Transactional Analysis: a review
Dr K S Kinmond k.kinmond@mmu.ac.uk
M A Watkin m.watkin@mmu.ac.uk
Manchester Metropolitan University

1.0 Introduction
This paper outlines the key features of Transactional Analysis (TA), briefly charting its history and development then proceeding to review the research on the concepts of TA, and finally, assessing the research evidence for TA’s efficacy as a therapeutic technique.

TA was developed by Berne (1957). It is both a theory of personality including theories of child development and psychopathology, which form the basis of a theory of psychotherapy, and also a theory of communication for understanding groups and organisations (Berne, 1963, 1966). It is used primarily as a therapeutic tool though it also has relevance for facilitating a deeper understanding of behaviours which affect relationships in a wide range of settings (Heyer, 1979). Theoretically, TA has roots in psychoanalysis but also cognitive behaviourist and humanist traditions, essentially combining some tenets of cognitive behaviourism and psychoanalytic insight “…within a Humanist value system” (Clarkson & Gilbert, 1988 p20).

1.1 Development of TA
The origins of TA as a recognised approach can be traced back to the 1950s to Berne’s public psychiatric and private practice and also, to his writing - by 1958 he had published articles which outlined the key terms and concepts of TA. By 1964 The International Transactional Analysis Association (ITAA) was founded and key seminal texts (Transactional Analysis in Psychotherapy and Games People Play) were published. Since the 1960s TA has developed with the ITAA and EATA (European Association of transactional Analysis) now comprising over 8000 members in over 100 countries and providing accreditation of practitioners, trainers and supervisors. In Britain TA has developed from the early days of seminar groups in the 1960s to a developing
and increasingly influential community comprising over 1,000 members with the approach finding resonance recently within the medical institutions again.

TA developed both theoretically and practically during the decades since Berne’s early work, such that by the 1970s there were three ‘recognised’ schools of TA: the Classical School, the Redecision School and the Catheksis School (Barnes, 1977); and until 2001 all qualifying accredited TA practitioners were required to be familiar with the principles and practice of all three Schools. Over the past four decades TA has developed further, specifically within the last decade to reflect the influence of debates in the psychotherapeutic community at large, and particularly the impact of work in neuroscience on psychotherapy. This impact re-ignites the long-standing debate over the epistemological location of TA; specifically whether or not it might be situated within a scientific epistemology, as Berne hoped. (This possibly relates to psychology and psychotherapy’s need to be perceived as scientific and thus achieve ontological security in acceptance by a ‘proper’ research community.)

1.2 Recent thinking in TA

A useful categorisation of current thinking and practice in TA is offered by Tudor (2007) wherein he identifies seven traditions giving the focus, theories and methods and external influences to each. Nonetheless, despite such categorisation it is a truism to state that TA is complex and certainly, the wide variation in traditions and resulting practices of TA makes general evaluation challenging. However, central to all traditions in TA are key concepts which offer a starting point for appraisal. First, is the emphasis on the contractual method – the commitment between therapist and client to a well-defined course of action “regarding their roles in achieving clearly stated objectives” (Nelson-Jones, 1982 p760). Next is the central theoretical concept of Ego states. TA is based on a model of three categories of ego states – Parent, Adult and Child, a theory of social interaction based on a model of transactions between the ego states, and a concept of a life script wherein the origins of the client’s problems reside. The premise is that the transactions engaged in by people stimulate the acquisition of ‘psychological disturbance’
and the psychological games they play then reinforce that disturbance and distress. TA also highlights the dynamic inter-action between client and therapist, and the need for the therapist to be fully conversant with the ego state they are employing in the transactional process.

1.3 Research evidence for TA
It was noted above that Berne sought scientific recognition for TA, though he argued that research and therapy should be separate endeavours. Perhaps in the current climate of reflexive research Berne would be more comfortable to link the two, since his reflective writing on the external manifestation of internal states is evidenced in both researchers’ and practitioners’ reflective accounts respectively. Additionally, TA concepts are largely appropriate to research and evaluation. For example, life-script is possible to research through questionnaires and ego-states through the Egogram (Tudor and Skills, 2006). Nonetheless, there remains a paucity of research evidence on the effectiveness of TA therapy. Certainly, the last decade has seen an increasing number of articles published in the Transactional Analysis Journal. However, much work in TA remains opinion and small-sale pieces. Arguably, there is little that the academic community would recognise as 'scientific'. Nonetheless, the articles that are available are reviewed below as a starting point for further research.

1.3.1 Studies on ego states
Methods used for identifying ego states have ranged from intuitive, and thus subjective, assessments of behavioural and linguistic indicators (Klein, 1980; Solomon, 2003; Steere, 1981) to the use of projective tests (Turner, 1988) and the development of instruments which empirically measure ego state functions (e.g. Heyer, 1979; Thorne & Fargo, 1980; Doelker & Griffiths, 1984). However it is noted (Loffredo & Harrington, 2008) that there has been little published literature on measures of ego states since the early 1980s. Therefore, this review is largely concerned with work between the 1970s and mid 80s.
1.3.2 The Egogram
A popular method for identifying ego states is the *Egogram* (Dusay, 1972), which presents an individual’s ego state energy distribution in the form of bar charts of five ego state functions (Nurturing Parent, Critical Parent, Adult, Free Child and Adapted Child). The *Egogram* is constructed (by either the practitioner or client) with a strong reliance on feeling and intuitive judgement about behaviours indicative of a specific ego state (Dusay, 1972). In defence of the client making a self-assessment, Dusay (1977) found high levels of concurrence when the *Egogram* was used with clients in group settings (Dusay, 1977). Also, Heyer (1979, p10) suggests that most clients can be ‘trained’ to judge the ego states they are employing. However distorted perceptions can result when a client is under pressure, when they use too much or to little Adult state, and when they have a lack of understanding of the attributes of the states (Dusay, 1977). Thus, arguably, the measure is highly subjective and has low reliability due to the inherent differences in the ability of individuals to correctly identify ego states (Heyer, 1979).

1.3.3 The Ego State Profile Questionnaire
As a response to the need for more objectivity, Heyer (1977) developed the Heyer *Ego State Profile Questionnaire* designed around the concept of *Egograms*. He suggests that the advantages of using an objective measure are to improve diagnoses, to provide a standardised measure which can be used to compare groups and individuals, and to help locate TA theory within a scientific paradigm (Heyer, 1979). The scale has been shown to have high validity and reliability (Heyer, 1979). Nevertheless, there remain several problems with it. First, Heyer (1976) initially gleaned descriptors of ego states from the TA literature, and a panel of expert practitioners classified the items. Yet, it must be remembered that this literature was often based on intuitive judgements about ego state behaviours which then, limits the objectivity of the scale. Further, the brevity of the 16 item scale has been suggested to limit its reliability and validity. Finally, there has been a lack of reporting of any statistical analyses to support Heyer’s claims (Loffredo & Harrington, 2008).
Heyer later developed a fifty item scale derived from his earlier research (Heyer, 1979), which was sampled on over 700 people aged 18-70 in a wide range of settings. However, once again there were problems with this scale. Heyer does not specify how these items relate to the 16 item scale. Additionally, the sample for his two test-retest measures (employed to assess the reliability of the scale), were questionable - one was a sample of the ‘normal’ population, and one was a group of San Quentin inmates. Correlations between the two tests was reasonably high for the ‘normal’ population ($r = .490 - .819$, $N = 22$), after a lapse of 4-6 weeks. Less consistency was revealed between the tests for San Quentin inmates ($r = .379 - .646$), after a 12-14 week period. This may be an artefact of the differences in re-presentation times, or an important indication that there may be differences dependant on the population group chosen in terms of tendencies to be inconsistent in responding to questionnaire items. (See Hare, 1983, for a fuller discussion of problems inherent in presenting questionnaires to assess the degree of psychopathy found in prison populations).

The scale is nonetheless, suggested to have convergent construct validity as Heyer (1979 p11) identified a significant correlation between it and a range of ‘related psychological constructs’ - a ten item self esteem scale based on Rosenberg, (1965); a 6 item test of acceptance of others (from Fey, 1955) and a ten item dogmatism scale (adapted from Trodahl & Powell, 1965). Thus, Heyer (1979) argues that these results support T.A. theory and practice. He subsequently reported that the Ego State Profile was administered over a five year period to several thousand participants in both clinical and organisational settings (1987). The findings suggest that the measure is able to differentiate ego state profiles of certain groups – e.g. alcoholics in treatment, anorexic women, criminal offenders and the abused which may make it useful for those working with these groups.

Participants were provided with 50 dichotomous statements (agree- disagree) designed to represent five ego states (Critical Parent, Nurturing Parent, Adult, Adult Child, FC). Data from the groups was separately analysed using factor
Almost exclusively items were assigned to factors with the same theoretical identity in both data sets (Heyer, 1987). Heyer (1987) suggests that these results confirm that ego states are “..distinct and operationally verifiable constructs in a general population” (Heyer, 1987 p286).

In conclusion, the *Ego State Profile Questionnaire* appears to meet the criteria to be considered a scientific attempt to identify ego states, which when used in practice would perhaps reduce the need to employ subjective assessments, which may be prone to error. However it must be remembered that although the questionnaire method is extensively used in a wide range of settings, self-report measures are prone to social desirability biases, and subjective self-assessments. Also the use of a dichotomous scale provides a limited perception of the complex nature of each ego state. A further issue with self-report questionnaires is interpretation. This is highlighted in this study, as Heyer (1987) reports that some items did not fit the predicted ego states. He attributed this to individual differences in perceptions of word meanings. For example, words such as –‘impulsive’, and ‘don’t like delayed gratification’, which had been assigned by judges to Free Child, were actually seen by participants as more closely associated with Adult Child.

### 1.3.4 The Ego State Inventory

A range of other scales have been developed to assess ego states. Doelker & Griffiths (1984) devised The *Ego State Inventory* (ESI), based on Heyer’s *Ego State Profile* (1979). The scale was developed using test items from two standardised personality tests – the *Personal Orientation Inventory* (POI) (Sholstrom, 1964), which assesses levels of self-actualisation and Cattell’s 16PF (1967), which identifies a wide range of personality dimensions.

Raters were provided with definitions of the five ego state functions identified by Dusay (1972) and asked to categorise questionnaire items into specific ego states. Inter-rater reliability was high for the 75 items which comprised the final questionnaire. However, the authors provide no explanations for using the POI and the 16PF in the development of the ego state categories; and thus no evidence that the chosen items do indeed measure Ego States.
Two other instruments were developed to assess the reliability of the ESI, the *Ego State Perception – Self Rating Scale* and the *Ego – State Perception Cohort Rating Scale*. The *Ego State Perception – Self Rating Scale* requires participants to rate statements in each ego state category on a scale of 1-10 (least like me - most like me). The *Ego – State Perception Cohort Rating Scale* then asks them to identify three individuals who knew them well and to make judgments about the extent to which these individuals would agree/ disagree with the same statements.

The ESI was piloted on 57 social science students, who one week later also completed the rating scales. The ESI was completed again eight weeks later to assess test-retest reliability. A significant positive relationship was found between the two administrations of the ESI. Significant relationships were also found between the Self and Cohort rating scales and the ESI. Although these results demonstrated that the method has high reliability, the Cohort rating scale is inherently problematic in terms of validity - the participant is being asked to make a subjective judgement, which may be prone to attribution biases. It may have increased the validity of the measure if the scale had been presented to the individuals chosen by the participant, although this could be logistically difficult to achieve. Also, the dichotomous nature of items from the POI and 16PF provides a limited choice of response.

Furthermore, in view of the small sample size results should be treated with caution. Clarke –Carter (1997) suggests a minimum of 100 participants are needed when employing questionnaire methods, to ensure less risk of making an incorrect assumption about the significance of the results. Additionally, the use of an opportunity sample of social science students may have impacted upon the results as they are neither representative of the general population nor the clinical population. Interestingly, the results suggested that these participants exhibited elevated Nurturing Parent and Free Child Ego states with the Adult elevated to similar levels. This may be a result of social desirability, or indeed it may confirm the validity of the ESI – at least in relationship to the stereotype of social science students.
In summary the ESI appears to be identifying some aspects of personality which may be related to specific ego states, however it remains uncertain whether the chosen items do indeed measure Ego States. Thus it is difficult to determine whether the Scale has construct validity.

1.3.5 The Group Ego State Measure
Swede (1978) also suggests that ego states and transactions are recognisable phenomena. He developed the Group Ego State Measure (GEM) to facilitate the identification of ego states and transactions in social interaction. The variables were operationally defined – although the author does not specify the rationale for the definitions employed. Ego states were then identified by four therapist judges from tape recordings from a therapy group of seven clients and two therapists (Swede, 1976). Each judge made a mean of 60 judgements for ego states and 18 each of transactions. Inter-rater reliability was high in identifying the three ego states category (P,A,C) (78% agreement) and six ego states categories (CP,NP, A, AC, LP, NC) (71%), and in identifying crossed and complementary transactions (81%)

However, there were no identifications of Controlling parent and only one of Natural child. Swede (1978) suggests that this may be either a result of unclear definitions for these categories, or that the group being assessed showed little evidence of these ego states. The author does not identify whether the GEM has been used in further empirical studies, and thus it is difficult to determine whether it has any applicability to clinical settings.

1.3.6 The Ego State Scale
Thorne & Faro (1980) suggest that earlier attempts to measure ego states in a systematic and quantifiable manner were fraught with methodological problems (e.g. deviation from accepted TA concepts, limited samples, and insufficient evidence of validity). They also suggest that although Heyer (1979) attempted to address these limitations, his scale did not particularly relate to pathological issues (Thorne & Faro, 1980), and therefore it may be limited in its clinical application. They thus, developed the Ego State Scale
(ESS) to measure ego states and to examine the relationship between ego states and pathology (e.g. depression, schizophrenia and hysteria).

Studies were then undertaken to assess whether the instrument could be used in clinical applications. The ego state rankings of clients in psychotherapy were noticeably different at the start of treatment from the rankings found in the standardisation group (\( r = .57 \)), but had become more similar by the end of the treatment programme (\( r = .81 \)). In order to assess whether the scale would be useful in an assessment of pathology, the MMPI (Minnisota Multiphasic personality Inventory) was used as a criterion. The MMPI (Hathaway & McKinley, 1942) is a widely used and validated measure of personality, which was developed to measure many different types of psychopathology. Results indicated that clients who showed most deviation from the norm on the MMPI also showed greatest deviation from the standardisation group on the ESS.

Thorne & Faro (1980) subsequently conducted a study to determine whether the ESS could be used to predict pathology, and to examine the relationship of each ego state to pathology. A sample of 173 clients at an on-campus clinic completed a battery of tests which included the MMPI and the Adjective Check List (ACL). (The ACL is a self administered scale consisting of three hundred adjectives and adjectival phrases which describe positive and negative perceptions of self, developed by Gough & Heilbrun, (1965). The ACL measures thirty seven traits, is considered to be highly valid, reliable and efficient, and is used in many applications.) The participants were divided into three groups, dependent on levels of pathology as assessed by the MMPI. A multivariate analysis of variance was performed to examine the severity of pathology and comparable ego states. A Pearson correlation coefficient was used to examine the relationship between each ego state and MMPI indices. The analysis identified four ego states which were significantly affected by pathology. As scores on the MMPI, increased so did scores on –Ac and – NP, with a concomitant decrease in scores on FC and A. Thorne & Faro (1980) suggest this is in line with TA theory.
1.3.7 Recent developments: The Ego State Questionnaire

In response to the lack of recent empirical studies identifying ego state functioning, Loffredo & Omizo, (1997) developed the *Ego State Questionnaire* (ESQ) based on content validity, and sampled on a ‘normal’ population of undergraduates. Test-retest reliability with a two week interval was high (r=.90). The ESQ is a forty item forced choice instrument using a five point Likert scale, and consisting of five subscales which separately measure NP, CP, A, FC and AC. Cronbach Alphas of over .7 confirmed the internal reliability of the measure. Loffredo & Omizo (1997) employed the ESQ to identify any differences in ego state functioning dependent on gender and ethnicity. A one way independent measures multivariate analysis of covariance (MANCOVA) was used to analyse the data. Significant differences were found between genders in the Nurturing Parent state, with females scoring significantly higher. However no significant differences were identified between ethnic groups in any of the ego state functions, or on two other measures included – dogmatism and locus of control.

A later study (Loffredo, Harrington & Okech, 2002) was conducted to explore the construct validity of the ESQ using a sample of 200 undergraduates. A VARIMAX rotation factor analysis was used to identify the pattern of factors representing the five ego states. The analysis revealed that the ESQ showed good construct validity as a measure of NP, AC and A, but poor construct validity as a measure of CP and FC. This supported the findings of Williams et al (1983) and thus, suggested that these theoretical constructs may need to be re-evaluated. The ESQ – Revised (Loffredo, Harrington, Munoz & Knowles, 2004) was later created to provide an ego state measure with construct validity for all five ego state functions.

Loffredo & Harrington (2008) report a further study using the ESQ-R. Most of the 1997 study was repeated, but with a new variable, college major, included. The purpose of the study was to assess gender differences on NP, assess differences in ego state functioning dependant on vocational interest, and to identify any functional ego state differences between different ethnic groups. Three hundred undergraduate and graduate students, 77 males and
223 females between 19 and 58 years old, and from a range of ethnic backgrounds, completed the ESQ-R online.

A statistically significant difference by gender was indentified, and a post hoc analysis revealed this centred on the NP ego state, with females scoring significantly higher than males. Although the initial analysis revealed statistically significant differences by college major on NP, CP and AC, due to small samples in each major a post hoc analysis could not be performed to identify which students in which majors differed on these. Therefore the differences identified were not deemed valid. No significant differences were discovered in relation to ethnic differences in ego state functioning. The authors note that this may be a result of using university students as participants, the small sample of non Caucasians, or evidence that no ethnic differences exist in terms of ego state functioning.

Loffredo & Harrington (2008) suggest that their finding that females tend to be higher in NP are consistent with Heyer (1979) but contradict Williams & Williams (1980) who found no gender differences on NP in their study. However it could be that as nurturing is considered a stereotypically female characteristic (Taylor, 2002: Rane & Draper, 1995) self report biases may have come into play. Also the uneven sample, which was more weighted to female participants, may have played a part – the authors do not indicate how they dealt statistically with the differences in participation by gender. However, they do note that therapists should be aware of gender issues that may lead to lower baselines in males in NP functioning.

1.3.8 Summary

In summary, there are real problems of validity with all the scales outlined above. Specific issues have been detailed in each section, but general points relate to sample size, inclusion and exclusion criteria and criteria for construction of the research tool. These all need to be taken account of in any research aiming for scientific credibility, and specifically before any claims relating to therapeutic practice can be made.
1.4 Empirical evidence for the effectiveness of TA as therapy

Novey (2002) suggests that the assessment of therapeutic interventions generally falls into two categories, effectiveness or efficacy studies. Efficacy studies involve the use of objective methods to determine the success or otherwise of a specific treatment programme. Clients with single diagnosis disorders are randomly assigned to either a treatment or control group, rigorous controls are implemented, outcomes carefully defined, and assessments are conducted at termination of treatment and at follow up.

Whilst having scientific credibility, efficacy studies have been criticised for being inadequate in taking into account a range of factors pertinent to the therapeutic process as it functions in practice (Novey, 1999): For example, the fixed time scale of the intervention used in efficacy studies, the active role of the client in choosing other therapeutic measures if they perceive one to be ineffective, and, importantly, that clients may not present with a clearly defined single ‘disorder’. Novey (1999) therefore posits that effectiveness studies, although considered more subjective in that they measure the clients self reported satisfaction with a particular measure, address some of these issues.

Khalil, Callaghan & James (2007), however, note a paucity of research within the TA paradigm to have employed either method with much rigour. In a report commissioned by the Berne Institute to determine the current evidence base for the positive effect of TA counselling, and its efficacy/ effectiveness compared to other therapeutic measures, Khalil et al (2007) identified relatively few studies which met their criteria for inclusion as valid pieces of research. Importantly, they also suggest that most of these were published in specialist journals and “….not subjected to quality reviews from the wider academic and health communities” ( Khalil et al, 2007 p20), thus limiting the ability of TA to gain credibility within these fields.

The authors conducted an extensive review of the literature published between 1950- September 2007, and noted that whilst there was ample material on TA theory and methods of application, much evidence about
outcomes was largely anecdotal, generally consisting of subjective accounts by therapists about an individual client’s progress through therapy. Furthermore much of the empirical research relating to outcomes was relatively dated, with little produced during the last decade.

Just thirty four studies were initially identified as relevant, however thirteen were excluded for a variety of reasons e.g. no standardised measures employed, no pre-test measures, not specifically relating to treatment or poor quality of reporting. A further two studies were unobtainable. Of the nineteen remaining, twelve reported positive effects of TA on client outcomes (Smith 1977; Prothero, 1978; Payton, Morris et al, 1979; Olsen, Ganley, Devine & Dorsey, 1981; McNeel, 1982; Talob, 1994; Wissinck,1994; Novey, 1999; Novey,2002; Boholst, 2003, Bledsoe & Grote, 2006). Four studies were inconclusive about the efficacy of therapeutic interventions (May & Tierney, 1976; Fetsch & Sprinkle, 1982; Greene,1988; Grunewald – Zemsch, 2002), and the remainder only indicated possible beneficial outcomes (Windell & Woollama, 1976; Childs – Gowell, 1977; Thunnissen, van Delft et al., 1998).

Five of the above studies affected a comparison of TA and other forms of therapy (Smith 1977; Prothero, 1978; Olsen, 1981; Novey, 1999, 2002; Bledsof, 2006). Of these, three suggested TA outperformed other therapeutic treatments (Prothero,1978:Novey,1999; 2002), whilst two determined TA to be less efficacious than other methods (Bledsof, 2006: Olsen,1981). The remaining study (Smith, 1977) indicated that whilst therapy generally appeared effective, no method outperformed any other.

1.5 Summary and Conclusion
In summary, this paper has outlined the key features of TA, charting its history and development then proceeding to review the research on the concepts of TA and finally, assessing the research evidence for TA’s efficacy as a therapeutic technique. In conclusion it should be noted that currently, there is a paucity of empirical research evaluating both the concepts of TA and also of its effectiveness as a therapy. Whilst certainly there are many practitioners in the field and the approach is growing both nationally and internationally and
across disciplines, there remains a lack of systematic evidence to support this growth. Thus, whilst Berne (1980, p244) claimed TA to be “A systematic phenomenology which could usefully fill the gap in psychological theory”, in order to substantiate such a claim and provide evidence for what Harris (1973, pxiv) identified as “A unified system of individual and social psychiatry that is comprehensive at the theoretical level and effective at the applied level” there is a need for a sound body of empirical research.

References


