

Realistic Evaluation and the Process-Tracing Method: A Combined Approach to Scrutinizing Causal Mechanisms

Salah eddine Bouyousfi and Mohamed Sabar

*Groupe Institut Supérieur de Commerce et d'Administration
des Entreprises, Casablanca, Morocco*

Abstract: *This article proposes a methodological contribution for causal mechanism scrutiny by combining realistic evaluation (RE) and the process-tracing method (PTM). To overcome some limits of RE's operationalization, the combination RE-PTM reinforces RE by breaking down the causal mechanisms and testing them with PTM. This combination was tested as part of a project for the professional integration of young people in difficulty. We developed a theoretical framework involving the framing of RE and used Bayesian logic to test the RE hypotheses. Semi-structured interviews served as the main evidence source. The results of the study support the fact that the RE-PTM combination can address some RE operationalization challenges. The RE-PTM procedure shows that PTM can provide RE with guidance on how to collect and assess data that support the contribution made by an intervention in achieving the expected outcomes.*

Keywords: *causal mechanisms, process tracing method, professional integration, realistic evaluation*

Résumé : *Cet article propose une contribution méthodologique à l'examen approfondi de mécanismes causaux en combinant l'évaluation réaliste (ER) et la méthode de traçage des processus (MTP). Pour surmonter certaines des limites de l'opérationnalisation de l'ER, la combinaison de l'ER avec la MTP consolide l'ER en désagrégeant les mécanismes causaux et en les mettant à l'essai avec la MTP. Cette combinaison a été mise à l'essai dans le cadre d'un projet pour l'intégration professionnelle de jeunes en difficulté. Nous avons créé un cadre théorique permettant d'encadrer l'ER et d'utiliser la logique bayésienne pour mettre à l'essai les hypothèses d'ER. La majorité des données a été recueillie par l'intermédiaire d'entrevues semi-structurées. Les résultats de l'étude appuient l'hypothèse que la combinaison d'ER et de MSP peut éliminer certains des problèmes liés à l'opérationnalisation de l'ER. La procédure d'ER-MSP montre que la MSP peut offrir à l'ER l'orientation nécessaire pour recueillir et évaluer les données qui appuient la contribution d'une intervention dans l'atteinte des résultats prévus.*

Corresponding Author: Salah eddine Bouyousfi, Groupe Institut Supérieur de Commerce et d'Administration des Entreprises, Km 7, Route de Nouasseur, Casablanca, Morocco; s.bouyousfi@evalconsulting.ma

Mots-clés : *mécanismes causaux, méthode de suivi des processus, intégration professionnelle, évaluation réaliste*

Although widely deployed, experimental evaluations are criticized for their simplified design of causality in focusing on whether an intervention worked, rather than on how and why it worked (Fletcher et al., 2016). To address these limitations, theory-based evaluations (TBE) explain how and why outcomes are achieved (Mackenzie et al., 2010). TBEs are within the fifth explanatory generation of evaluations that integrate contextual and causal mechanisms (Brousselle & Buregeya, 2018).

One such approach is realistic evaluation (RE) (Pawson & Tilley, 1997), an application of generative causation that provides a means of unraveling the context and mechanisms of causality (Pawson, 2013); however, RE also has operational limitations and conceptual challenges (Porter, 2015). In fact, RE does not give clear guidance on how to collect and assess empirical data to measure the intervention contribution in achieving outcomes (Befani & Stedman-Bryce, 2016). The process-tracing method (PTM), a qualitative data-analysis method that examines causal mechanisms, is another generative method that deals with causal inferences (Befani & Mayne, 2014). This article proposes merging both RE and PTM to contribute to RE's operationalization. This combination was tested on an evaluation of a project for the professional integration of young people in difficulty.

The article begins with an overview of RE and PTM and then proceeds to a description of the RE-PTM combination that has been tested within one study. The results of the research have increased our confidence in applying the principles of the PTM in RE; therefore, merging RE and PTM addresses some of RE's challenges.

OVERVIEW OF REALISTIC EVALUATION AND THE PROCESS TRACING METHOD

Realistic evaluation

Associated with positivist epistemology, experimental evaluations invalidate a mechanism if they do not produce the desired effect. In contrast, critical realism (CR) attempts to explain the invalidated mechanism through contextual conditions and compensatory mechanisms (Devaux-Spatarakis, 2014). CR emphasizes the generating mechanisms by proposing articulations responsible for the observed events (Bhaskar, 1979).

Associated with CR, RE is a theory-driven approach that describes what mechanisms produce which outcome, and in what context. It's abbreviated as Context-Mechanism-Outcome (CMO) configurations (Pawson & Tilley, 1997). CMO is based on the principles of causal explanation of the pathways connecting

causes and their effects. RE defines mechanisms as the underlying processes or structures that operate in particular contexts to generate outcomes (Astbury & Leeuw, 2010). It seeks to understand not only whether an intervention contributes to outcomes but also how, for whom, and in what circumstances (Pawson & Tilley, 1997). Pawson (2013) advocated creating a matrix with three CMO columns and linking them in such a way as to have $C+M = O$. The RE iterative process relies on identifying hypotheses about CMO configurations as well as on testing and refining the intervention theory (Pawson & Tilley, 1997).

Although RE generates operational results that improve the intervention (Punton et al., 2016), it has methodological limitations and implementation constraints. In addition to the lack of pre-existing data for the development of CMO configurations, the approach is time and resource-consuming (Salter & Kothari, 2014). Moreover, its concepts are still subject to various interpretations and operational difficulties (Pawson & Manzano-Santaella, 2012). Dissociation of context elements from the mechanisms themselves is therefore the major challenge for RE's operationalization (Ridde et al., 2012). According to Rolfe (2019), REs have difficulty conceptualizing and identifying mechanisms, separating context from the mechanism, and identifying contextual factors. Furthermore, implementing the approach is heavily time- and resource-consuming as well.

The process-tracing method

PTM is a research method that examines causal mechanisms using within-case empirical analyses of how a causal mechanism link causes and outcomes (Beach & Pedersen, 2019). It uses Bayesian logic¹ to confirm or disprove hypotheses about the path between cause and effect (Rohlfing, 2014). The process begins by predicting what will be empirically collected and then assessing the certainty and uniqueness of the evidence. Using empirical tests, the method adopts a cumulative process of collecting evidence by eliminating all other explanations (Beach & Pedersen, 2019). In the PTM, the mechanism is defined as a causal process that is triggered by causes and links those causes with outcomes (Beach & Pedersen, 2019). PTM supposes that the mechanism leaves traces that inform of its existence and also recognizes that not all of these traces are equally informative (Befani & Stedman-Bryce, 2016). PTM is carried out in five steps (Beach & Pedersen, 2013). First, it explains intervention theory through a hypothesis, that is, that mechanisms contribute toward the achievement of results. Each mechanism is broken down into entities that carry out activities. Next, it attempts to predict the manifestations of the causal mechanisms that have been theorized. In the third step, the degree of contribution of the mechanisms in the production of the results is evaluated through the collection of empirical evidences. Then the probative value of the evidence's inferential elements is assessed through the determination of a reasonable degree of confidence as to whether each part of the causal mechanism exists. PTM often uses the four tests of Van Evera (cited in Befani & Stedman-Bryce, 2016), who classified the different types of prediction according to the uniqueness and certainty of the evidence. The weakest test is straw-in-the-wind, where

the evidence collected does not represent certainty or uniqueness to confirm or disprove the hypothesis. The hoop test is certain but not unique; its failure reduces confidence in the hypothesis, but its success does not imply inference either. It is generally used to exclude the alternative hypothesis. The smoking gun test is strongly unique but not certain, and its success confirms the hypothesis. However, once more, its failure does not necessarily negate it either. The fourth test, doubly decisive, is certain and unique; it allows the hypothesis to be strongly confirmed or disproved. Lastly, based on the evidence collected and the tests applied, a degree of confidence in each part of the mechanism is determined to accept or reject the mechanism as a whole.

Nonetheless, the tracing and identification of processes prove to be difficult (Hay, 2016). This concern is shared by Stern and colleagues (2012) who have argued that, for complex interventions, PTM provides interpretations rather than evidence. However, this criticism seems to be shared by all the methods that seek to trace the complex causal path (Wadson et al., 2020). The added value of PTM is to allow a clear distinction between “absence of evidence,” which has little inferential power, and “evidence of absence,” which can strongly challenge a hypothesis (Befani & Stedman-Bryce, 2016).

PROCESS TRACING METHOD APPLIED TO REALISTIC EVALUATION

Recent theory-based approaches do not provide evaluators with guidance on how to collect and assess data to support contribution claims (Befani & Stedman-Bryce, 2016). Some studies have tried to establish guidance to evaluators for data collection and measuring confidence in the findings. Therefore, Befani and Mayne (2014) combined Contribution Analysis (CA) to PTM. CA-PTM uses probability to assess the strength of contextual and other factors influencing the intervention, through observations and evidence, and with regard to their relative contribution to the theory of change. Likewise, Lemire et al. (2012) pursue the same objective to guide the application of theory-based evaluation. They propose a framework to translate into practice some of the key steps in CA. To verify the pathway between intervention and outcome, they propose the Relevant Explanation Finder to measure the importance of the influencing factors and explanations.

From this perspective, this article aims to combine PTM and RE to contribute to the operationalization of the latter by responding to some of these challenges. PTM details hypothetical causal links in entity form, whereas TBE—of which RE is a part—merely makes assumptions of the whole mechanism (Leeuw, 2012). RE considers the mechanism as a causal force but not necessarily as an entity (Pawson, 2003). Otherwise, PTM does not simply formulate an overview but focuses on each link between a cause and a given effect, and it tests all parts of the mechanism (Beach & Pedersen, 2019). Befani and Stedman-Bryce (2016) advocate combining RE with PTM to collect data and to measure how much the evidence increases or decreases confidence in how a particular activity or

mechanism contributes to achieving outcomes. However, ER-PTM not only uses PTM principles to verify the quality of evidence with qualitative criteria, as CA does, but is also used to scrutinize the mechanism in CMO configuration.

This study sought to apply the principles of PTM to RE. This effort could enrich the evaluation research literature and guide RE implementation. Indeed, RE is a logic and not a method (Pawson, 2013), so PTM can inform RE about the evidence to be collected and the criteria for assessing it.

For the purpose of this study, the RE-PTM steps consisted of developing the initial theory of intervention (CMO configuration), decomposing mechanisms according to the PTM, collecting and analyzing data using PTM principles, testing the hypothesis, and refining the theory.

The combined RE-PTM procedure

As a qualitative study, this research adopted the realistic approach as its theoretical framework. So a RE-PTM combination was described and tested on a case study: a project that sought to improve youth employability. The methodological contribution of this article consists of decomposing the mechanisms of the RE-PTM configuration in order to test them by mobilizing the PTM principles, which provides qualitative criteria that judge the quality of the evidence collected through certainty and uniqueness tests (Beach & Pedersen, 2019). The different phases of the RE-PTM were carried out as follows.

Step one

Based on exploratory interviews with those in charge of the project, and on related documentation, an initial intervention theory was developed. At this level, the first elements of context and presupposed mechanisms were identified. According to Manzano (2016), interviews should begin with individuals who are more familiar with the intervention, such as those responsible for implementation.

Step two

In the second step, assumptions were developed about how the intervention would work in CMO configurations. The hypotheses were formulated based on a literature review on professional integration, the concepts emerging during the research, the exploratory interviews with the project managers, and project documentation. To identify how causal mechanisms may operate to generate outcomes, as advocated by RE, the mechanisms were conceptualized according to PTM. Thus, for each part of a mechanism hypothesis, a plausible alternative explanation, as well as observable manifestations that support or eliminate that hypothesis, were elaborated upon (Beach & Pedersen, 2013). Subsequently, the development of tests was based on the prediction of evidence that would be unique and/or certain to confirm or disprove each part of the mechanism. However, the goal is not to reach perfect certainty about mechanisms but to formulate hypotheses about their existence and to look for evidence in an attempt to increase or decrease confidence in such hypotheses (Befani & Stedman-Bryce, 2016).

Step three

The third step concerned the data collection to test the hypotheses. An interview guide was developed and semi-structured interviews were conducted. Data collection was also based on project documentation. The semi-structured interviews and the focus group topic-guides were based on both exploratory interviews with participants who are responsible for implementation, and a review of the literature on professional integration, as well as the various concepts identified and emergent during the breakdown of the mechanisms.

A realist interview method was adopted to answer the fundamental question of the RE: How, why, for whom, and under what circumstances did the intervention work? The realist interview approach aims to consider the theoretical structures under investigation to allow the research participants to construct meaning through the conceptual framework elaborated on by the interviewer (Pawson, 1997). However, realist interviewing requires that the interviewer critically examine the interviewee's responses by exploiting resources and information external to the interview context (Smith & Elger, 2014).

Van Evera's tests were mobilized to verify each part of the mechanism. The questions were formulated in such a way as to seek uniqueness and certainty of evidence.

Step four

As the last step, the hypotheses were tested, data analyzed, and theory refined. The evidence collected was verified before it was used in the data analysis. Based on PTM, the content and accuracy of the evidence were analyzed by using triangulation and project documentation. The statements of the different actors were grouped together to check their content and accuracy. Such analysis was not limited to the intervention but was also applied to what preceded its implementation. Indeed, Archer proposed enriching the $C+M = R$ formula by integrating historical data (cited in De Souza, 2015).

Data analysis was conducted in two phases. The first phase was the verification of the hypothetical mechanisms using the four Van Evera tests. In the second phase, the interview and focus group transcripts were analyzed using Nvivo software to identify some trends. Finally, the initial theory was updated and refined.

Setting up the case study

The RE-PTM was tested on the evaluation of the Najah project, which aims to enhance the employability of young people between the ages of 18 and 30 in difficult situations living in the city of Casablanca. The project was funded by the French Development Agency and implemented by the *Apprentis d'Auteuil*² (AA) association and *l'Heure Joyeuse*³ (HJ) association. This project seeks to train, accompany, and guide young people to find work. They are generally people from underprivileged backgrounds without any vocational training or qualifications. The research wasn't funded, and its purpose and design were fixed by HJ and the Management Research Department of Groupe Institut Supérieur de Commerce et d'Administration des Entreprises.

The choice of this project stemmed from its complexity, the substantial role of the context, and the importance of the reasoning of the actors in the project implementation. The relationship between cause and effect in this type of initiative is not always linear, and the mechanisms responsible for achieving results are not always obvious (Befani & Mayne, 2014). The RE-PTM does not aim to quantify the project results but to understand and explain why, how, and under what circumstances results are achieved. Its implementation proceeded in four steps, as depicted below. The RE-PTM application on the first theory component is presented at the end of the next paragraph as an illustration.

Step 1: Elaboration of intervention initial theory

The conceptualization phase of the theory was deductive insofar as it was based on existing theoretical knowledge to explain how the actions carried out within the Najah project had contributed to young people's professional integration. Development of the initial theory, as shown in Figure 1, was based on three exploratory interviews with those responsible for implementation (a director and two project managers from HJ), a literature review of professional integration, and the theoretical concepts identified during the research such as motivation, satisfaction,

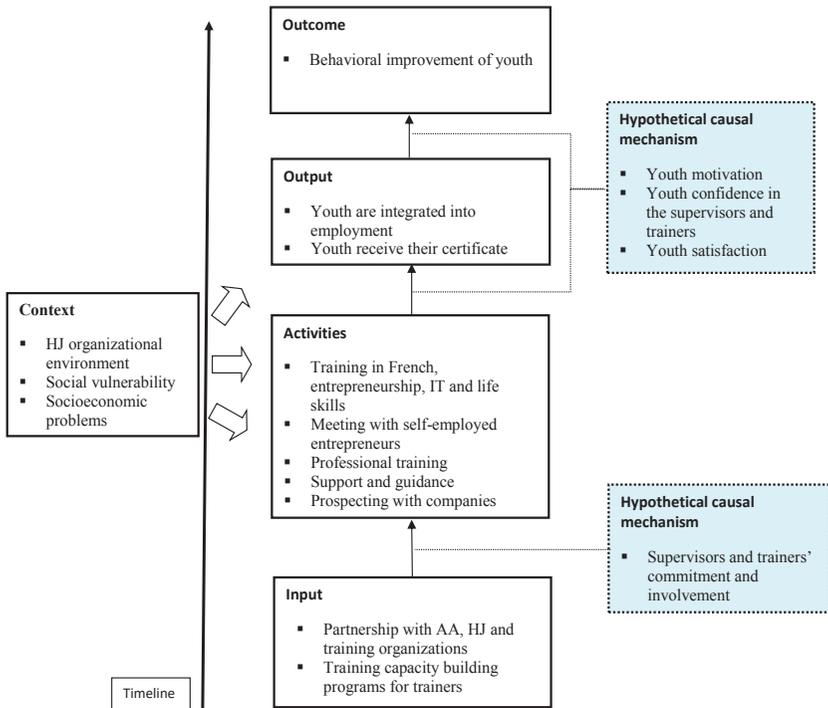


Figure 1. Initial theory intervention

engagement, and project documentation. Research materials consisted of project reports, a review of relevant documentation as young people's profile, and the rate of integration into the employment and evaluation tests to confirm beneficiaries' skill levels before and after the project. The initial theory model was informed by Mayne's Useful Program Theories (Mayne, 2015).

For example, the theoretical concept of professional integration and the factors that influence it have been defined; indeed, Fournier and Monette (2001) defined socio-professional integration as the successful transition from training to working life. They proposed that the transition from education to the labor market depends on the individual's adaptability to the working environment. Along similar lines, Fournier and Bourassa (2001) went beyond training and employability matching by advocating a training-employability-education-socialization logic. They proposed a proactive approach that integrates all stakeholders (individual, company, and state) in socio-professional inclusion. Furthermore, Allard and Ouellette (1995) grouped the factors most likely to influence socio-professional integration into three major dimensions: sociological (economic, political, cultural, geographic, and demographic factors), socio-psychological (family, school, work, friends, and media environments), and psycho-professional (factors related to the construction of personal and professional identity).

The initial theory intervention postulates that if supervisors and trainers are committed and involved, then the youth will be motivated and will develop confidence in the supervisors and trainers, which is arguably the key that will ensure their satisfaction and help improve their skills and behavior. These mechanisms, such as commitment, satisfaction, and motivation, are contingent upon particular contexts, such as social vulnerability, HJ organizational environment, socio-psychological factors, and socioeconomic environment.

Step 2: Establishment of hypotheses on how the intervention works and the mechanisms in CMO configurations are broken down according to the PTM

The mechanisms identified in the first step were broken down into entities engaged in activities that transmit causal forces. The practice of RE emphasizes mechanisms operating at the individual level rather than at the organizational level (Westhorp, 2018). This research focused primarily on the reasoning of actors; therefore, the analysis of mechanisms was more concerned with the micro-level (individuals). The project context analysis was based primarily on the four factors that shape context: people, interpersonal relationships, institutional rules, norms, and infrastructure (Pawson, 2013). Thus, after the development of the intervention theory, the CMO configurations were developed inductively and deductively (Funnell & Rogers, 2011) through exploratory interviews with the implementers, a systematic review of professional integration, and the different concepts emerging during the research such as trust, motivation, satisfaction, and commitment. The intervention was divided into three areas that correspond to three causal chains, namely, adhesion of young people to the project, youth training, and youth professional integration. Subsequently, each mechanism was broken down into parts that contain an entity and an action, respectively.

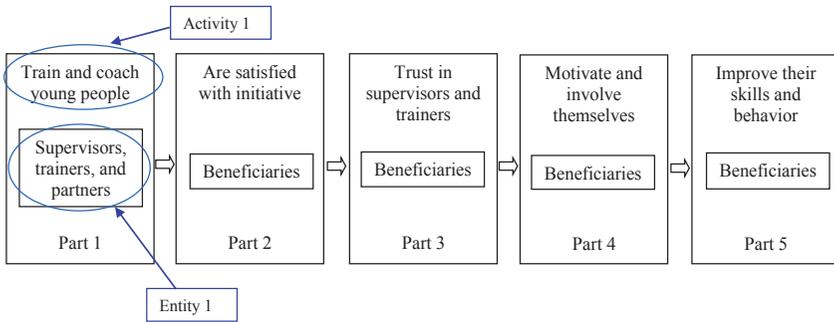


Figure 2. Hypothetical training causal mechanism

For example, a causal mechanism in process tracing based on training is conceptualized as being made up of several parts, composed of entities that engage in activities (Figure 2). Thus, all the steps between the integration of young people into the project (the hypothesized cause) and the improvement of their skills and behavior (the outcome of interest) were explained.

Furthermore, trust and motivation were identified as two parts of the mechanism that improved the skills and behavior of beneficiaries. Indeed, behavioral change depends on two factors: trust as a mediating mechanism and motivation as a moderating mechanism (Zhu & Akhtar, 2014). A mediating causal mechanism is identified as a component of a program that acts between two other components, while the moderating causal mechanism represents a link between program components that are conditioned by a third factor (Chen, 2005). Otherwise, trust is a key element in motivation within organizations. Morgan Hunt's implication-trust theory (cited in Heavey et al., 2011) states that trust is a prerequisite for involvement and motivation. Based on the above and the synthesis of the three exploratory interviews, we hypothesized that young people's motivation and involvement were due to their trust in supervisors and trainers. In other words, we were confident enough that this proposition was valid, and we expected that observable manifestations that support it (interviews with young people who attribute their motivation and involvement to the confidence they have in the project team) demarcated the most plausible explanations (i.e., that the motivation and involvement of young people are due to their desire to improve their well-being). All plausible options were considered, and observable manifestations that rule out the strongest alternative hypothesis were defined (little or no actions were taken by young people to improve their situation before or after they joined the Najah project). Therefore, tests were to be deployed through semi-structured interviews and documentation research that would confirm the uniqueness and/or the certainty of the evidence that supports the causal mechanism part. The test would establish that the motivation and involvement of young people are due to their trust in their supervisors and trainers and would exclude other alternatives that might generate them. If not, then the hypothetical causal mechanism would be rejected.

Table 1. CMO Configurations

Components of the theory	Context	Mechanisms	Outcomes
Youth buy-in to the project	<ul style="list-style-type: none"> • Precariousness • Social vulnerability • Socioeconomic problems • HJ organizational environment 	<ul style="list-style-type: none"> • Supervisors' and trainers' commitment and involvement 	<ul style="list-style-type: none"> • Integration of the young people in project
Youth training	<ul style="list-style-type: none"> • HJ organizational environment • Socio-psychological factors 	<ul style="list-style-type: none"> • Youth satisfaction • Youth confidence in the supervisors and trainers • Youth motivation 	<ul style="list-style-type: none"> • Obtaining a certificate • Improvement of skills and behavior of young people
Youth professional integration	<ul style="list-style-type: none"> • Socioeconomic environment 	<ul style="list-style-type: none"> • Business buy-in to the project 	<ul style="list-style-type: none"> • Youth employability improvement

In summary, the CMO configurations were developed according to RE and the mechanisms were cut out according to PTM. Practically, some theory-based predictions were developed in light of what is expected to be found empirically. At this level, the choice of tests to verify the hypotheses and the data to be collected was fixed. The summary of the CMO configurations is shown in [Table 1](#).

Step 3: Data collection

The third step concerned data collection to test hypotheses and to analyze the intervention context. Data collection from July to December 2019 was carried out through 17 semi-structured interviews and three focus groups and was completed by documentary research on the intervention. The interviews and focus groups involved all stakeholders (beneficiaries, trainers, supervisors, and representatives of partner companies). The selection of interviewees was based on purposive sampling, and the questions were adapted to the respondents. For example, the choice of beneficiaries was focused on young people who had completed their training and were still looking for work, those who were able to find a job, and those who had worked but had left their jobs. Interviews were completed after saturation and completeness of the data were reached ([Corbin & Strauss, 2008](#)).

The wording of the questions was based on the realist interview. The main idea was to expose the stakeholders to the elaborated CMO configurations and incorporate their reactions into the program theory. Therefore, questions were formulated to explore each part of the mechanism in the CMO configurations. For example: Can the engagement of trainers occur if there are other intervening mechanisms? Is there a single observation that explains the job retention levels amongst the youth? To explore the mechanisms, the questions included were the following: How do you think the expertise of the trainers influenced the youth's buy-in to the project? To explore the results: How do you think the project changed your behavior?

To reduce confirmation bias concerning interest in the data that support the hypothesis at the expense of those that do not (Fairfield & Charman, 2019), all evidence was ranked according to inferential power or probability (Beach & Pedersen, 2019).

Step 4: Data analysis, hypothesis testing, and theory refinement

The empirical material was evaluated. Thus, the content and accuracy of the interviews were analyzed using triangulation. Grouping the different stakeholders' statements allowed us to analyze the content before using it as evidence to improve our confidence in the absence or presence of causal mechanisms. In some interviews, more than one version emerged, so we tested both versions against reality. For example, we checked project reports and companies' documents to verify why some beneficiaries left their job just after a few working days. Some young people attributed this to poor work conditions, while supervisors and company representatives pointed toward adaptability problems of young people. An in-depth analysis was therefore conducted to check the inferential power of the two statements. More young people were sought who worked in the same company, and their reasoning and reactions were examined.

Data analysis was conducted in two phases. The first phase was the verification of the mechanisms using the four Van Evera tests. The hypotheses were accepted or rejected, and then the mechanisms and the context factors for each theory component were updated. In general, the hypotheses that have survived a hoop test are assessed with a smoking gun test. The passage of the last test strongly confirmed a hypothesis that needs to pass through the first one, which was often used to exclude alternative hypotheses (Beach & Pedersen, 2013). The second phase consisted of a content analysis of the semi-structured interview and focus group transcripts using Nvivo software.

The semi-structured interviews and focus groups were recorded, transcribed, and encoded. Based on the existing literature, the coding process was mainly deductive but was also flexible (Hsieh & Shannon, 2005). Subsequently, we developed 42 codes that were connected to seven themes, taking into account the CMO configurations (Table 2), that is, young people's expectations, the behavior of trainers and supervisors, the context, the supervision, the training, professional integration, and the beneficiaries' personality.

Table 2. Codes

Codes	Sources	References
Youth expectations	17	79
Well-being	9	13
Learn	9	13
Hiring	13	28
Financial independence	2	2
Set objectives	10	23
Trainers' and supervisors' behavior	17	74
Friendship	8	9
Benevolence	8	11
Good contact	8	12
Confidentiality	7	9
Active listening	6	7
Flexibility	3	4
Involvement and commitment	13	22
Context	21	129
Difficulty to reach workspace	4	5
Family and friends	18	42
The job market	14	19
Organizational	14	25
Social vulnerability	11	21
Reputation of Heure Joyeuse	10	17
Supervision	16	38
Tutoring	13	27
Guidance	3	5
Awareness	6	6
Training	16	55
Skills	16	33
French language	7	12
Life skills	9	10
Professional integration	19	38
Stakeholders	17	128
Beneficiaries	11	45
Supervisors	10	28
Representatives of corporates	7	14
Trainers	12	41
Personality of beneficiaries	21	171
Positive attitudes and behaviors	21	147
Adaptability	7	8
Mature	15	22
Self-confidence	11	13
Trust	13	25
Motivation	14	34
Satisfaction	7	8

(Continued)

Table 2. Continued

Codes	Sources	References
Feeling good	8	14
Willingness	11	23
Negative attitudes and behaviors	17	24
Dependence	4	4
Despair	7	7
Inactivity	6	7
Inward-looking	5	6

To illustrate how RE and PTM are integrated into one approach, we present below how RE-PTM was used in the first theory component.

For the first component, we hypothesized, according to a literature review of professional integration and exploratory interviews, that youth buy-in to the project was a mechanism that contributed to the achievement of project objectives. Accordingly, the mechanism was divided into three parts: HJ organizational support, supervisors' and trainers' commitment and involvement, and youth integration.

The second part of the mechanism hypothesized that the commitment and involvement of the supervisors were due to the HJ support. This hypothesis was tested through two tests. The first was a hoop test which consisted of checking whether the supervisors were involved and committed and whether the support of the HJ was necessary for their commitment and involvement. The second test, smoking gun, aimed to verify if this aspect was sufficient for their commitment and involvement. The test results showed that the hypothesis survived the first test and failed the second, confirming the existence of other causes responsible for supervisors' commitment and involvement. Passing the hoop test, the hypothesis wasn't excluded, but finding evidence does not enable inferences to be made. The evidence is necessary but not sufficient. In addition, the failure of the smoking gun test does not strongly undermine the hypothesis, which is neither confirmed nor undermined.

Indeed, the interviews we conducted with the supervisors showed that they were committed and involved. This conclusion is the result of the measurements made of their involvement and commitment according to the three elements of Thevenet (1995). First, we noted the great interest that the supervisors expressed in their work. Second, they affirmed that their work is valued. Finally, the supervisors confirmed that they benefit from a considerable degree of discretion in fulfilling their tasks. This satisfies the three conditions recommended by Thevenet.

In addition, the turnover rate is quite low among supervisors. This can be explained by their attachment to the HJ, their commitment, and their involvement. For the sake of triangulation, the beneficiaries also affirmed that the supervisors are involved and committed. This behavior was thus verified by the hoop test,

which also confirmed that the involvement of the supervisors is partly due to adequate working conditions in the HJ. Indeed, the organizational support for the Najah project creates favorable conditions to promote their involvement and commitment but remains insufficient to confirm our hypothesis.

For that, we took the second test, smoking gun, to see if other factors could explain the supervisors' conduct. However, this test failed because the organizational framework is not sufficient to explain the supervisors' commitment and involvement. Indeed, the supervisors stressed that their conduct is also driven by the desire to do good, to help young people in difficulty, and by their love for community-based work.

In summary, the hypothesis that attributes the supervisors' commitment and involvement to organizational support was not accepted. Indeed, the desire of supervisors to contribute to the well-being of the young people partly explains their conduct, which is favored by a positive working environment. Thus, the failure of the second test invalidated the hypothesis. However, the analysis highlighted the

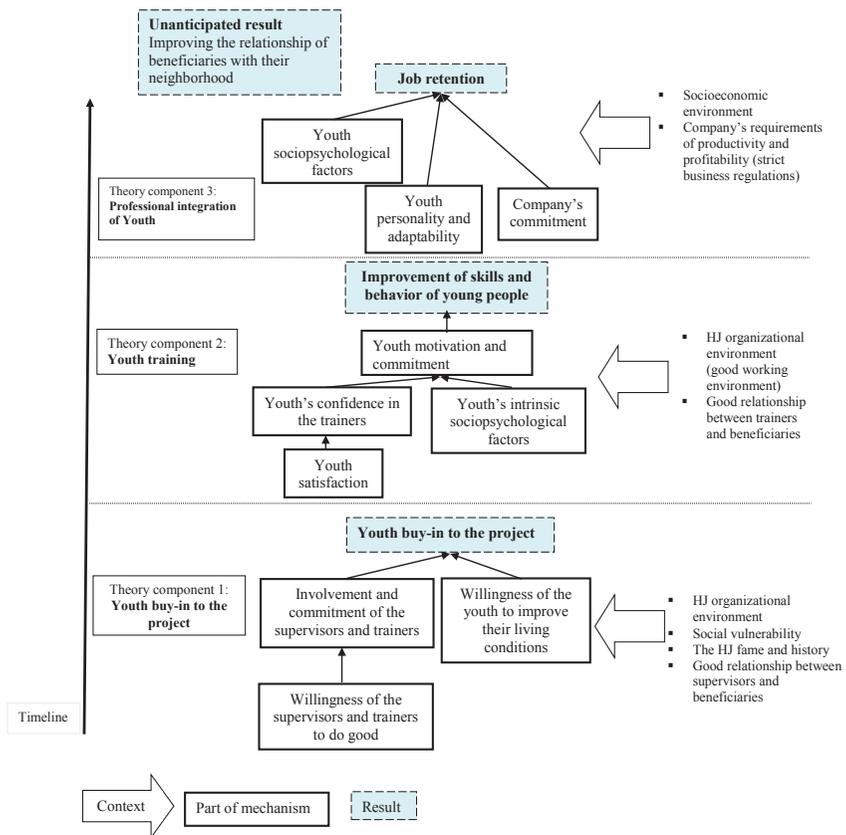


Figure 3. Refined theory intervention

crucial role of the HJ organizational environment as a contextual factor leading to youth buy-in to the project.

The findings of the study allowed for the update of intervention theory, as shown in [Figure 3](#). RE-PTM revealed that the involvement and commitment of the supervisors and trainers seem to explain the youths' buy-in to the project. Those mechanisms interact and meld with other mechanisms as young people's willingness to improve their well-being. Those underlying mechanisms are sensitive to the HJ organizational environment, which triggered youth motivation and confidence in the trainers' mechanism as they contribute to the improvement of their skills and behavior. Otherwise, socioeconomic constraints were essential for the company's commitment and young people's personality and adaptability to work as facilitating mechanisms that link prospecting activities to youth professional integration. It corresponds to the COM-B system that states that behavior change involves three essential conditions: capability, opportunity, and motivation ([Michie et al., 2011](#)). A behavior change is the result of the presence and adequacy of capabilities, opportunities, and motivation. Therefore, the enabling environment comprises all those events and conditions that are needed to make the intervention theory work, which [Mayne \(2018\)](#) named COM-B Based Theory of Change.

The research findings, as shown in [Table 3](#), were formulated as an answer to the questions how, why, for whom, and under what circumstances did the intervention work? Although the rationale for this evaluation didn't aim to quantify the project results, it revealed that the Najah project had generally attained both its objectives as well as an unanticipated result: the impact of the development of young people's skills on their neighborhood. It seemed that the relationships between beneficiaries and their family and friends had undoubtedly improved.

CONCLUSION

This article aspires to contribute to the operationalization of RE and the application of PTM principles and tests in evaluation. It aims to provide straightforward guidance on how to conceptualize and test the mechanism in CMO configurations to increase the robustness of the evaluation findings.

This research was structured according to the Pawson and Tilley model. The starting point was a theory derived from a hypothesis formulated through the literature and data from stakeholders. It was then tested and refined through a case study. The proposed framework was demonstrated in the evaluation of the project Najah, which aims to improve young employability. The methodological contribution of this article is to scrutinize the mechanisms of CMO configurations by applying PTM, since RE does not provide clear techniques to collect and assess the strength of data, nor to conceptualize and test the CMO configurations. Therefore, the combination of RE-PTM allowed us to reinforce RE operationalization. To our knowledge, this is the first study using PTM principles to examine the mechanisms in CMO configurations. Other research has tested the application of

Table 3. How, why, for whom, and under what circumstances the intervention worked

What works?	<p>The good contact between beneficiaries and supervisors and trainers strengthens the young people's confidence in the project team, which seems to be necessary for the learning process.</p> <p>The organizational environment, the reputation of the HJ, feedback from former beneficiaries, and the nature of the training (mainly life skills) encourage young people to join the project.</p> <p>Familiarization is necessary to keep the beneficiaries in the project. A friendly relationship between the young person and the trainer promotes learning.</p>
For whom?	<p>It seems that motivated, adaptable, and self-confident beneficiaries are more likely to obtain and maintain a job.</p>
How?	<p>Trainers and supervisors build a relationship of trust that encourages young people to get involved and improve their skills and behavior.</p> <p>The young person's personality, adaptability, and maturity are determining factors in the professional integration process.</p> <p>In addition to the values shared within the HJ, the supervisors and trainers are animated by a love for community-based work.</p> <p>Young people sometimes refuse a job because they cannot afford to pay the costs of mobility.</p> <p>The supervisors keep in close contact with parents to increase their involvement and to follow the young people's progress.</p> <p>The willingness of the supervisors and trainers to do good and their attachment to the HJ association reinforce their commitment and involvement.</p> <p>The adhesion of the young people to the project is due both to their willingness to improve their well-being and to the commitment and involvement of the supervisors and trainers.</p> <p>The quality of training and supervision and the motivation of young people are responsible for improving their skills and behavior.</p> <p>Maintaining employment is due mainly to the personality of young people and to working conditions.</p> <p>Maintaining employment depends on the commitment of companies as well as the personality, adaptability, and socio-psychological factors of young people.</p> <p>The social vulnerability and dependence of young people on their families hamper their professional integration and ability to stay employed.</p> <p>The partnership with the companies is not stable and depends strongly on the requirements of productivity and profitability.</p> <p>The HJ positive working environment contributes to the involvement and commitment of the supervisors and trainers and, therefore, promotes the training of young people.</p> <p>The reputation and history of HJ contribute to the adhesion of the young people to the project.</p>
Under what circumstances?	<p>The social vulnerability and dependence of young people on their families hamper their professional integration and ability to stay employed.</p> <p>The partnership with the companies is not stable and depends strongly on the requirements of productivity and profitability.</p> <p>The HJ positive working environment contributes to the involvement and commitment of the supervisors and trainers and, therefore, promotes the training of young people.</p> <p>The reputation and history of HJ contribute to the adhesion of the young people to the project.</p>

PTM in the field of evaluation. Indeed, contribution analysis was combined with PTM to explore how PTM could be used to strengthen a contribution analysis (Befani & Mayne, 2014). Likewise, the contribution-tracing approach involved the formulation and testing of competing hypotheses that could explain observed outcomes. It was inspired by the principles of process tracing and Bayesian updating (Befani & Stedman-Bryce, 2016). Finally, a realist contribution analysis was used as the main generative causation method, supported by contribution tracing with Bayesian updating (UK Department for Business, Energy & Industrial Strategy, 2017).

We conclude that it is possible to combine the two approaches without claiming that they address all the RE challenges. The study's aim is not to produce generalizable results but to demonstrate how the two approaches can be combined in a case study. It constitutes a test case for a proof-of-concept justification. Nevertheless, the theoretical conclusions can be nested within a broader research framework, as the results of RE can generate transferable lessons for similar projects (Wand et al., 2011). Indeed, RE-PTM revealed different contexts that triggered various mechanisms leading to outcomes of the Najah project. Therefore, we propose its application in social intervention where the reasoning of beneficiaries is crucial in achieving outcomes. For similar initiatives, managers should determine the specific needs of their partner companies during project conceptualization to define the appropriate training program. Also, it is considered more effective to select the beneficiaries who have more chances to be integrated: young people who have a more adaptive capacity; on the other hand, young people who have difficulty adapting must undergo prior support before their integration into similar intervention. In addition, more efforts should be made to gain involvement of families in the professional integration process. Furthermore, trainers should build a relationship based on trust with beneficiaries to promote their learning. Moreover, training should focus on the life skills sessions to promote both confidence and adaptive capacity within the young people. Finally, we can state that PTM provides RE with criteria to judge the strength of evidence, and therefore the proposed combination increases the internal validity and credibility of evaluation findings.

However, the research has some limitations. First is the bias of the face-to-face interviews with the different stakeholders. Second, families, whose influence is decisive in professional integration, were not interviewed due to the difficulty of reaching them. Generally, the young people involved in the study grew up in disruptive families who are living in poverty. Our efforts did not always result in getting an interview with them. Finally, it was time-consuming to rigorously assess the probative value of different pieces of evidence.

Otherwise, to make the assumptions deployed in the RE-PTM combination more robust and credible, we propose that the Van Evera tests should be combined with a mathematical tool, using the probability of evidence in Bayes's theorem, to quantify the probative value of the evidence as Bayesian confidence updating (Befani & Stedman-Bryce, 2016).

NOTES

- 1 Bayesian logic is a branch of logic applied to decision making and inferential statistics that deals with probability inference: using the knowledge of prior events to predict future events.
- 2 A French Catholic foundation recognized as a public utility that engages and acts with the most fragile young people and families.
- 3 The association l'Heure Joyeuse is a recognized Moroccan not-for-profit association of public utility, committed to the fight against social and professional exclusion

REFERENCES

- Allard, R., & Ouellette, J. G. (2002). Vers un modèle macroscopique des facteurs déterminants de l'insertion socioprofessionnelle des jeunes. *Carriérologie*, 8(3), 497-517.
- Astbury, B., & Leeuw, F. L. (2010). Unpacking black boxes: Mechanisms and theory building in evaluation. *American Journal of Evaluation*, 31(3), 363-381. <https://doi.org/10.1177/1098214010371972>
- Beach, D., & Pedersen, R. B. (2013). *Process tracing methods: Foundation and guidelines*. University of Michigan Press.
- Beach, D., & Pedersen, R. B. (2019). *Process-tracing methods: Foundations and guidelines* (2nd ed.). University of Michigan Press. https://www.press.umich.edu/10072208/process_tracing_methods%0A
- Befani, B., & Mayne, J. (2014). Process tracing and contribution analysis: A combined approach to generative causal inference for impact evaluation. *IDS Bulletin*, 45(6), 17-36. <https://doi.org/10.1111/1759-5436.12110>
- Befani, B., & Stedman-Bryce, G. (2016). Process tracing and Bayesian updating for impact evaluation. *Evaluation*, 23(1), 42-60. <https://doi.org/10.1177/1356389016654584>
- Bhaskar, R. (1979). *The possibility of naturalism: A philosophical critique of the contemporary human sciences*. Humanities Press.
- Brousselle, A., & Buregeya, J.-M. (2018). Theory-based evaluations: Framing the existence of a new theory in evaluation and the rise of the 5th generation. *Evaluation*, 24(2), 153-168. <https://doi.org/10.1177/1356389018765487>
- Chen, H.-T. (2005). *Practical program evaluation*. Sage. <https://doi.org/10.4135/9781412985444>
- Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Sage.
- De Souza, D. E. (2015). Critical realism and realist review: Analyzing complexity in educational restructuring and the limits of generalizing program theories across borders. *American Journal of Evaluation*, 37(2), 1-22. <https://doi.org/10.1177/1098214015605175>
- Devaux-Spatarakis, A. (2014). L'évaluation « basée sur la théorie » entre rigueur scientifique et contexte politique. *Politiques et Management Public*, 31(1), 75-91. <https://doi.org/10.3166/pmp.31.51-68>
- Fairfield, T., & Charman, A. (2019). A dialogue with the data: The Bayesian foundations of iterative research in qualitative social science. *Perspectives on Politics*, 17(1), 154-167. <https://doi.org/10.1017/s1537592718002177>

- Fletcher, A., Farah, J., Moore, G., Evans, R. E., Murphy, S., & Bonell, C. (2016). Realist complex intervention science: Applying realist principles across all phases of the Medical Research Council framework for developing and evaluating complex interventions. *Evaluation*, 22(3), 286–303. <https://doi.org/10.1177/1356389016652743>. Medline:27478401
- Fournier, G., & Bourassa, B. (2001). *Disparition ou transformation des formes de l'emploi ? Entretien du CRIEVAT-Laval avec José Rose*. Trajectoires professionnelles et marché du travail contemporain. Presses de l'Université Laval.
- Fournier, G., & Monette, M. (2000). *L'insertion socioprofessionnelle : un jeu de stratégie ou un jeu de hasard?* Presses de l'Université Laval.
- Funnell, S. C., & Rogers, P. J. (2011). *Purposeful program theory: Effective use of theories of change and logic models*. Jossey-Bass.
- Hay, C. (2016). Process tracing: A laudable aim or a high-tariff methodology? *New Political Economy*, 21(5), 500–504. <https://doi.org/10.1080/13563467.2016.1201806>
- Heavey, C., Halliday, S. V., Gilbert, D., & Murphy, E. (2011). Enhancing performance: Bringing trust, commitment and motivation together in organisations. *Journal of General Management*, 36(3), 1–18. <https://doi.org/10.1177/030630701103600301>
- Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288. <https://doi.org/10.1177/1049732305276687>. Medline:16204405
- Leeuw, F. L. (2012). Linking theory-based evaluation and contribution analysis: Three problems and a few solutions. *Evaluation*, 18(3), 348–363. <https://doi.org/10.1177/1356389012452051>
- Lemire, S. T., Nielsen, S. B., & Dybdal, L. (2012). Making contribution analysis work: A practical framework for handling influencing factors and alternative explanations. *Evaluation*, 18(3), 294–309. <https://doi.org/10.1177/1356389012450654>
- Mackenzie, M., O'Donnell, C., Halliday, E., Sridharan, S., & Platt, S. (2010). Do health improvement programmes fit with MRC guidance on evaluating complex interventions? *BMJ*, 340(c185). <https://doi.org/10.1136/bmj.c185>. Medline:20123834
- Manzano, A. (2016). The craft of interviewing in realist evaluation. *Evaluation*, 22(3), 342–360. <https://doi.org/10.1177/1356389016638615>
- Mayne, J. (2015). Useful theory of change models. *Canadian Journal of Program Evaluation*, 30(2), 119–142. <https://doi.org/10.3138/cjpe.230>
- Mayne, J. (2018). *The COM-B Theory of Change Model (V4)*. https://www.researchgate.net/publication/323868561_The_COMB_ToC_Model4
- Michie, S., West, R., & van Stralen, M. M. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, 6, article 42. <https://doi.org/10.1186/1748-5908-6-42>. Medline:21513547
- Pawson, R. (2003). Nothing as practical as a good theory. *Evaluation*, 9(4), 471–490.
- Pawson, R. (2013). *The science of evaluation: A realist manifesto*. Sage.
- Pawson, R., & Manzano-Santaella, A. (2012). A realist diagnostic workshop. *Evaluation*, 18(2), 176–191. <https://doi.org/10.1177/1356389012440912>
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. Sage.

- Porter, S. (2015). Realist evaluation: An immanent critique. *Nursing Philosophy, 16*(4), 239–251. <https://doi.org/10.1111/nup.12100>. Medline:26392234
- Punton, M., Vogel, I., & Lloyd, R. (2016). *Reflections from a realist evaluation in progress: Scaling ladders and stitching theory*. CDI Practice Paper 18. IDS.
- Ridde, V., Robert, E., Guichard, A., Blaise, P., & van Olmen, J. (2012). L'approche Realist à l'épreuve du réel de l'évaluation des programmes. *The Canadian Journal of Program Evaluation, 26*(3), 37–59.
- Rohlfing, I. (2014). Comparative hypothesis testing via process tracing. *Sociological Methods & Research, 43*(4), 606–642. <https://doi.org/10.1177/0049124113503142>
- Rolfe, S. (2019). Combining theories of change and realist evaluation in practice: Lessons from a research on evaluation study. *Evaluation, 25*(3), 294–316. <https://doi.org/10.1177/1356389019835229>
- Salter, K. L., & Kothari, A. (2014). Using realist evaluation to open the black box of knowledge translation: A state-of-the-art review. *Implementation Science, 9*(1), 115. <https://doi.org/10.1186/s13012-014-0115-y>. Medline:25190100
- Smith, C., & Elger, T. (2014). Critical realism and interviewing subjects. In P. K. Edwards, J. O'Mahoney, & S. Vincent (Eds.), *Studying organizations using critical realism: A practical guide* (1st ed., pp. 109–131). Oxford University Press.
- Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., Befani, B., & Thévenet, M. (2012). *Broadening the range of designs and methods for impact evaluations*. Report of a study commissioned by the Department for International Development. Working Paper 38. Editions d'organisation.
- UK Department for Business, Energy & Industrial Strategy. (2017). *Evaluation of the transitional arrangements: Phase 2*. <https://www.gov.uk/government/publications/evaluation-of-the-transitional-arrangements-phase-1>
- Wadeson, A., Monzani, B., & Aston, T. (2020, March 25). Process tracing as a practical evaluation method: Comparative learning from six evaluations. *Monitoring and Evaluation News*. https://mande.co.uk/2020/media-3/unpublished-paper/process-tracing-as-a-practical-evaluation-method-comparative-learning-from-six-evaluations/?utm_source=dlvr.it&utm_medium=twitter
- Wand, T., White, K., & Patching, J. (2011). Realistic evaluation of an emergency department-based mental health nurse practitioner outpatient service in Australia. *Nursing and Health Sciences, 13*(2), 199–206. <https://doi.org/10.1111/j.1442-2018.2011.00601.x>. Medline:21595814
- Westhorp, G. (2018). Understanding mechanisms in realist evaluation and research. In N. Emmel, J. Greenhalgh, A. Manzano, M. Monaghan, & S. Dalkin (Eds.), *Doing realist research* (pp. 41–58). Sage.
- Zhu, Y., & Akhtar, S. (2014). How transformational leadership influences follower helping behavior: The role of trust and prosocial motivation. *Journal of Organizational Behavior, 35*, 373–392. <https://doi.org/10.1002/job.1884>

AUTHOR INFORMATION

Salah eddine Bouyousfi is a researcher in Laboratoire de Recherche en Management du Groupe Institut Supérieur de Commerce et d'Administration des Entreprises

(ISCAE)- Casablanca. Bouyoufsi focuses on enhancing evaluation use with particular emphasis on realistic evaluation.

Mohamed Sabar is a professor in the Laboratoire de Recherche en Management du Groupe Institut Supérieur de Commerce et d'Administration des Entreprises (ISCAE)- Casablanca. His primary research interest focuses on supply chain management and logistics.