Program theory-driven evaluation science in a youth development context

Kelsey L. Deane [a,*], Niki Harré [b]

[a] Faculty of Education, The University of Auckland, Private Bag 92601, Auckland 1150, New Zealand
[b] School of Psychology, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

ABSTRACT

Program theory-driven evaluation science (PTDES) provides a useful framework for uncovering the mechanisms responsible for positive change resulting from participation in youth development (YD) programs. Yet it is difficult to find examples of PTDES that capture the complexity of such experiences. This article offers a much-needed example of PTDES applied to Project K, a youth development program with adventure, service-learning and mentoring components. Findings from eight program staff focus groups, 351 youth participants' comments, four key program documents, and results from six previous Project K research projects were integrated to produce a theory of change for the program. A direct logic analysis was then conducted to assess the plausibility of the proposed theory against relevant research literature. This demonstrated that Project K incorporates many of the best practice principles discussed in the literature that covers the three components of the program. The contributions of this theory-building process to organizational learning and development are discussed.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Youth development programs offer experiences that are designed to promote young people's holistic growth (Catalano, Berglund, Ryan, Lonczak, & Hawkins, 2004; Roth & Brooks-Gunn, 2003). They take many forms including adventure, arts, or culture-based programming; community involvement and service; life skills training and mentoring (Ministry of Youth Development, 2009). Many use a combination of novel or challenging activities and relationship building to build skills, confidence, and connection to others within an empowering, youth-centred context (Gallagher, Stanley, Shearer, & Mosca, 2005; Roth & Brooks-Gunn, 2003; Urban, 2008). Although programs vary in the specific outcomes they target and their efficacy in achieving these, there is evidence that they can help develop young people's practical, interpersonal, and intrapersonal skills, as well as the self-efficacy and motivation to engage with life and the challenges and opportunities it offers (Catalano et al., 2004; Deane & Harré, 2013; DuBois, Portillo, Rhodes, Silverthorn, & Valentine, 2011; Durlak, Weissberg, & Pachan, 2010).

Despite having ambitious goals and in many cases showing positive outcomes, there is variation in the effectiveness for participants within and across such programs (Deane & Harré, 2013; DuBois et al., 2011; Durlak et al., 2010) and the reasons for this variation are not well understood. For instance, few evaluations offer evidence of how outcomes are achieved and for whom they work best (Durlak et al., 2010; Larson, 2011; Riggs & Greenberg, 2004). Program theory-driven evaluation science (PTDES, Donaldson, 2007), a contemporary addition to the theory-driven evaluation genre, provides a useful framework for designing evaluation studies that can redress this gap.

Like all theory-driven evaluation approaches, the central task of PTDES is to make explicit the reasoning or "theory" behind how a program is presumed to produce positive change. This theory is then used to identify evaluation questions that assess how well the theory stands up in practice. PTDES is distinguished from other approaches in its emphasis on using scientifically rigorous methods to assess the theorized links (Donaldson, 2007). These methods could include unstructured or semi-structured interviews or observations, document reviews, quasi or true experimental designs. According to Donaldson (2007) multiple sources should be used to generate the program theory including discussions with stakeholders, observations of the program in action, and the social science literature. Stakeholders should also be involved in prioritizing the evaluation questions that result from the process. Finally, methodological

* Corresponding author. Tel.: +64 9 373 7513; fax: +64 9 623 8898.
E-mail addresses: k.deane@auckland.ac.nz, kelseylaurendeane@gmail.com (K.L. Deane), n.harre@auckland.ac.nz (N. Harré).

flexibility is needed, driven by the questions being asked and the context in which the program operates.

Donaldson's framework can be usefully employed to address the numerous calls for an increase in rigorous and diverse research methods that account for program processes and moderating influences within the youth development field (Durlak et al., 2010; Larson, 2011; Riggs & Greenberg, 2004). The current article aims to show how an approach based on PTDES can contribute to meeting this call. It is also a much needed case example of theory-driven evaluation as, to date, theoretical articles on this approach have far outweighed practice cases (Coryn, Noakes, Westine, & Schroter, 2011; Donaldson, 2007). This has led to some questioning the utility of the theory-driven evaluation genre (e.g. Stufflebeam, 2001), a question we hope to refute.

The focus for our study was the development of a program theory for Project K, a New Zealand youth development program that operates in seven centres throughout the country. Project K involves young people in their second year of high school who have been identified through a self-report survey and teacher ratings as having low self-efficacy. The program is run with groups of 12 students from a single high school. They begin with a 21-day *Wilderness Adventure*. This involves goal-setting, communication, and problem-solving activities at a residential outdoor camp followed by a 7- to 10-day journey in the wilderness in which they take turns being the group leader. Next, participants return to their daily routines but simultaneously undertake a *Community Challenge* in which they develop a project to contribute to their community. In the final program stage, each participant is matched with a trained adult mentor. The pair meets fortnightly and the mentor provides support with goal-setting and generally acts as a non-judgemental friend over a 12-month period (see www.fy-d.org.nz for more details).

Project K was developed by the Foundation for Youth Development (FYD). FYD licenses regional centres throughout New Zealand to implement the program. Each centre has a Regional Manager who oversees the administration and delivery of the program. All Regional Managers are provided with detailed training and program materials in order to promote program fidelity. Adventure programming providers are contracted to facilitate the Wilderness Adventure and, in some regions, the Community Challenge. As a result, internal regional staff members are not always directly involved in these components. Mentor coordinators are, however, employed as part of the regional team to oversee the Mentoring component.

A strong evaluation culture is also promoted by FYD, and in 2004 a longitudinal, randomized controlled trial (RCT) evaluation that focused on self-efficacy, academic achievement and several health and lifestyle behaviours was put in place. This study is ongoing and several other process evaluations have been conducted since the program commenced in 1995 (these are outlined in more detail further on). Despite the RCT and other process evaluations, at the time of the current study Project K had not gone through a systematic and detailed process of identifying the underpinning theory for the program, hence the theory of change process described here.

In the remainder of this article we outline the development of Project K's theory of change and show how it is being used for further research and program development. We finish by highlighting the benefits and limitations of our approach and providing recommendations to other evaluators and program developers, particularly those with a youth focus.

2. The theory-building process

Drawing on Donaldson's PTDES approach, we developed a four-step process to guide the development of Project K's theory of change. The steps were: (1) *generate the preliminary theory* grounded in the views of those close to the program; (2) *assess the level of consensus* across different program sources; (3) *conduct a direct logic analysis* (see Brouselle & Champagne, 2011) of the program theory by assessing whether or not the proposed links between program processes, influencing factors, and outcomes seemed plausible according to the social science literature; and (4) *assess the hypothesized links* by systematically investigating the relationships between the proposed linkages. Here we outline the first three steps and how we set the stage for the fourth. Ethical approval for the study was granted by The University of Auckland Human Participants Ethics Committee.

3. Method

3.1. Step 1: Generating the preliminary theory

For this stage the first author conducted a focus group interview with the two program founders, both of whom were still active in managing the program. In addition, she conducted one individual and four focus group interviews with a total of 15 staff members involved in program delivery from all but one regional centre (two to five people participated in the focus groups, including the Regional Managers). We will refer to all these data as "staff interviews" from here on. Detailed notes were kept throughout the interview process.

The interview questions were informed by Gugiu and Rodriguez-Campos' (2007) Semi-Structured Interview Protocol for conducting logic models. Questions relating to the *antecedent condition* (i.e. the situation that prompted the need for the program), the *participant profile, the essential program strategies, the influencing factors* (i.e. moderating variables that influenced the operation and success of the program), and the *key outcomes* were posed to the interviewees (see online Supplementary Material). Where appropriate, the first author also asked participants about interactions and causal links between the phenomena they discussed. For instance, if the interviewees drew attention to a positive outcome resulting from the program (e.g. self-confidence), she would ask what process within Project K generated the outcome and which factors might reduce or enhance the likelihood of this outcome occurring.

To generate a preliminary program theory, an inductive approach was taken where questions about the program components guided the analysis but no preconceived thematic categories were formulated prior to engagement with the interview notes. The notes from the staff interviews were examined in detail and themes produced by following Braun and Clarke's (2006) six-step guideline for thematic analysis: familiarization of data through multiple readings; identification of similarities and distinctions across transcripts; regrouping of similar themes into overarching thematic categories; review and revision of extracts to identify any miscategorization; reflection on the meaning assigned to thematic categories; and finalization of thematic labels.

Because the aim was to produce a theory that best represented the views of those closest to the program, both the prevalence of a concept and the experience-level of the person articulating it were considered. For instance, particular importance was given to the views of the program founders as they had been involved with the program since its inception. Another staff member had extensive involvement in the wilderness and community components of the program and found it easy to explain how these were linked to positive outcomes, whereas those involved in coordinating and monitoring the mentoring relationships evidently had more to offer in terms of the contribution of this component. It is important to note that although the first author attempted to elicit causal sequences from the focus group participants when discussing

program processes, she created the themes (i.e. the theoretical components) and was responsible for interpreting how these components could fit together to create a coherent program theory. Feedback on the appropriateness of the proposed theory was sought from staff members after consensus with the other data sources was assessed. This process is discussed further on.

3.2. Step 2: Assessing the level of consensus

The preliminary program theory was assessed against information from three comparative data sources: Project K youth participants' qualitative responses to the question "What did you gain from participating in Project K?" in an end-of-program survey; four key program manuals; and the findings from six research studies previously conducted on Project K. The program documents were used to substantiate aspects of how the program *should* operate and the factors that *should* be attended to facilitate program success. The previous research studies and the youth participant responses provided empirical evidence (albeit limited) of whether some of the theoretical components are supported in practice.

3.2.1. Project K survey respondent sample

The Project K participants came from the 40 groups involved in the program between the end of 2004 and the end of 2006. A total of 476 young people began the program in these 40 groups; however 65 did not complete it, leaving 411 available to answer the end-of-program survey. Of these, 351 responded to the question of interest here, representing 85% of the program completers. Females accounted for 47.3% (166) of the participants. The majority ($n = 280$) were of European descent; 82 participants identified as Māori (indigenous New Zealander); 43 as Pacific; 14 as Asian; and five did not fall into these categories. Note that individuals could be counted in more than one category if selecting multiple ethnicities. To analyse these data, each response was coded as to its fit with the outcomes identified in the preliminary model. Simultaneously, the responses were checked for any new outcomes or themes of interest.

3.2.2. Document review

Four key documents were used: The Project K Policies and Procedures Manual; the Project K Evaluation Manual; the Project K Liaison Manual; and the Project K Facilitator Manual. These data were analysed using a table that had a row for each component of the program theory and a column for each document. Each of the documents was then read for the presence or absence of components within the preliminary theory, and the table marked accordingly. At the same time, any material that suggested new components was noted.

3.2.3. Previous Project K research

As noted, Project K has been the focus of several other evaluation research projects; thus, as a final source for the second stage of the theory model development, the first author reviewed six relevant Project K research projects. A brief description of each is given below. Where relevant, the results are discussed in the next section.

1. Warren (2005) used a quasi-experimental pre-post design to evaluate changes in the happiness, family cohesion, and goal-directed activities of a selection of Project K participants compared with a comparison group. She also interviewed a subset of these Project K participants about their experiences at the end of each program stage.
2. O'Neill (2005) compared the content of the training program for Project K mentors against the international best practice

recommendations for mentor training and assessed the mentors' knowledge and skills before and after the training.
3. Qiao and McNaught (2007) analysed changes in self-efficacy, academic achievement and health behaviour outcomes for Project K participants in comparison to control participants from pre to post and one-year post-program in the early stages of the randomized controlled trial evaluation.
4. Somervell (2011) explored the Project K participant selection process. She analysed of the demographic characteristics and rates of young people excluded from the program for reasons of risk and interviewed program staff and relevant school personnel.
5. Zhang (2011) used the complete data set from the randomized controlled trial evaluation to examine health outcomes pre, post and one-year post-program using multilevel modelling statistical strategies. Zhang also interviewed Project K staff to assess how health behaviours were promoted within the program.
6. Hollis, Deane, Moore, and Harré (2011) used a narrative interview format to investigate the experiences of six Māori Project K graduates to assess what they felt were the critical components of the program and to ascertain how they felt the program accommodated them as Māori.

Although the above research contributed to our understanding of Project K, substantial knowledge gaps remained because the six projects focused on specific aspects of the program. In synthesizing the results from this research with the other data sources mentioned, we were able to produce a coherent picture of the program based on the information collected to date.

3.2.3.1. Results: Steps 1 (generating preliminary theory) and 2 (assessing consensus). As stated, Step 1 created a preliminary theory of change generated from interviews with the Project K founders and program staff. Step 2 assessed the level of consensus for this theory against the sources described in the previous section. In the interests of reducing repetition and not overwhelming readers with the detailed results from each part of the process, Fig. 1 shows Project K's theory of change as it stood after Steps 1 (generating preliminary theory) and 2 (assessing consensus).[1] As it happens, the theory did not change after Step 3 (direct logic analysis). This step did, however, produce further insights into the components within the theory as will be discussed later. We now give an overview of each component of the theory. As we describe each one, we also outline how it came to be part of the theory of change presented in Fig. 1.

3.2.3.2. Overview of Project K's theory of change. According to the staff interviews, New Zealand youth are plagued by "negative statistics" and community connectedness is lacking. Although there are multiple resources and opportunities for gifted or high-risk youth, those on the "cusp of risk" (i.e. at higher than average, but not extreme, risk) often fall through the cracks. This is the *antecedent condition* that Project K responded to by developing a program to support this group. From the beginning the founders considered "low self-esteem" to be an identifying characteristic of these young people; this was subsequently modified to "low self-efficacy". Thus almost all the staff interviewees talked explicitly about the program's target group being mid-risk 13–15-year-olds with low self-efficacy. Mention was also made of the attempt to exclude high-risk youth who have demonstrated behaviours such as substance abuse, violence or suicide attempts, regardless of their self-efficacy. This target group was also made evident in all program documents. However, a theme from the staff interviews

---

[1] Additional details on the results for each data source are available from the first author.
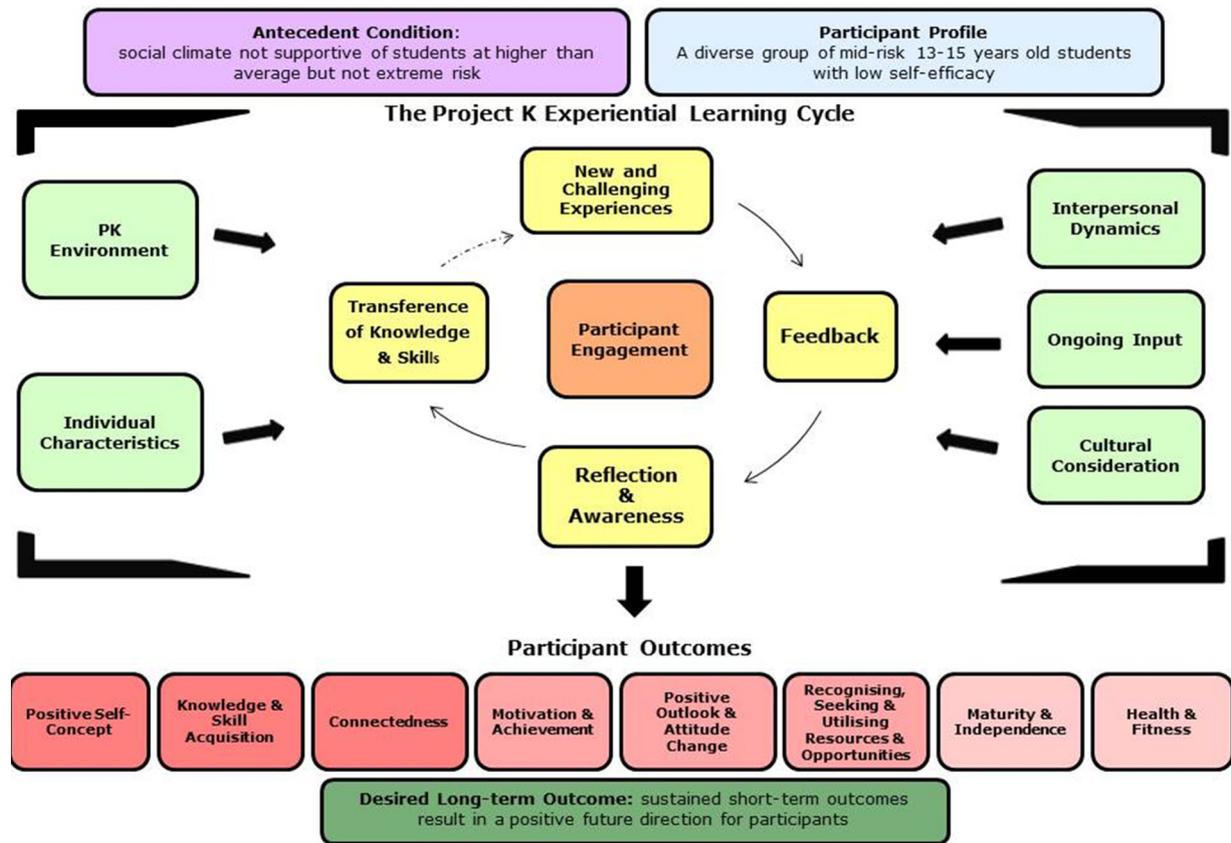
Fig. 1. The Project K theory of change depicting the antecedent condition (purple), characteristics of the target group (blue), key program processes (yellow), the influencing factors (light green and orange), short-term and intermediate outcomes (pink) and desired long-term outcomes (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

was that participants also seemed to lack "something" else (such as motivation or good decision-making skills), and in many cases had experienced disruptive events in their lives. So the *participant profile* was considered diverse.

There was considerable disagreement amongst the staff both with regard to the ideal participant characteristics and the characteristics of those who end up on the program, making this the least consistent part of the theory. Some staff argued that, although the target group should be mid-risk youth, in reality many high-risk youth participate in the program. Notably, Somervell's (2011) study on participant selection, which was consulted in Step 2 (assessing consensus), also revealed that the criteria for excluding high-risk youth is not applied in a standardized fashion across the regional centres.

Each component of the program (the Wilderness Adventure, Community Challenge and Mentoring) was described by the staff as providing *new and challenging experiences* in real-world settings with meaningful (authentic) consequences. These include physical tasks and leadership (especially in the Wilderness Adventure), teamwork (predominantly in the first two components of the program), exploration of their community and "giving back" (in the Community Challenge and Mentoring components), goal-setting (especially in the Mentoring component), and relationship building (throughout). Many staff then went on to describe a process of experiential learning. Some even explicitly talked about experiential learning cycles. Both these experiences and the program personnel were seen to provide *feedback* that encourages *reflection and awareness*, which in turn leads to *transference* of their new knowledge and skills to other settings. These themes were echoed in several of the responses from the Project K participants to the target question on the end-of-program survey ("What did

you gain from participating in Project K?). For example, one participant wrote, "*I've learnt more about myself and from talking to others that I don't know*", while another claimed, "*I have gained many valuable life skills that I can apply to everyday situations, including setting goals, socially, in work and other areas.*" In addition, the interviews that Zhang (2011) conducted with program staff and Hollis et al. (2011) and Warren (2005) conducted with youth participants corroborated several of the learning cycle components, including challenging mastery experiences, reflection, and facilitator and mentor support.

The change process was described by the staff as being influenced by *participant engagement*, which acts as a centrepiece to keep the cycle moving. A few noted this in relation to the lack of engagement in the Community Challenge, which they acknowledged influenced the impact of this program component. The findings from Warren's (2005) interviews with Project K participants supported their claims.

Several factors were assumed to support engagement in the learning cycle process. These were: the Project *K environment* (intense, safe and supportive of the young person's autonomy or agency with the provision of scaffolded support); the *interpersonal dynamics* (the interactions within the program group, between participants and program staff or mentors, and between program staff and the participants' families and schools); and the *individual characteristics* of participants (e.g. socioemotional history, behavioural conduct), as well as those of the facilitators and mentors (e.g. level of attunement). Finally, the *ongoing input* provided by many different individuals within the context of a three-component program of 14 months' duration was described as strongly influencing sustained participant outcomes. We included ongoing reinforcement and support from key people in the youth

participants' external networks (family members, teachers and peers) in this thematic category and a few staff members acknowledged this was lacking for some young people. The difficulties with the reintegration from the Wilderness Adventure to school life expressed by the interviewees in Warren's (2005) study reflected this concern as these young people perceived some teachers to be unwilling to support them to catch up on work missed. The potential problems associated with reintegration after the Wilderness Adventure due to a lack of understanding from family members, teachers and friends was also addressed in the Facilitator Manual. This was accompanied by suggested strategies for facilitating a smoother transition, but evidently this is not always easy.

The staff members claimed that if all goes well, the young person will experience one or more positive outcomes, including a more *positive self-concept, knowledge and skill acquisition, connection* with others, *stronger motivation* and flow-on *achievement*, a *positive outlook, the recognition and use of new resources and opportunities, maturity, independence*, and *greater mental and physical health and fitness*. In the long term, several staff described a *positive future direction for participants*. All of the immediate outcomes were substantiated by the youth participants' comments. More than 99% of the respondents reported gains as a result of Project K with the most prevalent themes relating to positive changes in self-concept (63%), knowledge or skill acquisition (46% primarily in relation to goal-setting and social skills), motivation and achievement (20%), and connectedness (19%). Notably, the program documents put a much higher emphasis on physical fitness than did the other sources. Although a small proportion of the youth participants acknowledged a sport- or fitness-related gain (5%), this outcome was not acknowledged to be of primary importance by the staff interviewed. Zhang's (2011) study also suggested the program had little direct effect on physical fitness, much of which appeared due to it not being emphasized by program personnel once the Wilderness Adventure ended.

Ultimately, there was high concordance across the four data sources. All of the components of the preliminary logic model from Step 1 (generating preliminary theory) were present in at least two of the four program manuals with the exception of the antecedent condition description (which was not contradicted, but not relevant to the program documents) and the additional nuances concerning the population reached revealed by the staff interviews. An additional component, cultural consideration, was added to the model at Step 2 (assessing consensus). Although this was almost absent from the staff interviews, the program documents clearly stated the need for culturally appropriate processes and sensitivity to the participants' different ethnic backgrounds. Cultural awareness was included in one of the mentor training modules evaluated by O'Neill (2005) and Hollis et al.'s (2011) findings also highlighted the need for cultural sensitivity in relation to Māori participants. Consequently, cultural consideration was included as an influencing factor in Fig. 1.

As outlined above, the youth participant data was based on a large sample of participants and were almost exclusively relevant to the program theory outcomes thus these findings were analysed in tandem with the outcomes generated from the staff interviews and contributed greatly to refining the outcomes component of the theory. These outcomes were reordered by prevalence (highest to lowest moving from left to right in Fig. 1) based on the frequency of youth responses relevant to each category. Staff were asked to distinguish between short, intermediate and long-term outcomes but they had difficulty confirming any outcomes that emerged after the youth had left the program, other than their expectations that the short-term outcomes would be sustained and youth would be heading in a positive direction. Because the expected timeframe for the outcomes was not described in any of the other data sources

we opted to collapse the short and intermediate-term outcomes and to allow the long-term outcome theme to reflect the vagueness of the staff comments. Further work will be needed to refine this aspect of the theory.

A diagram of the program theory was presented to a group of Project K staff, including contracted providers of the adventure and community-based components and the program founders at a national conference (some of whom had been interviewed, but not all). Staff members were given time to provide written feedback on each component of the model. No one disagreed with any of the major components. Most restated the aspects of the model that resonated strongly with their experiences. For instance, several staff reiterated that the target group *should* be mid-risk youth but that often high-risk youth were selected and this did create difficulties. Many strongly endorsed the idea that outcomes are maximized through ongoing input and scaffolded support. A few provided additional suggestions of important program features (e.g. that the outdoor adventure focus fit well with the New Zealand culture, the active methods of learning were well suited to young people with low academic self-efficacy, fun was an important element). Since each of the suggestions was put forward by a single individual and no one discredited the proposed theoretical components, we chose to retain the proposed version. A report summarizing each component of the model and the feedback obtained from individuals at the conference (including the additional suggestions) was produced and distributed to staff members. The report also pointed out that the program theory should remain open to change, be revisited and modified as new information comes to light.

3.3. Step 3: Direct logic analysis

A *direct logic analysis* ascertains if a program's theory is consistent with relevant academic theories, empirical evidence, or expert opinions (Brousselle & Champagne, 2011). For this step, each aspect of the Project K logic model was evaluated against the best practice guidelines, causal mechanisms, moderating influences, and outcomes discussed in academic literature on adventure, service learning, and mentoring programs. Table 1 provides a summary of the logic analysis findings.

As can be seen in Table 1, Project K's theory of change generally held up extremely well in light of the literature. Thus, no changes were made to the model as shown in Fig. 1. Here we discuss particular insights revealed in this third step. First cultural consideration, which was not identified by the staff interviews as a key part of the Project K theory of change, has also received relatively little attention within the youth programming literature. Although highlighted as important in the mentoring literatures (Darling, Bogat, Cavell, Murphy, & Sanchez, 2006; Evans, Jory, & Dawson, 2005; Sánchez & Colón, 2005), it is largely ignored in the fields of service learning and adventure programming and few empirical studies of youth programs have explicitly focused on differential responsiveness to and effectiveness of programs for youth of different cultural or ethnic backgrounds (Deane, 2012; Deane & Harré, 2013; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). Similarly, physical health, also underplayed by Project K staff, appears of limited interest as a positive outcome in the youth development programming literature. Rather physical health tends to be the focus of interventions explicitly targeting those outcomes (see Brown & Summerbell, 2009 and De Bourdeaudhuij et al., 2011 for reviews of school-based programs targeting youth obesity).

One factor considered to influence success in the service-learning literature, but absent from the Project K theory of change, was the importance of having strong connections with community members and organizations (Dymond, Renzaglia, & Chun, 2007;

Table 1
Summary of direct logic analysis results.

Theoretical component	Description	Relevant sources
Antecedent condition	Youth development programs are needed to supplement the lack of social supports and opportunities within young people's natural environment. These may be particularly effective for those at higher than average but not extreme risk, unless additional resources and support can be provided	Deane and Harré (2013), Benson (1997), DuBois et al. (2011), Herrera et al. (2013), Roth and Brooks-Gunn (2003), Schwartz et al. (2011)
Experiential learning cycle	Experiential learning results from new and challenging real-world experiences, feedback, reflection, awareness and transference of knowledge and skills. This learning influences and is influenced by participant engagement	Deane and Harré (2013), Dymond et al. (2007), Eyler (2002), McKenzie (2000, 2003), Sibthorp and Arthur-Banning (2004), Terry and Bohnenberger (2003, 2004)
The program environment	Positive growth is promoted by an intense, safe environment that encourages autonomy while providing suitable scaffolds	Deane and Harré (2013), Davis-Berman and Berman (2002), Dymond et al. (2007), Larson, Walker, and Pearce (2006), Li and Julian (2012), McKenzie (2000), Morrow and Styles (1995), Roth and Brooks-Gunn (2003), Terry and Bohnenberger (2003)
Individual characteristics	The characteristics of facilitators, mentors, and participants can both enable and limit engagement and learning	Deane and Harré (2013), Darling et al. (2006), DuBois et al. (2011), Grossman et al. (2012), Hamilton and Fenzel (1988), McKenzie (2000), Priest and Gass (2005), Sánchez and Colón (2005), Scales, Benson, Roehlkepartain, Sesma, and van Dulmen (2006), Schwartz, Rhodes, Spencer, and Grossman (2013), Sibthorp and Arthur-Banning (2004)
Interpersonal dynamics	Group dynamics have an important impact on individuals' experiences and in mentoring relationships, a strong the mentor–mentee bond is critical	Grossman and Bulle (2006), Grossman et al. (2012), Martin and Leberman (2005), McKenzie (2003), Neill and Dias (2001), Parra, DuBois, Neville, Pugh-Lilly, and Povinelli (2002), Rhodes and Lowe (2009), Spencer (2006), Terry and Bohnenberger (2003, 2004)
Ongoing input	A progression of staged program components and sustained sources of personal support encourages deeper learning and lasting change	Catalano et al. (2004), Durlak et al. (2007), Dymond et al. (2007), Grossman and Rhodes (2002), Grossman et al. (2012), Hattie, Marsh, Neill, and Richards (1997), Li and Julian (2012), Schwartz et al. (2013), Urban (2008)
Cultural consideration	Sensitivity to culture is important	Deane and Harré (2013), Darling et al. (2006), Evans et al. (2005), Fabrizio and Neill (2005), Lugg (2004), Purdie and Neill (1999), Riggs and Greenberg (2004), Roth and Brooks-Gunn (2003)
Outcomes	Positive self-concept; knowledge and skill acquisition; connectedness; motivation and achievement; positive outlook and attitude change; recognizing, seeking and utilizing resources and opportunities; maturity and independence; health and fitness	Deane and Harré (2013), Billig (2004), Catalano et al. (2004), Conrad and Hedin (1991), DuBois et al. (2011), Hattie et al. (1997), Wilson and Lipsey (2000), Yates and Youniss (1996)

Note. The participant profile is not shown, as this derives from the "antecedent condition". For the purposes of this summary, the components within the learning cycle are not broken down nor are the participant outcomes. A more detailed version of the table is available from the authors on request.

McBride, Pritzker, Daftary, & Tang, 2007). Connection with participants' families is also seen to enhance likelihood of success in youth development programs (Catalano et al., 2004; Durlak et al., 2007). While the idea of connecting with families was an implicit part of the *ongoing input* theme, as we know from several years of working with FYD, staff are aware of a gap regarding family involvement. Nevertheless, it is not explicitly built into the theory of change for a number of reasons. One concerns resource implications, particularly the time it would take for the Regional Centre in each region to establish and maintain these relationships. However, in relation to family, staff members have also talked about the desirability in some cases for Project K to be separate from the participants' family dynamics, particularly when these involve conflict.

It was interesting to note that mentoring programs have recently been shown to be particularly effective for a middle-risk group (DuBois et al., 2011; Schwartz, Rhodes, Chan, & Herrera, 2011), which is the target group for Project K. Further, mentors working with high risk youth may face additional challenges and require more extensive training and support (Herrera, DuBois, & Grossman, 2013). DuBois et al. (2011) argued that intensive and therapeutic support is often more appropriate for high-risk individuals than mentoring, or mentoring alone. It is not surprising then that in Project K regions where young people at high risk were selected for participation, staff encountered substantial challenges and expressed the need for additional resources and support.

The desirability of a relatively long mentoring period is also notable (MENTOR, 2009). Although Project K offers this, some participants leave the program during this stage which some studies show may be detrimental (Grossman & Rhodes, 2002; Spencer, 2007). Thus, following up with those that do drop out of Project K to evaluate why they do so and implementing strategies to mitigate any iatrogenic effects is recommended.

Grossman, Chan, Schwartz, and Rhodes' (2012) study on the impact of re-matching youth to new mentors after early termination of the initial match in school based mentoring programs illustrates that it is not simply the dosage of mentoring that matters, but the integrity of the relationship that really impacts positive outcomes. Ongoing input, as depicted in the Project K theory of change, extends the common notion of program dosage and highlights a similar point. Rather than simply a matter of frequency of contact with program services and program length, ongoing input captures the benefit of a progression of scaffolded learning experiences accompanied by multiple sources of support (from the program group, program facilitators, directors and mentors, and key persons in the young people external support networks) over a lengthy duration. This allows participants to transfer new lessons and skills to consecutive program settings, firmly establishes the learning process before program completion, and increases the chances that changes will be sustained after program completion. In addition, it provides greater and more varied sources from which to draw personal support. Interestingly, this idea aligns strongly with Li and Julian (2012) recent conceptualization of *developmental relationships*, which they argue is the primary "active ingredient" that distinguishes effective child and adolescent interventions from ineffective ones.

3.4. Step 4: Test the hypothesized links

The first three stages build the conceptual theory of the program. Although the staff and youth perceptions and the findings from previous research provide some empirical support for the theoretical components, in a sense the theory is still an idealized and descriptive version of how the program operates. Thus, the final stage of the process is to explicitly test the links hypothesized by the theory. As noted, this article covers only the first three stages in detail. However, here we outline the links that are currently being tested as well as other evaluation questions raised by the theory-building process.

It is worth noting that Donaldson (2003, 2007) recommends involving various stakeholders in using the program theory to prioritize evaluation questions for further evaluation. In our case, Project K stakeholders (FYD and Project K staff, board members, and some mentors and parents) were involved in voting on the most important evaluation questions *prior* to the theory-development stage. The reason for the reversal of Donaldson's recommended steps stemmed from our desire to ascertain if a theory-driven approach to the evaluation of Project K was of interest to our stakeholders. That is, the entire theory-driven approach was one of the evaluation options they could vote on. The results of the voting exercise revealed a great interest in understanding how and for whom the program works thus we proceeded with a theory-driven approach as a first step in answering these questions.

Subsequent to the development of the theory of change, the researchers worked with a smaller group from FYD to establish priorities for further evaluation, based on the theory. A quasi-experimental study (single-group, pre-post program design) is now under way that investigates links between aspects of the Project K environment, the interpersonal dynamics participants experience, their individual characteristics, and key outcomes. This study investigates participants' self-reported experiences of autonomy support, competence, relatedness, and intrinsic motivation within each of the program stages using measures derived from self-determination theory (Ryan & Deci, 2002) and additional outcomes to those used in previous Project K evaluations (i.e. global motivation, Guay, Mageau, & Vallerand, 2003; relatedness, Harré & Bullen, 2010; and sense of community, Peterson, Speer, & McMillan, 2008).

There is also great interest in a future observational study which would aim to develop a structured observational schedule based on the components of Project K's model. An external observer would then use this to code the characteristics of the programs in action. The findings could then be tested statistically to obtain empirical support for the theoretical links. The strongest predictors, mediators and moderators of program outcomes could also be assessed.

Variations of the program theory development process have also been employed by FYD and other research colleagues to clarify the theory underlying other FYD programs. The theory-building process was also useful for identifying other areas for further research. As noted, the process revealed a mismatch between the official target group and that defined by some program staff. This raises the question of whether these staff members are simply using different language or whether the program is targeting different young people in different regional centres. If different young people are being targeted, this provides an excellent opportunity to investigate how the program processes and outcomes are related to a wide variety of individual differences, as the current literature offers few empirically based conclusions. Similarly, further study on how young people from different cultures respond to the program is warranted. Finally, the issue of cultural sensitivity clearly needs to be examined. Was it overlooked by those delivering the program because it is not considered useful or for other reasons? Is cultural sensitivity of particular importance in this context or not? Intriguingly, although the Māori participants in Hollis et al.'s (2011) study of Project K noticed some cultural insensitivity, they also liked that the program generally ignored their ethnicity because it allowed to them to escape the negative stereotypes commonly associated with their ethnicity and thus feel equal to the other participants.

4. Lessons learned

This theory-building process allowed us to assess organizational consensus about the program's theory of change and served as a platform from which the legitimacy of the theory could be assessed against the academic literature. It also provided the groundwork for testing the proposed links, and raised several other evaluation questions as noted in the outline of Step 4.

In addition it provided a strong platform for program development because the process of creating the theory revealed issues related to how the program should or could be improved. For instance, four development opportunities were illuminated. These related to the challenges or discrepancies between data sources described in the results for Step 1 (generating preliminary theory) and 2 (assessing consensus), namely: creating a more engaging Community Challenge experience; reintegrating participants into supportive home and school contexts following the Wilderness Adventure; explicitly attending to the cultural backgrounds of the participants; and revisiting the position of the physical health and academic outcomes as primary program objectives. Consequently, the first author facilitated action-oriented brainstorming with a range of stakeholders to identify strategies that could be implemented to address these opportunities. Although we do not know if they were particularly receptive because of the inclusive nature of the theory-building process, it seems feasible that this was the case.

At a more general level, this study helps bridge the gap between evaluation theory and practice. As stated in the introduction, detailed examples of how theory-driven evaluations are put into practice are scarce (Coryn et al., 2011; Donaldson, 2007). We have provided one and hope in doing so have defended criticisms that the model lacks real-world utility (Stufflebeam, 2001). We also hope that youth program developers and evaluators are interested in adapting our version of the PTDES for the evaluation and development of their own programs. Not only does this process reveal interesting and useful information about the program and how it fits within current knowledge about best practice, it also gets program stakeholders talking with each other and engaging with a wide variety of views.

Although we suggest this approach has numerous benefits and recommend its use, there are limitations. A common criticism of theory-driven models of programs outlined by Donaldson (2003) is that they do not reflect the complex systemic nature of reality. Although Donaldson also argues that this is beneficial as it helps focus communication on the core elements of the program, it is certainly the case that any pictorial model of a program's theory of change is partial and serves to obscure some opinions about and features of the program. As we constructed this model, we made numerous decisions about what to include, what to leave out, how to organize the components and so on. We intentionally avoided the more standard inputs-activities-outputs-outcomes logic model format of depicting a program theory as we felt this would be a gross simplification of the program theory but there are no precise rules that can be followed at every step of the way that result in a fully objective and complete model. Instead it is a living document and should change as the program evolves. We also kept detailed notes expanding the content of each component within the

diagram, only some of which is included here. We recommend this to others as a way to help mitigate this limitation.

A further limitation concerns the difficulty of assessing the model against the literature. With numerous overlapping concepts called by slightly different names in the youth development field and every program being unique and operating in a unique context, it is impossible to directly verify the theory of the program being examined with existing literature. Nevertheless, the process is arguably still an important part of developing a program that seems as if it should work.

This can also be a time consuming undertaking, especially in a large organization with varied stakeholders dispersed across multiple geographical regions, as was the case with Project K. It would have been ideal to include a more varied range of stakeholder group representatives in the development of the preliminary program theory but the complexity of FYD's organizational context presented numerous challenges in this regard. The hierarchical nature of the organization meant that access to downstream stakeholders such as the mentors, youth participants and their families was only possible through indirect communication via the Regional Coordinators, many of whom already felt over-extended in their program responsibilities and, due to this, were already hesitant to commit to participation in the focus groups. Consequently, the program theory is heavily weighted towards the staff members' perspective and important voices are absent. We acknowledge this limitation but felt justified in focusing on this stakeholder group because this combination of individuals were well placed to discuss how the program was first conceptualized and how it actually unfolds in the here and now.

In our case, consensus across the organization was high (with regards to the program theory) thus few tensions needed to be resolved. This seems unusual for an organization of this size but may be because the program was developed in careful consultation with youth development and educational experts from the beginning, the program documents are thorough, and FYD's National Support Office provides training and support to all regions. The theory development process could, however, generate tensions between stakeholders if and when their perspectives differ. However, such tensions can be reduced by giving people multiple opportunities for input and providing feedback on how their views have been incorporated. Regardless, we believe the opportunities for consciousness-raising through critical reflection on the important features of a program created by a PTDES process is part of high quality service provision.

The process also may not directly lead to the type of empirical evidence that some funders are seeking. However, it may increase the chances of finding favourable empirical outcomes in subsequent studies by more clearly identifying what the program can achieve and potential confounds that, when unaccounted for, mask true program effects (Cook, 2000; Donaldson, 2007). For example, the randomized controlled trial evaluation set up prior to this process measured many variables that, in retrospect, were unlikely to show positive outcomes, and indeed failed to do so (see Zhang, 2011). Given the time, expense and effort that goes into a robust empirical evaluation, using a PTDES approach as the starting point may serve to actually speed up the process of demonstrating success.

Finally, Step 4 (testing the hypothesized links) cannot be realistically completed within one study. Weiss (2000) advised that one must be selective about the questions pursued at this point in the theory-driven evaluation process and a responsive evaluator will be attentive to which questions and methods are most appropriate for the specific context. The follow up studies proposed for Project K in Step 4 are complex in design and require a substantial time commitment and methodological expertise and still, only a portion of the theoretical links will be explored.

The PTDE process need not be as extensive or time-consuming as that illustrated in our case study of Project K, however. More limited variations may be desirable in many circumstances. These include when the aims of the program are very clear and easily measured, and when the program is new or small and so less socially and structurally complex. The first three steps can feasibly be accomplished in a much shorter time frame if stakeholders can come together for one or two sessions to develop and review the proposed content for the preliminary theory. The time frame for the fourth step should then depend on nature of the questions and evidence desired by the stakeholders and the logistical constraints associated with these.

5. Conclusion

To conclude, we recommend a PTDES approach to program practitioners and developers and evaluation researchers, especially those working with youth. We believe it provides interesting insights, promotes informed discussion amongst stakeholders and ensures that programs attend to relevant empirical evidence. It also allows for well-planned ongoing evaluation that involves stakeholders.

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.evalprogplan.2014.03.009.

References

Benson, P. L. (1997). *All kids are our kids. What communities must do to raise caring and responsible children and adolescents.* San Francisco, CA: Jossey-Bass.

Billig, S. H. (2004). *Head, hearts, hands: The research on K-12 service-learning.* St. Paul, MN National Youth Leadership Council.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101 http://dx.doi.org/10.1191/1478088706qp063oa

Brouselle, A., & Champagne, F. (2011). Program theory evaluation: Logic analysis. *Evaluation and Program Planning, 34,* 69–78 http://dx.doi.org/10.1016/j.evalprogplan.2010.04.001

Brown, T., & Summerbell, C. (2009). Systematic review of school-based interventions that focus on changing dietary intake and physical activity levels to prevent childhood obesity: An update to the obesity guidance produced by the National Institute of Health and Clinical Excellence. *Obesity Reviews, 10,* 110–141 http://dx.doi.org/10.1111/j.1467-789X.2008.00515.x

Catalano, R. F., Berglund, M. L., Ryan, J. A. M., Lonczak, H. S., & Hawkins, J. D. (2004). Positive youth development in the United States: Research findings on evaluations of positive youth development programs. *Annals of the American Academy of Political and Social Science, 591*(1), 98–124 http://dx.doi.org/10.1177/0002716203260102

Conrad, D., & Hedin, D. (1991). School-based community service: What we know from research and theory. *Phi Delta Kappa International, 72*(10), 743–749.

Cook, T. D. (2000). The false choice between theory-based evaluation and experimentation. *New Directions for Evaluation, 2000*(87), 27–34 http://dx.doi.org/10.1002/ev.1179

Coryn, C. S., Noakes, L. A., Westine, C. D., & Schroter, D. C. (2011). A systematic review of theory-driven evaluation practice from 1999 to 2009. *American Journal of Evaluation, 32*(2), 199–226 http://dx.doi.org/10.1177/1098214010389321

Darling, N., Bogat, G. A., Cavell, T. A., Murphy, S. E., & Sánchez, B. (2006). Gender, ethnicity, development, and risk: Mentoring and the consideration of individual differences. *Journal of Community Psychology, 34*(6), 765–780 http://dx.doi.org/10.1002/jcop.20128

Davis-Berman, J., & Berman, D. (2002). Risk and anxiety in adventure programming. *Journal of Experiential Education, 25*(2), 305–310.

Deane, K. L. (2012). *Project K in black & white: A theory-driven & randomised controlled trial evaluation of a youth development programme (Unpublished doctoral thesis).* Auckland, New Zealand: University of Auckland.

Deane, K. L., & Harré, N. (2013). The youth adventure programming model. *Journal of Research on Adolescence.* http://dx.doi.org/10.1111/jora.12069

De Bourdeaudhuij, I., Van Cauwenberghe, E., Spittaels, H., Oppert, J.-M., Rostami, C., Brug, J., et al. (2011). School-based interventions promoting both physical activity and healthy eating in Europe: A systematic review within the HOPE project. *Obesity Reviews, 12*, 205–216 http://dx.doi.org/10.1111/j.1467-789X.2009.00711.x

Donaldson, S. I. (2003). Theory-driven program evaluation in the new millennium. In S. I. Donaldson & M. Scriven (Eds.), *Evaluating social programs and poblems: Visions for the new millennium* (pp. 109–141). Mahwah, NJ: Erlbaum.

Donaldson, S. I. (2007). *Program theory-driven evaluation science. Strategies and applications.* New York, NY: Lawrence Erlbaum Associates.

DuBois, D. L., Portillo, N., Rhodes, J. E., Silverthorn, N., & Valentine, J. C. (2011). How effective are mentoring programs for youth? A systematic assessment of the evidence. *Psychological Science in the Public Interest, 12*(2), 57–91 http://dx.doi.org/10.1177/1529100611414806

Durlak, J. A., Taylor, R. D., Kawashima, K., Pachan, M. K., DuPre, E. P., Celio, C. I., et al. (2007). Effects of positive youth development programs on school, family, and community systems. *American Journal of Community Psychology, 39*(3–4), 269–286 http://dx.doi.org/10.1007/s10464-007-9112-5

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*(1), 405–432 http://dx.doi.org/10.1111/j.1467-8624.2010.01564.x

Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American Journal of Community Psychology, 45*, 294–309 http://dx.doi.org/10.1007/s10464-010-9300-6

Dymond, S. K., Renzaglia, A., & Chun, E. (2007). Elements of effective high school service learning programs that include students with and without disabilities. *Remedial and Special Education, 28*(4), 227–243.

Evans, I. M., Jory, A., & Dawson, N. (2005). International: Australia and New Zealand. In D. L. DuBois & M. J. Karcher (Eds.), *Handbook of youth mentoring* (pp. 408–421). Thousand Oaks, CA: Sage Publications Ltd.

Eyler, J. (2002). Reflection: Linking service-learning and learning – linking students and communities. *Journal of Social Issues, 58*(3), 517–534.

Fabrizio, S. M., & Neill, J. T. (2005). Cultural adaptation in outdoor programming. *Australian Journal of Outdoor Education, 9*(2), 44–56.

Gallagher, K. M., Stanley, A., Shearer, D., & Mosca, C. (2005). Implementation of youth development programs: Promise and challenges. *Journal of Adolescent Health, 37*, S61–S68 http://dx.doi.org/10.1016/j.jadohealth.2005.05.006

Grossman, J. B., & Bulle, M. J. (2006). Review of what youth programs do to increase the connectedness of youth with adults. *Journal of Adolescent Health, 39*(6), 788–799.

Grossman, J. B., Chan, C. S., Schwartz, S. E. O., & Rhodes, J. E. (2012). The test of time in school-based mentoring: The role of relationship duration and re-match on academic outcomes. *American Journal of Community Psychology, 49*, 43–54 http://dx.doi.org/10.1007/s10464-011-9435-0

Grossman, J. B., & Rhodes, J. E. (2002). The test of time: Predictors and effects of duration in youth mentoring relationships. *American Journal of Community Psychology, 30*(2), 199–219 http://dx.doi.org/10.1023/a:1014680827552

Guay, F., Mageau, G. A., & Vallerand, R. J. (2003). On the hierarchical structure of self-determined motivation: A test of top-down, bottom-up, reciprocal, and horizontal Effects. *Personality and Social Psychology Bulletin, 29*(8), 992–1004 http://dx.doi.org/10.1177/0146167203253297

Gugiu, P. C., & Rodriguez-Campos, L. (2007). Semi-structured interview protocol for constructing logic models. *Evaluation and Program Planning, 30*, 339–350 http://dx.doi.org/10.1016/j.evalprogplan.2007.08.004

Hamilton, S. F., & Fenzel, L. M. (1988). The impact of volunteer experience on adolescent social development: Evidence of program effects. *Journal of Adolescent Research, 3*(1), 65–80 http://dx.doi.org/10.1177/074355488831006

Harré, N., & Bullen, P. (2010). Personal project commitment in adolescence: The role of relatedness, competence and integrity. *European Journal of Social Psychology, 40*(2), 235–251 http://dx.doi.org/10.1002/ejsp.616

Hattie, J., Marsh, H. W., Neill, J. T., & Richards, G. E. (1997). Adventure education and outward bound: Out-of-class experiences that make a lasting difference. *Review of Educational Research, 67*(1), 43–87 http://dx.doi.org/10.3102/00346543067001043

Herrera, C., DuBois, D. L., & Grossman, J. (2013). *The role of risk: Mentoring experiences and outcomes for youth with varying risk profiles.* New York, NY: A Public/Private Ventures project distributed by MDRC.. http://www.mdrc.org/sites/default/files/Role%20of%20Risk_Final-web%20PDF.pdf

Hollis, H., Deane, K., Moore, J., & Harré, N. (2011). Young Māori perceptions of a youth development programme. *Kotuitui: New Zealand Journal of Social Sciences Online, 6*(1–2), 50–61 http://dx.doi.org/10.1080/1177083x.2011.615847

Larson, R. W. (2011). Positive development in a disorderly world. *Journal of Research on Adolescence, 21*(2), 317–334 http://dx.doi.org/10.1111/j.1532-7795.2010.00707.x

Larson, R., Walker, K., & Pearce, N. (2006). A comparison of youth-driven and adult-driven youth programs: Balancing inputs from youth and adults. *Journal of Community Psychology, 33*(1), 57–74 http://dx.doi.org/10.1002/jcop.20035

Li, J., & Julian, M. M. (2012). Developmental relationships as the active ingredient: A unifying working hypothesis of "What Works" across intervention settings. *American Journal of Orthopsychiatry, 82*(2), 157–166 http://dx.doi.org/10.1111/j.1939-0025.2012.01151.x

Lugg, A. (2004). Outdoor adventure in Australian outdoor education: Is it a case of roast for Christmas dinner? *Australian Journal of Outdoor Education, 8*(1), 4–11.

Martin, A. J., & Leberman, S. I. (2005). Personal learning or prescribed educational outcomes: A case study of the Outward Bound experience. *Journal of Experiential Education, 28*(1), 44–59.

McBride, A. M., Pritzker, S., Daftary, D., & Tang, F. (2007). Youth service. *Journal of Community Practice, 14*(4), 71–89 http://dx.doi.org/10.1300/J125v14n04_05

McKenzie, M. (2003). Beyond "The Outward Bound Process": Rethinking student learning. *Journal of Experiential Education, 26*(1), 8–23.

McKenzie, M. D. (2000). How are adventure education program outcomes achieved?.: A review of the literature. *Australian Journal of Outdoor Education, 5*(1), 19–27.

MENTOR. (2009). *Elements of effective practice for mentoring* (3rd ed.). Alexandria, VA: MENTOR.. http://www.mentoring.org/downloads/mentoring_1222.pdf

Morrow, K. V., & Styles, M. B. (1995). *Building relationships with youth in program settings: A study of big brothers/big sisters.* Philadelphia, PA: Public/Private Ventures.. http://www.issuelab.org/resource/building_relationships_with_youth_in_program_settings_a_study_of_big_brothersbig_sisters

Ministry of Youth Development. (2009). *Structured youth development programmes: A review of evidence ('A report undertaken for the Ministry of Youth Development, September 2009').* Wellington: Ministry of Youth Development.. http://www.myd.govt.nz/documents/resources-and-reports/publications/structured-youth-development-programmes-review-of-evidence-full-report-.pdf

Neill, J. T., & Dias, K. L. (2001). Adventure education and resilience: The double-edged sword. *Journal of Adventure Education & Outdoor Learning, 1*(2), 35–42 http://dx.doi.org/10.1080/14729670185200061

O'Neill, M. (2005). *A mentor training programme: Evaluation of effectiveness* (Unpublished master's thesis) Auckland, New Zealand: University of Auckland.

Parra, G. R., DuBois, D. L., Neville, H. A., Pugh-Lilly, A. O., & Povinelli, N. (2002). Mentoring relationships for youth: Investigation of a process-oriented model. *Journal of Community Psychology, 30*(4), 367–388 http://dx.doi.org/10.1002/jcop.10016

Peterson, N. A., Speer, P. W., & McMillan, D. W. (2008). Validation of a brief sense of community scale: Confirmation of the principal theory of sense of community. *Journal of Community Psychology, 36*(1), 61–73 http://dx.doi.org/10.1002/jcop.20217

Priest, S., & Gass, M. A. (2005). *Effective leadership in adventure programming* (2nd ed.). Champaign, IL: Human Kinetics.

Purdie, N., & Neill, J. (1999). Japanese students down-under: Is Australian outdoor education relevant to other cultures? *Australian Journal of Outdoor Education, 4*(1), 48–56.

Qiao, C., & McNaught, H. (2007). *Evaluation of Project K.* Wellington, New Zealand: Ministry of Social Development.

Rhodes, J. E., & Lowe, S. R. (2009). Mentoring in adolescence. In R. M. Lerner, & L. Steinberg (Eds.), Hoboken, NJ: John Wiley & Sons.

Riggs, N. R., & Greenberg, M. T. (2004). After-school youth development programs: A developmental–ecological model of current research. *Clinical Child and Family Psychology Review, 7*(3), 177–190 http://dx.doi.org/10.1023/B:CCFP.0000045126.83678.75

Roth, J. L., & Brooks-Gunn, J. (2003). Youth development programs: Risk, prevention and policy. *Journal of Adolescent Health, 32*(3), 170–182.

Ryan, R. M., & Deci, E. L. (2002). An overview of self-determination theory. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 3–33). Rochester, NY: University of Rochester Press.

Sánchez, B., & Colón, Y. (2005). Race, ethnicity, and culture in mentoring relationships. In D. L. DuBois & M. J. Karcher (Eds.), *Handbook of youth mentoring* (pp. 191–217). Thousand Oaks, CA: Sage Publications Ltd.

Scales, P. C., Benson, P. L., Roehlkepartain, E. C., Sesma, A., Jr., & van Dulmen, M. (2006). The role of developmental assets in predicting academic achievement: A longitudinal study. *Journal of Adolescence, 29*(5), 691–708 http://dx.doi.org/10.1016/j.adolescence.2005.09.001

Schwartz, S. E. O., Rhodes, J., Chan, C. S., & Herrera, C. (2011). The impact of school-based mentoring on youth with different relational profiles. *Developmental Psychology, 47*(2), 450–462 http://dx.doi.org/10.1037/a0021379

Schwartz, S. E., Rhodes, J. E., Spencer, R., & Grossman, J. B. (2013). Youth initiated mentoring: Investigating a new approach to working with vulnerable adolescents. *American Journal of Psychology, 52*, 155–169 http://dx.doi.org/10.1007/s10464-013-9585-3

Sibthorp, J., & Arthur-Banning, S. (2004). Developing life effectiveness through adventure education: The roles of participant expectations, perceptions of empowerment, and learning relevance. *Journal of Experiential Education, 27*(1), 32–50.

Somervell, J. (2011). *Participant selection and exclusion in a New Zealand youth development programme* (Unpublished doctoral dissertation) Auckland, New Zealand: University of Auckland.

Spencer, R. (2006). Understanding the mentoring process between adolescents and adults. *Youth & Society, 37*(3), 287–315 http://dx.doi.org/10.1177/0743558405278263

Spencer, R. (2007). It's not what I expected": A qualitative study of youth mentoring relationship failures. *Journal of Adolescent Research, 22*(4), 331–354 http://dx.doi.org/10.1177/0743558407301915

Stufflebeam, D. (2001). Evaluation models. *New Directions for Evaluation, 2001*(89), 7–98 http://dx.doi.org/10.1002/ev.3

Terry, A. W., & Bohnenberger, J. E. (2003). Service learning: Fostering a cycle of caring in our gifted youth. *Journal of Secondary Gifted Education, 15*(1), 23–32.

Terry, A. W., & Bohnenberger, J. E. (2004). Blueprint for incorporating service learning: A basic, developmental, K-12 service learning typology. *Journal of Experiential Education, 27*(1), 15–31.

Urban, J. B. (2008). Components and characteristics of youth development programs: The voices of youth-serving policymakers, practitioners, researchers, and adolescents. *Applied Developmental Science, 12*(3), 128–139 http://dx.doi.org/10.1080/10888690802199400

Warren, J. K. (2005). *Adolescent well-being: Effects of time and intervention* (Unpublished doctoral dissertation) Auckland, New Zealand: University of Auckland.

Weiss, C. H. (2000). Which links in which theories shall we evaluate? *New Directions for Evaluation, 87,* 35–45.

Wilson, S. J., & Lipsey, M. W. (2000). Wilderness challenge programs for delinquent youth: A meta-analysis of outcome evaluations. *Evaluation and Program Planning, 23*(1), 1–12 http://dx.doi.org/10.1016/s0149-7189(99)00040-3

Yates, M., & Youniss, J. (1996). A developmental perspective on community service in adolescence. *Social Development, 5*(1), 85–111 http://dx.doi.org/10.1111/j.1467-9507.1996.tb00073.x

Zhang, X. (2011). *The effectiveness of Project K on participants' health and lifestyle behaviours* (Unpublished master's thesis) Auckland, New Zealand: University of Auckland.

Kelsey L. Deane, PhD is a Lecturer in the School of Counselling, Human Services and Social Work, Faculty of Education at the University of Auckland. Her major research interests include positive youth development, program evaluation research, and applied research methods.

Niki Harré, PhD is an Associate Professor in the Department of Psychology at the University of Auckland, New Zealand. Her recent research projects have focused on positive youth development, sustainable communities and schools, and political activism.