Goal Attainment Scaling and Coaching Client Core Self-Evaluations

David Tee, David Shearer and Gareth Roderique-Davies

**Objective:** With a strengthening of the evidence base for coaching as an effective developmental intervention, greater understanding is needed regarding the factors that may contribute to this effectiveness. This study sought to determine the effect client ‘core self-evaluations’ (CSEs) may have on the attainment of contracted coaching goals. It also adopted a revised approach to Spence’s (2007) goal attainment scaling (GAS) procedure.

**Design:** A quasi-experimental design was employed. 135 participants were recruited from undergraduate awards within a UK university. Seeking to improve on previous client trait research, a homogeneous, objectively measurable criterion variable (overall year end grade), was adopted.

**Method:** Each of the four facets in the CSE construct were measured using established psychometric scales. Participants received training in core coaching skills and then took part in six peer coaching sessions, conducted at monthly intervals over a period of half a year.

**Results:** A multiple regression analysis found a non-significant relationship between the four CSE facets and coaching goal attainment ($R^2 = 0.03\%$, $F_{4,130} = 1.04$, $p < .39$).

**Implications:** The amount of outcome variance attributable to participant CSE scores was 66% lower than that from a comparable study using self-report measures.
(Tee, Shearer & Roderique-Davies, 2017), suggesting that the use of objective measures of client goal progress has a marked effect. The rigorous GAS procedure used in this study allowed more consistency in comparing goal attainment across clients. Finally, these findings may imply that ‘core-self evaluations’ may not be the key set of client traits that predict coachability.

**Keywords:** coaching psychology, client factors, coachability, active ingredients, core self-evaluations, goal attainment scaling

**Introduction**

Early coaching researchers argued the need for evidence to determine the effectiveness of coaching as a developmental intervention (see Fairhurst, 2007; Leonard-Cross, 2010). However, the threefold increase in the number of scholarly coaching publications since 2010 (Grant & O’Connor, 2019) has produced sufficient studies to allow for their aggregation and, thus, the generation of systematic reviews and meta-analyses (see Lai & McDowell, 2014; Theeboom, Beersma & van Vianen, 2014).

A number of these reviews (Grover & Furnham, 2016; Jones, Woods & Guillaume, 2016) have determined from the emerging evidence base that coaching is an effective intervention that does generate benefits. While further research is needed concerning coaching effectiveness, these determinations legitimise researchers now broadening the agenda to identify variables that explain how coaching succeeds.

**Active ingredients**

One model has been used as the framework for a number of coaching psychology
studies, despite having its origins outside coaching in a distinct but related ‘helping by talking’ intervention. The ‘Active Ingredients’ framework is based on the ‘common factors’ model in counselling and psychotherapy (Asay & Lambert, 1999), where decades of research were distilled into four ‘factors’ (‘therapeutic relationship’, ‘expectancy, hope and the placebo effect’, ‘theory and technique’ and ‘client/extratherapeutic factors’). Of these, ‘client/extratherapeutic factors’ has been determined to have the greatest influence on the effectiveness of psychotherapy, predicting 40% of systematic outcome variance. This framework was first advocated by McKenna and Davis (2009) as having utility within the context of coaching and has been cited many times since (see de Haan, Duckworth, Birch & Jones, 2013; Smith & Brummel, 2013; Tee, Shearer & Roderique-Davies, 2017).

The number of client characteristics that potentially inform the process and outcome of psychotherapy is argued by Clarkin and Levy (2004) to be almost limitless. They cite studies concerning factors such as personality traits (e.g. ego strength), objective demographic characteristics (e.g. age or gender), biological factors (e.g. REM sleep characteristics) and external characteristics (e.g. social support). Hubble, Duncan, Miller and Wampold (2010) argue that clients are the most neglected factor in studies of psychotherapy, with a greater number of studies concerning models, techniques and the practitioner-client relationship. An equivalent bias may arguably exist within coaching psychology research.

Core self-evaluations

Where there has been research concerning coaching client factors, numerous characteristics, including developmental readiness (Bsharah, 2018), learning goal
orientation (Bozer, Sarros & Santora, 2013) and narcissism (Mansi, 2009) have been theorised as being of especial relevance. However, several coaching psychology studies have examined the broad dispositional trait ‘Core Self-Evaluations’ (CSE), defined by Judge, Locke and Durham (1997) as fundamental, subconscious conclusions that individuals reach about themselves, other people and the world. They posit CSE as comprised of four specific facets: self-esteem, neuroticism, locus of control and generalized self-efficacy, each of which are claimed to focus on global evaluations an individual might make about themselves or how they relate to their environment (Judge, Locke, Durham & Kluger, 1998).

Libri and Kemp (2006) used a cognitive behavioural coaching intervention to increase CSEs. A single case design was used, with an Australian finance executive as the sole coaching client. Whilst an increase in CSE ratings was realised at the end of the 12 week study and also at a six month follow-up evaluation point, the nature of the design and sample size makes it difficult to determine any causal relationships between CSEs and coaching interventions.

Tee, Shearer and Roderique-Davies (2017) used a behavioural (goal-focused) coaching approach and made use of full scales for each of the four CSE facets, allowing for future factor analysis of any specific scale item that might predict coaching success. Participants (N=45) were randomly assigned to the experimental or control condition. Whilst not statistically significant, 9.7% of the variance in goal attainment was explained by CSEs. However, progress on the criterion variable was reliant on participant self-report and there was also marked variance in the nature and difficulty of the personally relevant goals that each participant set, making
meaningful conclusions problematic.

Mackie (2015) adopted a strengths-based coaching approach and hypothesised that CSEs would predict coaching outcomes and would also increase as a result of participating in the coaching intervention. In common with Libri and Kemp (2006), this study made use of a 12 item CSE scale (Judge, Erez, Bono & Thoresen, 2003) and featured a sample of thirty coaching clients, with 17% dropping out during the course of the programme. Whilst CSEs were a significant predictor variable, this was for only the second of the two cohorts. Mackie points to the sample size and logistical issues as notes of caution when seeking to generalise from his findings.

This present study therefore sought to use a larger sample size than previous CSE studies and to avoid any reliance on self-report measures of the criterion variable in order to obtain data concerning the impact of client CSEs.

**Goal Attainment Scaling (GAS)**

Goals are argued to be a dominant focus in the coaching process (Grant, 2006; Ives, 2008) and therefore were used to represent coaching effectiveness in this study. Grant, Passmore, Cavanagh and Parker (2010), recognising that a wide range of idiosyncratic measures will prevent a coherent body of knowledge from developing, advocate the broader adoption of GAS (Kiresuk & Sherman, 1968) in coaching research. Therefore, the extent to which the client’s contracted goal was achieved was operationalised as the criterion variable using GAS. Previous studies (Tee, Shearer & Roderique-Davies, 2017; MacKie, 2015) encouraged clients to set realistic goals of personal relevance, resulting in a heterogeneity of topic, difficulty
and importance of goals, compromising the validity of goal attainment as a meaningful construct. This study therefore required each participant to set the same coaching goal: ‘Target percentage grade achieved for current year of academic study’.

Spence (2007) detailed an eight-step procedure for using GAS in coaching impact evaluation studies. Spence does discuss shortcomings, mainly focusing on MacKay and Lundie’s (1998) request for data to be treated as ordinal for analysis purposes. However, Spence’s procedure does allow latitude in the percentage range any individual participant might allocate to each of the five GAS goal bands (which range from ‘Worst expected outcome’ to ‘Best expected outcome’). Some clients might set a precise percentage score (such as ‘65%’) as a target for a particular band, whereas others might place, for example, a twenty percentage point range within that same band (such as ‘50-70’%). This variance in the use of GAS potentially means coaching clients deviating from their target outcome to an identical extent would therefore end up in different bands on the GAS scale.

To ensure a more consistent measure of goal attainment in this study, a prescriptive use of GAS as the criterion variable scale was introduced, detailed below. The focus on academic attainment as the criterion variable also allowed greater objectivity in this study. Previous CSE coaching studies had a reliance on self-report data for the extent to which the client’s contracted goal had been accomplished. For this study, written consent was obtained from the participants for year-end grades to be obtained from the host university, allowing for an objective measure of goal attainment to be incorporated into the research design.
Study aims and hypothesis

To determine whether CSEs might significantly predict coaching success, operationalised as client progress towards contracted goal attainment, this study sought to work with a larger sample size than previous CSE coaching studies. In addition, it introduced greater homogeneity to the criterion variable by working with a population who held a common – and therefore comparable – goal. It introduced a consistent formula to working with bands within the GAS instrument. Finally, it used an objective measure of goal attainment to reduce reliance on client self-report of progress. The hypothesis was that a statistically significant positive correlation would exist between client CSE scores and coaching goal attainment.

Method

Participants

A cluster sampling strategy was adopted, with participants drawn from undergraduates studying a range of academic subjects within a UK university. Participants \((N = 135)\) were enrolled as participants in a co-coaching programme throughout the autumn and spring term of an academic year. 63% of the participants were female \((n=85)\) and the mean age was 20.36 years \((SD = 5.75)\).

Measures

For the CSE predictor variable, each of the four facets were operationalised using individual established measures, as follows:

Self-Esteem Scale (Rosenberg, 1965). This has a four-point Likert-type scale ranging from strongly disagree to strongly agree. All ten items were used. An
example item is ‘I wish I could have more respect for myself’ (α ranges from .77 to .88). In the current study, the Cronbach alpha coefficient was .87.

**Neuroticism sub scale (Eysenck, Eysenck & Barrett, 1985).** All twelve items were used to measure neuroticism on a two point Yes/No scale. An example item is ‘Does your mood often go up and down?’ According to Eysenck et al. (1985), the ‘Neuroticism short scale’ has a reported reliability of .80. In the current study, the Cronbach alpha coefficient was .80.

**Generalized Self-Efficacy Scale (Schwarzer & Jerusalem, 1995).** All ten items from were used on a four-point ‘not at all true’/‘exactly true’ scale. An example item is ‘I can solve most problems if I invest the necessary effort’. According to Schwarzer and Jerusalem (1995), in samples from 23 nations, Cronbach’s alphas range from .76 to .90. In the current study, the Cronbach alpha coefficient was .77.

**Locus of Control Scale (Rotter, 1966).** All 29 items were used in this study. An example forced choice item is ‘There are certain people who are just no good/There is some good in everybody’. Internal consistency scores ranged between .65 and .79. In the current study, the Cronbach alpha coefficient was .69.

**Goal Attainment Scale (Kiresuk & Sherman, 1968).** Participants set a precise target grade (plus/minus two percentage points), creating a five percentage point range for Band Three (‘Target outcome’) on the GAS scale. Bands Two (‘Worse than expected outcome’) and Four (‘Better than expected outcome’) then each had ten percentage point ranges, with bands One (‘Worst possible outcome’) and Five (‘Best possible outcome’) finally allocated grades covering the remaining range of possible outcomes down to zero and up to one hundred respectively. Table 1 illustrates an example range of GAS band scores.
Table 1: An example of the GAS target mark bandings

<table>
<thead>
<tr>
<th>Band</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band 5</td>
<td>Best possible outcome: 79-100% (The gap from Band 4 to 100%)</td>
</tr>
<tr>
<td>Band 4</td>
<td>Better than expected outcome: 69-78% (10 % point range above Band 3)</td>
</tr>
<tr>
<td>Band 3</td>
<td>Target outcome: 64-68% (Target grade of 66%, +/-2%)</td>
</tr>
<tr>
<td>Band 2</td>
<td>Worse than expected outcome: 54-63% (10 % point range below Band 3)</td>
</tr>
<tr>
<td>Band 1</td>
<td>Worst possible outcome: 0-53% (The gap from 0% to Band 2)</td>
</tr>
</tbody>
</table>

**Procedure**

Ethical approval for this study was granted by the host university’s Faculty Ethics Panel. At the start of the UK academic year participants gave written informed consent to take part in the study, then completed the CSE measures. From September to March, participants engaged in monthly sixty minute coaching skills workshops, based on Grant and Greene’s (2001) programme. Each workshop was followed by a sixty minute supervised co-coaching session. During the initial session, participants considered their aspirations for life beyond graduation and, for that to be realised, what target overall grade they would need to achieve for their current year of study. The five subsequent monthly coaching sessions were used for participants to co-coach their partner on progress towards these target grades.
The researchers obtained the actual academic grades achieved by each participant at the end of the academic year, allowing an objective comparison to be made against their contracted coaching goal.

**Results**

Mean and standard deviations show that each of the four CSE variables are significantly correlated but that none of these predictor variables are significantly correlated to the criterion variable (Table 2).

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Neuroticism</td>
<td>5.72</td>
<td>3.26</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Generalised Self Efficacy</td>
<td>18.84</td>
<td>3.60</td>
<td>.43**</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Locus of Control</td>
<td>13.00</td>
<td>3.71</td>
<td>-.22*</td>
<td>-.23*</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>4. Self Esteem</td>
<td>18.59</td>
<td>4.73</td>
<td>.69**</td>
<td>.49**</td>
<td>-.19*</td>
<td>-</td>
</tr>
<tr>
<td>5. Goal attainment</td>
<td>2.74</td>
<td>1.25</td>
<td>-.09</td>
<td>-0.05</td>
<td>-.01</td>
<td>.05</td>
</tr>
</tbody>
</table>

** = Correlation is significant at the 0.01 level (2-tailed)
* = Correlation is significant at the 0.05 level (2-tailed)

Multiple regression was conducted to determine the effect of the four CSE facets on coaching goal attainment. Preliminary analyses ensured the data set featured no violation of assumptions of normality, multicollinearity, linearity and homoscedasticity. A statistically non-significant result was produced, with the $R^2$ value indicating that 3.1% of variance in goal attainment can be explained by
variances in the four predictor variables. \( F_{4,130} = 1.04, p < .39 \). The analysis suggested that Self-Esteem (\( \beta = .21 \)) and Neuroticism (\( \beta = .20 \)) were the most influential predictors and Locus of Control (\( \beta = -.03 \)) was the least influential predictor in the model. Neuroticism (\( t = 1.69, p = .09 \)), Self Esteem (\( t = 1.75, p = .08 \)), Generalized Self Efficacy (\( t = -.78, p = .44 \)) and Locus of Control (\( t = -.29, p = .78 \)) were shown not to be statistically significant predictors of goal attainment.

**Discussion**

This study sought to explore whether the core self-evaluations of coaching clients impacted on the attainment of contracted coaching goals. The hypothesis predicted that there would be a significant correlation between client CSE scores and coaching goal attainment. As this prediction was not supported by the data, the null hypothesis cannot be rejected.

Comparing these findings with previous studies, MacKie (2015) did report a statistically significant relationship between client CSEs and coaching effectiveness. Although that study required clients to set three goals (concerning a realised strength, an unrealised strength and a learned behaviour or weakness) and to track these during subsequent coaching sessions, the impact of CSEs on coaching effectiveness was operationalised by using changes in score on a transformational leadership questionnaire, rather than by progress against the three contracted goals.

Considering the other coaching client CSE studies, Libri and Kemp (2006) used a single case study design and Tee, Shearer and Roderique-Davies (2017) did not find a significant relationship between CSEs and coaching effectiveness. Given the
findings from this present study are also non-significant, there remains a lack of evidence that client CSEs are meaningful predictors of coaching outcome variance.

Focusing on the criterion variable, any common approach amongst coaching researchers to measuring goal attainment will allow comparisons across individual studies. The ability of the GAS process, as an established technique that aids coaches in converting any client goal into empirical data, justifies Grant et al.’s (2010) call for its wide adoption. In addition, the refinement in the application of Spence’s (2007) GAS procedure in this present study may partially account for the lower predictive validity of client CSEs (3.1%) compared with Tee, Shearer and Roderique-Davies (2017), where CSEs accounted for 9.7% of the variance in goal attainment. Given the difference in findings across these two studies in particular, an argument can be made for using a more consistent score band parameter in GAS coaching studies, particularly when asking clients to set their ‘Band 3’ targeted outcome: the wider adoption of the refinement to the GAS procedure used in this study should allow conclusions and inter-study comparisons to have greater validity.

Previous coaching client CSE studies allowed clients to choose their own goal focus. This has ecological validity, as it is in keeping with the dominant approach in workplace coaching practice. In contrast, this present study dictated the focus of the client goal. Whilst asking undergraduate students to determine and then work towards their target grade for their current year of study has high face validity as a relevant focus, the denying of choice as to the topic for their contracted coaching goal may have impacted on their engagement and effort in the process. It can be
speculated that some of the student participants were not strongly motivated to achieve a given academic standard of performance and, were it not for this study, would not otherwise be setting themselves target grades nor be unduly concerned by whatever grades they subsequently went on to achieve.

Finally, this study featured an objective criterion variable measure, using actual academic grades achieved by the participants and obtained from their host university. As de Haan (2012) reports when surveying coaching outcome research between 1993-2012, field studies using client reports of success consistently reported much larger effect sizes than field studies with objective measures. The reduction in variance explained by core self-evaluations from 9.7% to 3.1%, comparing these findings against the self-report study conducted by Tee, Shearer and Roderique-Davies (2017), is consistent with de Haan’s conclusions. Although the limitations of self-report criterion measures in coaching research have long been argued, the fact that this study recreates many of the design features of Tee, Shearer and Roderique-Davies (2017) does allow a direct comparison and therefore serves as compelling evidence for the distorting effect that self-report measures may have on any coaching impact research.

Conclusions

None of the studies that have used CSEs as a predictor variable, regardless of any design strengths or limitations, have suggested these might predict an outcome variance approximating the 40% attributed to client and extratherapeutic factors in therapy research. It can therefore be concluded that there is merit in future research
instead identifying non-CSE client factors that might contribute to coaching
effectiveness. It may also be the case that the variance attributed to
‘client/extratherapeutic factors’ by Asay and Lambert (1999) is not transferable from
a therapeutic context and that different ‘common factors’ are at play in coaching.

Progressing from self-report to making use of objective levels of client goal
attainment in future coaching research should increase confidence in the validity of
any findings.

Finally, the more prescriptive use of the five bands within the goal attainment scaling
technique when setting targeted outcomes with coaching clients in future research
should allow more meaningful comparisons and conclusions to be reached across
coaching effectiveness.

Correspondence

David Tee

Wales Coaching Centre, University of South Wales

Email: david.tee@southwales.ac.uk
References


