICC demonstrated good inter-rater reliability and scores ranged from $ICC(2, 6) = .60$ ($p < .001$) to $ICC(2, 6) = .99$ ($p < .001$). The average ICC across 6

rater pairs was .796 and all ICC values reached significance at $p < .001$. Once adequate agreement was reached, one coder was randomly assigned to code each poster.

Poster quality summary scores were then computed based on coders' scores. These include (1) Alignment between values (elements 1-4) and future aspirations (elements 8a-8d); (2) Total poster alignment (alignment across all poster elements); (3) Consistency of poster response with PYD and Inspire>Aspire principles; (4) How well developed the Values are in the first panel; (5) How well developed the pupils' character/story, quotes, and inspirational figure are (elements 5, 6, and 7); (6) How well developed the future aspirations are in the last panel (elements 8a-8d); and (7) a broad summary score, consisting of the mean of the quality summary scores 3, 4, 5, and 6. For most analyses, the broad summary score was used to measure poster quality.

Goal setting/selection. The Selection, Optimization, and Compensation (SOC) questionnaire developed by Freund and Baltes (2002) and modified by Gestsdottir, Bowers, von Eye, Napolitano, and Lerner (2010) was used to measure goal setting/selection. Participants completed a 24-item scale (Wave 1: $M = 14.70$, $SD = 3.53$, $\alpha = .59$; Wave 2: $M = 14.70$, $SD = 4.56$, $\alpha = .80$). Each item presented two statements that included examples of two different ways people might behave: one describes use of SOC skills, and the other does not. A sample item involved the two statements such as "I am always working on several goals at once" (which is not indicative of selection) or "I always focus on the one most important goal at a given time" (which is indicative of selection). After reading the two statements the pupil must select the statement that best describes him or her. The SOC is frequently divided into 4 subscales of 6 items each: Selection, Optimization, Compensation, and Loss-based Selection. For early adolescents, a 9-item subscale (the SOC-9) is commonly used (Wave 1: $M = 6.35$, $SD = 1.96$, α

= .61; Wave 2: $M = 6.31$, $SD = 2.34$, $\alpha = .76$; Gestsdóttir, Bowers, von Eye, Napolitano, & Lerner, 2010).

In addition, the Aspiration Index (Kasser & Ryan, 1993) was used to assess both intrinsically and extrinsically motivated goals across four domains: self-acceptance, affiliation, community feeling, and financial success. Aspiration is rated on two dimensions: personal importance (Wave 1 ASPI: $M = 3.03$, $SD = .48$, $\alpha = .84$; Wave 2 ASPI: $M = 3.11$, $SD = .42$, $\alpha = .83$) and chances of attaining them (Wave 1 ASPC: $M = 2.91$, $SD = .53$, $\alpha = .88$; Wave 2 ASPC: $M = 2.95$, $SD = .49$, $\alpha = .88$). Participants were presented with 21 future states or items to be ranked on a 5-point Likert-type scale; with 0 = not important/very low chances of attainment and 4 = very important/very high chances of attainment. Sample subscale items include: “You will know and accept who you really are” for self-acceptance, “You will know people that you can have fun with” for affiliation, “You will help people in need” for community feeling and “You will have a job with high social status” for financial success.

The SOC and Aspiration Index were supplemented with qualitative measures from pupil interviews. The qualitative interview included a measure of possible selves (Oyserman, Johnson, & James, 2011; Oyserman & Markus, 1990). Possible selves involve hoped for and feared versions of an adolescent’s future self. Conceptualizing possible selves allows adolescents to visualize and consider themselves in a multitude of future roles and situations, as well as contemplate ways to achieve these future selves (Massey, Gebhardt, & Garnefski, 2011).

Future Mindedness. Future-mindedness was assessed with a subset of 10 items from Steinberg’s Future Orientation Scale (Wave 1: $M = 1.68$, $SD = .52$, $\alpha = .71$; Wave 2: $M = 1.61$, $SD = .36$, $\alpha = .50$; Wave 2 10-item: $M = 1.66$, $SD = .53$, $\alpha = .74$; Steinberg et al., 2009). The full scale includes three subscales: (a) time perspective, (b) anticipation of future consequences, and

(c) planning ahead. Participants were presented with two statements that included examples of two different ways people might behave: one describes a future orientation, and the other does not. A sample item is “Some people like to plan things out one step at a time” (which is indicative of future orientation) or “Other people like to jump right into things without planning them out beforehand” (which is not indicative of future orientation). After reading the two statements the pupil must select the statement that best describes him or her and state whether it is “really true for me” or “sort of true for me.”

Sense of Purpose. To measure pupils’ sense of purpose, we employed the Stanford Youth Purpose Survey (SYPS), developed by Bundick et al. (2006) as part of a John Templeton Foundation-supported project on purpose conducted by William Damon. This 17-item scale has two subscales, Internal Directed Aims (with 12 items), and Externalized Aims (with 5 items) (Wave 1: $M = 4.74$, $SD = .61$, $\alpha = .83$; Wave 2: $M = 4.74$, $SD = .61$, $\alpha = .82$). Youth were asked to rate the extent to which they agreed or disagreed with a set of statements about the purpose of their life. Response options ranged from 0 = strongly disagree to 6 = strongly agree. For the present report, we created a new subscale based on the 5 items most salient to Inspire>Aspire. These 5 items all complete the following sentence stem: “The purpose of my life is to...” and include these items: Help others; Make the world a better place; Do the right thing; Discover new things about the world; and Support my family and friends. The first two items are from the Externalized Aims subscale, and the last three items are from the Internal Directed Aims subscale (Wave 1 5-item: $M = 4.97$, $SD = .72$, $\alpha = .73$; Wave 2 5-item: $M = 4.98$, $SD = .71$, $\alpha = .70$).

Positive Youth Development (PYD). PYD was assessed using the PYD-Very Short Form (PYD-VSF; see Geldhof et al., 2014). Sixteen of the 17 items were used in Wave1 ($M =$

8.54, $SD = 1.38$, $\alpha = .76$); one item in the Character subscale was replaced with a similar question in a process described below. All 17 items of the PYD-VSF were used in Wave 2 ($M = 8.48$, $SD = 1.47$, $\alpha = .77$). The PYD-VSF includes three items each from the Caring, Competence and Confidence subscales, and four items each for the Character, and Connection subscales. All items use four or five-point Likert response options, with higher scores reflecting higher levels of the construct. Caring items measure this construct globally. A sample item from the Caring subscale is, “When I see someone being taken advantage of, I want to help them.” Connection items assess relationships with neighborhood, family, school, and peers, and Competence items index academic, social, and physical facets of this construct. Sample items from these subscales include, “I get a lot of encouragement at my school,” for Connection, and for Competence, the comparison of two statements such as, “Some kids have *a lot* of friends,” and “Other kids *don't* have very many friends.” Confidence items index self-worth, positive identity, and physical appearance. A sample item for the Confidence subscale includes the comparison of two statements such as, “Some kids are *happy* with their height and weight,” versus, “Some kids are *not happy* with their height and weight.” Character items assess support for diversity, performance, and personal values and include sample items such as, “Accepting responsibility for my actions when I make a mistake or get in trouble.” A conduct behavior item asks respondents to select which of two statements is most true for them. To address social conscience in the Character subscale for Wave 1, an item from the Revised Stanford Youth Purpose scale (Bundick et al., 2006; Steger, Frazier, Oishi, & Kaler, 2006) was included using a Likert-scale ranging from 0 to 6 to index agreement from *strongly disagree* to *strongly agree*. As with the conduct behavior item, all Competence subscale items followed the two-statement format and seek respondent ratings on each as either *really true of me* or *sort of true of me*.

Items were transformed to ensure each was allotted equal weight, including the proxy Character item for the Wave 1 scale. For analyses, this item was transposed from a 6-point scale to a 5-point scale to match the other items. For more detailed information about how each subscale was transformed and coded, see Geldhof et al. (2014). Higher scores on all items reflect higher amounts of each C.

Preliminary Analyses

Aim 2 focused on pilot testing quantitative measures of key outcome variables. Our primary objectives were to determine whether the piloted measures were appropriate for a sample of young Scottish adolescents (most measures were developed for American adolescents) and to assess whether the quantitative measures were sensitive enough to detect change. To this end, we began by assessing the psychometric properties of our quantitative survey measures. For each survey scale, we assessed the “fit” of each item in relation to its larger scale or subscale. We used a structural equation modeling technique, where each scale (or subscale) was considered a latent construct, and each item as a measured indicator of the latent construct. As such, we considered the weak and strong invariance of each item. Weak invariance implies that the model (or scale) works the same way across different groups (e.g., boys and girls; pupils in different schools); in the structural equation model, the factor loadings will be identical across groups. Strong invariance measures whether the scores from different groups have the same origin, or in fact measure the same construct (Wu, Li, & Zumbo, 2007).

SOC. First, we tested the SOC scale, focusing on the subsale of 9 items that typically works well in young adolescent populations. One item (“When something doesn’t work as well as usual, I look at how others do it”) did not load significantly at either wave; we removed it and re-ran the analyses. This model fit adequately. However, the 8-item model did not pass weak

invariance (equality of factor loadings) with or without this item. We investigated which items were causing the problem. Table 2 demonstrates the items where there was a large discrepancy between the factor loadings at Wave 1 and Wave 2. Most of the factor loadings were moderately to highly unequal and partial invariance (with one or two loadings freed) did not work. Thus, computing scale scores and testing for differences across time was not possible because the constructs cannot be assumed to be measuring the same thing at both time points.

Table 2. Factor Loadings for SOC

	W1	W2
SOC4	0.490	0.747
SOC8	0.420	0.437
SOC11	0.645	0.567
SOC12	0.614	0.864
SOC14	0.477	0.760
SOC18	0.545	0.799
SOC20	0.620	0.826
SOC23	0.600	0.544

Because of the unreliability of this scale, we then conducted Exploratory Factor Analyses (EFA) in Mplus with one through 5 factors, using the entire item set, to see if the SOC measure could be used with a different factor structure. According to the EFA results, a three factor model provided the best fit, but none of the resulting factor solutions had a sound theoretical basis.

Sense of Purpose. First, we tested the 2 subscales (externally and internally directed aims) suggested by the creator of the scale. This two subscale model did not fit well. The two constructs were highly correlated at each time point, but each subscale had several items that did not load strongly (or significantly) and several items for which the model did not explain significant variance. Because of the poor fit of this original model, we examined the items individually as well as based on our knowledge of Inspire>Aspire, the pupils' responses to interviews, and the pathway model. We selected the five items that best represented sense of

purpose in the context of Inspire>Aspire. The items included, “The purpose of my life is to...: help others, make the world a better place, do the right thing, discover new things about the world, and support my family and friends.” These five items combine internally and externally directed aims. This constructed subscale easily passed the test of weak invariance, but did not pass strong invariance (due to the first item - help others).

Aspiration Index. As described above, there are two constructs represented in the Aspiration Index: Importance (e.g., the importance of each goal to the pupil) and Chances (e.g., the chances the pupil will reach the goal). Each category has four subscales/domains: self-acceptance, affiliation, community feeling, and financial success. For the Importance items, all subscales fit well and passed both tests of invariance (weak and strong). For the Chances items, three of the four subscales fit well and passed both tests of invariance. One subscale (affiliation) passed both tests of invariance, but the subscale itself did not have a good fit. Overall the subscales of the Aspirations Index worked well; items hung together as subscales, the items measured what we hoped they were measuring, and they performed well across different groups.

Future Orientation. When we examined the 10 item form of the future orientation scale that was available at both Wave 1 and Wave 2, the full model fit was passable and passed both weak and strong tests of invariance.

PYD. The overall PYD-SF did not fit well and did not pass invariance tests; when we examined the 5 Cs separately, there was a much better fit, but similar to the findings from the SOC scale, the factor loading patterns are extremely different across time points. One potential reason for this (and it could be the case for SOC as well) is that the data are organized by time point rather than by age. The age range in the sample is fairly small (12-14) but it is possible the

pattern of loadings could be different across 12, 13, and 14 year olds (however there are not enough pupils in each age group to test for invariance by age).

Given our thorough examination of the psychometric properties of the survey measures, we have some concerns about the use of some of these measures with a sample of Scottish S2 youth. A search for more appropriate measures may be warranted. For example, the measures we used were developed for American adolescents, generally aged 12-20. Our sample was at the low end of this age range and although English is the common language, the phrasing and use of typically American vocabulary may not have been as accessible to Scottish youth. Therefore, we recommend using measures that are geared toward a younger age group (e.g., 10-14 year-olds). In addition, cognitive interviewing should be conducted with a small group of 12-14 year-old Scottish youth to assess how the children are understanding and interpreting the measures.

Results

The first set of evaluation questions addressed the goals of the process evaluation and specifically **Aim 1** which focused on understanding teacher variation in program implementation as well as variations in teacher characteristics (e.g., teaching experience, experience implementing Inspire>Aspire, gender). The second set of evaluation questions addressed **Aim 2** (pilot testing quantitative measures of key character outcomes) and **Aim 4** (assessing the relationship between variations in program implementation and poster quality) by examining poster quality. The third set of evaluation questions addressed **Aim 2** by examining pilot quantitative measures' association with pupil participation in Inspire>Aspire. This set of evaluation questions also addressed **Aim 3** by examining findings from the pupil qualitative interviews. The last set of evaluation questions addressed **Aim 4** by exploring the link between teacher implementation and pupil outcomes.

Variation in Program Implementation (Aim 1). To get a general idea of the experience the participating teachers were bringing to Inspire>Aspire, we assessed teachers' experience levels. Teachers' experience varied, both in total experience as well as in experience with implementing Inspire>Aspire. Teachers reported an average of 19 years ($SD = 10$ years) of total teaching experience. On average, teachers have been implementing Inspire>Aspire for 4.0 years ($SD = 1.3$ years).

To assess program implementation, we relied on teachers' responses to the online survey and telephone interview. On average, teachers spent 8.2 hours ($SD = 4.1$ hours) dedicated to Inspire>Aspire poster related activities. Teachers reported dedicating an average of 4.6 hours ($SD = 2.9$ hours) to supplemental Inspire>Aspire activities, and 7.0 hours ($SD = 4.5$ hours) to activities directly linked with the poster during class time. Forty-one percent of teachers assigned supplemental activities as homework. On average, teachers reported assigning 0.8 hours ($SD = 1.1$ hours) of homework for supplemental poster activities, and 1.2 hours ($SD = 1.3$ hours) of homework on activities directly linked with the poster. Teachers, on average, implemented Inspire>Aspire over a period of 8.2 weeks ($SD = 2.9$ weeks).

We also examined the data with a "person-centered" lens, in addition to our "variable-centered" approaches discussed previously. To this end, we conducted a cluster analysis of teacher program implementation using four variables, all of which are described as percentage of time spent on any given activity (the denominator used to calculate the percentages was the total time across all four activities): class time spent on poster, class time spent on supplemental activities, poster assigned as homework, and related supplemental activities assigned as homework.

A three-cluster solution appeared to represent the data most parsimoniously with a clear interpretation of each of the clusters (see Figure 11). Four teachers were missing data on half of the components so percentages could not be calculated for them (accordingly, the cluster analyses had a sample size of 19). The “All Class, (Nearly) All Poster” cluster ($n=4$) consists of teachers who only use class time (e.g., they do not assign poster activities as homework) and spend nearly all (or all) of their class time on the poster (93% on average). Teachers in this cluster spend very little or no time on supplemental activities. All Inspire>Aspire activities are delivered in the class and not as homework. The “Mostly Class, Mostly Poster” cluster ($n=11$) includes teachers who primarily (but not exclusively) deliver Inspire>Aspire during class time (about 75%) and most (but not all) of their time is dedicated to poster related activities (as opposed to supplemental activities; about 65%). Teachers in this group are using a mix of poster-related and supplemental activities both of which are primarily delivered during class time. Teachers in this cluster are also assigning some homework both related directly to poster completion and to supplemental activities. Most of the teachers in this study were in this cluster. The “Mostly Class, Supplemental” cluster ($n=4$) included teachers who primarily deliver Inspire>Aspire during class time (about 95%) and devote considerable time in class to supplemental activities (67%) rather than poster completion.

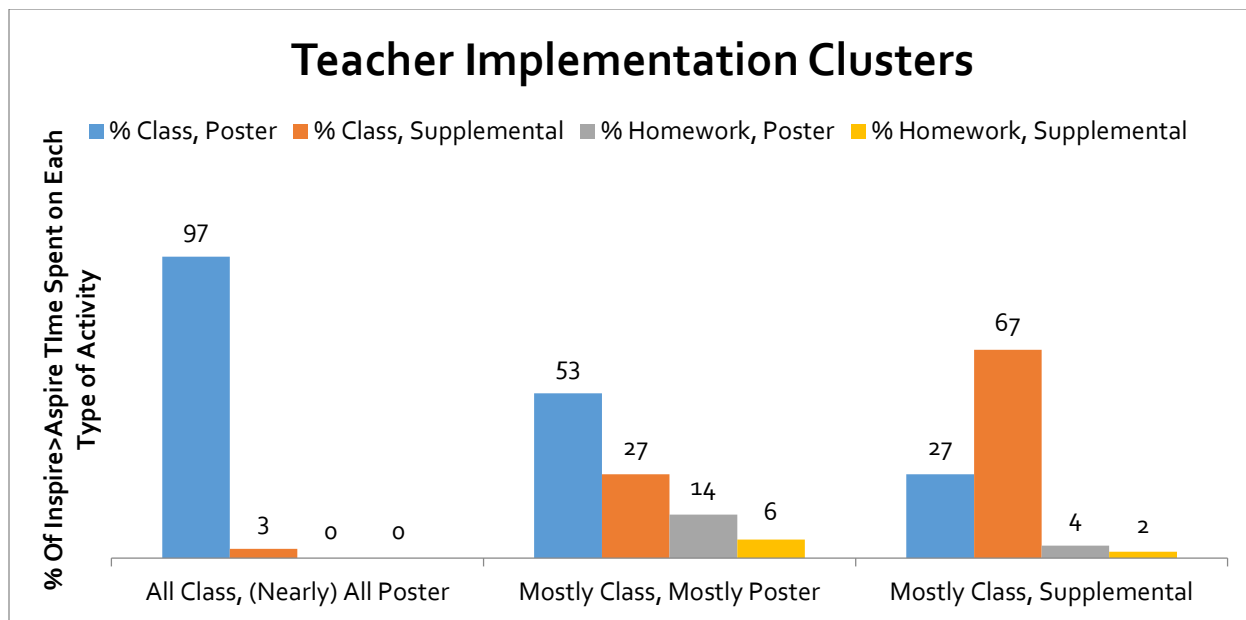


Figure 11. Teacher Implementation Clusters

Qualitative Analysis. In addition to a higher-order analysis of teacher implementation, we also conducted a more detailed content analysis using an open-coding process. There were a variety of reasons that teachers offered for participating in Inspire>Aspire. First, teachers stated that the program gives students a chance to reflect on who they are and what kind of person they would like to be:

Well, the first thing they do is learn about themselves. Very rarely do they get time to think about themselves in this type of way. It makes them realize that, I think some of them don't really think about their strengths enough either, I mean they do it often in class, you know "I do this well, I do whatever" but not about themselves as an individual person. I think it adds to their self-worth, I think it helps them to recognize that we're all individuals and all, I mean we talk about this word "respect" quite a lot, but this takes them far beyond that as well, and makes them really think about themselves a lot more deeply. ...I participate in it because having done it for a number of years now, kids don't often get the chance to really reflect and think about what sort of person they are, what sort of person they want to be, just even thinking about inspiring quotes. They get a chance to talk to each other as well, about things that they don't normally get a chance to. I think the poster is brilliant. I like the way it focuses on values and qualities. I think sometimes we don't talk about that enough, because we're too busy, like in English they might talk about two novels and they might talk about the qualities of the characters and

stuff like that, but they never really get a chance, I think, to go deeply into things that really do affect them....

(T502, 2015)

T506 suggested that the program inspires students to begin to think about their futures:

I think it is a fantastic program for the kids to do and they get a lot from it. ...It provides a lot of variety and also gets them to think about the future, and what they're planning to do in the future and what they would like to be themselves.

(T506, 2015)

Teachers suggested that it offers students an opportunity to investigate values as a vehicle for developing themselves as individuals:

...I like the idea that it maybe contributes to developing character, I think for the kids it largely gets them to think about what is actually important in their lives, sort of values that they may or may not, you know, ascribe to. I really don't think that they're given that kind of an education, to that extent, anywhere else within the school. ...

(T515, 2015)

Teachers also suggested that students' probing of values exists within a larger examination of morals. In the following excerpt, the teacher suggests that Inspire>Aspire can help teachers engage students in lessons around moral issues that make them suitable across and amongst various demographics:

...I brought it into the school because we don't have RE, we don't have Religious Education as part of the time table subjects. And, I felt that I wanted to bring something into school that actually was teaching a little bit about morality. I wanted it to be a secular form so that no particular individual or anyone from a different cultural background, that they weren't comfortable and I found that this program, which to me was all about exploring core values, was able to cut across all barriers in terms of religion, or culture. So that's why I brought it into the school when it was called "Learning for Life" and of course then it developed into Inspire>Aspire. So that's why I brought it into the school.'

(T509, 2015)

The importance of teaching morality is deepened when its applicability is extended beyond the individual and toward a more global perspective. T509 suggests a connection between moral education and good citizenship and implies that a program that asks students to think about issues around morality could help cultivate good individuals and good citizens:

Because we don't have RE (Religious Education). Because I want them to think about what makes a good person. What it is it, and I want them to be good people, and I want them to be responsible citizens, and I want them to be non-judgmental, so because this is the only way I can get any form of morality teaching into them, and if I can get it into them at a young age, because the average age in S2 is 13, I just think I have to do something.

(T509, 2015)

While the previous example references the program's potential to understand values in the service of citizenship, the following excerpt overtly acknowledges the program's ability to support the school's efforts toward cultivating good citizens:

...It fits in really well in terms of what we are trying to do in terms of citizenship, in terms of values and in terms of recognizing. We are really working hard on our pupils on recognizing values and recognized strengths and that sort of thing. It fits in well with what we do. ...

(T514, 2015)

T507 echoes the impact that the program has on global citizenship and its role in helping students understand their place in the world:

I think it is increasingly beneficial for their personal development, it gives them specific space and time to explore who they are and look at their strengths and weaknesses and what they would like to improve on and I think looking at the bigger picture as well, the last page, The Arc of Destiny, I think it really helps them to think about their place in the world. I think this is one area they get the opportunity to do that for an extended period of time and I think that's really valuable.

(T507, 2015)

Teachers identified understanding values, the relationship between good values and good citizenship and the ways in which they can cultivate individuals who recognize these concepts as a significant pedagogical goal. Thus, teachers' responses contributed to an important trajectory: teaching core values as an educational goal → teaching core values as part of a moral education focus → implementing a focus on moral education with the intention of cultivating global citizenry.

Teachers appreciated the way in which the program's focus on core values in all its iterations (moral education, citizenship), which emerged as an important asset to the program, is enhanced by its customizability:

I think it hits all of the core values we are trying to teach them at school. It absolutely, I mean it covers everything, if I were looking for an ideal set of lessons, they are covered by Inspire>Aspire, and you can add to it, you can extend it or you can compress it but the core values are covered in whatever you do. So for me it just sums up what we need to do and character building is very much at the heart of what we want to do in PSE in particular. ...

(T521, 2015)

T521 highlights the flexibility that the program offers teachers who want to enhance or truncate a lesson. This is a particularly important finding within the context of the range of a teacher's experience and expertise. A veteran teacher who has experience navigating certain concepts or a teacher who is limited for time might elect to deliver an abridged version of a lesson. A less-experienced teacher, who might require more support, or a teacher whose students require additional scaffolding could utilize the lesson plan in its entirety.

The program's flexibility is one feature that contributed to the reason why teachers like to use it.

....The other thing as I mentioned is that I like the fact that it is all there. There was very little extra planning that I had to do to make it work terms of because of the nature of my management role. So that was one part of it....

(T515, 2015)

Teachers stated that they ‘cherry-pick’ and often customize the resource offered on the Website based on their own discretion, as well as their population of students:

I’ve looked at them [resources] all and adapted those that I wanted to adapt and I don’t really want to damn anything because I think everything is useful because they’re there to give you ideas. And sometimes I’ve used them, sometimes I’ve not because each year I’ve taught it, it’s been different. And even within this year, I’ve got 3 groups and that’s the most I’ve ever had, and they’re all very different groups. Some things work with some kids and some things don’t work, you know, for dynamics and, you know. So I would say that everything is useful.

(T516, 2015)

The program’s inherent flexibility allows teachers to customize resources provided on the Inspire>Aspire website, It also allows teachers to use materials they create:

Just because we’ve developed our own approach and it works for us, and I think it always depends on the class I think as well, if you have a really good relationship with a class where you can debate and talk about things and they’re quite happy to talk about that, you don’t need resources behind you. If you want a bit more prompts or stimulus to get that discussion going, then absolutely you could use them. So we would take on an adaptation and tailor the needs to the class in front of us.

(T500, 2015)

T500 stated that she crafted her own materials but suggests that resources from the Website could help scaffold a teacher with less experience or offer suggestions that can help to motivate a group of students who require inspiration. Thus, resources provided in the Website can be mined by both experienced and less-experienced teachers.

The website emerged as an important resource not only for teachers but for students, as well:

Yes, the website is still available to the pupils. If they really are toiling to come up with an inspirational figure, we have some pupils, we live in a very, very poor area and our school has a large number of children who receive free school meals and are very poor and some of these kids find it very, very difficult to identify someone who inspires them, other than their mom and dad or grandparents, and for these kids, sometimes we'll sit down with them and look at some of the people who are on the list and see if they can get any inspiration out of them. So although we don't use it for everyone, we do use it for some pupils, the poorer pupils. We really are struggling to get them motivated and thinking what we want them to be thinking.

(T517, 2015)

T517 suggests that the Website is not just a tool for completing a section of the poster but functions as a means by which to make the poster program equitable for students regardless of socioeconomic level. This is a significant finding because it suggests that the program offers opportunities for success for all students through fairness and equity.

Like the program's flexibility, ease of use is seen as another one of the program's assets as suggested in the following three excerpts:

I like the logistics, it's very clearly laid out.

(T505, 2015)

It's easy to deliver, it comes in a box, it's glossy, you know, it looks good for the kids. And every child, irrespective of ability, can gain out of it. I think it gives pupils the opportunity to reflect on not only what's happening here and now, but what are they thinking about for the future? And they don't have to actually speak out, they can write it, which is perhaps more personal and you know sometimes they think they're going to get laughed at or whatever in class if they have ideas about changing the world and what they want to do with their lives. Sometimes it's not cool to be eager to do good things because of your peer culture. I just think it's very good and... is that all? Very good, very good, very good.

(T516, 2015)

I don't need to spend very much time. That is one of the good things about it, the way it is broken up into sections. There is not a lot...an hour I guess for all the lessons, it is very straightforward to deliver. It doesn't need too much preparation.

(T514, 2015)

While some teachers found the program easy to use and stated that they do not spend 'very much time' on either delivery or preparation, T502 suggests that a teacher's familiarity with the program can inform the amount of time that teacher will spend on it:

Well, it depends, it really depends if I've already got an idea for it, I mean if I've taught it last year and it worked really well I'll just go with that. Or, I've got a new idea this year, so I might try something else out, but it's hard to say. Imagine a younger person, like my colleague next year, who's only just started teaching it this year, he's just out of probation and it's his first teaching year, he's teaching it for the first time, he will spend a lot more time than I am only because I've taught it several times and I know it at all so I may be a bit quicker.

(T502, 2015)

It is unclear from T502's excerpt whether 'time spent' refers to preparation time (preparing for the lesson by reviewing the content, gathering supplemental materials, etc), class time (the amount of time spent delivering the content since the speaker's implication is that experience teaching the content makes it easier *and faster* to teach) or both. It is also unclear if T502 created her own materials or if she used (and/or modified) materials from the Inspire>Aspire website. Her utterance does suggest, however, that external resources are an important element, especially for teachers who are using the program for the first time.

Program support through specific guidelines may be as useful as external resources. For example, T524 suggested that the Arc of Destiny is one of the poster's most challenging components to complete.

...all the values stuff is really good, but also the arc of destiny, they think about it quite hard, they find that probably the most challenging section, to write about themselves, and

how they can make the world a better place, they find that challenging. That is often where the most amazing stuff comes up, so probably that, actually.

(T524, 2015)

T524 suggests that this is a challenging element because students begin to consider more deeply their role as global citizens. As a result, it seems imperative to offer teachers guidance for helping them to lead their students through the section. The following excerpt offers a recommendation for one possible strategy:

The way that we do the “Arc of Destiny” is I’ve suggested to the pupils to imagine that they’ve got a magic wand and if they could use that magic wand what would they see themselves doing or what would they do in their future life that would make the world a better place. And I’ve found that this has worked quite well because it gives the pupils sort of like an area to start from, because when they look at the “Arc of Destiny” they just didn’t have a clue, they had no idea what they were supposed to do in that and whenever I tried to say to them, “just imagine yourself in the future, and you can do whatever you want to do, you can become whatever you want to become, and what would it be? And as a result of that, we’ve had over the last couple of years, well since the “Arc of Destiny’s” been put into the poster anyway, we’ve had some really, really good thoughts written down on the page and some really good information written down.

(T503, 2015)

Teachers appreciated the poster for a variety of different reasons.

I think the poster is brilliant. I like the way it focuses on values and qualities.

(T500, 2015)

T500 cited the posters emphasis on ‘values and qualities, as does T522, stating how the poster is a culminating product of students’ values education:

...I see this poster in itself as very, very different because it’s very personal, but it links back, and they’ve already got a good grounding in values education, wanting to be the best, and how they can improve, to look up to people being all they can be, and doing that themselves, so you know within our course 3 mark, we do a lot of things like that, that touches on it, and so, so I’m already finding this throughout other lessons, which, when it comes to, which I think makes the poster such a good project because there sort of almost

there, they get the poster and it's just like putting it all together, like putting the pieces of a jigsaw together, so in terms of what we've already had.

(T522, 2015)

In addition to the poster's positive focus on values, teachers commented on specific sections of the poster. T500 refers to the back section of the poster, which seems to promote prosocial behavior amongst students:

I like the way the poster has a section on the back and which says it asks the pupil to evaluate themselves and then it asks them to ask someone else, like peer evaluation. So I get them to swap posters, to share, to read each other's. Sometimes, if somebody has read a really good one and said, "Oh, now that is really good," they'll say "can I read it."

(T500, 2015)

By inviting students to complete the peer evaluation portion of the poster, students appear to exhibit an interest in their peers' work, expressing their respect and admiration.

While teachers responded positively to the poster, they also offered some critiques. T500 echoes teachers' appreciation of the poster but suggests that its 'format is...limiting,' that 'it could be...more creative' and that 'new versions...need to be drastically different:'

I think the Inspire>Aspire poster is great, I think it's really good. I think some things I think the poster format is maybe quite limiting but I like the structure, it's essential. Sometimes, I think it could be a lot more creative, maybe like to look at that in a bit more moralistic but certainly I think the Inspire>Aspire poster itself is good. When they're bringing in new versions of it, they need to be drastically different.

(T500, 2015)

In order to complete the poster, teachers stated that they engage their students in supplemental activities, such as class discussions. Teachers suggest that the discussions can help students connect with one another:

They get a chance to talk to each other as well, about things that they don't normally get a chance to.

(T502, 2015)

Teachers also suggested that discussions can be a pedagogical approach that invites students who may not otherwise participate to engage in the discussion:

Because I enjoy it, the kids enjoy it, because it gives us structure, and it gives us something to talk about in the school, you know when the pupils have been successful. And some pupils are successful while others are not really making much success elsewhere in the school. And in fact, they're turned off by subjects like math and science. It becomes a nice sense of freedom with the topics, and the free discussion and the fact that they can listen to others and say anything they like, and they can think anything they like, because they can prove that they've reasoned their way to that position. So we spend a lot of time in the first year teaching them the rudiments of arguments and the philosophy introduction that when most kids are in, by the time they get to this they're thinking about making the case or something. They think about actually coming to a conclusion if you like, and they try to support that with reasons and they can weigh it out to traditional and philosophical arguments or not. You can draw a conclusion on your own basis, basically. Sometimes the pupils have sub-equations and sub-reasons to support the sub-conclusions, and it can get pretty messy, but the thinking through everything logically. And for every statement that they make of opinion, they can go right back to their two reasons, or more. I think it does a lot for the kids' self-esteem.

(T510, 2015)

T510 stated that both she and her students enjoy participating in class discussion and that the discussions provide a degree of structure to the lesson. She also suggested that these discussions give students the freedom to formulate arguments and creates an arena within which they need to listen to one another.

Poster Quality (Aims 2 and 4)

Pupils who participated in Inspire>Aspire generally completed the entire poster; 94.4% of pupils completed all elements (1-11) of the poster template (see Figure 12 for a depiction of the

element number assigned to each poster section; see supplemental poster codebook for more information on poster codes and process). We assessed alignment of each element of the poster to other elements of the poster. Alignment refers to how well aligned or “matched” various components of the poster were to each other. For example, we coded how well each of the values and qualities reported in the poster (elements 1-4) were aligned with the description of the inspirational figure selected by the pupil (element 6). All reported alignment scores are the percentage of posters where alignment between the named poster elements existed. Higher percentages indicate that more pupils had posters where the elements were aligned. The alignment scores are reported in Figures 13-19 below.

<p>Values and Qualities: Elements 1-4</p> <p>Active & Confident Individuals</p> <p>Connected & Responsible Citizens</p> <p>Sustainable & Effective Contributors</p> <p>Flourishing & Successful Learners</p>	<p>Inspirational Figure: Element 6</p>	<p>Arc of Destiny: Element 8A What kind of person do you want to be?</p>
		<p>Arc of Destiny: Element 8B What do you want to achieve in your life?</p>
		<p>Arc of Destiny: Element 8C What is your vision for a better world?</p>
<p>Favourite Story or Fictional Character: Element 5</p>	<p>Inspiring Quotes: Element 7</p>	<p>Arc of Destiny: Element 8D What will you personally do to bring this vision to life?</p>

Figure 12. Depiction of the elements of the Inspire>Aspire poster (not to scale)

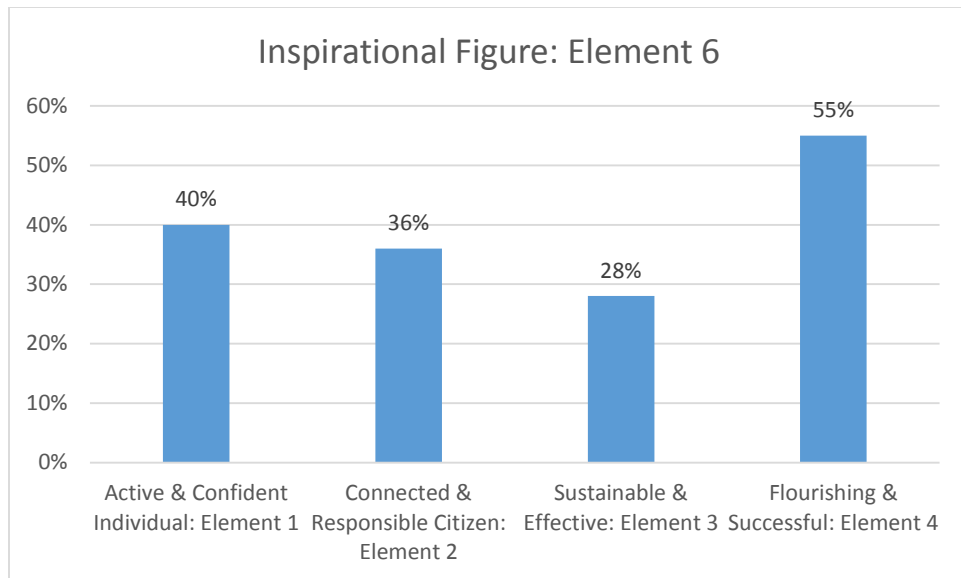


Figure 13. Alignment of Element 6 with Elements 1, 2, 3, and 4

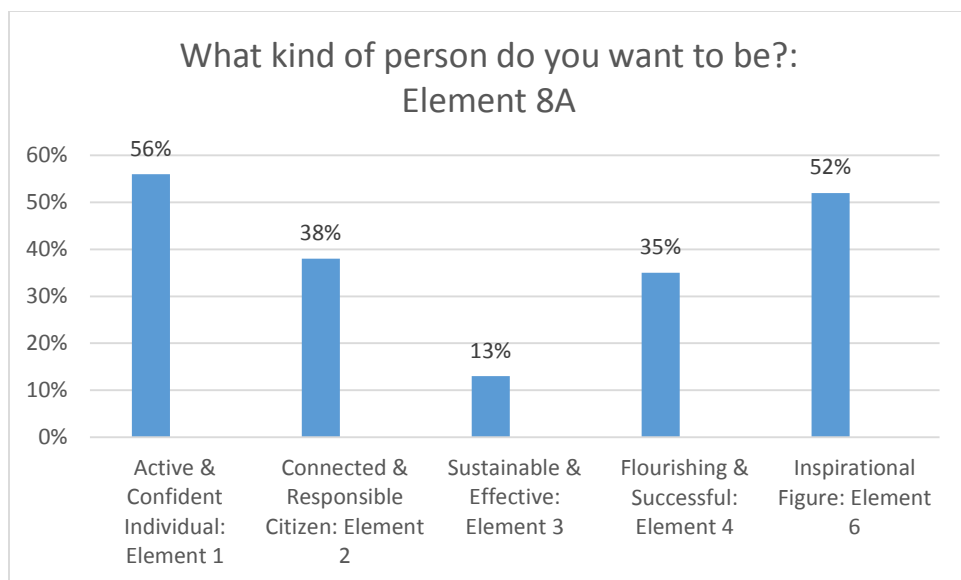


Figure 14. Alignment of Element 8a with Elements 1, 2, 3, 4, and 6

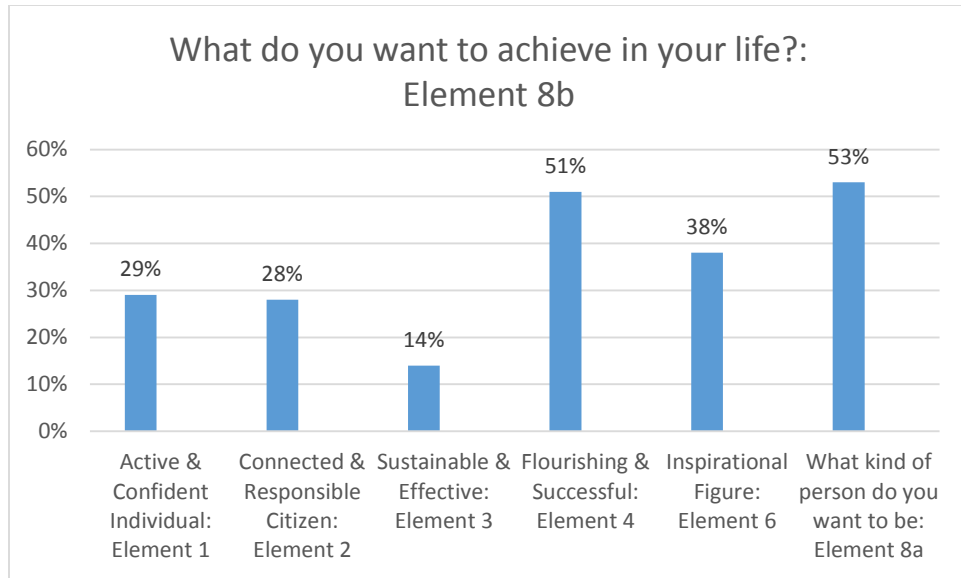


Figure 15. Alignment of Element 8b with Elements 1, 2, 3, 4, 6, and 8a

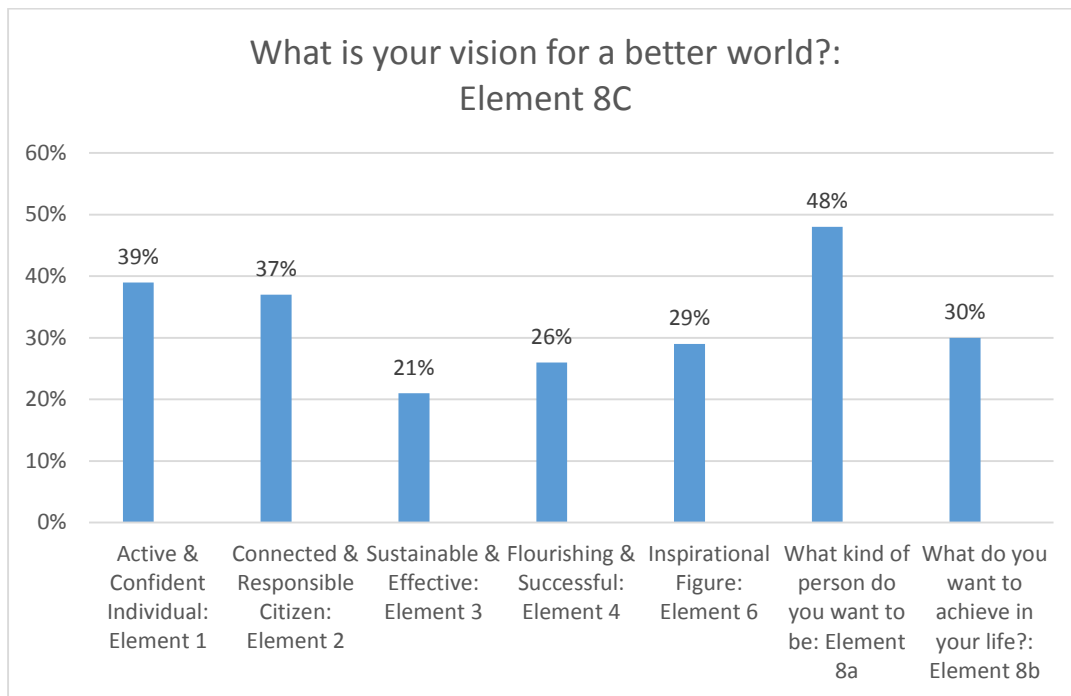


Figure 16. Alignment of Element 8c with Elements 1, 2, 3, 4, 6, 8a, and 8b

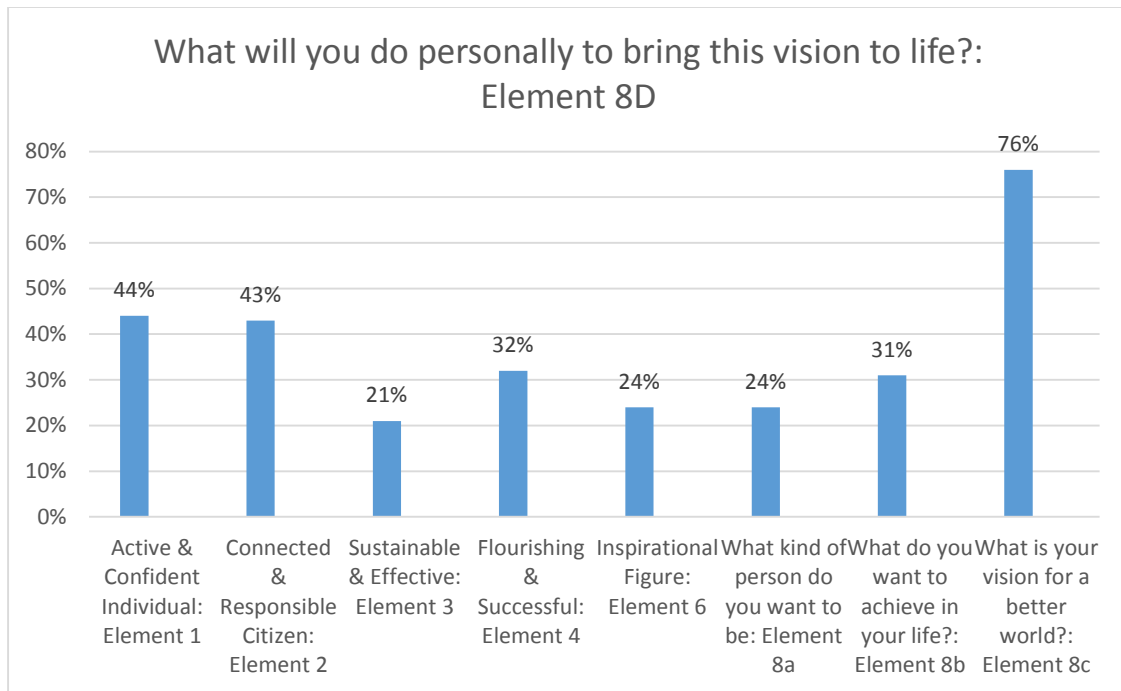


Figure 17. Alignment of Element 8d with Elements 1, 2, 3, 4, 6, 8a, 8b, and 8c

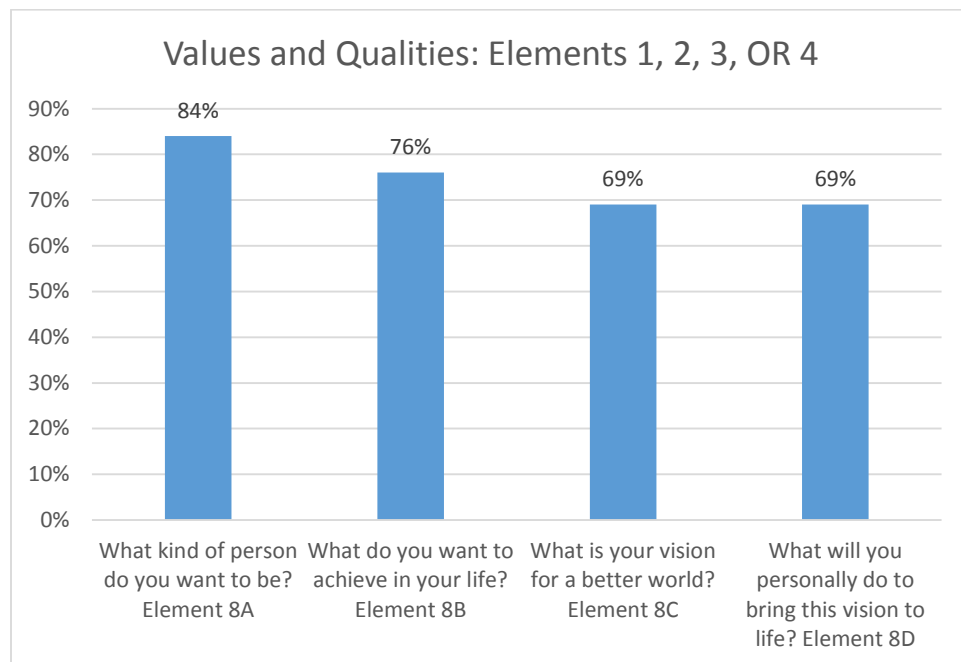


Figure 18. Alignment of First and Last Panels: Element 8a-d separately aligned with at least one of Elements 1, 2, 3, OR 4

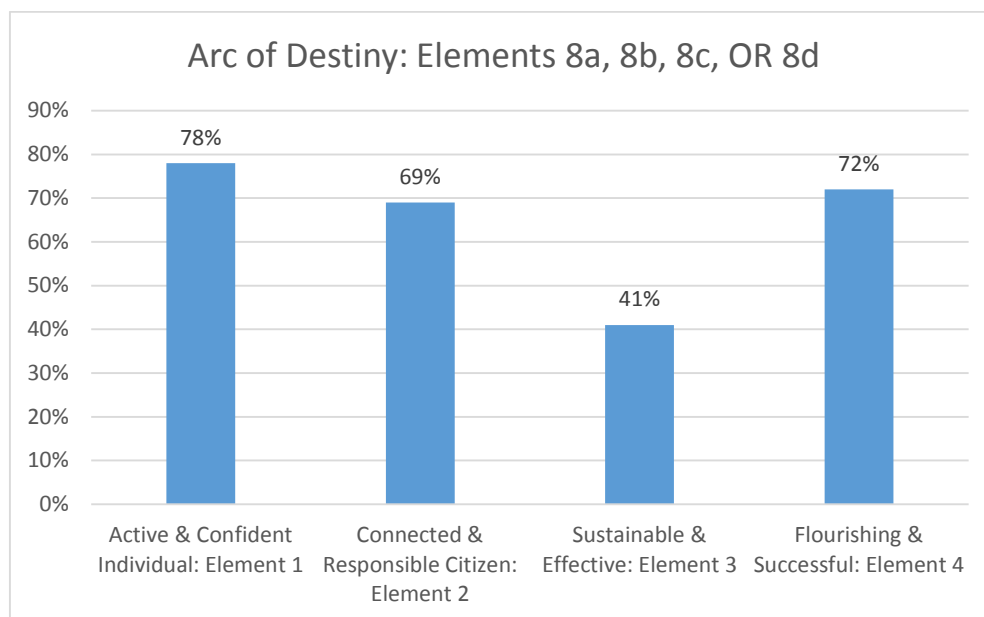


Figure 19. Alignment of First and Last Panels: Elements 1, 2, 3, & 4 separately aligned with at least one of Elements 8a, 8b, 8c, OR 8d

Most notably, there was good alignment between pupils' vision for a better world and what they will do to personally bring this vision to life (76%) indicating a clear connection between what youth envision and what they plan on doing personally to enact that vision. Even with this high percent of alignment between these two poster elements, 24%, or 48 pupils did not demonstrate alignment of these elements. Further analysis indicates that 28 of the 48 pupils showed no alignment between 8d (what they will do to personally bring this vision to life) and any other element (e.g., the virtues, inspirational figure, or other last panel elements). Twenty of the 48 pupils showed some alignment with at least 1 of the other elements (element 1, 2, 3, 4, 6, 8a or 8b).

There was also evidence of alignment between the inspirational figure and the type of person the pupils want to be in the future (52%). The general summary alignment scores (Figures 18 and 19) indicate strong alignment between values selected and future aspirations. These

summary alignment scores ranged from 41% (for the alignment of sustainable and effective

values with any of the Arc of Destiny [future aspirations] elements) to 84% (for the alignment of the future aspirations for what kind of person the pupil wants to be and any of the values selected). Most of these summary alignment scores were in the high 60% to mid-70% range.

Pupil Poster Quality Profile Analysis. Poster quality was defined and assessed in several different ways: (1) Degree of alignment between the first and last poster panel; (2) Degree of poster consistency with principles of positive youth development and Inspire>Aspire (i.e., language consistent with the 5 Cs was used to describe the inspirational figure, the figure was described as inspiring, contribution was mentioned in the third panel); (3) The extent to which the pupil gave thoughtful and well developed responses to the reflection on personal strengths and areas in need of improvement (values); (4) The extent to which the pupil gave thoughtful and well developed responses to the fictional story, inspirational figure, and inspirational quotes sections of the poster; and, (5) The extent to which the pupil gave thoughtful and well developed responses to the questions regarding future aspirations (third poster panel). As expected, these measures of quality are correlated (pupils who score high on one measure of quality tend to score high on other measures of quality). In particular, the alignment between the first and last panels, and the development of aspirations, are fairly highly correlated (see Table 3). However, there is variation between pupils in terms of which of these poster quality elements is most prominent. Therefore, we conducted a profile analysis of pupil poster quality, excluding the alignment quality variable.

Table 3. Intercorrelation of Poster Quality Scores

	AlignFirstLastPanels	ConsistPYD	ValuesDev	FicQuotesFigDev
AlignFirstLastPanels	1			
ConsistPYD	.424**	1		
ValuesDev	.228*	.293**	1	
FicQuotesFigDev	.448**	.539**	.488**	1

AspirationsDev	.656**	.383**	.484**	.517**
----------------	--------	--------	--------	--------

This person-centered analysis of pupil poster data was computed using four of the poster quality variables (Consistent with PYD/Inspire>Aspire; Values Well Developed; Fictional Quotes Well Developed; and Aspirations Well Developed). A four cluster solution had the best fit with the data and the following cluster labels were selected to describe the four poster quality profiles (Figure 20): High Aspirations ($n = 63$); Consistently Average ($n = 80$); Low ($n = 41$); and High PYD ($n = 14$). Pupils in the High Aspirations cluster had posters that scored relatively high in terms of the development of their responses for the third poster panel (future aspirations). Pupils in the Consistently Average cluster had posters that were generally rated as average in quality across all of the measures of poster quality. Pupils in the Low cluster had posters that were generally rated as low in quality across all of the measures of poster quality. Finally, pupils in the High PYD cluster had posters that scored high in terms of their alignment with PYD and Inspire>Aspire principles. Interestingly, pupils in this profile also had posters that scored high in terms of the development of their responses in the third panel (future aspirations).

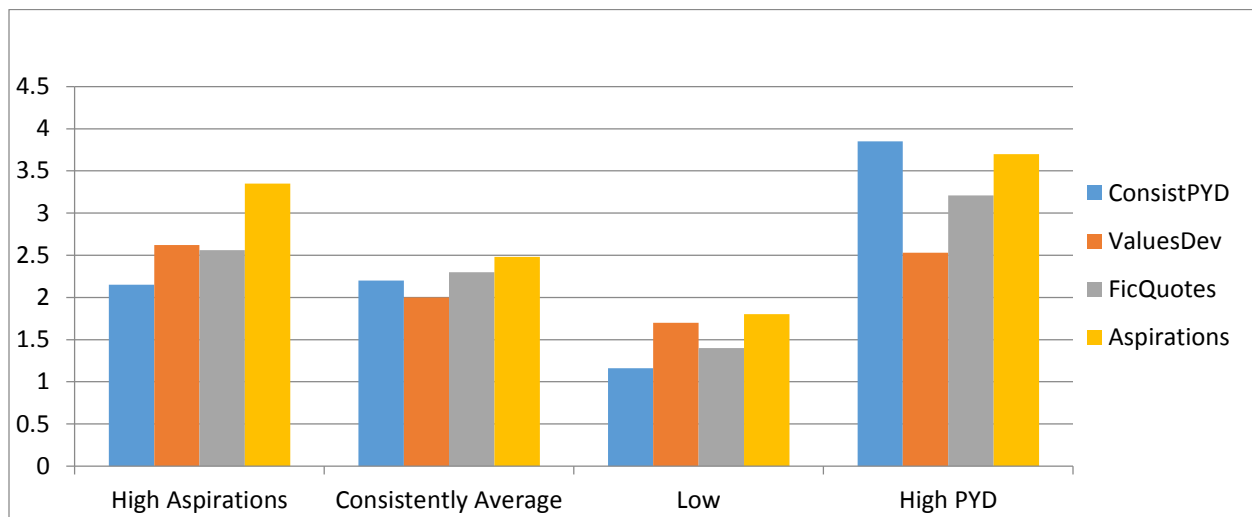


Figure 20. Profiles of Pupil Poster Quality

We also looked at the relationship between pupil poster quality profile membership and teacher variables (Aim 4). The results of these analyses are discussed below.

Pupil Outcomes: Comparing Wave 1 and Wave 2 Surveys (Aim 2)

Analysis of pupil survey data revealed very little change in outcome scores. The constructs measured at pre-test and post-test appear to be very stable; there is not much change being reported by participants. Specifically, there is no statistical change in mean scores for goal selection (as measured by the SOC-9), Positive Youth Development subscales or full scale, sense of purpose, or future orientation. There were statistically significant mean-level differences between Wave 1 and Wave 2 for some subscales of the Importance of Future Aspirations scale. A paired *t*-test revealed the Self-Acceptance subscale of the Importance of Future Aspirations scale is significantly higher at Wave 2 ($t = -2.91, df = 107, p < .01$); The Affiliation subscale demonstrated a trend-level increase from Wave 1 to Wave 2 ($t = -1.8, df = 107, p < .10$). All other future aspirations total and subscale scores showed no statistically significant difference between Wave 1 and Wave 2.

We also examined differences between Wave 1 and Wave 2 survey outcome scores by pupil poster profiles. Very few differences were found in these subgroup analyses as well. Figure 21 illustrates Future Orientation scores at both waves by poster profiles. It appears there is a slight decrease between Wave 1 and Wave 2 for the “High PYD” subgroup, but it is important to keep in mind the sample for this subgroup in this analysis (e.g., those pupils who participated in both Wave 1 and Wave 2 of the survey who were members of that profile) is very small ($n = 8$).

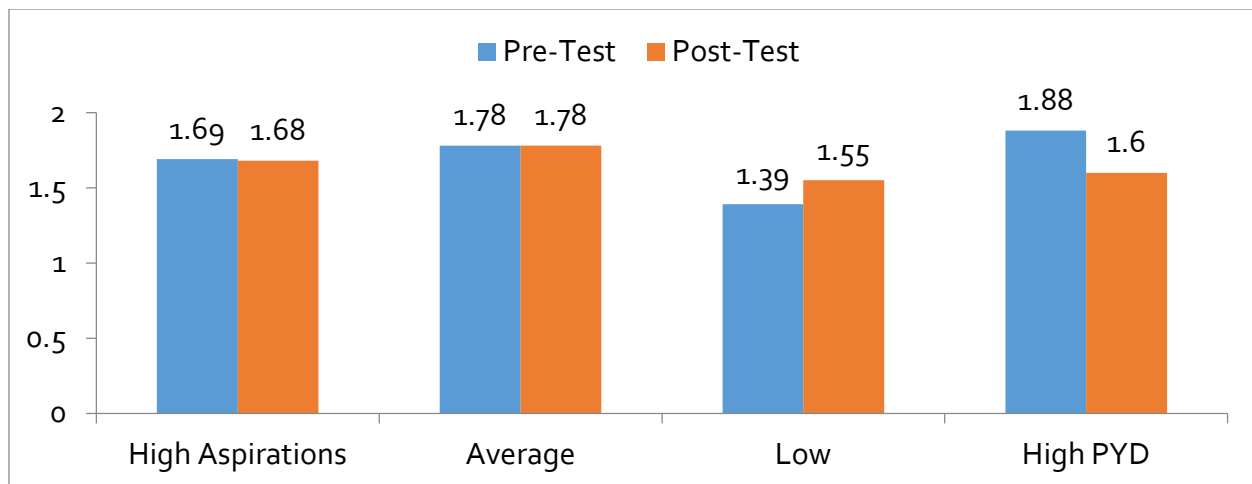


Figure 21. Future Orientation Scores by Profile

Figure 22 shows scores by profile for the five purpose items (measured on a 0-6 scale): Help others, Make the world a better place, Do the right thing, Discover new things about the world, and Support my family and friends. None of the paired-sample *t*-tests were statistically significant and the overall purpose scores were high for all groups at both time points indicative of ceiling effects.

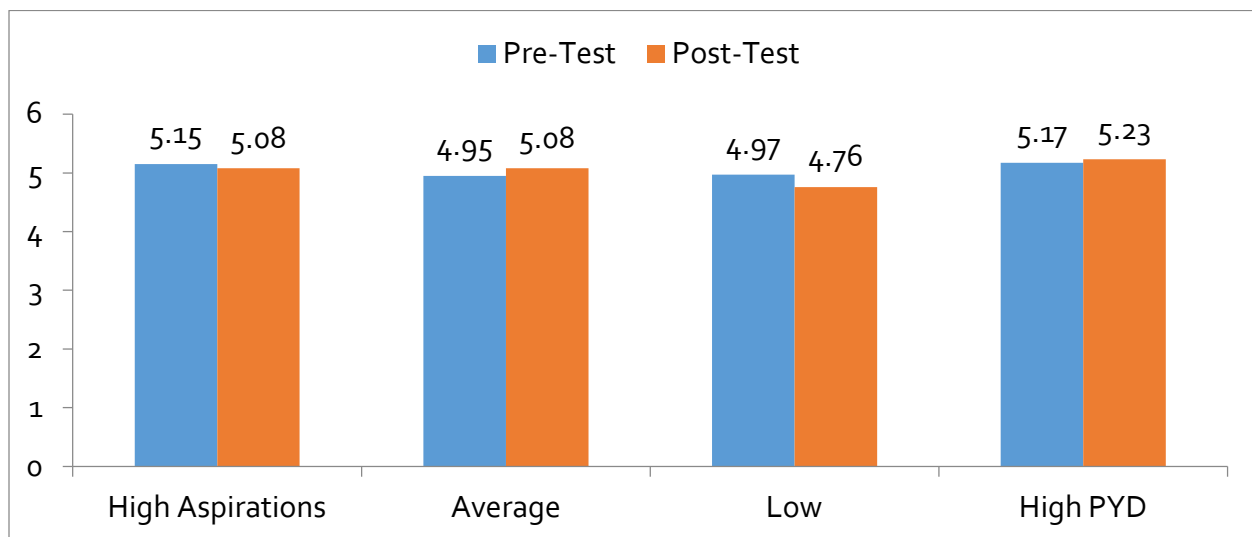


Figure 22. Purpose Scores by Profile

Figure 23 illustrates the scores for the 9-item composite of the SOC scale. Comparing pre-test and post-test scores for SOC must be done with caution because the psychometric analyses showed that the magnitude of factor loadings was not the same for the items across the two time points (e.g., the loading for item 1 at pre-test was not equivalent to the loading for item 1 at post-test). This means that at both time points, the items are measuring a general “SOC” factor, but that the importance of specific types of items may be different at the two time points. For example, SOC at pre-test had the highest factor loadings for “I think about exactly how I can best realize my plans”, “I make every effort to achieve a given goal”, and “I always pursue goals one after the other.” At post-test, the items with the highest loadings were “I make every effort to achieve a given goal”, “When I decide upon a goal, I stick to it”, and “When things don’t work the way they used to, I look for other ways to achieve them.” Even though different items were more important for SOC at pre-test compared to post-test in the latent framework of CFA, we created scale scores (in which all items are treated equally). Thus, the scale scores do not entirely accurately represent the importance of all of the questions at each time point. The paired-samples *t*-test for the comparison in the Low profile was statistically significant and represents a decrease –albeit a very small one – from pre-test to post-test.

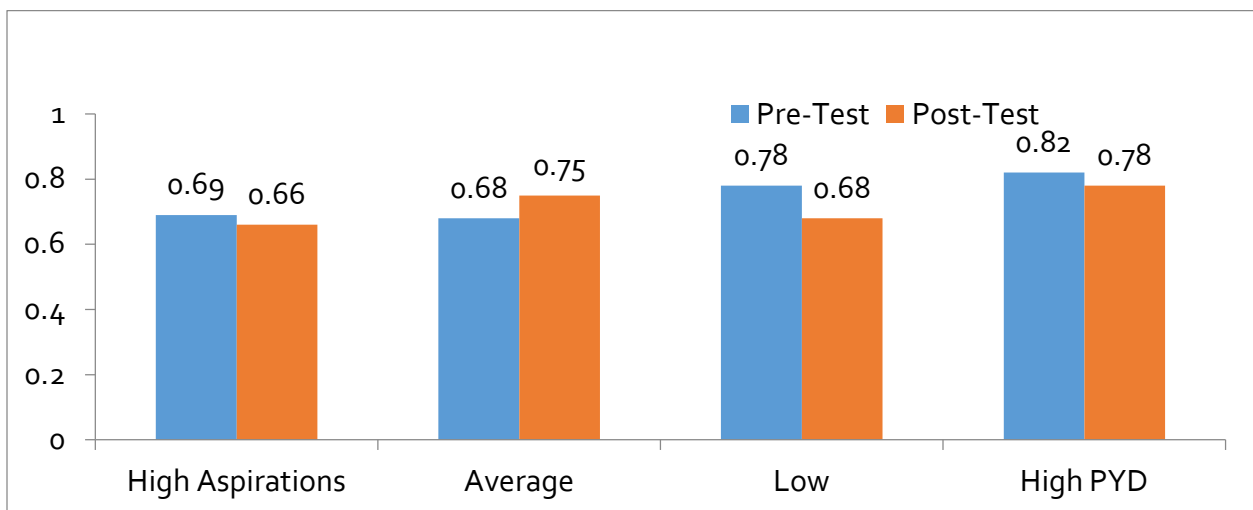


Figure 23. SOC Scores by Poster Profile

Figures 24 through 27 show the average scores for each profile group on the four subscales of the Aspirations – Importance scale. Figure 24 illustrates the results for Aspirations – Importance Self-Acceptance. All groups showed increases in average scores although none were statistically significant.

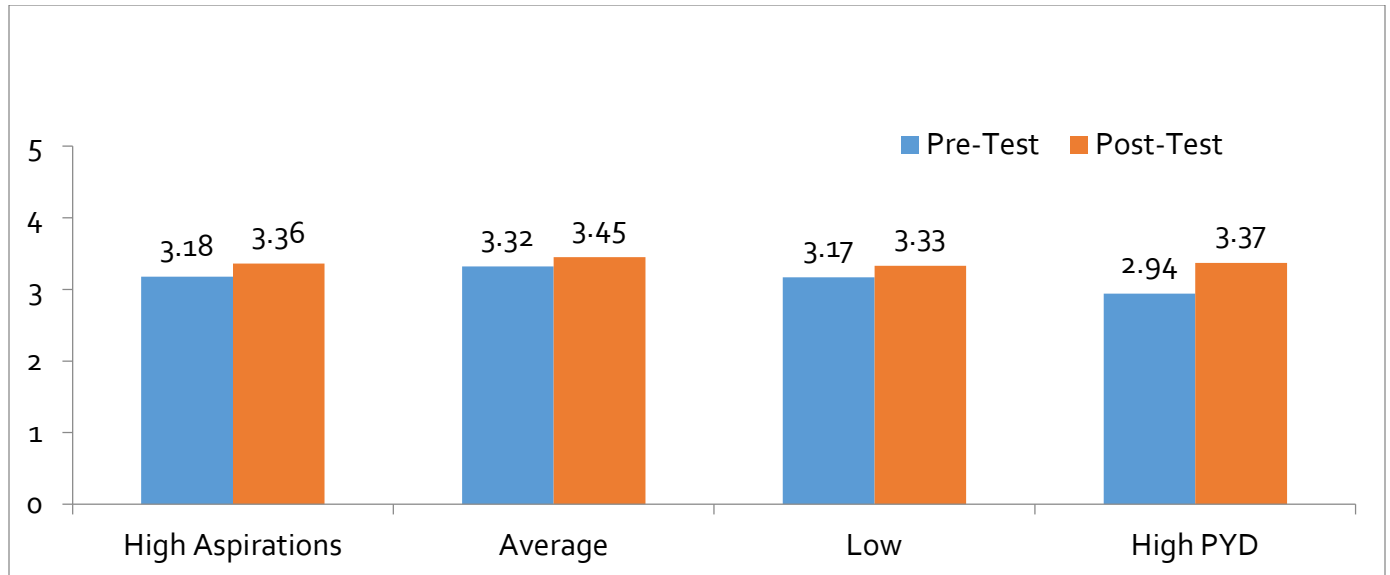


Figure 24. Aspirations - Importance Self-Acceptance Scores by Profile

Figure 25 shows the results for Aspirations – Importance Affiliation. None of the paired *t*-tests were statistically significant.



Figure 25. Aspirations - Importance Affiliation Scores by Profile

Figure 26 shows the results for Aspirations – Importance Community Feeling. None of the paired *t*-tests were statistically significant.

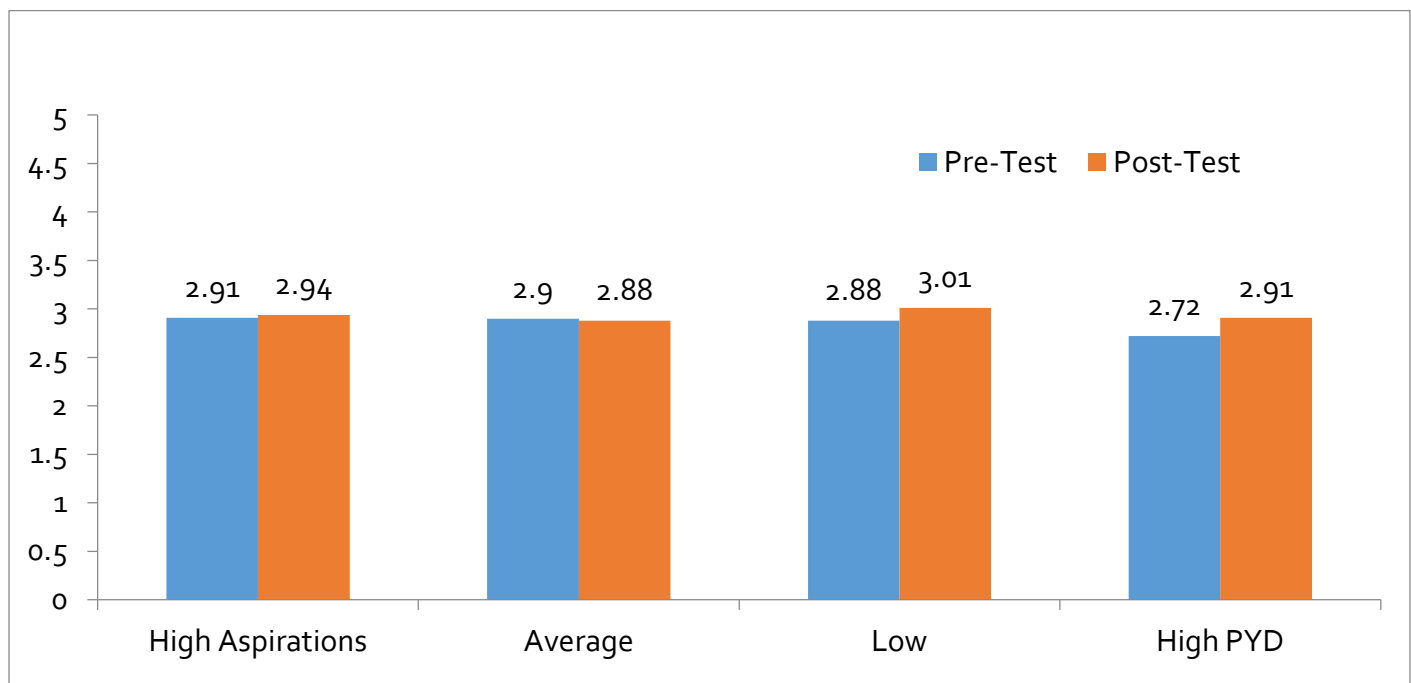


Figure 26. Aspirations - Importance Community Feeling Scores by Profile

Figure 27 shows Aspirations – Importance Financial Success. None of the paired *t*-tests were statistically significant.

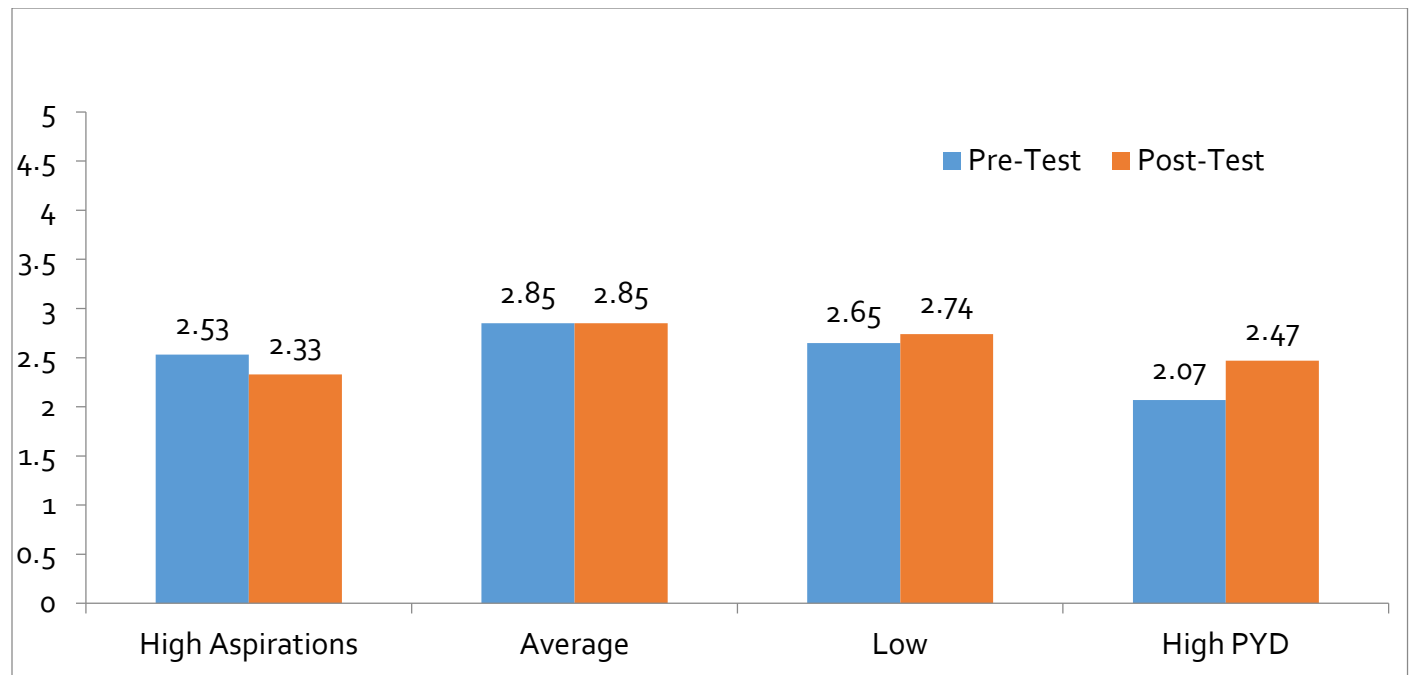


Figure 27. Aspirations - Importance Financial Success Scores by Profile

Figure 28 shows the results for Aspirations – Chances Self-Acceptance. The increase in scores in the High PYD profile was significant at $p < .10$ (with 8 participants in that group).

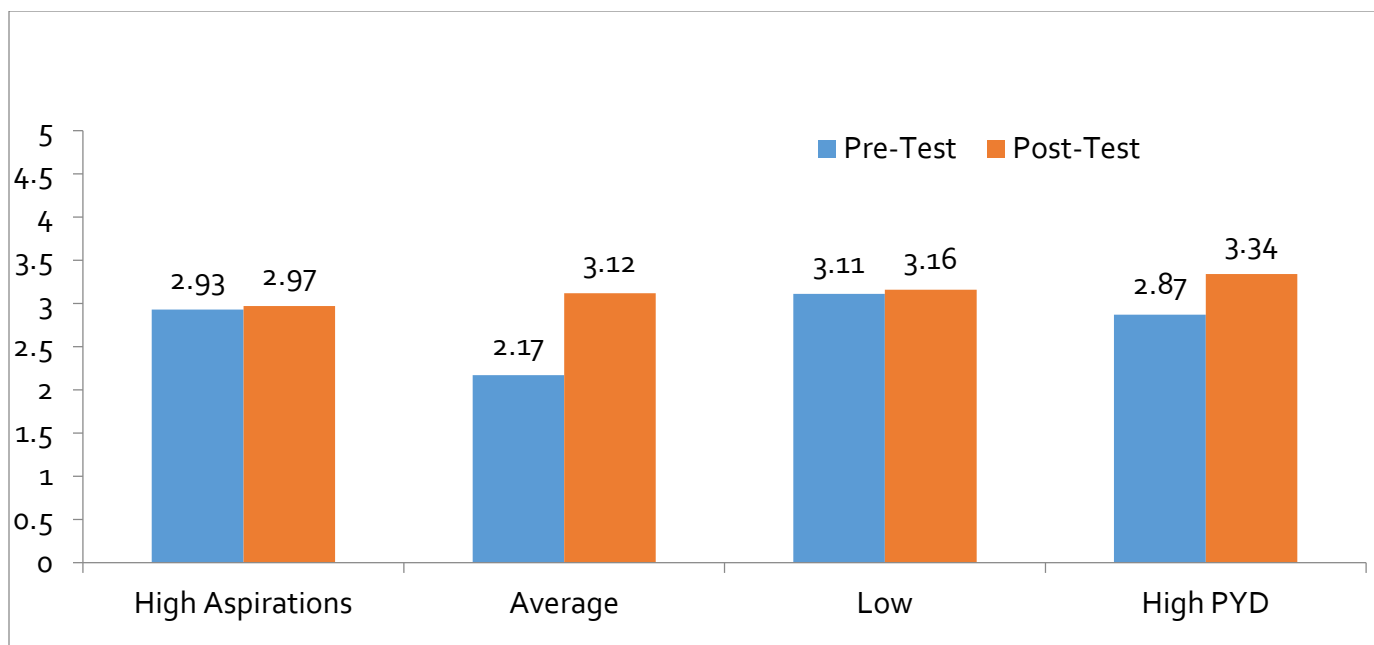


Figure 28. Aspirations - Chances Self-Acceptance Scores by Profile

Figure 29 shows Aspirations – Chances Affiliation scores. As with the subscale of Chances Self-Acceptance, the increase in scores in the High PYD profile was statistically significant at $p < .10$ (with 8 participants).

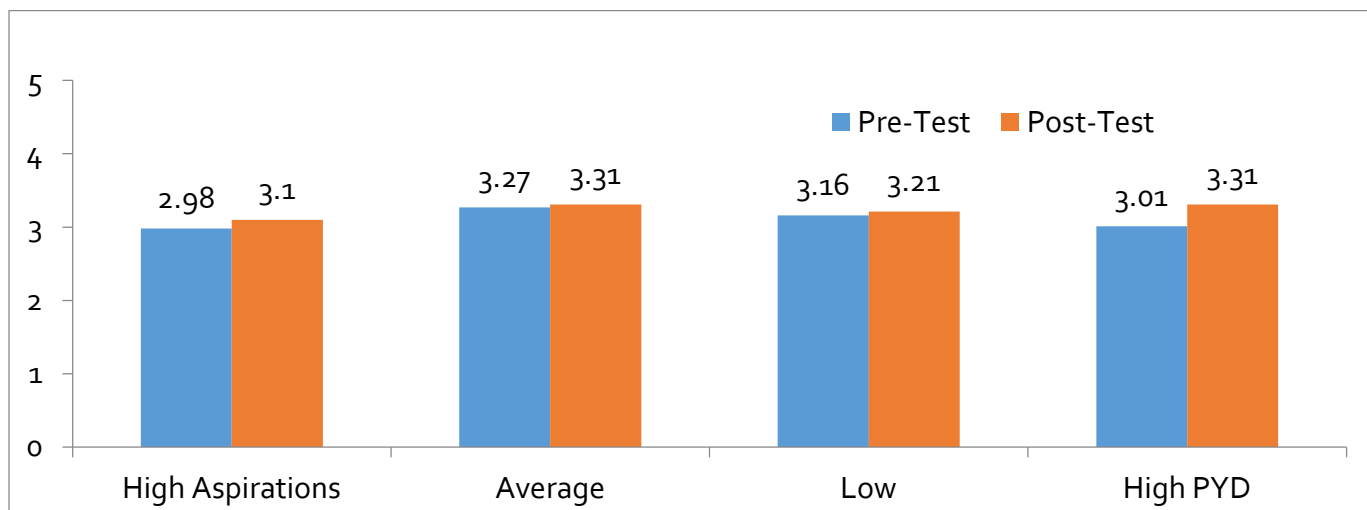


Figure 29. Aspirations - Chances Affiliation Scores by Profile

Figure 30 shows scores by profile for Aspirations – Chances Community Feeling. None of the paired samples t -tests were statistically significant.

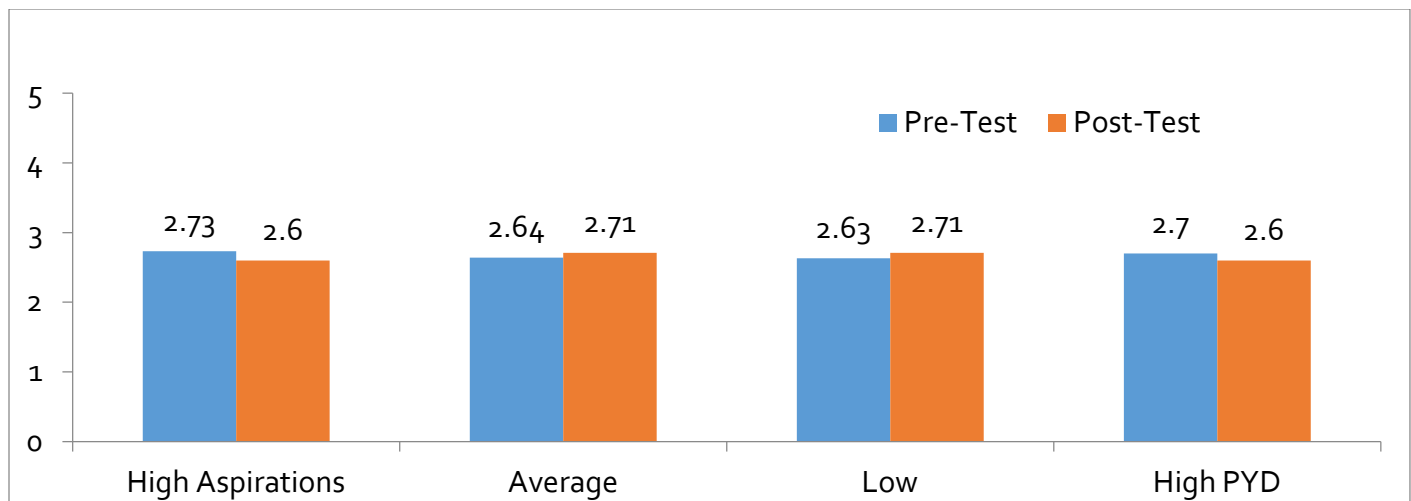


Figure 30. Aspirations - Chances Community Feeling Scores by Profile

Finally, Figure 31 shows scores for Aspirations – Chances Community Feeling by profile. The increase in scores for the High PYD profile was statistically significant at $p < .05$.

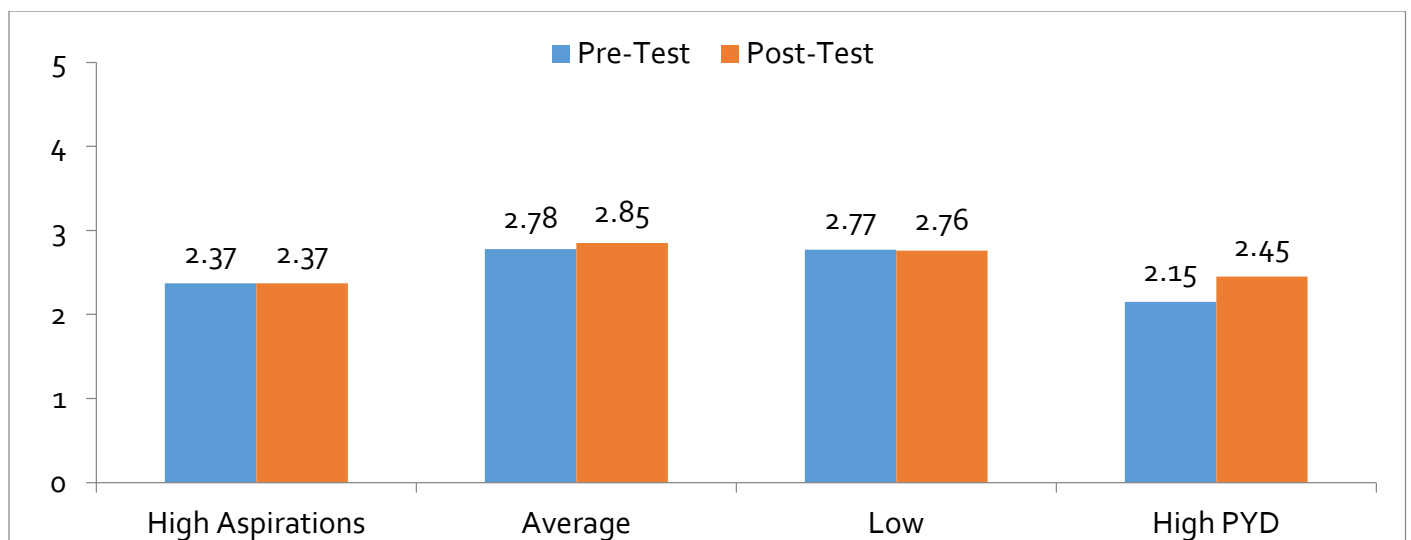


Figure 31. Aspirations - Chances Financial Success Scores by Profile

There are several potential reasons for the general stability of pupil survey outcome constructs, including measurement error (which introduces noise into the analyses and makes change more difficult to detect). With a larger sample size, these analyses could be conducted within a latent variable framework that would help account for measurement error. It is also

possible that the measures are not sensitive to changes in the actual constructs, or that the time interval between pre-test and post-test was not long enough to detect the change.

Accordingly, without much change from pre-test to post-test, there is not much that can be predicted by the quality of the poster that the participant made (regardless of whether it is defined by the poster profile membership, broad quality score, or alignment score, or even some other measure that was not included in these analyses).

Table 4 below shows the correlations between Quality and Alignment and the pre-test variables. Pre-test purpose scores were positively associated with poster Alignment scores. Poster Quality scores were negatively associated with both Importance and Chances for Financial Success subscales of the Aspirations Index.

Table 4. Correlations among Poster Quality and Pre-Test Variables

Pre-Test Variable		Quality	Alignment
SOC	Pearson Correlation	-.015	.141
	N	95	95
Purpose	Pearson Correlation	.078	.209*
	N	98	98
Aspirations – Importance Self Acceptance	Pearson Correlation	-.076	-.047
	N	98	98
Aspirations – Importance Affiliation	Pearson Correlation	-.020	.044
	N	98	98
Aspirations – Importance Community Feeling	Pearson Correlation	.071	.162
	N	98	98
Aspirations – Importance Financial Success	Pearson Correlation	-.236*	-.180
	N	98	98
Aspirations – Chances Self Acceptance	Pearson Correlation	-.097	-.076
	N	97	97
Aspirations – Chances Affiliation	Pearson Correlation	-.123	.058
	N	98	98
Aspirations – Changes Community Feeling	Pearson Correlation	.100	.172

	N	98	98
Aspirations – Chances Financial Success	Pearson Correlation	-.310**	-.172
	N	98	98
Future Orientation	Pearson Correlation	.179	.123
	N	80	80

Table 5 shows the same relations except with the post-test variables. The same pattern was observed as for the pre-test variables.

Table 5. Correlations among Poster Quality and Post-Test Variables

Post-Test Variable		Quality	Alignment
SOC	Pearson Correlation	.028	.090
	N	85	85
Purpose	Pearson Correlation	.184	.227*
	N	85	85
Aspirations – Importance Self Acceptance	Pearson Correlation	-.005	.072
	N	85	85
Aspirations – Importance Affiliation	Pearson Correlation	-.168	-.076
	N	84	84
Aspirations – Importance Community Feeling	Pearson Correlation	-.053	-.023
	N	84	84
Aspirations – Importance Financial Success	Pearson Correlation	-.215*	-.023
	N	85	85
Aspirations – Chances Self Acceptance	Pearson Correlation	-.029	.137
	N	85	85
Aspirations – Chances Affiliation	Pearson Correlation	-.026	.067
	N	85	85
Aspirations – Chances Community Feeling	Pearson Correlation	-.055	-.012
	N	85	85
Aspirations – Chances Financial Success	Pearson Correlation	-.231*	-.136
	N	85	85
Future Orientation	Pearson Correlation	.087	.077
	N	81	81

Pupil Outcomes: Comparing Wave 1 and Wave 2 Interviews (Aim 3)

We conducted a qualitative content analysis by coding 24 student pre- and post-test interviews using both a priori codes derived from the literature and the Pathways Model (e.g., Goal Setting, Future Mindedness, Caring/Compassion), as well as codes that emerged organically from the data set (Knowledge of Values, Virtues and Qualities; Low-level Caring/Compassion; Self-care). Some of the codes fall into natural clusters, so they have been grouped accordingly. For example, Confidence, Connection, Competence, Caring/Compassion and Character constitute the category called the 5 Cs (Lerner et al., 2005), and the precursor categories to the 5 Cs that emerged during open coding have also been grouped. Goal Setting is a necessary precursor to Future Mindedness, so these have been grouped together. We coded some of the questions in pairs (1 & 2, 3 & 4, 6 & 7, 9 & 10, 12 & 13, 15 & 16, 19 & 20) since there was tendency for students to (1) begin answering the second question in the grouping in response to the first question or (2) require prompting or *think time* while formulating the response to the first question, thus answering in the second part. Therefore, we compared each of these groups pre-post in order to capture, as accurately as possible, any change that had occurred and to avoid penalizing students for offering their response earlier or later within the grouping. (For the sake of clarity, only those questions within each grouping that offered a substantive response were included in the following narrative.) We coded questions 5, 8, 11, 14, 17 & 18 as stand-alone items.

There were instances in which a negative code would not have properly represented the respondent's utterance. For example, in the following example (Y001, Q11), the fall transcript was coded for academic competence. However, the student does not offer an utterance that would qualify as academic competence in the spring.

What is one concern or a self to-be-avoided in the next year?

To be honest, I don't want to fail most subjects that I'm enjoying; I don't want to be the person with the lowest confidence in the world.

(Fall, Q11, Y001)

What is one concern or a self to-be-avoided in the next year?

To not be living a healthy lifestyle, I would like to live a healthy lifestyle.

What about this is important to you?

I don't want to go down the stages where I can't do much sports and everything, I'd like to stay active.

(Spring, Q11, Y001)

If the spring utterance had been coded as negative for academic competence, it would potentially misrepresent the student's utterance, which may simply be a shift in focus. Since the student does not mention the academic component again in the spring and we cannot assume the student is not thinking about it, the student earns a no change for having said it in the fall. In addition, we coded each utterance for its adherence to the definition; we did not code the quality of the goal. For example, if a student said they wanted to be a badminton player, we did not judge that as a good or a bad career choice. If in the next or prior interview the student stated he or she wanted to be a physicist, the utterances were coded as No Change. We were interested qualitatively in the steps, thought and/or preparation that a student has made or plans to make with respect to their goal.

Goal Setting and Future Mindedness. Our definition of goal setting was consistent with the Selection, Optimization, and Compensation (SOC) model (Freund & Baltes, 2002). Selection (S) involves the development of and commitment to goals and includes Elective Selection (ES), or the focus on attaining a desired state, and Loss-Based Selection (LBS), or the restructuring of a goal system when there is a loss in goal-relevant means. Optimization (O) involves the actions

taken to reach goals and Compensation (C) involves taking alternative actions to reach goals when there is a loss in goal-relevant means. Although LBS and C both refer to the actions taken after goal-relevant means become unavailable, LBS indicates a change in goals, while C indicates a change in means of reaching the original goal (Freund & Baltes, 2002).

There were many more instances of positive Goal Setting than Negative Goal Setting or Goal Setting - No Change. The instances of Goal Setting that were coded often exhibited plans for optimization. There was only one example of goal selection, one example of loss-based selection and one utterance that exhibited the potential for loss based-selection. The larger variety and distinction amongst these utterances is in the quality. In the following excerpt, the student's response lacks specific objectives for how to attain his goal:

[1st Goal] What are you doing this year to attain or reach that goal?
I'm trying hard to do it and yea.

(Fall, Q7, Y033)

In the spring, the student suggests more concrete strategies for attaining his goal:

[1st Goal] What are you doing this year to attain or reach that goal?
I'm trying to, like, be nice to other people and try to talk to them more and help them. I try to just save my money so I can go to cinema and stuff with them.

(Spring, Q7, Y033)

We coded this positive for Goal Setting because one of the student's goals is concerned with a connection to and interaction with others. If, however, the student had simply responded with, '*I try to just save my money so I can go to cinema and stuff with them,*' we might have appreciated the increased level of specificity but taken into greater consideration the *quality* of the response. Compared against the weak response in the fall, if the latter portion had been the only response to this question, this may have been coded 'No Change.' Ultimately, both the fall and the spring responses were deemed weak.

There are instances in which No Change can be assigned to two qualitatively comparable responses. The following grouping was coded as No Change but simply represents a shift in focus from academic to social. In this excerpt, the student's goal is academic in nature:

[1st Concern] Are you currently doing something so this will not happen next year?
Just trying to keep up with like, make sure you're still doing things like handing your homework in on time and things like that.

[1st Concern] What are you doing this year to reduce the chances this will describe you next year?
Is there anything else you're doing?
Just in general keep doing stuff and not get sloppy or anything.

(Fall, Qs 12 & 13, Y037)

In the spring, the student offers an answer of equal quality in the realm of goal setting but focuses on the social:

[1st Concern] Are you currently doing something so this will not happen next year?
Just trying to be nice.

[1st Concern] What are you doing this year to reduce the chances this will describe you next year?
I try to treat people how I would like to be treated.

(Spring, Qs 12 & 13, Y037)

It is difficult to account for the shift from academic to social or vice versa. It is important to recognize that these utterances are subject to the personal, context-specific experiences of the speaker. Therefore, there are numerous explanations for a student's particular focus.

There were many instances of positive Goal Setting, and these were frequently accompanied by strategies for optimizing the goal. In the following excerpt from the spring, optimization is present but may not yet be fully realized.

What do you hope to accomplish in the next two years?

Um I'd definitely like to, cause with my drama group I'm trying to aim for a grade thing, so yeah. I'm just trying to aim for this school grade thing that I can get to help me in college. That would be of a bit of a kind of stamp of like "I'm good at this" and that can help me out in the future. So I'm trying to achieve that.

You said grade thing?

Yeah, it's like, it's a grade for drama. I've been trying to reach it for a few months now, but I'm doing the actual performance thing next year with the drama club and if I get a good judge, a kind of good mark on it, then I'll get a grade and if it's good then it can help me out in the future.

So this grade thing that you're talking about, somebody judges your performance and then if you get a good mark...

Yeah.

What happens if you get a good mark?

Uh I get kind of like a grade thing I get to put my mark on my CV and like college stuff. I'm not sure exactly how it works yet, but I'm sure, my drama teacher told me it would help me out with like jobs and college.

So why is that important to you?

Cause I do want to be an actor when I'm older. Maybe something in Doctor Who, the TV show. That would be pretty cool.

Do you have another goal for the next two years?

Uhh not really. I guess it's kind of to be happy and less stressed as I said, but yeah not really.

(Spring, Q17, Y028)

There were also instances of optimization in utterances that were coded Negative and No Change for Goal Setting. This is a significant finding because it suggests that optimization is emergent. However, because it is still in its early stages, it may not manifest on a quantitative measure. This is thematically consistent with the different stages that students seem to occupy for a variety of the categories that we coded. It is unclear, however, whether positive, negative or neutral changes can be attributed to Inspire>Aspire or if they are a result of maturation.

In the following excerpt from the fall, the student states a clear goal of trying to learn in school and from his teacher:

[1st Goal] Are you currently working on that goal or doing something about that expectation?

Well basically just going to school is kind of trying, you know, to learn information. Just trying to pick up as much info from the teacher as I can.

(Fall, Q6, Y014)

In the spring, the student's response offers a more focused, elaborated goal and delineates a concrete strategy for optimizing that goal:

[1st Goal] What are you doing this year to attain or reach that goal?

Just going to school or doing outer school activities and just trying to maybe do, if I don't feel like I'm keeping up on something maybe doing something that I wasn't told to do, like just look at it again or do another task on it maybe.

And is there anything else you're doing to reach this goal?

Probably I go to a like a, I don't know if you have heard of it but air-cadet like air basically but cadet and that just helps me kind of because they teach you to be just like to listen and good kind of values so they teach you all that kind of things, so that's probably something.

Air cadet (flying basically)

(Spring, Q7, Y014)

By virtue of the definition of Future Mindedness, instances of the latter could not be coded without instances of the former. Responses that reflected Future Mindedness articulated a clear pathway with at least 3 steps:

...if you don't do well on your exams you cannot go to university and if you can't go to university, it is harder to get a job.'

(Spring, Q18, Y68)

Fifteen of the 24 interviewees formulated an utterance that captured Future Mindedness.

Coding for Goal Setting and Future Mindedness when grouped revealed two distinct categories:

Goal Setting Only and Goal Setting with Future Mindedness. Within the latter category, the instances of Future Mindedness ranged from one to four coded utterances of Future Mindedness within an interview set. The interesting finding emerged along this trajectory. Specifically, interviews with more instances of Future Mindedness were coded as either Positive or No Change with very few, if any, instances of Negative. This suggests that students engaged in setting a pathway comprised of at least three distinct steps and those who did so 3 or more times during their interviews have begun to manifest Future Mindedness by the spring (Positive Change) or had already manifested Future Mindedness in the fall and offered a comparable response in the spring (No Change). In the instances when interviews were coded once or twice for Future Mindedness and at least one of those instances of Future Mindedness was coded as negative, this suggests an emerging ability to create a three-step path. Students may be ‘testing’ the parameters of Future Mindedness and are vacillating between a quality articulation of pathways. In these cases, students’ utterances were coded No Change for Goal Setting but Negative for Future Mindedness. In the following two excerpts, the student’s initial goal is equally high-quality in both the fall and the spring. In the fall, the student offers three distinct steps toward her goal:

What do you hope to accomplish in the next two years?

To pass my exams.

Do you have another goal?

To have done more 5Ks and things.

Why are these goals important to you?

Because I want... I want to do 5Ks because I want to get better so I can do marathons and things. And just so I can get a good job as well. And have your life *unknown*.

(Fall, Q17, Y047)

While the fall excerpt satisfies the definition of Future Mindedness with the correct number of steps, the following excerpt from the spring lacks both the requisite three steps and is much less focused:

What do you hope to accomplish in the next two years?

I hope to have gotten my qualification.

Ok, and what about this is important to you?

Just so I have a possible career path, and I'm able to succeed in life.

(Spring, Q17, Y047)

There seems to be an interesting connection between students' inability to demonstrate Future Mindedness or the tentativeness with which they approach it, if they do. On their posters, students identified patience as a virtue to which they needed to pay more attention. One might suggest that patience is required in thinking through a goal path and seeing that path through, instead of setting a goal that is either quickly or easily attainable. Students seem to understand what it means to be patient but appear to be grappling with the operationalization of it, and this conflict appears to manifest itself in Goal Setting and, more directly, Future-Mindedness.

No Change also occurred with Goal Setting. This simply implies that students kept a goal of the same quality as they had in the spring. In other words, there was no notable improvement because the quality of the fall interview was already high.

Goal Setting: Connection-Sense of Purpose. In addition to its connection with Future Mindedness, Goal Setting appears to share a relationship with Connection and Sense of Purpose. We used Lerner et al.'s (2005) definition of Connection, which is recognized as 'the positive bonds with people and institutions that are reflected in bidirectional exchanges between the individual and peers, family, school and community in which both parties contribute to the

relationship’. We followed Damon (2008), who defines Sense of Purpose as ‘the way in which people develop character and a sense of purpose in their work, family, and community relationships; how young people can approach their careers with a focus on purpose, imagination, and high standards of excellence.’ In some instances, Goal Setting was simply paired with Connection. The following excerpt from the fall was coded for Goal Setting only:

What is a second thing that you expect you will be like or that you expect to be doing next year?

That’s quite tough actually. Well, I’ll definitely have a job by next year.

(Fall, Q8, Y10)

In the spring, the student has a positive change for Goal Setting because his focus shifts from a job to participating in a positive youth development program. This is an important shift because there is more value for a child to be involved in an organization like the scouts than there is in a job that simply earns him money. In addition, the student addresses the importance of the connections he will make through his experience with this organization, thus pairing Goal Setting and Connection:

What is a second thing that you expect you will be like or that you expect to be doing next year?

I’ll be in explorers which is like scouts.

Ok, and what about this is important to you?

The explorers and scouts is really big because I meet a lot of new people doing it, like going on the international camps, and the national camps, and other districts are really nice.

(Spring, Q8, Y10)

The earlier partnering of Goal Setting-Connection expanded to include Sense of Purpose. The following excerpt features Goal Setting and Connection as the student suggests that the work he hopes to do in the future will offer him a connection:

What is one thing that you expect you will be like or that you expect to be doing next year?

Umm, probably just kind of like the same thing as I'm doing like living as I am right now. Just always coming home to the same house, still with my parents obviously. Yea just kind of at the same, doing the same thing.

What about this is important to you?

Well it is important to me because coming home and where I feel safe and having mum is someone I can trust and that is really important to me. Being able to come home and do that kind of stuff.

(Fall, Q5, Y29)

By the spring, the student not only associates the work he wants to do with a connection but evolves that into a more global sense of purpose that expands his environment from beyond his home to his interaction with people other than his family:

What is one thing that you expect you will be like or that you expect to be doing next year?

Umm probably still have a job. More or less the same, but I'll probably just be like doing different things at school.

Okay, so you're saying you would still be working and still be in school?

Yeah.

And why are these things important to you?

My work is important to me because I get to do my work and I get to interact with different people. I get to speak to different people and you just kind of get more confidence through it as well. And then it's not all about the money side either, it's more I really enjoy doing my job and stuff. And umm school I'll probably just be like doing harder work and stuff, but still enjoying it as I do as well.

This student also manifested positive change for social competence at several points during his spring interview, which would suggest the potential to expand Goal Setting-Connection-Sense of Purpose to include Social Competence.

Sense of Purpose seemed to exist on a continuum, as well. There were instances in which Sense of Purpose was an individual pursuit:

What about this is important to you?

Because it follows what I want to, it's something I need to be able to do to be able to succeed in my like big big goal, like the one that I'm trying to work towards for my whole life really.

And what is that big goal?

Well it's a bit farfetched but I'd like to become either a test pilot or an astronaut like Chris Hadfield.

(Spring, Q1, Y014)

There were other instances in which Sense of Purpose seemed to move along the continuum to capture a more global impact:

Why would you like to have this job?

Because it just, well I just think it would be something, it would be incredible being able to go to space and it would give me such a responsibility or have something to work towards. As well because being able to pretty much represent you know like Britain, and the whole world really up in the international space station. It would just be such a good feeling and so much fun. So that's why I'd like to become an astronaut.

(Spring, Q3, Y014)

This seems to suggest that Sense of Purpose and Goal Setting can co-exist without Connection and/or Social Competence only when Sense of Purpose is at its infancy, or at the lower end of the developmental continuum. Once it evolves into a construct that includes a contribution that is greater than that of the individual, it starts to embody the qualities consistent with Connection and Social Competence.

The 5 Cs. We derived our understanding of the 5 Cs from Lerner et al. (2005).

Confidence is 'an internal sense of overall positive self-worth and self-efficacy; one's global self-regard as opposed to domain-specific beliefs. Connection is recognized as the positive bonds with people and institutions that are reflected in bidirectional exchanges between the individual and peers, family, school and community in which both parties contribute to the relationship.

Competence is a positive view of one's actions in domain specific areas including social, academic, cognitive, and vocational. *Social competence* pertains to interpersonal skills (e.g., conflict resolution). *Cognitive competence* pertains to cognitive abilities (e.g., decision making). School grades, attendance, and test scores are part of *academic competence*. *Vocational competence* involves work habits and career choice explorations. Caring/Compassion is defined as a sense of sympathy or empathy, and Character is respect for societal and cultural rules, possession of standards for correct behaviors, a sense of right and wrong (morality), and integrity.

There were no cases in which a student's interviews were coded positive, negative or no change for all 5 Cs. Connection and Social Competence have been addressed in the context of Goal Setting and Sense of Purpose in acknowledgment of the relationships that exists amongst these categories. The most frequently coded 'C' across student interviews was Confidence. This was consistent with teacher reports. When asked what teachers thought pupils gained by participating in Inspire>Aspire, many teachers cited increased confidence. The most significant finding we encountered while coding for the 5 Cs was our inability to categorize some of the utterances within the 5 Cs. This required us to create new categories within which to subsume what appeared to be low-level iterations of the 5 Cs. As a result, we added pre-cursor codes, such as Low-level Caring/Compassion and Low-level Connection, which we viewed as potential predictors of Caring/Compassion and Connection, respectively. It became necessary to acknowledge the emergence of these characteristics, which could not be given 'full credit' since they were not fully developed and, thus, did not satisfy our definitions for the full iteration (i.e., Caring/Compassion and Connection). Thus, there was a difference in quality amongst the

categories that comprise the 5 Cs. However, these differences were categorized as a more mature version (i.e., one of the 5 Cs), or as a less-developed version (i.e., one of the Low-level codes).

The following utterance, which was coded Positive for Caring/Compassion, captures a response that acknowledges the speakers empathy toward others:

How do your values guide the decisions or choices you make?

Well I think I try and try my best to do all of them so that everybody feels respected and that my friends know that they're important to me and that they know that I can support them and stuff so they can come to me about anything.

(Spring, Q4, Y029)

The following excerpt is an example of what we considered Low-level Caring/Compassion:

Why would you like to have this job?

I like helping other people and I like being like a being quite calm in a stressful situation like that if anything went wrong.

(Spring, Q20, Y89)

While the student expresses consideration toward people, the focus is on people 'in general,' and the response does not overtly express sympathy or empathy. In fact, the reason for wanting to help people seems to be the speaker's ability to stay calm in stressful situations.

While this is an important attribute, it does not meet the requirements for Lerner et al.'s definition of caring. Therefore, in order to be able to categorize these types of utterances, we created a Low-level Caring/Compassion category. The utterances that were coded Low-level Caring seem to suggest a pre-cursor to Caring/Compassion as defined by Lerner. Once again, this is an example of codes that reside on a continuum and to which different levels of development are attributed.

The same situation presented itself with Connection. We identified several cases of Low-level Connection that suggested a pre-stage because it did not fully satisfy Lerner's definition.

Following is an example of Connection, as defined by Lerner:

What is a second thing that you expect you will be like or that you expect to be doing next year?

I think I will be going to Germany to visit my old friends.

You think you'll be going back to Germany to visit old friends?

Yea.

What about this is important to you?

I think it is important to me to, like, umm to uphold to stay in touch with them or to stay friends with them.

(Spring, Q8, Y033)

The respondent articulates that maintaining a connection with friends is important to him. By comparison, the following excerpt constitutes a lower-level version of Connection:

What do you hope to accomplish in the next two years?

Probably finish my national 5's and get good grades in them, maybe learn how to do basic flying and basic gliding and maybe get ranked up in cadet maybe like a corporal so that would be, yea and just keep on a good basis with people I know, that would be it.

The interviewees' response to 'just keep on a good basis with people I know' is vague and unfocused. It also lacks the sense of commitment manifest in the previous excerpt from Y033 that articulates the importance of these relationships to the speaker.

These low-level categories were an important addition to our coding categories because there were many more instances of Low-level Caring/Compassion and Low-level Connection than there were of Caring/Compassion and Connection. These offer further substantiation that the interview respondents are beginning to work with these concepts but have not fully formed their understanding of them. This seems to support the students' self-reports from the posters

themselves. When asked which attributes they need to work on, the most frequently cited by both boys and girls were appreciation, commitment and understanding. All three of these fit squarely within the parameters of Connection and/or Caring/Compassion. The students' recognition of these potential areas of improvement also substantiates the suggestion that they are aware of these attributes but are in the beginning stages of developing them.

Self-knowledge and Self-care. Another category that we coded for is Self-knowledge. We defined Self-knowledge as, 'knowing that something can/will have an impact on me (i.e., my health; understanding of self, of one's own self-worth, what's good for me and what's not (reflection)). This can also function as a precursor to self-awareness.' Following is an example of Self-knowledge:

What is one thing that you expect you will be like or that you expect to be doing next year?

I hope to be working towards bettering myself in the subjects I'm not that good at, like math. Umm and I really want to try and focus on things other than my body image.

And why are these important to you?

Because I think that as long as I'm completely happy with myself I'll be fine with everything else in my life. And umm if I want to do physiotherapy as a subject when I'm older I need to be able to get good grades in math, so I'd like to do that.

(Spring, Q5, Y019)

Consistent with our definition of Self-knowledge, this excerpt demonstrates the students' desire to work toward 'bettering' herself, and those goals are associated with her future occupation. In fact, the respondent overtly acknowledges a shift away from focusing on body image and toward the steps she needs to take in order to become a physiotherapist.

The following excerpt is another example of Self-knowledge:

Like smoking and drugs and stuff, I don't want to go near that.

What about this is important to you?

Because it will affect my health.

(Spring, Q11, Y089)

In this example, the respondent talks about the impact that smoking and drugs will have on her health. Recognizing the impact that one's decisions will have is the definitional requirement of Self-Knowledge. There were other instances that did not reach that bar, but we wanted to be able to capture those since we saw those as pre-cursors to Self-knowledge. As a result, we created the category Self-care.

What are you doing this year to reduce the chances this will describe you next year?

I'm trying to like be more active and like any time that I get the chance to go out somewhere and do something different I will do it and like not be holding myself back because I can't be bothered or something like that.

(Spring, Q13, Y029)

In the previous example, the student makes reference to being more active but does not articulate the potential impact this could have. This distinction warranted the sub-category Self-care, which resides in the same theoretical supposition as we have been positing that the interviewees are at different stages of many of these codes. In many cases, these different stages required us to create new categories to accommodate the data.

Self-awareness. There is a clear distinction between self-knowledge and self-awareness. We defined Self-awareness as 'knowing my strengths and weaknesses.' Utterances that qualified for this code might include phrases comparable to 'I know I'm really good at...'

There were very few coded instances of Self-awareness and these were coded Positive Change, Negative Change or No Change.

The following example set illustrates Positive Change for Self-awareness:

What do you hope to accomplish in the next five years?

Have a job. I don't know.

What about this is important to you?

Well, it is the main way you can be able to live your life and everything.

(Fall, Q18, Y065)

What do you hope to accomplish in the next five years?

Well I guess I'd like to be at University and studying engineering or something.

Ok, and how did you choose this as something you'd like to do in the next 5 years?

I'm interested in science and stuff, and I'm pretty good at math, so it just interests me in that kind of way, that's kind of why I chose it.

(Spring, Q18, Y065)

In addition to Positive and Negative Change, there were also examples of No Change.

Why would you like to have this job?

Because I like working with kids and I like being helpful and think like people that need me like that can count on me to make sure I help them.

(Fall, Q20, Y089)

Why would you like to have this job?

I like helping other people and I like being like a being quite calm in a stressful situation like that if anything went wrong.

(Spring, Q20, Y089)

The previous example set from Y089 offers two strong responses that clearly address the respondent's strengths. Because the spring response is qualitatively comparable to the fall response, the utterance was coded No Change for Self-awareness.

Because this was coded so infrequently, it is difficult to draw any conclusions based on the data set. One possible insight is that Self-awareness requires a level of personal understanding and recognition that students this age have not yet fully cultivated and that may be at the beginning stages. This explanation is consistent with the trends that have emerged, thus

far, from many of the other coding clusters that suggests that students' self-awareness resides at the beginning stages along a larger continuum and, along with many of the other codes, is in early stages of development.

Being Inspired. There were clear instances of this code, particularly in cases when the student directly expressed that sentiment in their response by including the phrase he/she is inspirational because... In all cases, there was either Positive Change or No Change. In the cases where no change occurred, the answers were qualitatively consistent from the fall to the spring as illustrated in the following two excerpts:

Can you name a person who you would consider to be “inspirational”?

Umm, probably like Beyoncé.

What about this person inspires you?

Just like, how she like, has her confidence and everything and how much confidence she got. And I think as well she's done a couple things for charity and it is good to see she is giving something back.

(Fall, Qs 1 & 2, Y029)

Although the student chooses a different inspirational figure in the following excerpt from the spring, the reason why the figure inspires her remains comparable:

Can you name a person who you would consider to be “inspirational”?

Emma Watson.

What about this person inspires you?

I think the way that she stands up for her beliefs and like the way she can kind of, like the way that she manages to speak up about what she believes in and she's not scared to do that.

(Spring, Qs 1 & 2, Y029)

There were many instances of positive change. In the following example, the respondent delivers a very straightforward response in the fall:

Can you name a person who you would consider to be “inspirational”?

My grandpa.

What about this person inspires you?

He flew planes a lot. He flew lots of planes in the Air Force.

(Fall, Qs 1 & 2, Y009)

In the spring, however, he offers a much more sophisticated articulation of the impact his grandfather had on him. Even though the program asked students not to choose a family member, this response is significant because it could suggest that the program helped the student more thoughtfully deconstruct why he chose his grandfather as his inspirational figure.

Can you name a person who you would consider to be “inspirational”?

So when you say inspirational do you mean to other people as well or just to me?

Just to you personally.

My grandpa.

What about this person inspires you?

Well that he started school at the age of 9 and taught himself to read and write and then joined the RAF.

Ok and what's the RAF?

Royal Air, I don't know what it stands for but it kind of like, It's a British flying planes.

Kind of like the Air force?

Yea, air force.

(Spring, Qs 1 & 2, Y009)

Knowledge of Values, Virtues and Qualities. We defined Knowledge of Values, Virtues and Qualities as ‘an articulated recognition or demonstrated understanding of values, virtues and/or qualities; this does not necessarily need to reflect values, virtues or qualities from the poster but must demonstrate a shift from fall to spring that illustrates an increased understanding of what values, virtues and/or qualities are.’

The following exchange from the fall suggests that the student lacked an understanding of what values are, even after the interviewer offers an explanation more than once:

What are some values that you consider to be desirable or important?

Um, could you repeat that I don't exactly understand.

So, what are some values or desirable traits that you think are important.

Um, sorry, I really don't understand the question.

Ok, that's okay. Um, so what are characteristics of people that you think are important or traits that they have that you think would be beneficial for them?

Just be confident and stuff like that, is that what you mean?

(Fall, Q3, Y001)

In the spring, the student expresses a recognition of what the term *value* means after the interviewer offers clarification. In fact, it seems that the student struggles with the term and not the concept:

What are some values that you consider to be desirable or important?

I'm sorry can you explain that question again I'm not understanding what you are saying?

So we're asking about your values, and your values would be like things that you think are important for a person to be like or qualities that you think are important to have. Like just ways for a person to be, something that you would think is important or beneficial to be in life.

Oh right well obviously to be a nice person, to be like nice to other people, respect them.

Ok so being nice to other people and respecting them?

Yeah.

(Spring, Q3, Y001)

The spring response seems to be an advancement over the fall. Specifically, during the fall the student did not demonstrate an understanding of the term or the concept. Although the student

does not initially recognize the term ‘values’ in the spring, he demonstrates an understanding of the concept once the interviewer defines the term ‘values.’

More compelling is the student’s response when he is asked what values he considers to be desirable or important. The following two excerpts, which were coded for Positive Change in Knowledge of Values, Virtues and Qualities may suggest that the student was unclear about what values are in the fall, but has a more solid understanding in the spring:

What are some values that you consider to be desirable or important?

Umm I don’t know. Honesty I guess, I don’t know. Values...

(Fall, Q3, Y065)

What are some values that you consider to be desirable or important?

Trust. Honesty.

(Spring, Q3, Y065)

In all cases of Knowledge of Values, Virtues and Qualities, there was either Positive Change or No Change. In the cases where no change occurred, the answers were qualitatively consistent from the fall to the spring.

Relationship between Teacher Characteristics and Poster Quality (Aim 4)

The next set of analyses focused on gaining a better understanding of the relationship between program implementation, teacher characteristics, and the quality of student posters. Teachers who invested more time in Inspire>Aspire had pupils with better aligned posters, and higher quality posters. For those teachers who took more time to complete the poster program (e.g., spent more weeks implementing the program), posters demonstrated higher alignment as well as higher quality scores. Pearson’s r correlations were significant when examining the “Start to Finish” time of Inspire>Aspire for all quality outcomes, including broad quality score ($r =$

.31; $p < .01$), Consistency with PYD ($r = .16$; $p < .05$); Values well developed in 1st panel ($r = .27$; $p < .01$), Fictional Story, Quotes, Inspirational Figure well developed ($r = .28$; $p < .01$), and Aspirations Well Developed ($r = .19$; $p < .05$).

Time spent doing the poster at home was negatively correlated with the quality of value development (first panel; $r = -.17$, $p < .05$). In other words, the more time teachers estimated the pupils spent completing the poster at home, the lower the quality of value development in the pupils' posters. This is an important finding. It may be that for youth who complete the poster at home, their parents are exerting undue influence on the poster content, so the youth do not benefit as much from the program as they could. This suggests that perhaps completion of the posters should not be assigned as homework. Other time variables were not associated with poster quality, either positively or negatively.

In terms of teacher experience, we found no correlation between teachers' overall teaching experience and pupil poster quality; we also found no correlation between teachers' experience with Inspire>Aspire and pupil poster quality. Thus, new teachers should be able to pick up Inspire>Aspire and get results (whatever that may be) the first time implementing it. It may be that broad dissemination of the program could be accomplished without much of a learning curve in terms of implementation.

We also examined pupil outcomes by teacher implementation clusters. In other words, did pupils who experienced a different implementation style have different goals, aspirations, future orientation, or sense of purpose? For all outcomes, we report scores at both Wave 1 and Wave 2. Group differences are indicated by letter superscripts above each bar. Figure 32 illustrates differences in SOC scores for the three clusters. There were no statistically significant differences in SOC scores across the three teacher implementation groups.

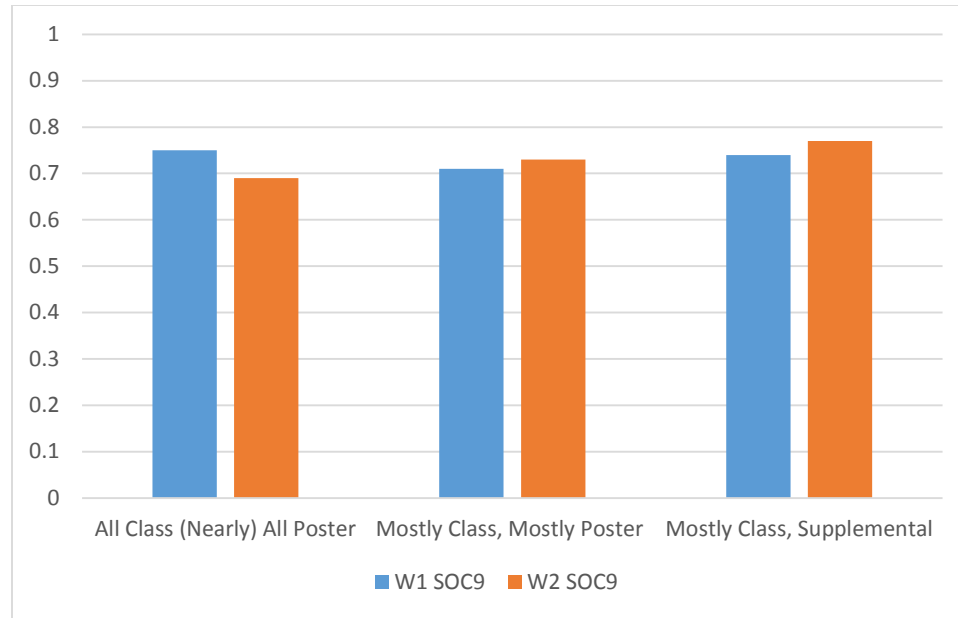


Figure 32. Differences in pupils’ SOC scores by teacher implementation cluster

Figure 33 shows the differences in youth purpose across the three implementation style cluster groups. Here, we see group differences in the contrasts that correspond to the letters above the bars. The ‘All Class (Nearly) All Poster’ cluster is significantly different than the ‘Mostly Class, Mostly Poster’ cluster (both have the superscript “a”) and the ‘All Class (Nearly) All Poster’ cluster is also different than the ‘Mostly Class, Supplemental’ cluster; these findings are based on ANOVA post-hoc difference tests.

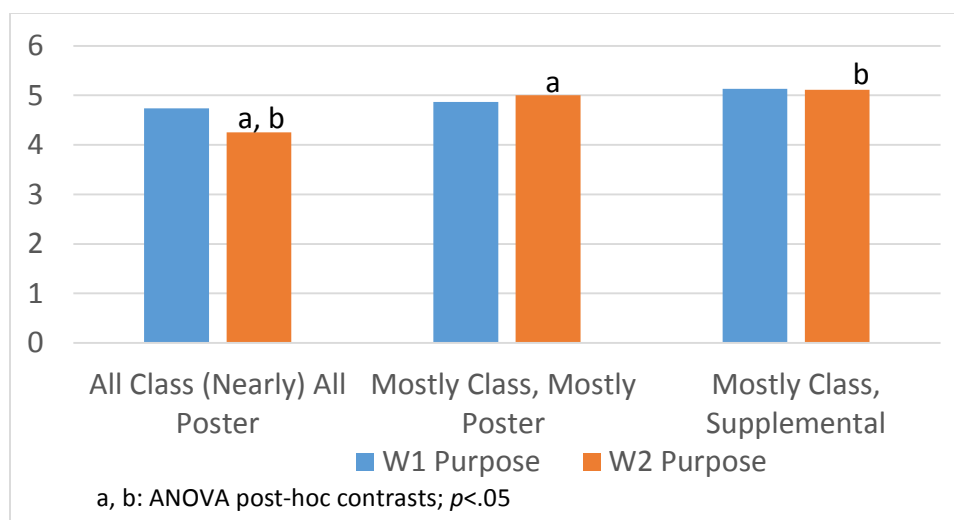


Figure 33. Differences in pupils' sense of purpose scores by teacher implementation cluster

Next, we examined differences in future orientation by teacher implementation cluster (see Figure 34). For this outcome, there were no cluster differences in the Wave 2 outcome, but one difference in Wave 1. Pupils with teachers in the 'All Class (Nearly) All Poster' cluster had significantly higher future orientation scores at Wave 1 compared to pupils with teachers in the 'Mostly Class, Supplemental' cluster.

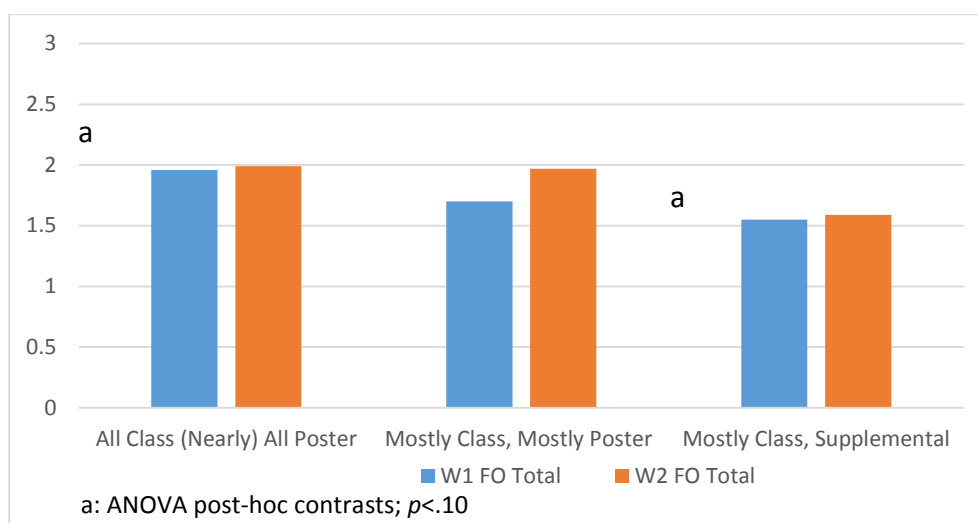


Figure 34. Differences in youth future orientation scores by teacher implementation cluster

Youth aspirations (importance) differed both at Wave 1 and Wave 2 by teacher implementation cluster (see Figure 35). At both Waves 1 and Wave 2, pupils whose teachers were in the ‘Mostly Class, Mostly Poster’ cluster attributed less importance overall to their aspirations compared to pupils with teachers in the ‘Mostly Class, Supplemental’ cluster.

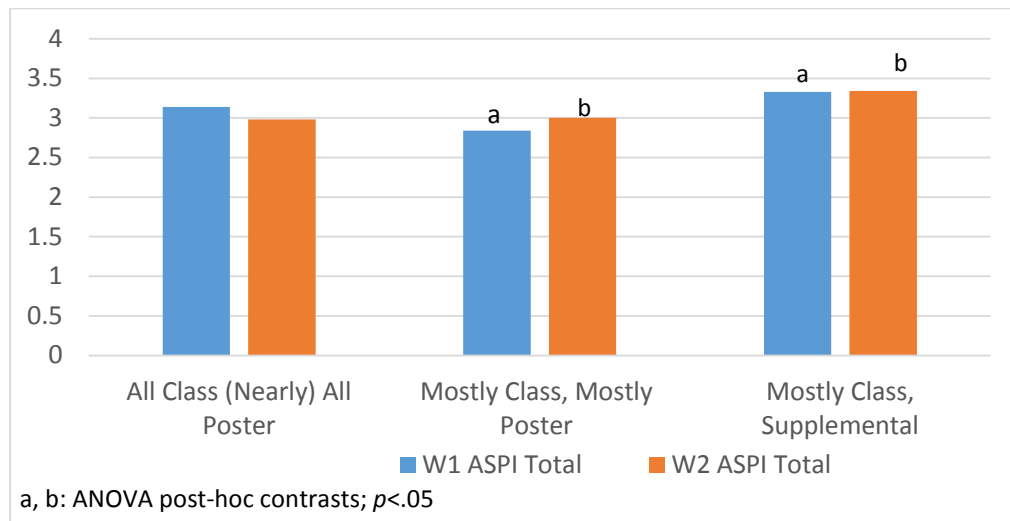


Figure 35. Differences in pupils’ aspirations (importance) scores by teacher implementation cluster

Figure 36 illustrates difference in pupils’ aspirations (chances of) scores by teacher implementation cluster. For pupils in both waves, those with teachers in the ‘All Class (Nearly) All Poster’ cluster or the ‘Mostly Class, Mostly Poster’ cluster reported significantly lower chances of aspirations compared to pupils with teachers in the ‘Mostly Class, Supplemental’ cluster.

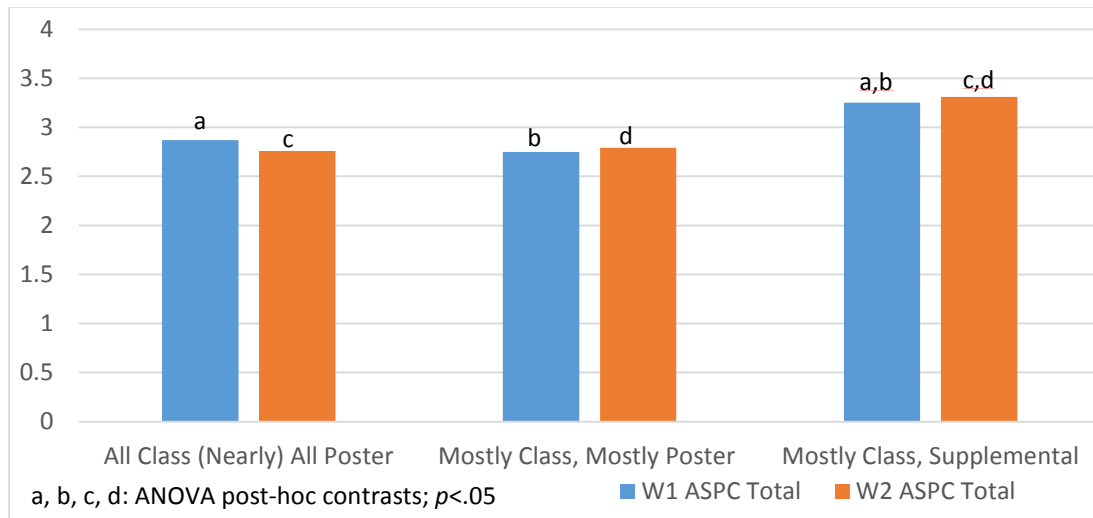


Figure 36. Differences in pupils' aspirations (chances of) scores by teacher implementation cluster

Recall that in addition to creating teacher implementation profiles, we determined pupil profiles based on the pattern of poster quality. We then examined differences in pupil poster quality profile membership based on teachers' implementation cluster (see Figure 37). The patterns were similar for the first two teacher implementation clusters: 'All Class (Nearly) All Poster' and 'Mostly Class, Mostly Poster'. Pupils who had teachers that used either of these implementation styles tended to have the highest percentage of youth in the 'High Aspirations' cluster, the lowest percentage of youth in the 'Average' cluster, lower still percentages of youth in the 'Low' cluster, and the lowest percentage of youth in the 'High PYD' cluster (there were only 14 posters in the 'High PYD' cluster, so these results must be interpreted with extra caution). Teachers who were in the 'Mostly Class, Supplemental' cluster tended to have the least favorable pupil poster quality outcomes. Most posters in those classrooms were in the 'Average' cluster. Based on these findings, Character Scotland may want to consider developing additional high quality activities that could be assigned as part of poster completion.

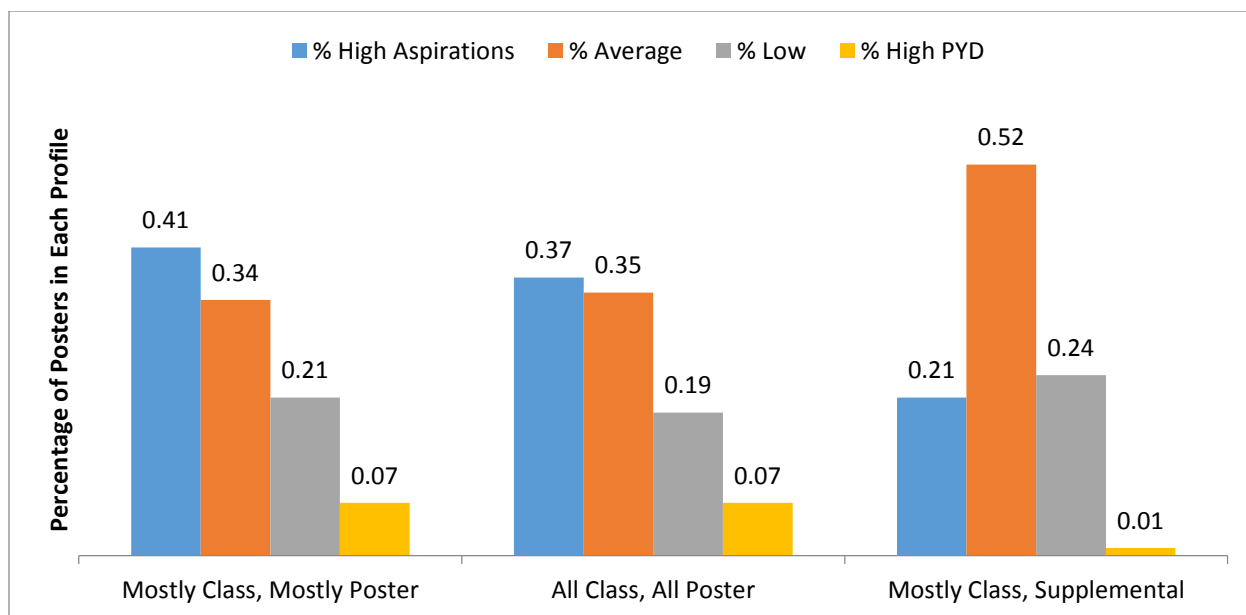


Figure 37. Percent of posters in each profile by teacher implementation cluster

We examined how integration of Inspire>Aspire into the broader curriculum differed by years of experience teaching Inspire>Aspire. Some teachers reported that they used Inspire>Aspire as a standalone program that was not integrated with other curricula. Other teachers reported that they were able to fully integrate Inspire>Aspire with the broader curriculum, while other teachers reported that they did a mix of integrating Inspire>Aspire with other course/curriculum materials and used some aspects of Inspire>Aspire as a standalone program. Figure 38 illustrates that, not surprisingly, those teachers who demonstrated the highest level of integration with the broader curriculum had the most years of experience (on average) implementing Inspire>Aspire.

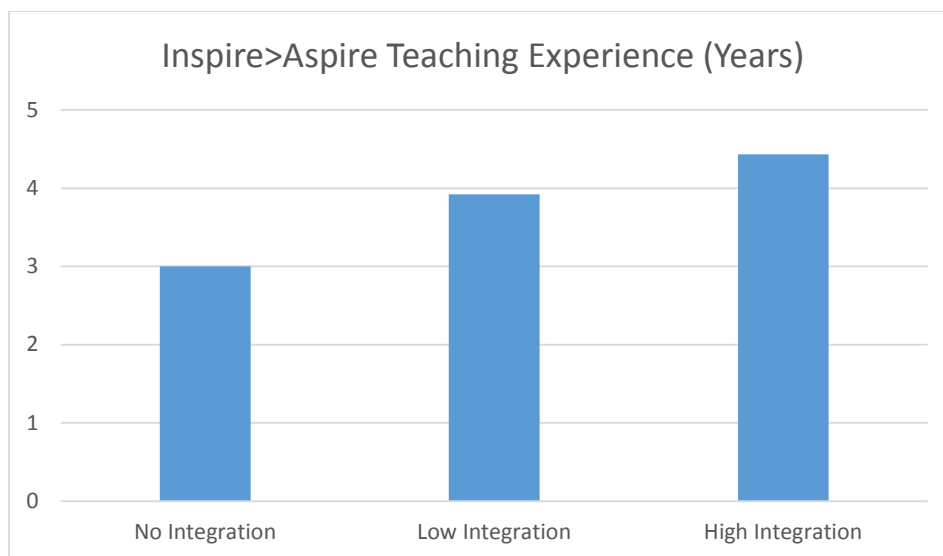


Figure 38. Integration of Inspire>Aspire with broader curriculum by years of experience teaching Inspire>Aspire

We then looked at differences in poster quality based on the extent to which Inspire>Aspire was integrated with the broader curriculum (see Figure 39). Teachers who implemented Inspire>Aspire as a standalone curriculum had, on average, the highest percentage of posters in the ‘Average’ cluster. Teachers who fully integrated Inspire>Aspire with other course/curriculum materials had, on average, the highest percentage of posters in the ‘High Aspirations’ cluster. Teachers who did a mix of integrating Inspire>Aspire with other course/curriculum materials and using some aspects of Inspire>Aspire as a standalone curriculum had a blend of posters from all of the profile groups.

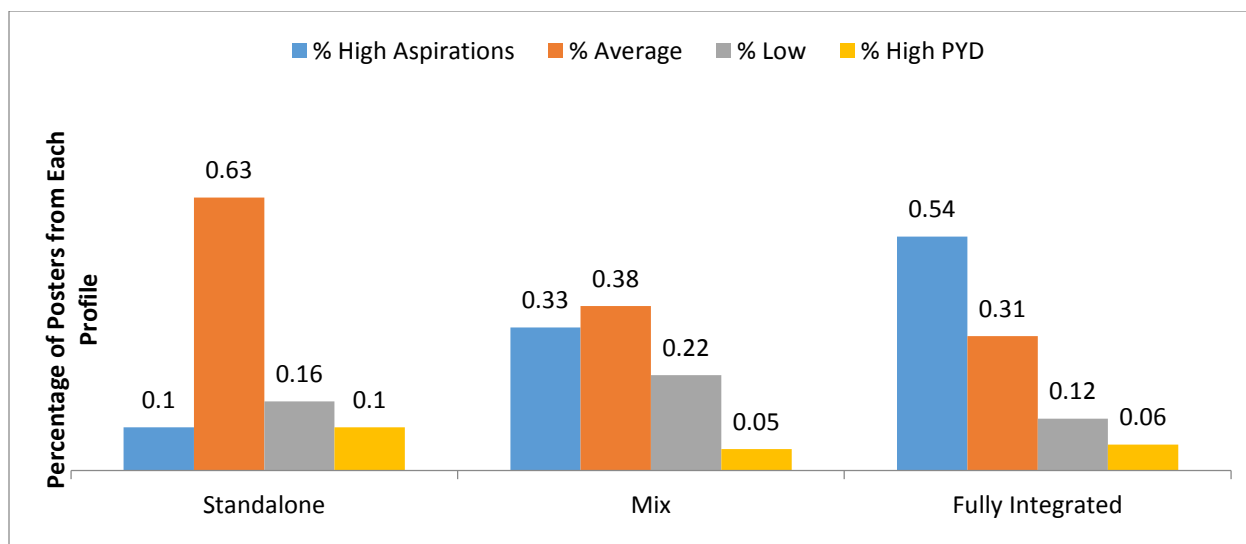


Figure 39. Percent of posters in each profile by integration

Despite the small sample size of teachers, it does appear that we can tentatively say that some teacher characteristics may be related to student poster quality. In particular, teachers who used more of the supplementary materials seemed to have lower average poster quality (as measured by both poster profile membership, and broad quality score) compared to those teachers in the other two implementation clusters (those who focused more on the poster itself). In other words, teachers who focus more on the poster had students whose posters were rated more highly. Teachers who report that they have fully integrated the program also had students with higher poster quality (as measured by both poster profile membership and broad quality score). This indicates that Inspire>Aspire is most effective (in terms of poster quality) when teachers fully integrate it with other aspects of the curriculum. Inspire>Aspire is less effective when it is used as a standalone program. It may be that Inspire>Aspire is best used as a supplemental program when it is well aligned with other curricular efforts.

CONCLUSIONS

The summary of results is organized by each Aim and then the Big Questions.

Aim 1: Assess variations in program implementation

There is a fair amount of variation in the ways in which teachers implement Inspire>Aspire. Teachers who have more experience with Inspire>Aspire tend to focus more on supplemental activities. Pupils whose teachers focused on poster activities and completed these activities almost entirely in class spent the most total number of minutes on Inspire>Aspire (807 minutes on average). Teachers who fully integrated Inspire>Aspire with the broader curriculum tended to have pupils with higher quality posters. The language level of the template and instructions is challenging for many S2 pupils. While teachers lauded the program's flexibility, this often came from experienced teachers.

Teachers vary in what they consider to constitute the "program." Some teachers view the poster itself as the program, whereas other teachers see the program as the process (and accompanying activities) leading up to completion of the poster. In the latter, the poster is viewed as a culminating activity that reinforces the previous lessons.

Aim 2: Pilot test quantitative measures of key character outcomes

Pupils did not show significant differences between pre- and post-test on most of the quantitative measures. Some significant results were found for future aspirations. This may be due to several reasons: the poor psychometric properties of the measures; the level of sophistication of the measures for this population (i.e., the measures are designed for adolescents but the pupils in this study were at the younger end of the age range; not enough time had passed for the constructs to emerge, and/or even shorter-term outcomes need to be measured (e.g., precursors to caring).

Aim 3: Validate the theory of change using qualitative interviews of youth

The qualitative results provided some support for the theory of change. Pupils often demonstrated precursors to many of the constructs that were measured but many pupils did not manifest the construct itself (e.g., caring/compassion). This indicates that shorter-term outcomes need to be added to the theory of change to capture the immediate effects of participating in Inspire>Aspire.

Some students struggled with the word ‘values.’ It is important to consider how important it is for students to understand what the word ‘values’ itself means. Is it, instead, more important for students to understand what the meaning of the values themselves are?

Aim 4: Assess the relationship between variations in program implementation (e.g., differences in dosage) and poster quality

Teachers who fully integrated Inspire>Aspire with the broader curriculum (as opposed to using it as a standalone program) tended to have pupils with higher quality posters. The highest quality posters are produced by teachers who spend significant time on poster related activities in class but do not immediately have pupils work on the poster template. Preparing the students to work on the poster by using activities that lay the foundation for the poster elements is effective. Posters are of higher quality when the poster is completed in school rather than at home.

Big Question 1: Can reflecting on and writing about the virtues espoused in the Laws of Life improve adolescent character development?

Perhaps. Evidence from quantitative measures is inconclusive. The specific outcomes of interest need to be reconsidered (perhaps looking at shorter-term outcomes or other outcomes such as self-esteem, confidence and self-reflection which emerged from the qualitative data).

Qualitative evidence suggests some indications of personal growth after participating in

Inspire>Aspire but it is unclear if this is due to program participation or normative maturation of the youth.

Big Question 2: What is the overarching theory of change for Inspire>Aspire: Global Citizens in the Making?

This is represented in the pathway model. Significant strides were made in articulating the theory of change. The pathway model will be revisited and revised in light of the findings from this pilot study.

Big Question 3: How is the theory of change best assessed?

A mixed methods approach is critical. Teacher report measures should also be used. Some quantitative measures did not have good psychometric properties and will need to be refined and/or replaced. Cognitive interviewing would help with testing and refining measures.

Big Question 4: Does preliminary evidence support the theory of change?

Some elements of the theory of change show preliminary support. Based on these findings, it is important to revisit whether earlier markers of change should be included in the theory of change. Variations in teacher implementation style have an impact on pupil poster quality. Ceiling effects were also present. Therefore, it is important to recruit a more diverse group of pupils to participate in Inspire>Aspire in order to adequately test for program effects.

Recommendations

To advance the mission of Inspire>Aspire, the REDSS Lab leadership team synthesized the findings to create design and implementation recommendations.

Create a Foundation for Success. First, the program developers should *clarify the program goal and the role of the poster*. Character Scotland should determine whether youth character development is meant to occur by simply completing the poster or whether it occurs

through classroom activities that provide foundational lessons that prepare pupils to complete the poster. **Is the poster a tool that reflects student character knowledge gained from preceding classroom lessons or is poster completion itself the vehicle for character education development?** Currently, some teachers seem to understand the program goal as poster completion, with little need for intensive classroom instruction on character elements. However, the current findings raise the question of whether the program goal should be learning what is necessary *in order to* complete the poster. With this approach, the poster would be a culminating project reflecting student learning from program participation. For teachers who spend a lot of time discussing the poster components that foster character development (including values and inspirational figures), the poster serves as a tool to complement and reinforce lessons. For teachers who simply give the posters to their students with little or no direction, simply completing the poster *is* the program. As such, there are currently two different implicit and competing program goals co-occurring. If the goal of the program is NOT only to complete the poster, the program guidelines and website should make it clear that a thoughtful approach to the preceding character content is the crux of character development. To create a foundation for programmatic success, it is critical that teachers have a shared understanding of the program goal.

Provide Best Practices to Teachers. Teachers are instrumental to the success of Inspire>Aspire, so providing them with guidance on activities and best practices is critical. The findings suggest that Character Scotland should ***recruit and encourage teachers who are able to integrate Inspire>Aspire with their broader curriculum.*** Findings indicate that students of teachers that used classroom time as a central implementation space to hold discussions of values, inspirational figures, and complete the poster with other course/curriculum materials had,

on average, the highest percentage of posters classified as High Aspirations. Thus, utilizing class time to integrate the program is important to the program impact.

Similarly, Character Scotland should *strongly encourage teachers to not assign Inspire>Aspire activities as homework*. Findings indicate that pupils who had teachers that primarily used homework time for Inspire>Aspire work tended to have the least favorable poster quality outcomes.

When establishing teacher expectations regarding program delivery time, Character Scotland should *strongly encourage teachers to spend about 800 minutes on implementing Inspire>Aspire*. Pupils whose teachers focused on poster activities and completed these activities almost entirely in class for about 800 minutes had the best outcomes. Character Scotland should provide guidance on how that time is best used to optimize pupil experiences.

Cultivate More Character. Once teachers are informed of the ideal structure for implementing Inspire>Aspire, the focus should turn next to content. The data suggests Character Scotland should *advise teachers to include (more) discussions* on character elements, ensuring pupils understand and reflect on specific character values to enhance their comprehension. Discussion can also be used to enhance other poster elements. To help facilitate these discussions, Character Scotland should *add discussion prompts to the Website*, ensuring that all teachers, regardless of experience level, are equipped to implement successful lessons that deliver the intended objectives.

Expand Guidelines and Resources. To ensure that all teachers deliver the program content with adequate depth and pupil engagement that leads to character development, Character Scotland should *provide more specific guidelines about teaching the program*. Techniques for creating fun, engaging lessons should be shared, along with specific tips and

guidelines about meeting learning objectives for each aspect of the program. To do so, ***create approach strategies for each section of the poster***. Teachers who have completed the poster program and have classroom-tested strategies may be excellent resources for compiling new resources. Moreover, gathering great lessons and strategies from prior implementers may serve as an opportunity for Character Scotland to recognize stellar teachers and foster program buy-in. Character Scotland can also ***offer guidelines on making appropriate adjustments to lessons and materials by offering suggestions for abridging or expanding them***. Doing so helps maintain the program quality and adherence to the program goals while meeting real world demands.

Create Additional Activities. In addition to expanding classroom activities, Character Scotland should ***develop additional supplemental activities that help pupils create high quality posters***. Character Scotland should ***invite teachers who have customized and adapted existing materials or created their own supplemental materials to contribute to a growing database of resources***. As previously mentioned, teachers experienced in Inspire>Aspire delivery may have resources to share, which highlights their work and encourages others to use those resources in their implementation. Resources should adequately focus on all poster elements, so Character Scotland should ***provide resources for each section of the poster (not just Inspirational Figures and Quotes) on the Website***.

Extend Access to Younger Participants. As the program is being delivered to younger pupils, the materials must be age- and ability-appropriate. Character Scotland should ***revise the template and Website materials to be more accessible to a younger age group*** by revisiting the vocabulary, reading level and specificity of directions. The newly developed and collected resources and teaching materials recommended above should also be age-appropriate to reach target pupils.

Revisit the Programmatic Theory of Change. The findings of the current study produced new insights and these should be incorporated into a revised theory of change (pathway model). To integrate these findings, the team should: *Revise the pathway model to reflect shorter-term outcomes or precursors and measure these pre-cursors*. The data suggest a number of shorter-term outcomes or precursors may be present in the program pathway, so adding them to the model and ensuring they are sufficiently measured will allow advance understanding of how the program works and better serve new generations of pupil participants.

References

- Bundick, M., Andrews, M., Jones, A., Mariano, J. M., Bronk, K. C., & Damon, W. (2006). *Revised youth purpose survey*. Unpublished Instrument. Stanford Center on Adolescence. Stanford, CA. Retrieved from <http://coa.stanford.edu/publications/revised-youth-purpose-survey>
- Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five traditions*. (2nd ed.). Thousand Oaks, CA: Sage.
- Damon, W. (2008). *The path to purpose: Helping our children find their calling in life*. New York: The Free Press.
- Freund, A. M., & Baltes, P. B. (2002). Life-management strategies of selection, optimization and compensation: Measurement by self-report and construct validity. *Journal of Personality and Social Psychology*, 82(4), 642-662.
- Geldhof, G. J., Bowers, E. P., Boyd, M. J., Mueller, M., Napolitano, C. M., Schmid, K. L., . . . Lerner, R. M. (2014). The creation and validation of short and very short measures of PYD. *Journal of Research on Adolescence*, 24, 163-176.
- Gestsdóttir, S., Bowers, E., von Eye, A., Napolitano, C. M., & Lerner, R. M. (2010). Intentional self regulation in middle adolescence: The emerging role of loss-based selection in positive youth development. *Journal of Youth and Adolescence*, 39(7), 764-782.
- Kasser, T., & Ryan, R. M. (1993). A dark side of the American dream: Correlates of financial success as a central life aspiration. *Journal of Personality and Social Psychology*, 65(2), 410-422. doi: 10.1037/0022-3514.65.2.410
- Lerner, R. M., Lerner, J. V., Almerigi, J., Theokas, C., Phelps, E., Gestsdóttir, S., . . . von Eye, A. (2005). Positive youth development, participation in community youth development

- programs, and community contributions of fifth grade adolescents: Findings from the first wave of the 4-H Study of Positive Youth Development. *Journal of Early Adolescence*, 25(1), 17-71.
- Massey, E. K., Gebhardt, W. A., & Garnefski, N. (2011). Goals. In R. J. R. Levesque (Ed.), *Encyclopedia of Adolescence* (pp. 1188-1196). NY: Springer.
- Oyserman, D., Johnson, E., & James, L. (2011). Seeing the destination but not the path: Effects of socioeconomic disadvantage on school-focused possible self content and linked behavioral strategies. *Self and Identity*, 10(4), 474-492. doi: 10.1080/15298868.2010.487651
- Oyserman, D., & Markus, H. R. (1990). Possible selves and delinquency. *Journal of Personality and Social Psychology*, 59(1), 112-125.
- Steger, M. F., Frazier, P., Oishi, S., & Kaler, M. (2006). The Meaning of Life Questionnaire: Assessing the presence of and search for meaning in life. *Journal of Counseling*, 53(1), 80-93.
- Steinberg, L., Graham, S., O'Brien, L., Woolard, J., Cauffman, E., & Banich, M. (2009). Age differences in future orientation and delay of discounting. *Child Development*, 80(1), 28-44.
- Trochim, W., Urban, J. B., Hargraves, M., Hebbard, C., Buckley, J., Archibald, T., . . . Burgermaster, M. (2012). *The guide to the systems evaluation protocol*. Ithaca, NY: Cornell Digital Print Services.
- Urban, J. B. (2008). Components and characteristics of youth development programs: The voices of youth-serving policy makers, practitioners, researchers, and adolescents. *Applied Developmental Science*, 12(3), 128-139.

Urban, J. B., Hargraves, M., Hebbard, C., Burgermaster, M., & Trochim, W. M. (2011).

Evaluation in the context of lifecycles: 'A place for everything, everything in its place'.

Paper presented at the American Evaluation Association, Anaheim, CA.

Urban, J. B., Hargraves, M., & Trochim, W. (2014). An evolutionary perspective on program evaluation. *Evaluation and Program Planning*, 45, 127-139.

Urban, J. B., & Trochim, W. M. (2009). The role of evaluation in research-practice integration: Working toward the "golden spike". *American Journal of Evaluation*, 30(4), 538-553.

Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research and Evaluation*, 12, 1-26.