# Theoretical Corpus for an Empirical Subject

*The Identification of Problems and Gaps in Pygmalion Effect Research*

# Elméleti korpuszt egy empirikus alanynak

*Problémák és rések azonosítása a Pygmalion-hatás kutatásában*

**Author and affiliation:** Sejla ALMADI, Assistant Professor, Institute of World Economy, Corvinus University of Budapest

**Abstract:** Theoretical corpus or body refers to the theory gathered on an issue, topic or subject. Despite having generalisable findings on the theoretical corpus and the scientific work with it daily in all disciplines, there is a lack of guidance on how to build it up. Therefore, the research objective was to provide a methodological perspective on its construction through an empirical subject with the aim of identifying the problems and gaps within the corpus. As case study, the Pygmalion effect at work was chosen, since it was a complex phenomenon with a long, extensive and intensive research history. Five steps were implemented to follow: text database, literature review technique(s), corpus overview, theoretical corpus, research problems and gaps. For the chosen subject, the methodological review technique was applied, that helped to identify six research problems and gaps (i.e. availability, trust, harm, research objects, settings, naturalness) on the angles of phenomenon limitations and ethical questions. As a response to these, it presented the author's own arguments and claims that led the trajectory of the paper. Parallel to the case study, the research in experimental psychology was also introduced for better understanding. Finally, the paper gave methodological suggestions on how to construct a theoretical corpus in general.

**Keywords:** methodological review, Pygmalion affect, research problem, research gap, theoretical corpus

**Szerző és hovatartozása:** Almadi Sejla, Egyetemi adjunktus, Világgazdasági Intézet, Budapesti Corvinus Egyetem

**Absztrakt:** Az elméleti korpusz vagy test egy adott ügyben, témában vagy tárgyban összegyűlt elméletanyagra utal. Bár az elméleti korpuszról és a hozzá kapcsolódó, napi szintű, valamennyi tudományterületen zajló munkáról általánosítható megállapítások tehetők, mégis hiányzik az útmutatás a szakszerű felépítéséhez. Ekképpen a kutatás célja az volt, hogy egyfajta módszertani nézőpontot nyújtson az elméleti korpusz felépítéséhez, egy empirikus kutatási alanyon keresztül, és azért, hogy azonosítsa a korpuszban megtalálható kutatási problémákat és réseket. Esettanulmányként a munkahelyi Pygmalion-hatásra esett a választás, amely egy komplex jelenség, hosszúra nyúló, extenzív és intenzív kutatási előzményekkel. A cél érdekében öt módszertani lépést vezettem be: szövegadatbázis elkészítése, szakirodalmi áttekintő technika kiválasztása és alkalmazása, korpusz áttekintése, elméleti korpusz kidolgozása, kutatási problémák és rések azonosítása. A választott alanyhoz módszertani áttekintő technikát alkalmaztam, amely elősegítette a hat kutatási probléma és rés (úgymint elérhetőség, bizalom, sérelem, kutatási tárgy, helyzet, természetesség) azonosítását a jelenség

korlátai és etikai kérdései mentén. Ezekre válaszul vagy megoldási javaslatként jelentek meg a szerző saját érveit és állításai, amelyek egyben a cikk ívét is vezették. Az esettanulmánnyal párhuzamosan a kísérleti pszichológiában végzett kutatások alapjai is bemutatásra kerültek az alany jobb megértése érdekében. Végül pedig, a cikk módszertani javaslatokat fogalmazott meg az elméleti korpusz általános felépítésére vonatkozóan.

**Kulcsszavak:** elméleti korpusz, kutatási probléma, kutatási rés, módszertani áttekintő, Pygmalion-hatás

## Introduction

An empirical subject or empirically researched subject refers to when a phenomenon[1] and its processes (e.g. memory, perception, personality, emotion, intelligence, etc.) are measured, and the behavioural manifestations of inferred processes are observed through e.g. tests, questionnaires, interviews, focus groups, observations, etc. (Christensen et al., 2015). And the theoretical corpus or body refers to the theory gathered on an issue, topic or subject. That theory is built up by both theoretical and above outlined empirical researches, their results, findings and conclusions. That is due to the complexity of all theories (Christensen et al., 2015), and the constant interaction between theory and empirics which has two directions: passing from empirics to theory and from theory to empirics. The former one is "the logic or context of discovery" or induction, the latter one is "the logic or context of justification" or deduction. Therefore, both generating and testing a theory are valuable processes in science, where the research outcomes add to the existing theory and guide the future research directions, and where the theoretical predictions are either confirmed or disconfirmed by the empirics that provide evidence on the usefulness or inaccuracy of the theory.

Building up accurately the theoretical corpus of a subject has the goals from overviewing, interpreting and apprehending to arguing and criticising it, and most importantly identifying the problems and gaps in its research. This identification can establish a solid ground for research claims, and lead to both theoretical and empirical papers. The construction of the body, and therefore this identification, should happen properly by the means of literature review techniques. The University of Southern California (2014) defined five of them: (1) "integrative: critiques and synthesises representative literature on a topic in an integrative way to generate new frameworks and perspectives on a topic; (2) historical: examines the evolution of research on a particular topic over a period of time to place it in an historical context; (3) theoretical: examines the body of theory that has accumulated in regard to an issue, concept, theory or phenomenon - these are often used to establish a lack of appropriate theories or reveal that current theories are inadequate for explaining new or emerging research problems; (4) methodological: focuses on research approaches, strategies, data collection techniques or analysis procedures, rather than the research findings - these are often used to provide a framework for understanding a method or methodology and to enable researchers to draw on a wide body of methodological knowledge; (5) systematic: uses a comprehensive pre-planned strategy for locating, critically appraising, analysing and synthesising existing research that is

[1] Phenomenon refers to any publicly observable behaviour, such as actions, appearances, verbal statements, responses to questionnaires and physiological recordings (Christensen et al., 2015).

pertinent to a clearly formulated research question to allow conclusions to be reached about what is known." These techniques may be implemented individually or in combination and should be applied in line with the attributes of the subject and the research goals.

The density of a theoretical corpus depends on the history (e.g. time interval, breaks, etc.), the extensity (e.g. number of involved disciplines or geographical areas) and intensity (e.g. type and number of data collection and analysis methods) of the investigation. Therefore, a fresher and narrower topic, such as the e-cigarettes' impact on human health or the trading with crypto currencies, has accumulated relatively less knowledge in the past few decade(s) than a more conventional and broader topic, such as the changes in the art of politics or the examination of the atom. Nevertheless, the denser a body is, the more difficult is to identify problems and gaps in its research.

Despite having generalisable findings on the theoretical corpus and the scientific work with it daily in all disciplines, there is a lack of guidance on how to build it up. Therefore, the research objective is to provide a methodological perspective on its construction through an empirical subject with the aim of identifying the problems and gaps within the corpus.

As case study, the Pygmalion effect (PE) will be introduced, that occurs "when a person's expectations for another person's performance (unwittingly) induces such a behaviour from the expecter, that it impacts the expectee's achievement to meet the original expectation" (Almadi, 2019, p. 14). It can appear anytime and anywhere, herein its "at work" dimension is taken under scrutiny that is between leader and follower. The PE at work is a complex phenomenon, a self-fulfilling prophecy (SFP) and a leadership style (LS) at the same time, with a long,[2] extensive and intensive[3] research history. Therefore, a dense body is expected with a strong interaction between theory and empirics. Moreover, since its theory has been widely criticised along the empirical attributes of its research (e.g. deceptions, focus, settings, etc.), a methodological review technique is going to be implemented to construct its theoretical corpus and to identify the research problems and gaps within, through the following five steps (Figure 1):

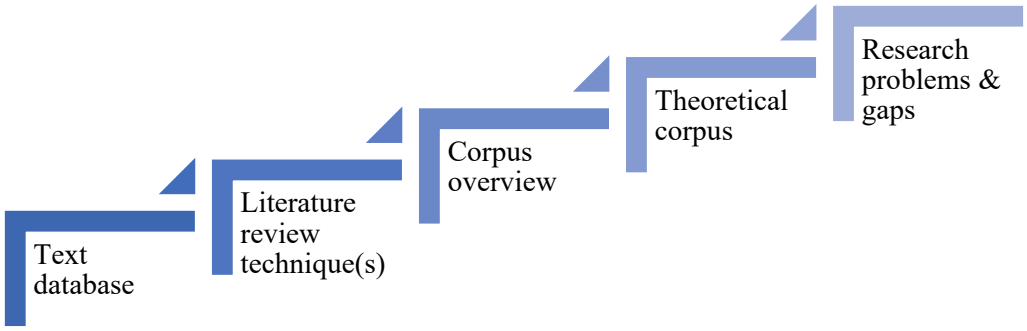
(1) Creating text database
(2) Choosing and implementing literature review technique(s)
(3) Overviewing the theoretical corpus of the subject
(4) Constructing the detailed theoretical corpus
(5) Identifying the research problems and gaps within

Figure 1

Five Steps to Construct a Theoretical Corpus (Own work, 2020)

[2]The story of PE originates in Ovid's Metamorphoses, first published in 8 AD. In the scientific world, it was first associated to Pfungst's attempt on Clever Hans in 1911 (Poornima & Chakraborty, 2010; Rosenthal, 2002). And it was first successfully experimented on humans by Rosenthal and Jacobson in 1968 as „Pygmalion in the Classroom".

[3] PE has been researched in several disciplines, mainly on the border of psychology, management and educational research, that is organisational behaviour. The latest text database on its research was composed of 102 research materials (Almadi, 2019).




## 1-2. Text Database and Literature Review Technique(s)

The two steps of creating text database and choosing and implementing literature review technique(s) are presented under one chapter. Since the entire text database of Pygmalion effect has been recently created, providing an explanation to its methodology (see Almadi, 2019), herein the empirical researches are focused on, and the text database is going to directly serve the purposes of the methodological review. The text database is the collection, categorisation and classification of sources on the research subject. Herein, the matrix character of the database is going to be the primary tool of the review that needs to outline the methodological research traditions of the subject. These might contain naming the relevant researches (authors, years), settings or contexts and their use of deception, sampling techniques and sizes, variable names and roles, data collection and analysis methods, hypotheses and their confirmations or results. For Pygmalion effect, the matrix on its research traditions can be found in Annex 1.

## 3. Corpus Overview

That step allows to have a first and overall view on the theoretical corpus, its main attributes. It establishes the use of the chosen literature review technique, hence its structure and content already reflect the goal and/or the main thread of the body. Herein, the PE research has been overwhelmingly experimental, only one non-experimental field study has been recorded (Whiteley et al., 2012)[4]. Therefore, a little knowledge on these needs to be gained before the overview specific to the chosen subject. The experimental approach is quantitative and designed to discover the effects of presumed causes (Christensen et al., 2015), hence they aim to eliminate potential bias and control conditions (or extraneous variables). Psychological experiments particularly, such as the ones on PE, refer to the "objective observation of phenomena which are made to occur in a strictly controlled situation in which one or more factors are varied and the others are kept constant" (Zimney, 1961 in Christensen et al., 2015, p, 52).

For the cause and effect measurement, manipulation is applied to one or more independent variables (IVs), that are expected to produce changes in the dependent variable (DV); and control is implemented over the experimental situation and conditions, such as the setting,

[4] Whiteley et al.'s (2012) research is considered as a non-experimental field study because its research did not involve manipulation.

sampling, counterbalancing, participant and experimenter effects. Therefore, these can involve e.g. the use of deception or double-blind placebo, the randomisation of participants or the avoidance of expectancy errors (Christensen et al., 2015). Manipulation can be applied in all kinds of settings, field, laboratory, internet (Christensen et al., 2015). The first one is conducted in real-life, where the variables are actively manipulated by the researcher and the most extraneous variables as possible are carefully controlled in the situation. The second one is conducted in laboratory, where one or more IVs are accurately manipulated by the researcher and the influence of nearly all extraneous variables are controlled. The third one is conducted through the Internet, it is a relatively new and emerging type of experimental settings. Experiments divide the participants into experimental or treatment and control groups. The manipulation is applied on the IV in the experimental group, where it is aimed to produce the effect, while the other group receives the active level of IV, so the amount that it would receive outside the research. Therefore, the control group serves both as a source of comparison and as a control for rival hypotheses.

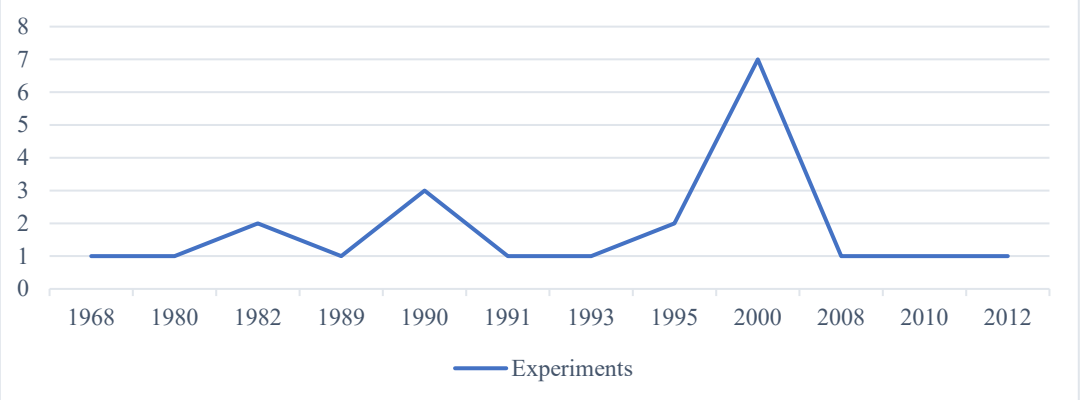
With regard to these, the overview on PE's corpus is presented along three themes with more characteristics: (1) contexts (timing, manipulation, deceptions, settings, researchers), (2) designs and methods (sampling sizes and techniques, constructs, variables, data collection and analysis methods, effect sizes, time intervals), (3) results (hypotheses, confirmation); the experiments lacking at least four of these characteristics were excluded from the research traditions matrix, hence from the overview as well.

### *3.1 Contexts of PE Experiments*

PE experiments' timing ranged between 1968 (Rosenthal & Jacobson) and 2012 (Whiteley et al.) (Figure 2). There was a little intensity in 1982 (Eden & Ravid; Eden & Shani), in 1990 (Eden a-c; Jenner; Learman et al.) and in 1995 (Dvir et al.; Eden & Zuk). The experiments peaked in 2000 by seven field experiments (Eden et al.). That peak actually segments the two phases in PE research: 'Phase 1' that tested the effect as a subconscious phenomenon, and 'Phase 2' that targeted to turn it into a conscious one. By the failure of Phase 2, no further phases were marked by the researchers of Pygmalion effect to date.

Figure 2

PE Experiments' Timing

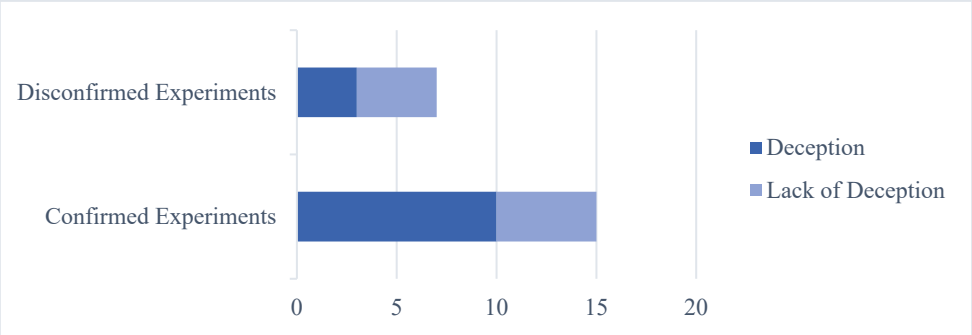


In PE research, manipulation was applied to influence the leaders' expectations for their followers' performance (as IV) in most cases, except in two cases: by Hezkiau-Ludwig and Eden (2010) who tested the reversed PE, and manipulated the followers' expectations for their leaders' performance; and by Whiteley et al. (2012) who used a non-experimental design to see if the leaders' expectations occurred naturally, without manipulation. The manipulations in PE research involved deception in 13 cases.

Deception is a deceit to control participant effect (Christensen et al., 2015). In PE research, it was used to make leaders believe in their followers' capabilities for better future performance - to believe the experimental group had higher, and the control group had lower capabilities. Figure 3 shows that the two-third of confirmed experiments used deceptions, while more than the half of disconfirmed (or only partially confirmed) experiments lacked deceptions. The experiments lacking deceptions were the 'Pygmalion-at-sea project' (Crawford et al., 1980), the 'Pygmalion Leadership Style' or PLS (Eden et al., 2000) and the 'Leaders' implicit followership theories and naturally occurring PE' or LIFTs (Whitely et al., 2012). In overall, the results on PLS were not significant.

Figure 3

Deceptions in PE Experiments




Among the three available experimental settings, only field experiments were implemented (real-life settings), even though the control on the conditions and settings varied. Figure 4 shows that from the seven different fields in PE research, rehabilitation and hospital management resulted solely in disconfirmed, while navy and retail resulted solely in confirmed experiments. The banking had an equal share in confirmed and disconfirmed experiments, while education and military were the most popular and successful settings for a confirmed PE experiment.

Figure 4

Settings in PE experiments

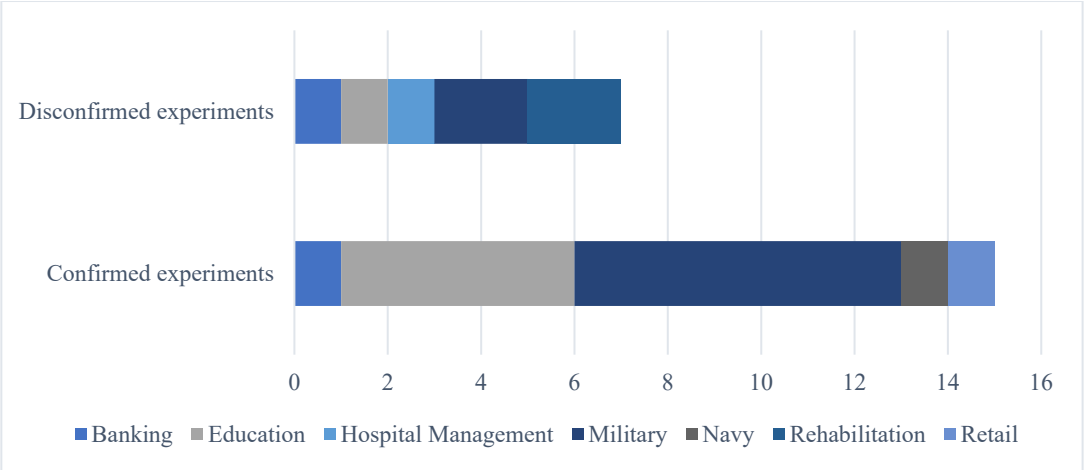
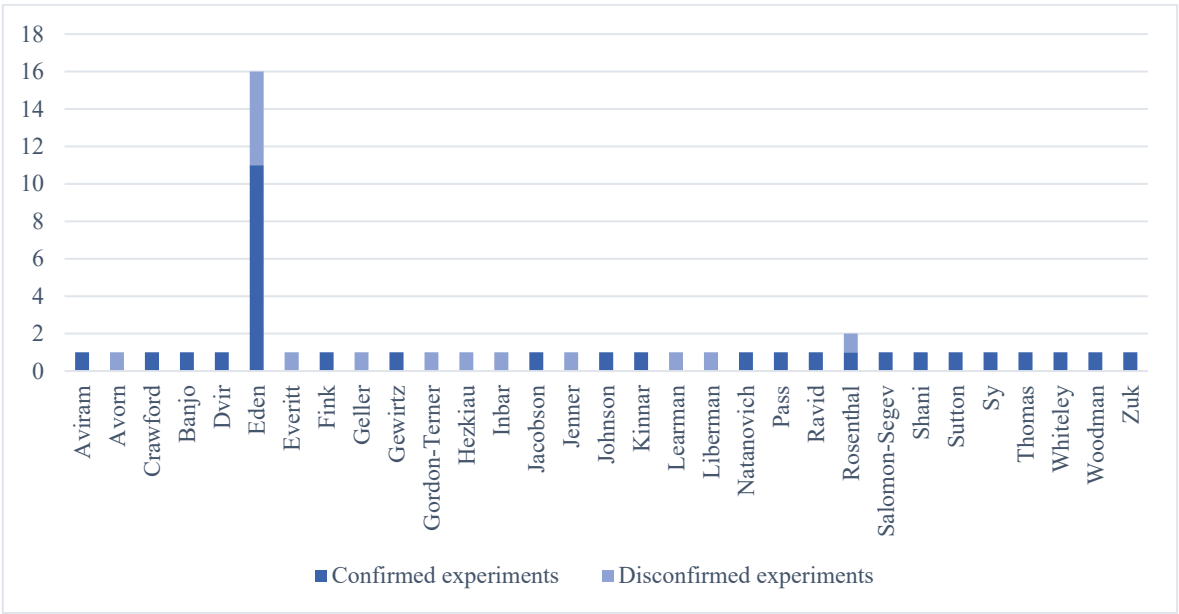



Figure 5 shows 31 researchers who conducted PE experiments. Only a little accumulation can be seen at Rosenthal with two experiments, while his work actually exceeded that number, particularly in the research of SFPs, thus he and Jacobson were the first to successfully experiment PE in 1968. He also reanalysed King's 1970 data in 1974. The largest accumulation appears at Eden with a total of 16 experiments, among which 11 was confirmed. His published works and contributions to the knowledge on the PE are immeasurable.

Figure 5

Researchers in PE Experiments




## *3.2 Designs and Methods of PE Experiments*

Among the overviewed empirical researches, the sample size of 4 was non-applicable. The rest had a sample ranging from 25 to 1000 participants with a mean at 319,89. The sampling method was either random (e.g. Eden & Shani, 1982) or purposeful that targeted the initially low

performer followers, to make the leaders believe that they had the potential for higher performance (e.g. Crawford et al., 1989).

The measured constructs had varying number of items. A part of them relied on only already existing and validated scales (e.g. Eden & Aviram, 1993; Eden & Kinnar, 1991; Eden & Zuk, 1995), a part of them used already existing scales with own-built ones, and another part of them did not add reference to the named constructs, hence their scales are assumed to be constructed completely. The Cronbach's alpha for internal reliability of scales was mostly measured, and confirmatory factor analysis for internal validity of constructs was also mostly applied. While the results of these were revealed, the set of items (on which they were carried out) were excluded from the annexes in each case. However, a few example questions were included in certain studies (e.g. Dvir et al., 1995; Eden et al., 2000; Eden & Kinnar, 1991; Hezkiau-Ludwig & Eden, 2010; Natanovich & Eden, 2008).

35 different variables could be identified in the experiments, and separated into two groups: (1) targeting the leader e.g. leadership behaviour, performance expectations, managerial self-efficacy (MSE), Pygmalion Attitude Index (PAI), Pygmalion Climate Supervisory Index (PCSI); (2) targeting the subordinate e.g. self-expectancy or self-expectations, motivation, performance, general (GSE) and specific self-efficacy (SSE), job satisfaction, perceptions of leaders' behaviour. The most used variables along their roles were: the manipulated IV: leaders' expectations for their followers' performance; the DV: followers' performance; the mediating variables: leadership behaviour and follower self-expectation; the moderator variables: gender, setting, and GSE. Hence, PE is not measured by a special formula, a single unique variable or a well-defined construct - that could definitely state that PE has been measured and not just a collection of variables that could refer to other effects at work -, but indirectly, through either the effect of the manipulated IV or the effect of the mediating and moderating variables on the DV, whose results has been incorporated into the Pygmalion hypothesis.

3 different data collection methods were found: tests, questionnaires and observations.[5] Similarly, McNatt (2000) delegated the PE performance measures into 3 categories: multiple-choice and essay exams testing specific materials covered during training; traditional performance appraisals filled out by supervisors or peers; actual task performance or observations of behaviour.

11 different analysis methods were found: descriptive statistics, correlation (3 types), ordinary least squares (OLS), confirmatory factor analysis (CFA), binomial effect size display (BESD), meta-analysis, variance (2 types), power analysis. ANOVA[6] has been used in 12, BESD[7] in 10, meta-analysis[8] in 7 cases. The effect sizes were calculated by omega or eta squared (both ANOVA), r (BESD) or d coefficients (meta-analysis). However, on effect sizes as well as on other methodological details such as standard deviations, not all researches provided reliable results. That was formerly highlighted by McNatt (2000) too who carried out

[5] Tests measure the participants' personality, aptitude, achievement, performance or specific variables; questionnaires self-reportedly measure the participants' demography, opinions and perceptions; observations are recorded by the researcher and contrast the participants' perceptions (Christensen et al., 2015).
[6] ANOVA is used to test whether three or more groups act different (Saunders et al., 2016).
[7] BESD "displays the change in success rate attributable to a certain treatment procedure" (Rosenthal & Rubin, 1982).
[8] Meta-analysis is used to describe the relationship among variables across multiple research studies (Christensen et al., 2015).

the first meta-analysis on PE in management (at work), and found a relatively large overall (population mean) effect coefficient (d=1.13). Even more, without the accurate information, the validity of results can be questioned: the above outlined parametric tests were applied to mostly ordinal data - instead of non-parametric ones.

The information given on the time interval of the experiments was also insufficient. According to the records, they varied from one-day workshops (Eden et al., 2000), through seven-month 24/7 navy performance (Crawford et al., 1980), to one school year (Rosenthal & Jacobson; 1968). Eden et al. (2000) concluded that the most intense the treatment of leaders and subordinates was, and the most frequent and intense the follow-ups were, the greater the significance was for the occurrence of PE.

### *3.3 Results of PE Experiments*

The results of PE experiments are based on their hypotheses and confirmations. The Pygmalion at work hypothesis was defined by Eden (1990b) as: *raising leaders' expectations for their subordinates' performance causes improved subordinate performance*. Based on that, herein, 11 different assumptions were identified in the former researches. However, to scientifically understand a phenomenon - such as PE - and its evolution, psychological research has four specific objectives or milestones to follow (Christensen et al., 2015): (1) description that identifies the phenomenon's characteristics (variables) and determines the extent or degree to which it exists; (2) explanation that identifies the antecedent conditions resulting in the phenomenon' occurrence; (3) prediction that enables to anticipate an event prior to its actual occurrence; (4) control that manipulates the conditions determining the phenomenon. Hence, the results, the hypotheses and confirmations of PE experiments should be evaluated under these milestones.

Description: the basic PE occurrence experiments involved the navy (Crawford et al., 1980) and military (Eden & Ravid, 1982; Eden & Shani, 1982) performance measurements. Somewhat more special approaches were given by the hypotheses of achieving reemployment (Eden & Aviram, 1993), increasing volunteering (Eden & Kinnar, 1991), and curing seasickness, which could be confirmed (Eden & Zuk, 1995), or by the other hypotheses on rehabilitation, which were partially or totally disconfirmed (Jenner, 1990; Learmann et al., 1990).

Explanation: the mechanism-related hypotheses further specified the "Model at Work" (Eden, 1990a), and confirmed that the contrast effect (Eden, 1990b) did not mean a threat to it, hence it could operate on entire groups without the need to differentiate followers within. While they disconfirmed the reversibility of the mechanism (Hezkiau-Ludwig & Eden, 2010) defining the PE as a solely top-down phenomenon passing from leader to follower.

Prediction: the trainability appeared among both confirmed and disconfirmed experiments. However, the meta-analysis on the seven field experiments, as a whole, disconfirmed the hypothesis, and concluded that the PE was a non-trainable phenomenon, which needed leaders' unawareness as ingredient (Eden et al., 2000). Hence, it had a strictly subconscious mechanism. Although it could seem that hypothesis was the most tested (Annex 1), in real, the gender generalisability was the most recurring one - with and without the researchers' intention. In the 1980s (Eden & Ravid, 1982; Eden & Shani, 1982; Sutton & Woodman, 1989), findings already

suggested that there were differences between genders in replicating PE, which were confirmed in 1995 (Dvir et al.) - that also referred to an appearing PE among men, and Golem effect[9] among women. Hence, the question of gender generalisability suggested experiments on the Golem effect (Davidson & Eden, 1997 and 2000; Eden & Davidson, 1997), which concluded on the margin of PE researches as well. Finally, in 2008, Natanovich and Eden closed the case by replicating the PE among both genders equally - however, only in the precised research context. Since the objective of prediction did not bring (complete) success, the PE research was stuck in the deception-based experiments, and the fourth, the upcoming objective of control could never really take place.

Instead of continuing that path, the latest non-experiment from Whiteley et al. (2012), that hypothesised the relationship between LIFTs and PE, led back the PE at work research to the stage of explanation by focusing on the naturally occurring possibility of the phenomenon. Prior to that, discussions already embarked on the existence of natural Pygmalions (Eden, 1990c), but not on their measurement or observation.

## 4. Theoretical Corpus

The theoretical corpus, as the main section, should receive the most attention and detail, along the continued use of the chosen literature review technique, both in its structure and content. Regardless the type of the subject, conceptual or empirical, it should focus on the findings and conclusions (objective elements), as well as on the arguments and counter-arguments (subjective elements) of former researches. However, the present author's own, individual but justified, scientific opinion and claims should also receive a great and explicit emphasis - i.e. I or We argue, claim, believe, suppose, etc., respectively -, to lead the trajectory of the actual paper. Herein the PE models and their theoretical perspectives are discussed along the research problems and gaps identified in the literature: (1) availability, (2) trust, (3) harm, (4) research objects, (5) settings, (6) naturalness.

These appear on two research angles: phenomenon limitations and ethical questions. Their focal point is the use of deception - prevalent tool in PE research - that shares the opinions since it is counter to the requirement of fully informing research participants about the nature of the research, and to the basic moral principle of trust that psychological researches should adhere to (Christensen et al., 2015). However, the use of deceptions can be justified to meet fidelity and scientific integrity. APA Ethical Principles and Code of Conduct (2003) limits the use of deception to studies where alternative procedures (e.g. role playing, inadequate substitutes, etc.) are not available, and have the potential to contribute to the understanding of a behaviour; while it is not allowed in studies that are expected to cause harm or severe emotional distress on the participants. Through the use of deception, all these ethical questions (i.e. availability, trust, harm) concern PE research. Moreover, its experiments were criticised for treating subordinates as passive participants or research objects. Although manipulations could be considered as an essential adage to experiments, their constant implementation limits the benefit of a phenomenon to the academics and does not allow to see what naturally happens, or how it could be implemented in everyday life, hence the results can be practically questionable. Not

[9] Golem effect is the reversed PE, hence, also a SFP, that occurs when the followers underachieve at work due to the leaders' low expectations (Oz & Eden, 1994).

surprising, the PE research was further criticised for implementing field experiments in atypical work settings, while drawing conclusions on the margin of PE at work or in management, in general - while the settings appeared as moderators of the effect.

Scientific attempts were aimed to respond to these ethical and practical issues in PE research, but just reached the limitations of the phenomenon's mechanism (i.e. non-trainability, fade-out, contrast effect, non-generalisability, irreversibility). These back and forth attempts, arguments and counter-arguments together construct the theoretical body of the subject.

### *(1) Availability*

The use of deception to induce PE was even criticised by researchers who applied it in their experiments (e.g. Eden, 1984) and who did not (McNatt, 2000; White & Locke, 2000). The principle of availability (of an alternative way to induce PE) led Eden et al. (2000) to carry out seven field experiments. It aimed to turn the PE mechanism (passing through the leaders) from subconscious to conscious, but since it failed, the use of deceptions seemed to be unavoidable. Upon that, Eden et al. suggested to combine the PLS with other LSs (i.e. transformational, charismatic, visionary) - Avolio et al. (2009) did not find these to be compatible with PLS but recommended to add new directions to the PE at work model, to confirm its theoretical concept and to support its induction. Eden et al. also suggested to implement more intense and frequent follow-ups upon the treatments (manipulations) to boost PE and to stop it from fading out, even in natural settings - Rosenthal's (2002) results, that explained the role of goal, learning opportunities and feedback at PE, confirmed that. McNatt and Judge (2004), who examined the generalisability of Galatea effect (GE)[10] to businesses, also argued the fade-out of SFPs. They replicated first a non-deceived, non-training GE on actual job performance of new and established professionals but the improved performance faded out shortly, similarly to PE experiments.

In contrast to Eden et al.'s suggestions, White and Locke (2000) believed that with different training techniques proven to build follower self-efficacy, the subconscious operation could be turned to conscious. Since their recommendation had already been prevented by PE researches on GSE and SSE (see Corpus Overview), I cannot see trainings of any kind as an available alternative way to induce PE, respectively. However, in line with their findings, I argue that GE (without deceptions) should be examined as an integral part to PE, which could involve more research on GSE and SSE. More research on PE versus other leadership concepts and on follow-ups or long-term interventions could also be considered.

### *(2) Trust*

Trust is the willingness to be vulnerable to the actions of someone, while expecting the other to perform a particular action that is important to the trustor, regardless the ability to monitor or control that other party (Mayer et al., 1995). At PE research and mechanism, it appears among:

[10] Galatea effect is also a SFP defined for the increased follower self-expectations for own performance. It can be induced directly by manipulations when deception is not required, or indirectly as a result and part of a deceived PE mechanism (Eden, 1990c).

the person conducting the manipulation and the other participants (mainly leaders), leaders and their followers. The leaders could change their expectations for their followers' performance, because they blindly believed and found credible the person who told them they should (e.g. psychologists, trainers). And the followers could change their expectations for their own performance, because they truly believed their leaders.

Although Kelman (1967) assumed that the constant use of deception would cause research participants to become distrustful; Sharpe et al. (1992) found that the participants accepted the arguments justifying the use of deception; Soliday and Stanton (1995) found that mild deception had no effect on the participants' attitudes toward researchers; Fisher and Fyrberg (1994) found that most of the student participants believed that the deception made studies scientifically valid and valuable, hence, its use was methodologically important, even when alternative ways, such as role playing or questionnaires, were also available. Despite these findings, in real life, the use of deceptions may not be regarded such valuable or important. Using them in organisations just to raise leaders' expectations for their followers' performance, 'may not be the ideal way to build trust' (Eden et al., 2000, p. 174). But without trainability, at the current state of our knowledge on PE, the organisations would have no other way to induce PE as a managerial tool - which was its main research objective (Eden, 1990a).

Thus trust and trustworthiness go in one hand: leaders' trustworthiness means a sufficient degree of trust placed into the leaders, who show supportive behaviour for their followers to perform (Karakowsky et al., 2012). As an evidence of trustworthiness, followers engaged more in risk-taking behaviours (Mayer et al., 1995). Both, the leaders' supportive and the followers' risk-taking behaviours were communication tools, where demographic attributes (e.g. age, gender, race) and expertise also affected the credibility of the source (Karakowsky et al., 2012). Herein the followers' perceptions for their leaders' ability, benevolence and integrity were improving in parallel with the leader-follower dyad, and were noticed by the leaders. Such positive leader-follower dyad, that could maintain high leader credibility, included accountability and even loyalty to the followers (Burke et al., 2007).

Hence, trust is a key component in inducing PE at all dimensions, including work. So I argue, respectively, that deceptions should be avoided in the future. Since the aim of research is the implementation of PE in real life, their use at organisations would raise ethical and practical concerns, moreover, it would be limitless since the fade-out requires frequent follow-ups. Similarly, White and Locke (2000) found deceptions at work complicated because: employing a person to deceive leaders might be destructive for the organisational trust; the required number of deceptions are unknown; the procedure would have serious organisational justice and legal implications.

### *(3) Harm*

The use of deceptions could cause harm on the research participants. That is detected by the debriefing sessions, post-experimental interviews or discussions where the research is revealed in details, including the use of deception (Christensen et al., 2015). These sessions are required by the ethical principles and codes of conduct. Incorrectly, PE experiments until the '80s did not report debriefing, which could have raised another ethical issue. On whether deceptions could cause harm, Christensen (1988) summarised the relevant literature, in which participants

did not perceive harm or having been misled upon experiments, in general. Moreover, Smith and Richardson (1983) found deceived participants to enjoy more the procedure and receive more educational benefit from it, even to feel the participation being more satisfactory than their non-deceived counterparts.

In PE experiments, deceptions impacted the leaders and the two groups of followers (experimental and control) differently. The leaders were told about having a better potential of followers in their team, hence, they were not exposed to harm. The debriefing revealed that leaders were unaware of treating the experimental group differently, and even remained unaware of it upon the session (Eden & Shani, 1982). The experimental group, that was supposed to have the better potential in its performance, benefitted from the deception, because it was treated better by the leaders, and indeed achieved a better performance at the end of the experiment. The control group, that was implicitly supposed to have the less potential in its performance, could be harmed, because it was not treated well or treated anyhow at all by the leaders, and indeed achieved a lower performance. Since the experimental group was of randomly selected participants or low performers and disadvantaged people, the control group was of possibly (initially) high performers, talented people. And no measurements were taken on their harm or ill impact. That was confirmed by Rosenthal (2002), when he looked back to the first successful PE experiment in 1968. He explained that the children who were expected to blossom intellectually (experimental group), became intellectually more alive and autonomous - at least in the eye of their teachers, while the children who were not expected to blossom (control group), were perceived as less well-adjusted, interesting and affectionate, especially when they achieved "unexpectedly" better.

Establishing a control group served two functions: it helped to make comparisons and to control rival hypotheses. However, the leaders' differentiated behaviour toward the two groups questioned if PE was simply a contrast effect. That possibility enforced to note for Eden (1990b, p. 394) that: "if creating the effect requires the presence of control subjects who do not gain, application may not be practical"; "if experimental subjects gain at the expense of their control counterparts, it would not make sense to create Pygmalion effects for the sake of organizational effectiveness". As a result, a no control group experiment was implemented that increased the leaders' expectations for all of their followers, and induced PE in entire groups. Hence, it eliminated the contrast effect from interpersonal but not from intergroup level. Eden argued that having a contrast at an intergroup level was acceptable and a normal attitude at work, in real life, so when the leaders have a higher expectation for their entire group of followers in contrast to other leaders' groups. In line with that, White and Locke (2000) suggested leaders to be trained for developing high expectations for all of their followers at work, by so-called Pygmalion training managers, who would focus "on particular behaviours in which they should engage and bring the effects of the expectations into conscious awareness" (p. 396), which would eliminate the need for deceptions at workplace.

Herein I have to respectfully disagree with the former statements. Concerning Eden (1990b), I argue that having or, more accurately, maintaining high expectations for an entire group of followers was not possible for the leaders, with regard to the fade-out of effect, and to the followers' on-going achievements that leaders perceive every day at work. Thus followers' self-image hardens along their career records, and while doing that, they also have to face a contrast to their peers and their performance within the group. Hence, the interpersonal contrast is more likely to appear at work, in real life, and could generate either solidary or rivalry,

cooperation or competition within groups - which possibility has not been discussed in relation to PE, to date yet. Concerning White and Locke (2000), I find their suggestion to employ Pygmalion trainers at work as a naïve consideration for more reasons: it would increase costs for the organisations; it would be limitless due to the fade-out; the seven field experiments (Eden et al., 2000) already concluded that PE could not be trained, to turn conscious for leaders.

### *(4) Research Objects*

Treating participants as active or passive elements in the experiment is another ethical issue. PE experiments treated leaders as active participants, whose expectations were directly manipulated (IV), and followers as passive participants or research objects, who reacted with their performance to the leaders' expectations (DV). After the failure of PLS, Eden et al. (2000) suggested to lessen the focus on leaders and to take the followers' contribution and behaviour into account. They shed light on the leader-focus that excluded to see GE, and suggested to examine leaders' beliefs in their followers' capabilities, followers' perceptions for their leaders' beliefs, and other performance measures. Therefore, the experiment had a "reversed" design that focused on the followers and their power in leader-follower dyads, particularly in PE (Hezkiau-Ludwig & Eden, 2010). Finally, it concluded that PE was a one-way, top-down, irreversible mechanism, passing from leader to followers. Although that result seemed to be a somewhat disappointing phenomenon limitation, Karakowsky et al. (2012) still claimed the followers to be the focus point and not only research objects.

Besides agreeing with the former statements, I have to remind the researchers respectfully to Eden's 1990c research that underlined the impregnation of organisational culture upon its individual members, the situation when the organisations' assumptions and expectations became their members' own assumptions and expectations, which explained how leaders could hold positive or negative expectations for their followers in natural cases. Moreover, a synergetic effect was found between leaders and their followers in organisations when creating PE (Poornima & Chakraborty, 2010). Hence, it was suggested to set individual goals in line with the supervisory goals, that were the extensions of organisational objectives, and to implement training and development programs, where employees' supportive perceptions, attitudes, behaviour could contribute to the organisational environment. And since PE is a top-down mechanism, I find it naturally fitting into the organisations' top down hierarchy. Hence, I argue that organisational culture and objectives should reflect the presence of PE through its members' perceptions, attitudes, behaviours, or by saying that a supportive work environment is a necessary condition for inducing PE - hence, these should also be measured.

### *(5) Settings*

PE at work experiments were implemented as field research with varying settings, that moderated the effect, particularly among genders. Besides, these settings were criticised for being atypical for work measures. But how could the most researched settings, education and military-navy, be considered as atypical? Without picking a sample country, it can be generally stated that employment rate in education is usually high, thus, both education and military are public spheres, so their employment rate should show less volatility than the private sector's.

Therefore, these settings could be claimed as typical for work. However, other settings, such as public or governmental organisations, and private organisations could also be claimed for PE research settings.

The question on gender generalisability was seemingly closed by Natanovich and Eden (2008), who concluded that the PE "may be produced without regard for sex" (p. 1381), it could operate equally among female and male participants. But the preventing critics raised attention to the moderating effect of settings on genders in the confirmed experiments: McNatt (2000) computed a larger effect size in military than in civilian settings, similarly to Gold and Kierein (2000), and White and Locke (2000) noted that only a few female Pygmalion leaders existed, such as in educational settings, where systematic gender differences were not found. Therefore, I must highlight that the Natanovich-Eden experiment took also place in education. The moderating effect of settings was explained by two reasons. (1) The way to achieve a leadership position, where women had to become superior to their male counterparts (Dvir et al., 1995). PE was found among female teachers, who did not have to compete as hard to become school teachers, while it was not found among female leaders, whose competition was difficult at work (White & Locke, 2000). (2) The genders' different LSs (Dvir et al., 1995), where females were more likely to lead in an egalitarian style, treat their followers equally and involve them in decision-making. Female leaders rather communicated to establish relationships, while male leaders' communication was task-oriented. Hence, in the view of that one successful experiment that could induce PE among female participants in education, I must argue that gender generalisability in PE mechanism is a paralogism, and the varying settings can still result in differences among male and female participants. Besides, I respectfully claim the more careful generalisation of results with regard to settings.

Thus, generalisability of older versus younger participants, and newcomers versus established work groups was also a difficult issue at PE (Eden, 1988; Livingston, 1969). White and Locke (2000) and McNatt and Judge (2004) found that PE had the least operation among established work groups in contrast to newcomers, while GE could operate well among both. Adding to their argument, becoming the part of a work group and leaving the newcomer status behind is an inevitable process. Because working along peers at an organisation, especially during the newcomer status, is a natural way of learning, handling tasks, providing-accepting support or forming teams, which structure has an important role in defeating challenges like market competition and technological innovation - which process leads to become the part of an established work group at the end. Not surprisingly, the model of work group socialisation was found to be consistent with PE and GE (Chen & Klimosky, 2003). PE described the interpersonal processes through which team expectations accounted for newcomer effectiveness, while GE enhanced the motivational processes through which newcomer expectations led to newcomer effectiveness. PE's and GE's role in such team processes remind for the gap in PE research for solidary or rivalry, cooperation or competition (see Harm).

*(6) Naturalness*

Unless research results can be implemented in real life, they stay for the benefit of academics. Hence, field research has to be conducted "in a manner that makes all variables operational in real-world terms" (Christensen et al., 2015, p. 58). A study of real-world terms is called as naturalistic inquiry, and shows how the phenomenon would unfold naturally, in a non-manipulative, non-controlling manner, being opened to whatever emerges without the predetermined constraints on findings. For that, Tunnell (1977) identified the naturalness and its three dimensions, the behaviour, setting and treatment (event), that would occur without the presence of research or researcher. PE field experiments were strongly controlled, and these dimensions appeared at only one-one of them.

Behaviour: 'Frogs to Princes' concept involved men lacking normal qualification for military service with either borderline IQ or mild psychological problems, who were supposed to be induced at a special training course but the organisation delegated them to the boot camp of the Israel Defence Forces (Eden, 1990c). The training instructors at the camp were not told about the delegated men's special case. However, they had access to their files. At the end of the training, the special case participants, who were treated as regular trainees by the instructors, responded as ordinary soldiers. Hence, the case could be regarded as an early natural experiment, revealing that natural Pygmalions existed.

Setting: industrialisation in a developing country was regarded as a SFP, where the people expected it to progress rapidly and successfully (Eden, 1990c). The speed of progression was promoted by the expectations that induced behaviours. Expectations had a key role in influencing productivity at both, individual and organisational levels, hence in industrialisation too. Although the research did not involve empirical measures, its concept can be regarded as the first and only approach in PE research, where the setting was supposed to be uncontrolled, natural. Thus, the term 'natural occurrence' in relation to PE was first coined here.

Treatment: the "naturally occurring Pygmalion effect" was hypothesised in a study that did not involve experimental treatment, manipulation (Whiteley et al., 2012). However, the summary noting that the results "supported a model of naturally occurring Pygmalion effects" (p. 822), is somewhat inaccurate, and rather, the conclusion noting that they provided evidence for a model delineating the relationships between positive LIFTs and naturally occurring PE, is more accurate - since the study did not measure PE, but leaders' followership theories, expectations for their followers' performance, liking, and leader-member exchange quality. The second from these variables is the usual IV in PE research, but the others are uncommon. Nevertheless, that again not careful generalisation on research results raises the question whether a single variable from PE mechanism would be enough to conclude on the margin of PE research? In more general, what is enough to conclude on its margin? When can it become sure that PE was replicated if, just as the Corpus Overview highlighted, PE is not measured directly by a special formula, a single unique variable or a well-defined construct? That uncertainty could mislead several PE researches, that was even admitted by a few experimenters (e.g. Dvir et al., 1995). Hence, beside the more careful generalisation of results, I claim PE research to incorporate elements that make the phenomenon unique and unmistakable among others, particularly among SFPs and LSs. Even more, the overview also highlighted that the mostly ordinal data were analysed by parametric tests in the traditional PE research, that could endanger the validity of results. Hence, I claim a methodological precision and improvement, even toward the direction of non-parametric tests.

Regarding the naturalness and its three dimensions, I argue that their broad, innovative and real-life-oriented approach would mean a great initiative toward a new perspective in (traditional) PE research that has been stuck in the prediction stage, hence I claim their implementation.

## 5. Research Problems and Gaps

In general, research problems can be identified through the problems the former studies referred to, and the stated research limitations; and research gaps can be identified through the stated future research directions uncovered to date, and the theoretical and methodological possibilities uncovered yet. Herein, the corpus was already structured along the six identified research problems and gaps, therefore, the responding arguments and claims are summarised below.

Availability: GE (without deceptions) should be examined as an integral part to PE, which could involve more research on GSE and SSE; more research on PE versus other leadership concepts and on follow-ups or long-term interventions could also be considered.

Trust: deceptions should be avoided in PE research with regard to the ethical and practical concerns at organisations, in real life.

Harm: interpersonal contrast is more likely to appear at work, in real life, and could generate either solidary or rivalry, cooperation or competition within groups, which should be examined in relation to PE. The suggestion to employ Pygmalion trainers at work seems to be a naïve consideration for more reasons.

Research objects: organisational culture and objectives should reflect the presence of PE through its members' perceptions, attitudes, behaviours, or by saying that a supportive work environment is a necessary condition for inducing PE - hence, these should also be measured.

Settings: gender generalisability in PE mechanism is a paralogism, and the varying settings can still result in differences among male and female participants, hence, results in PE research should be more carefully generalised.

Naturalness: PE research and measurement should incorporate elements (e.g. a special formula, a single unique variable or a well-defined construct, etc.) that make the phenomenon unique and unmistakable among others, particularly among SFPs and LSs. Thus, traditional PE researches should be methodologically precised and improved, even toward the direction of non-parametric tests. The broad, innovative and real-life-oriented approach of the naturalness and its three dimensions would mean a great initiative toward a new perspective in (traditional) PE research that has been stuck in the prediction stage.

## Conclusions

The present study aimed to provide a methodological perspective on how to construct a theoretical corpus, particularly for an empirical subject, such as the Pygmalion effect, that was a complex phenomenon with a long, extensive and intensive research history. To this end, it implemented five steps to follow: text database, literature review technique(s), corpus

overview, theoretical corpus, research problems and gaps. It applied the methodological review technique for the subject and drew up a matrix for its methodological research traditions. It overviewed 21 experiments and 1 non-experiment between 1968 and 2012 (modus=2000), from 31 researchers (modus=Eden). Their sample ranged between 25 and 1000 (mean=319,89), and collected by random or purposeful technique. The most frequently used variables were the leadership behaviour, types of expectations and efficacies. Data were collected by tests, questionnaires, observations, measured through 11 kinds of hypotheses (modus=trainability), and by 11 kinds of analysis methods (modus=ANOVA). 13 experiments used deception, among which 11 was confirmed.

The theoretical corpus had to focus on both, the objective and subjective elements of PE research, hence, on the findings and conclusions and the arguments and counter-arguments as well. The PE models and their theoretical perspectives were discussed along the research problems and gaps identified in the literature, such as the availability, trust, harm, research objects, settings, naturalness, which appeared on the angles of phenomenon limitations and ethical questions. As a response to these, the present author's own arguments and claims also received an emphasis and led the trajectory of the paper - which were summarised in the last section.

While the methodological guidance on the construction of theoretical corpus became evident, the research in experimental psychology was also introduced for better understanding. However, the guidance was designed for a specific case study. Hence, it should be interpreted carefully and implemented flexibly. The following methodological suggestions are given on how to construct a theoretical corpus in general:

- All necessary sources but not more should be included in the text database.
- The literature review technique(s) should be chosen with regard to the density of body. Hence at a lighter body, one and simple technique is enough to be applied (e.g. historical), especially when there is an easy-to-follow thread in the evolution of the subject; while at a denser body, a more difficult technique (e.g. systematic) or the combination of two techniques (e.g. historical with integrative) is needed.
- The corpus overview should focus on the theoretical attributes of research at a (mainly) theoretical subject, and on the empirical attributes at a (mainly) empirical one, their mixture is less recommended since it would possibly exceed the scopes of the overview, otherwise it would not be rigorous.
- The theoretical corpus should be constructed from both objective and subjective elements of former researches, while giving an emphasis to own arguments and claims, which should be clearly stated.
- The research problems and gaps should be deduced directly from the main corpus and justified scientifically.

## References

Almadi, S. (2019). The research method of conceptual frameworks. The case of Pygmalion effect: A leadership style or self-fulfilling prophecy?. *Hungarian Educational Research Journal (HERJ), 9*(3) 1-24. DOI:10.1556/063.9.2019.3.41.

Avolio, B. J., Reichard, R. J., Hannah, S. T., Walumbwa, F., & Chan A. (2009). A meta-analytic review of leadership impact research: Experimental and quasi-experimental studies. *The Leadership Quarterly, 20*(5), 764–784. doi:10.1016/j.leaqua.2009.06.006

American Psychological Association (APA). (2003). *Ethical Principles of Psychologists and Code of Conduct.* Retrieved from https://www.apa.org/ethics/code/

Burke, C. S., Sims, D. E., Lazzara, E. H., & Salas, E. (2007). Trust in leadership: A multi-level review and integration. *The Leadership Quarterly, 18*, 606–632.

Chen, G. & Klimosky, R. J. (2003). The impact of expectations on newcomer performance in teams as mediated by work characteristics, social exchanges, and empowerment. *Academy of Management Journal, 46*(5), 591–607.

Christensen, L. B., Johnson, R. B., & Turner, L. A. (2015). *Research methods, design, and analysis* (12th ed.). Upper Saddle River, NJ: Pearson.

Crawford, K. S., Thomas, E. D., & Fink, J. J. (1980). Pygmalion at sea: Improving the work effectiveness of low performers. *Journal of Applied Behavioral Science, 16*(4), 482–505. doi:10.1177/002188638001 600405

Davidson, O. B., & Eden, D. (1997). *Golem effects among underprivileged women: Mediating role of leadership, subordinate expectations, and motivation.* Paper presented at the Annual Meeting of the Academy of Management, Organizational Behavior Division, Boston, MA.

Davidson, O. B., & Eden, D. (2000). Remedial self-fulfilling prophecy: Two field experiments to prevent Golem effects among disadvantaged women. *Journal of Applied Psychology, 85,* 386–398.

Dvir, T., Eden, D., & Banjo, M. L. (1995). Self-fulfilling prophecy and gender: Can women be Pygmalion and Galatea? *Journal of Applied Psychology, 80*(2), 253–270. doi:10.1037/0021-9010.80.2.253

Eden, D. (1984). Self-fulfilling prophecy as a management tool: Harnessing Pygmalion. *Academy of Management Review, 9*(1), 64–73. doi:10.5465/amr.1984.4277938

Eden, D. (1988). Pygmalion, goal setting, and expectancy: Compatible ways to boost productivity. *Academy of Management Review, 13*(4), 639–652. doi:10.5465/amr.1988.4307530

Eden, D. (1990a). *Pygmalion in management: Productivity as a self-fulfilling prophecy*. Lexington, MA: Lexington Books.

Eden, D. (1990b). Pygmalion without interpersonal contrast effects: Whole groups gain from raising manager expectations. *Journal of Applied Psychology, 75*(4), 394–398. doi:10.1037/0021-9010. 75.4.394

Eden, D. (1990c). Industrialization as a self-fulfilling prophecy: The role of expectations in development. *International Journal of Psychology, 25*(3–6), 871–886. doi:10.1080/00207599008247933

Eden, D., & Aviram, A. (1993). Self-efficacy training to speed reemployment. *Journal of Applied Psychology, 78*(3), 352–360. doi:10.1037/0021-9010.78.3.352

Eden, D., & Davidson, O. B. (1997). *Remedial self-fulfilling prophecy: Preventing Golem effects among disadvantaged women*. Poster presented at the Annual Conference of the Society for Industrial and Organizational Psychology, St. Louis, MO.

Eden, D., Geller, D., Gewirtz, A., Gordon-Terner, R., Inbar, I., Liberman, M., Pass, Y., Salomon-Segev, I., & Shalit, M. (2000). Implanting Pygmalion Leadership Style through workshop training: Seven field experiments. *The Leadership Quarterly, 11*(2), 171–210. doi:10.1016/S1048-9843(00)00042-4

Eden, D., & Kinnar, J. (1991). Modeling Galatea: Boosting self-efficacy to increase volunteering. *Journal of Applied Psychology, 76*(6), 770–780. doi:10.1037/0021-9010.76.6.770

Eden, D., & Ravid, G. (1982). Pygmalion versus self-expectancy: Effects of instructor- and self-expectancy on trainee performance. *Organizational Behavior & Human Performance, 30*(3), 351–364. doi:10.1016/ 0030-5073(82)90225-2

Eden, D., & Shani, A. B. (1982). Pygmalion goes to boot camp: Expectancy, leadership, and trainee performance. *Journal of Applied Psychology, 67*(2), 194–199. doi:10.1037/0021-9010.67.2.194

Eden, D., & Zuk, Y. (1995). Seasickness as a self-fulfilling prophecy: Raising self-efficacy to boost performance at sea. *Journal of Applied Psychology, 80*(5), 628–635. doi:10.1037/00219010.80.5.628

Fisher, C. B., & Fyrberg, D. (1994). Participant partners: College students weigh the costs and benefits of deception research. *American Psychologist, 49*, 417–427.

Gold, M., & Kierein, N. 1998. *Pygmalion in organizations: A meta-analysis.* Poster presented at the Annual Conference of the Society for Industrial and Organizational Psychology, Dallas, Texas.

Hezkiau-Ludwig, R., & Eden, D. (2010). *Bottom-up Pygmalion: A field experimental test of the effects of follower expectations on leader performance* (Working Paper, Research No. 06780100). Tel Aviv, Israel: Henry Crown Institute of Business Research.

Jenner, H. (1990). The Pygmalion effect: The importance of expectancies. *Alcoholism Treatment Quarterly, 7*(2), 127–133. doi:10.1300/J020V07N02_09

Karakowsky, L., DeGama, N., & McBey, K. (2012). Facilitating the Pygmalion effect: The overlooked role of subordinate perceptions of the leader. *Journal of Occupational and Organisational Psychology, 85*(4), 579–599. doi:10.1111/j.2044-8325.2012.02056.x

Kelman, H. C. (1967). Human use of human subjects. *Psychological Bulletin, 67*, 1–11.

King, A. S. (1970). *Managerial relations with disadvantaged work groups: Supervisory expectations of the underprivileged worker* (Unpublished doctoral dissertation). Texas Tech University, Lubbock, TX.

Learman, L. A., Avorn, J., Everitt, D. E., & Rosenthal, R. (1990). Pygmalion in the nursing home: The effects of caregiver expectations on patient outcomes. *Journal of the American Geriatrics Society, 38*(7), 797–803. doi:10.1111/j.1532-5415.1990.tb01472.x

Livingston, J. S. (1969). Pygmalion in management. *Harvard Business Review, 81*(1), 97-106.

Mayer, R., Davis, J., & Schoorman, F. (1995). An integration model of organizational trust. *Academy of Management Review, 20*, 709–734.

McNatt, D. B. (2000). Ancient Pygmalion joins contemporary management: A meta-analysis of the result. *Journal of Applied Psychology, 85*(2), 314–322. doi:10.1037/0021-9010.85.2.314

McNatt, D. B., & Judge T. A. (2004). Boundary conditions of the Galatea effect: A field experiment and constructive replication. *Academy of Management Journal, 47*(4): 550-565.

Natanovich, G., & Eden, D. (2008). Pygmalion effects among outreach supervisors and tutors: Extending sex generalizability. *Journal of Applied Psychology, 93*(6), 1382–1389. doi:10.1037/a0012566

Oz, S., & Eden, D. (1994). Restraining the Golem: Boosting Performance by changing the interpretation of low scores. *Journal of Applied Psychology, 79*(5), 744–754. doi:10.1037//0021-9010.79.5.744

Poornima, S. C., & Chakraborty, D. (2010). The dynamics of Pygmalion effect in organisations. *The IUP Journal of Soft Skills, 4*(1–2), 49–56. Retrieved from http://search.ebscohost.com/login.aspx?direct= true&db=edb&AN=51265916&site=eds-live

Rosenthal, R. (1974). *On the social psychology of the self-fulfilling prophecy: Further evidence for Pygmalion effects and their mediating mechanisms*. New York: MSS Modular Publications.

Rosenthal, R. (2002). Cover communication in classrooms, clinics, courtrooms and cubicles. *American Psychologist, 57*(11), 839–849. doi:10.1037/0003-066X.57.11.839

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development.* New York, NY: Holt, Rinehart, & Winston.

Rosenthal, R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology, 74* (2),166-169.

Saunders, M. N. K., Lewis, P., & Thornhill, A. (2016). *Research methods for business students* (7th ed). UK: Pearson.

Sharpe, D., Adair, J. G., & Roese, N. J. (1992). Twenty years of deception research: A decline in subjects' trust? *Personality and Social Psychology Bulletin, 18*, 585–590.

Smith, S. S., & Richardson, D. (1983). Amelioration of deception and harm in psychological research: The important role of debriefing. *Journal of Personality and Social Psychology, 44*, 1075–1082.

Soliday, E., & Stanton, A. L. (1995). Deceived versus non-deceived participants' perceptions of scientific and applied psychology. *Ethics & Behavior, 5*, 87–104.

Sutton, C., & Woodman, R. (1989). Pygmalion goes to work: The effects of supervisor expectations in a retail setting. *Journal of Applied Psychology, 74*(6), 943–950. doi:10.1037/0021-9010.74.6.943

Tunnell, G. B. (1977). Three dimensions of naturalness: An expanded definition of field research. *Psychological Bulletin, 84*, 426–437.

University of Southern California. (2014). *Writing Guide - Literature Review*. Retrieved from http://libguides.usc.edu/writingguide/literaturereview

White, S. S., & Locke, E. A. (2000). Problems with the Pygmalion effect and some proposed solutions. *The Leadership Quarterly, 11*(3), 389–415. doi:10.1016/S1048-9843(00)00046-1

Whiteley, P., Sy, T., & Johnson, S. K. (2012). Leaders' conceptions of followers: Implications for naturally occurring Pygmalion effects. *The Leadership Quarterly, 23*(5), 822–834. doi:10.1016/j.leaqua. 2012.03.006

Zimney, G. H. (1961). *Method in experimental psychology*. New York, NY: Ronald Press.

## Annexes

Annex 1. Matrix for Pygmalion effect's methodological research traditions

| *Researchers, Date* | | *Setting, Deception* | | *Sample (n)* | *Variables* | *Analysis Methods* | *Hypotheses, Confirmation* | |
|---|---|---|---|---|---|---|---|---|
| Crawford, Thomas & Fink | 1980 | navy | - | NA | NA | NA | PE | x |
| Dvir, Eden & Banjo | 1995 | military | x | 435 | Leadership, subordinate self-expectancy, trait anxiety, GSE, self-task compatibility, performance | ANOVA power analysis | gender generalis ability | x |
| Eden | 1990 | military | x | 1000 | NA | ANOVA ($\omega^2$) | contrast effect | x |
| Eden & Aviram | 1993 | military | x | 154 | GSE | ANOVA ($\eta^2$) | reemploy ment | x |
| Eden & Geller | 2000 | military | - | 70 | GSE, MSE, motivation, state expectancy, crew efficiency, absence, PAI | NA + meta-analysis & BESD on 7exp. | trainabili ty | - |
| Eden & Gewirtz | 2000 | education, ideology | - | 233 | Leadership, teamwork, commitment to youth movement, Jewish identity, fun | ANOVA + meta-analysis & BESD on 7exp. | trainabili ty | x |
| Eden & Gordon-Terner | 2000 | hospital manageme nt | - | NA | MSE, PAI, PCSI | NA + meta-analysis & BESD on 7exp. | trainabili ty | - |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Eden & Inbar | 2000 | banking | - | 291 | GSE, MSE, PAI, service self-efficacy, personal-professional- fast- and accurate service, collective efficacy, PCSI | NA + meta-analysis & BESD on 7exp. | trainability | - |
| Eden & Kinnar | 1991 | military | x | 556 | GSE, SSE | ANOVA BESD | volunteering | x |
| Eden & Liberman | 2000 | military | - | 326-358 | Leadership, motivation, specialty and physical exams | descriptive stat., ANOVA, BESD + meta-analysis & BESD on 7exp. | trainability | - |
| Eden & Pass | 2000 | education | - | 267 | PAI, MSE, job satisfaction, GSE | ANOVA + meta-analysis & BESD on 7exp. | trainability | x |
| Eden & Ravid | 1982 | military | x | 70 | Performance, self-expectancy, equity | ANOVA ($\eta^2$) | self-expectancy | x |
| Eden & Salomon-Segev | 2000 | banking | - | 154 | PAI, MSE, GSE | NA + meta-analysis & BESD on 7exp. | trainability | x |
| Eden & Shani | 1982 | military | x | 110 | Learning performance, attitudinal effects, leadership, subordinate perceptions of supervisor behaviour, job satisfaction, expectancy | ANOVA ($\eta^2$) | PE | x |
| Eden & Zuk | 1995 | military | x | 25 | GSE, SSE | descriptive stat., regression | seasickness | x |
| Hezkiau-Ludwig & Eden | 2010 | education | x | 259 | Subordinate expectation for supervisors, supervisor self-efficacy, supervisor self-expectations, supervisor motivation, supervisors' followership, performance of both | ANOVA, BESD | reversibility | - |
| Jenner | 1990 | rehabilitation | x | NA | NA | NA | rehabilitation | - |
| Learmann, Avorn, Everitt & Rosenthal | 1990 | rehabilitation | x | NA | NA | NA | rehabilitation | - |
| Natanovich & Eden | 2008 | education | x | 534 | leadership, self-efficacy | ANOVA, BESD | gender generalisability | x |
| Rosenthal & Jacobson | 1968 | education | x | 524 | Age, sex, ability, intellectual growth | descriptive stat., correlation | PE and SFP | x |
| Sutton & Woodman | 1989 | retail | x | 259 | Leadership behaviour, subordinate self- | descriptive stat., | PE and expectations | x |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | expectations, individual sales | ANOVA, correlation | | |
| Whiteley, Sy & Johnson | 2012 | education /research assistants in workgroups | - | 453 | leadership variables e.g. positive LIFTs, leaders' performance expectations, leaders' liking for their followers | descriptive stat., CFA, common method variance, partial correlation procedure | Leaders' positive conceptions of followers and Pygmalion effects | x |