Title Page

Title:

A tool for measuring therapeutic jurisprudence values during empirical research

Author name and affiliations:

Dr. Anna Kawalek, Lecturer in Law

Leeds Beckett Law School

Portland Building,

Leeds Beckett University,

Portland Way,

Leeds

LS1 3HE

UNITED KINGDOM

a.kawalek@leedsbeckett.ac.uk

Word Count: 13,072 including footnotes and appendices, excluding reference list.

Abstract

Since its inception in the 1980s, the growing body of Therapeutic Jurisprudence scholarship has continued to challenge and optimise not only our understanding of what the law "is", but also how we can leverage its agency to improve the ways in which we apply, observe and evaluate the law. By focusing on the emotional, human and psychological consequences of legal processes, Therapeutic Jurisprudence empowers practitioners to design emotionally intelligent and remedial strategies to either minimise harmful consequences or enhance restorative legal goals and outcomes. As the influence of Therapeutic Jurisprudence scholarship and its practical applications has continued to gain traction in rapid and organic growth, collaborations brokered with a wide range of social science disciplines have called for a more robust focus on validated measurement scales. To that end, there is a clear and growing need for a suite of Therapeutic Jurisprudence specific tools for empirical evaluation. As there has been no formal attempt within the Therapeutic Jurisprudence research community to begin this process, this paper breaks new ground by providing a validated tool for empirical measurement of the therapeutic quality of judicial officers' behavioural and interactional styles in problem-solving court jurisdictions. Using original empirical data recently collected at a problem-solving court in England, the paper takes the reader through the journey of statistically validating the levied scaling systems by performing Principal Component Analysis and Cronbach's Alpha. In doing so, the paper offers an original contribution to Therapeutic Jurisprudence methodology.

Key Words: Therapeutic Jurisprudence; Problem-Solving Courts; Empirical Measurement; Principal Component Analysis; Cronbach's Alpha; Applied Research.

1. Introduction

Therapeutic Jurisprudence ("TJ") is a legal philosophy concerned with the human effects of the law. It is known for drawing upon insights from the behavioural sciences (e.g., criminology, sociology, psychology, social work, and motivational interviewing) to instil therapeutic values within legal theory, practice, and analysis (Gal & Wexler, 2015). Findings from the social sciences, such as therapeutic interpersonal styles from psychology and counselling, help to articulate what works in communication and behavioural change, and TJ adapts these findings for use in law processes.

TJ is most commonly distinguished by its observations of the judges from problem-solving courts, and scholars have traditionally been fascinated by the interactional and behavioural styles implemented by judicial benches to engage progress and enhance rehabilitative outcomes for service-users (Winick & Wexler, 2003). TJ does, however, have far broader application and scope than this and, as the movement has gained traction, the interests of freshly involved scholars, including empiricists, have taken the paradigm into new realms. A popular current topic is how TJ can be invoked as a methodology (Stobbs, Bartels & Vols, 2019a). Although TJ research has historically been defined by doctrinal methodologies, TJ has expanded in rapid and organic growth; this has led to an upsurge of interest from empirical researchers, who have raised questions regarding the measurability of TJ.

In response, Stobbs et al. (2019a) recently released a handbook offering a laudable collection of chapters to assist with TJ research. Despite the advancements made by the volume, such as Stobbs' (2019) therapeutic imperative, there remains no rectified instrument for empirically measuring TJ values. The lack of standardisation has led to problems with research replicability, coherency and precision, and it was recently stated that: 'although the practice of applying the principles of TJ to both criminal and civil law settings appears to be gaining significant momentum in many countries, it has not yet been documented systematically in evaluation studies nor have its impacts been reported in detail' (Cooper, 2019: p. 290).

Lending itself to TJ's interdisciplinary nature, this paper implements statistical techniques from the social sciences (principal component analysis ("PCA") and Cronbach's alpha) to compound statistically validated TJ scales, using original data. To break new ground in this area, it

proposes a tool for measuring judicial interactional and behavioural styles in problem-solving court jurisdictions to be further piloted, modified, and refined. The tool can be used in both qualitative and quantitative studies and can be adapted for different settings involving interactions, such as between between lawyers, police officers, and clients. As such, this paper seeks to respond to the methodological gap in TJ. Although the ambitions for this paper are wide, the aim is to make a constructive contribution to the literature, rather than a polemical commentary on existing research.

2. Literature review

2.1. TJ: background & the research problem

TJ is a legal philosophy rooted within humanist discourse (Winick & Wexler, 2003). It rationalises that the operation of the law and its associated legal rules, procedures, roles, and institutions, impact the wellbeing of all people who encounter them, whether that be in practitioner, victim, witness, offender, or friend and/or family capacity (Winick, 2003). Through its interdisciplinary work, TJ draws upon insights brokered by the behavioural sciences to examine how the law impacts health (Winick & Wexler, 2003). As such, TJ has reconceptualised the law into a social force which, through its energy and agency, by-produces therapeutic or countertherapeutic outcomes, whether they are intentional or not, 'know it or not, like it or not' (Wexler, 2003: p. 3). To this end, TJ scholarship explores ways to diminish damaging repercussions on emotional, psychological, physical, relational, economic, and social personhood, and investigates how to implement law as a restorative, remedial, and healing instrument (Stobbs, 2019).

TJ began its life with a modest agenda in the early 1980s, aiming to generate positive reform to the experiences of those subjected to the mental health system in the United States (Wexler & Winick, 1991). However, it very quickly attracted attention from the embryonic drug treatment court movement (Hora, Schma & Rosenthal, 1999). Arising coincidentally at the same time, drug courts shared similar visions with TJ; seeking to assimilate criminal justice and the drug treatment services to achieve therapeutic outcomes for drug-using offenders, they emphasised a unique court environment where lawyers, judges and service providers worked as a supportive team, tasked with addressing the underlying causes of offending (Hora et al, 1999). TJ analyses have allowed drug courts (and latterly the growing problem-solving court

movement) to anchor their practice within an array of well-matured TJ canons (KPMG Consulting, 2014; Hora, 2002; McIvor, 2009; Wexler & Winick, 1999). It is difficult to deny the widespread impacts of this; that there are now 3,142 drug courts in the United States alone speaks its own truth in terms of prosperity (National Institute of Justice, 2012) and problem-solving courts have been piloted across jurisdictions worldwide, including: the UK (Kerr et al., 2011; McIvor, 2009), Belgium (Dekkers, Beerens, Wittouck, & Vanderplasschen, 2016) Canada (National Crime Prevention Centre, 2018), and Australia (KPMG Consulting, 2014).

Of the drug court components (Ashcroft, Daniels & Herraiz 2004), TJ scholarship indicates that judicial leadership, in which judges redefine their role from neutral arbitrators of the law to managers of treatment programs, carries the most gravity (Kerr, et al. 2011; McIvor, 2009; Hora, 2002). A broad suite of TJ research has identified that when judicial benches forge therapeutic relationships with service-users, this increases motivation to stay drug-free and helps maintain compliance with the court programme (Kerr et al. 2011; McIvor, 2009, Petrucci 2002). This relationship is facilitated by therapeutic interactional and behavioural styles of judicial engagement during court conversations (Winick & Wexler, 2003). As quoted by TJ's co-pioneer, Winick (2003: p. 127):

> even though they [offenders] have engaged in wrongdoing, a special sensitivity to the individual's pain, shame, sadness and anxiety in coming to terms with the existence of psychological or behavioural problems that have produced criminality and the victimisation of others is called for in the judge-offender interaction.

TJ research indicates that many values come under the purview of therapeutic judicial interactional and behavioural styles. In line with TJ's interdisciplinary nature, many of these originate from other subject domains such as criminology, sociology, and psychology. These styles are articulated in the key text: Judging in a Therapeutic Key as: empathy, acceptance, warmth and self-expression, hope and expectancy, a future focus, and empowerment and possibility (Clark, 2001), and neutrality, respect, participation and trustworthiness (Warren, 2002). Elsewhere, research by Petrucci (2002) demonstrated that respect was the cornerstone of a therapeutic judicial interaction in a domestic violence court, which is a form of problem-solving court. Meanwhile, Winick (2003) has argued that empathy is paramount to a therapeutic court approach and Perlin (2013a, 2017, 2019, 2020) has consistently observed the importance of dignity. Skills in the therapeutic court are often contextualised by the doctrine of procedural justice, apropos to manners that judges 'apply procedures that fully respect the

individual's participatory and dignitary interests' (Winick, 2003: 1088). Key skills here include voice, validation and voluntariness (Lynch & Perlin, 2016). Hopkins and Bartels (2019) recently identified compassion and compassionate motivation as core of for solution-focused judging. Lens, Katz, & Suarez (2016) observed the importance of respectful, empathetic, and supportive environments in a family court. In the context of UK family drug and alcohol courts, Harwin, Alrouh, Ryan, & Tunnard (2014) measured behavioural styles by way of the following indicators: talking to the parent directly, inviting parent views, expressing an interest in progress, commenting on family strengths, praising parents, stating the courts aims, explaining decision making, urging parents to take responsibility, and using a problem solving approach.

As such, various theorisations, toolkits, and measurement systems have emerged to substantiate the core meaning of therapeutic interpersonal, engagement and behavioural skills. Perhaps most comprehensively a judicial bench book, authored by Goldberg (2011) for the Canadian National Judicial Institute, characterised these skills as: empathy, respect, active listening, a positive focus, non-coercion, non-paternalism, and clarity. The importance of the Canadian National Judicial Institute's work was affirmed by Hora (2011: p. 46) for being: 'the first international body to discuss taking problem-solving courts to scale.' As such, the guidance by Goldberg arguably provides something more encyclopaedic than any previous attempt, a point that will be returned to.

Whilst differing somewhat in their approaches, each of the above-mentioned adaptations is hallmarked by a therapeutic spirit and is premised in the notion that the legal system, particularly judges, has the power to act as a therapeutic agent. To make many of these inroads, TJ researchers have utilised a whole host of valuable, but diverse, tools and scaling systems to perform empirical measurement. However, this has left behind a legacy of predicament for the devoted empiricists. The problem is that the varying perspectives on the fundamental meaning of a "therapeutic judicial bench" has resulted in somewhat loose articulation of the very core values that earmark and define TJ as a discipline (Roderick & Krumholz, 2006). The impact on research is that TJ's theoretical tenets translate themselves inconsistently onto empirical measurement leaving a somewhat scattered, externally invalid, and unduplicable set of results. Not only does this make it difficult to compare research outcomes across international projects causing critics to discredit findings for lacking replicability (Roderick & Krumholz, 2006), but suspicions amongst newcomers often resolve themselves as refusal to exploit TJ in research designs where it would be perfectly suitable (Stobbs, 2019). The defensive position is that a

loose coding of TJ principles should not be viewed as a shortcoming but rather an opportunity for measurement to take creative forms (Daicoff & Wexler, 2003). Still, the concerns for standardised measurement remain outstanding.

2.2. What is TJ?

TJ research problems are symptomatic of deeper-seated critiques of the TJ paradigm claiming that it is rudderless and undertheorized (Slobogin, 1995). Over the years, discussions as to what TJ is and what it is TJ not has foregrounded debate (Hudson, 2017; Cooper, 2017; Wexler, 2018; Perlin, 2018). These discussions question whether TJ is a theory or practice (King, Freiberg, Batagol & Hyams, 2014), multitude of normative principles (Stobbs, 2019), lens (Cattaneo & Goodman, 2009), philosophy (Roderick & Krumholz, 2006), method (Stobbs et al, 2019), a way of thinking, a 'set of procedural guidelines, protocols and techniques' (Stobbs, 2019: 45) or even an adjective (Slobogin, 1995) or community (Stobbs, 2019). These debates are often concluded, and insurgents tempered, by the assertion that TJ is a heuristic tool that can have all the above-mentioned applications depending on the way that it is applied (Stobbs, 2019; Wexler, 1993; Wexler 1995). As recently articulated by Stobbs et al. (2019b: 18), 'TJ can now legitimately be conceived of as a theory, as a research method, as a conceptual lens, as a community of scholars and practitioners, or as a paradigm.' Whilst its broadness would to be viewed auspiciously by proponents, TJ is often left without clear demarcation in many of these applications, leaving it vulnerable the critique that it is too broadly imagined (Slobogin, 1995).

Similarly, TJ's subject matter is wide; distinguished originally by criminal law and justice expertise (Wexler and Winick, 2003), it has since been applied to torts (Shuman, 1992), contract (Harrison, 1994), estates (Glover, 2012), family law (Lens, 2016), mental health law (Perlin, 2012, 2013b; Perlin & Lynch, 2016), disability law (Perlin, 2013c), workplace bullying (Yamada, 1999, 2008), administrative and private law (Cramer and Vols 2016; Koch & Diesen, 2016) and refugee reunion (Marson, Ferris & Kawalek, 2019a, 2019b), *inter alia*. Whilst its wide scope is not problematic per se, and can in fact be regarded advantageously, it has further exposed TJ to the critiques of being abstract, especially when application of the terms "therapeutic" "countertherapeutic" and "wellbeing" to these topics are seemingly boundless (Slobogin, 1995; Roderick & Krumholz, 2006). Here it has been argued that the theoretically modalities of therapeutic and anti-therapeutic should be better explained (Slobogin, 1995).

Perhaps muddying the waters further, Vols (2019) articulates that there are broadly four types of theory: doctrinal, micro-level, macro-level, and normative, but posits that TJ straddles each of these. Vols (2019) carries out a literature review revealing that TJ scholars have generally had five different interpretations of TJ as a theory. Results show that the first two responses altogether avoid the terminology "theory" although allude to TJ as providing a theoretical foundation; the first response replaces theory with the word lens and the second with jurisprudential (Vols, 2019). Within the third category, Vols (2019: p. 65) reveals that some do not perceive TJ to be a theory at all, or at least not 'true theory', as this would require a clearer set of testable indicators. According to this group 'the concepts of therapeutic and anti-therapeutic are not specific and precise enough to be used in a theory that can be tested' (Vols, 2019: 65), which rings home to the problems as articulated by critics (Roderick & Krumholz, 2006; Slobogin, 1995). Fourthly, Vols (2019) identifies that some regard TJ as wholesale theory, clear enough to be understood and distinguished in theoretical terms, largely by its normative agenda; for instance, Birgden and Ward (2003: 336) see TJ as a 'legal theory that utilizes psychological and other social science knowledge to determine ways in which the law can enhance psychological well-being of individuals who experience the law' (Vols, 2019: p. 66).

Vols' (2019) final interpretation is that TJ theory is pluralistic, that its meaning is peculiar to individuals, and is shaped by the political or sociological debates that characterise scholars' background disciplines. By ascribing to the final interpretation, Vols (2019) argues that TJ can be regarded as a theory with both descriptive (observing the law's effect on people) and normative (prescribing how legal systems should be designed and applied) components. However, Vols' review only serves to highlight the complexity of cultivating TJ as a theory. If this is true of just one area of its many applicable forms, and TJ can also be implemented as a practice, set of principles, method, philosophy (etc.), it is easy to see why TJ ontology is so widely contested.

The amorphous, subjective, and evolving nature of TJ stems from reluctance to shoehorn the doctrine into a singular ontological form or subject matter, or to singlehandedly finetune its theoretical tenets, which would risk eclipsing its full potential and scope (Wexler, 1995; Winick, 2005; Wexler, 1992). As the various characterisation of a therapeutic judicial bench show, this has allowed practitioners and academics to enjoy significant amounts of imagination

during interpretation. Nonetheless, scepticism that exacts clearer theoretical, ideological, empirical and conceptual bases still stand, as 'issues have been inconsistently defined within a largely confusing and un-integrated theoretical framework' (Roderick & Krumholz, 2006: 219). As such, Stobbs has recently suggested responding by issuing a 'comprehensive robust "apology of TJ"' (Stobbs, 2019: p. 34) to reconcile these problems. A key part of this is better articulation of TJ methodology.

In coming back to the research problem, it is easy to see why TJ, as a paradigm and within its theoretical modalities, has struggled to be implemented in consistent epistemological and methodological terms when its ontological form is founded on shaky grounds. However, if TJ is an unfettered doctrine, then TJ 'does not lend itself to a one-size-fits-all method' (Stobbs, 2019: p. 18). This at least partially explains why empirical results have been piecemeal. Whilst it is useful to assort instruments, and existing studies have been paramount in shaping how TJ should be framed in methodological and analytical terms, this does not mean that calling for robust and validated measurement tools should be disparaged. Quite the opposite is true; there now is an appetite for standardisation (Richardson, 2019; Cooper, 2019). As such, there is a clear and growing need for a suite of TJ specific tools to guide systematic empirical observation, evaluation and measurement.

2.3. TJ as a methodology

In response to these criticisms and requests, Stobbs et al. (2019a) recently released a handbook handling structured advice, principles, and examples to assist with TJ research and to usher the development of TJ methodology. They volume recognised the need for a resource that 'identifies, illustrates and explains examples of best practice for conducting TJ research and practice' (Stobbs, 2019: p. 26). Given the significance of the volume, it is important to flag how this paper fits within it.

In chapter 3, Stobbs (2019) distinguishes between two types of TJ research: pure and applied. The former refers to theoretical research 'about TJ', and the latter to practical or empirical research that moves beyond conceptual thinking (Stobbs, 2019: p. 43). Similarly, in chapter 4, Vols (2019: p. 68) marks TJ research as either 'fundamental' or 'applied/empricial'. Fundamental research is theoretical since it 'compares and juxtaposes TJ as a theory and their core underpinning principles with other (meso and macro) theories that are used to analyse the

law' (Vols, 2019: p. 60). Applied is empirical research as typically understood (e.g., data collection) (Vols, 2019). As such, the Vols and Stobbs research typologies are tantamount, with the first category broadly referring to doctrinal methodology, and the latter to empirical. This paper is concerned with applied research, as conceptualised by Vols (2019) and Stobbs (2019).

To provide empirical research with a clearer link to TJ, Stobbs (2019: p. 44) postulates the 'therapeutic imperative', a three-pronged structure for modelling TJ research summarised as: 'command', 'duty', and 'rule'. The idea is to generate fresh insight that 'cannot be produced by applying a well-worn methodology' (Stobbs, 2019: p. 30). The 'command' aspect of the therapeutic imperative offers guidance on how TJ theory can infiltrate data collection and analysis phases, and thus concerns technical aspects of research such as: ontology, epistemology, and research paradigms (Stobbs, 2019). It aims to identify anti-therapeutic and therapeutic aspects of legal processes to understand where legal systems, rules, and practices can be improved to positively impact wellbeing. Within this, Stobbs (2019) discusses the importance of outset explanations of what is or is not therapeutic. Since these modalities have no coherency in the existing literature, and only harks back to problems already outlined, this paper seeks to substantiate the first (command) component of the therapeutic imperative by proposing consistent definition of the terms therapeutic and countertherapeutic for defining and measuring the behaviour and interactions of problem-solving court benches.

The second component of the therapeutic imperative is duty; this refers to an obligation for TJ researchers to argue for modifications to legal practice if the previous command stage has identified that a context by-produces countertherapeutic outcomes. According to Stobbs (2019), if there is an argument to affect therapeutic change, in the absence of any countervailing normative considerations, then this argument should be made. This sits in line with TJ's modest goal, which does not aim to interrupt and trump existing values of legal systems, but rather to operate with subscription to them (King, et al. 2014). As consistently quoted throughout the work of Lynch & Perlin (2016, p. 15): 'TJ asks whether legal rules, procedures, and lawyer roles can or should be reshaped to enhance their therapeutic potential while not subordinating due process principles'. As such, although researchers should endeavour to necessitate therapeutic change according to empirical results, they should do so with respect to existing values (Stobbs, 2019). Although the tool in this paper is devised primarily to bolster the duty of the therapeutic imperative, if the command stage elucidates that a bench unsatisfactorily meets a therapeutic standard, then a case for change should be made. If for instance a judicial

judge fails to engage therapeutic dialogue as per the second measurement scale, a researcher could suggest training in motivational interviewing. However, these suggestions should only be carried out in the absence of countervailing normative considerations.

Finally, according to the rule, if countervailing normative considerations are identified, TJ should not be invoked as a method for redetermining decision-making (Stobbs, 2019). The rule should be followed by researchers using the tool. Whilst the tool might cast fresh insight onto areas of the law and legal process that might otherwise have been overlooked, TJ should not be solicited as a method for determining decisions that compel legal change if a counterargument is identified (Stobbs, 2019). As such, this paper has a place for all three aspects of the therapeutic imperative, but the clearest link is to the command (Stobbs, 2019).

Elsewhere, Cooper (2019) acknowledges Wexler's (2015) seminal analogy by exploring how bottle and wine level considerations can be translated into research questions, evaluation, and findings. The former refers to analysis of macro-level structures, including policy, organisational strategy and legalisation; the latter refers to application of TJ principles in practice, including judicial skillsets and methods of engagement. Both tiers contain therapeutic and anti-therapeutic value (Wexler, 2015). Building on this, Vols (2019) puts forward an approach for constructing applied research comprising four components that integrate wine and bottle components:

1. **Analysis of the therapeutic design of the law**: scoping the relevant legislation and case law through doctrinal analysis for anti-therapeutic and therapeutic responses at "bottle" level.
2. **Analysis of therapeutic application of the law**: reviewing the law in real-life by collating raw data that 'investigates the workings of, for example, a court in real life' (p. 70), and whether these systems produce therapeutic or anti-therapeutic consequences at "wine" level.
3. **Unravelling the weight given to different interests:** upholding the current values of that legal system whilst promoting therapeutic outcomes. This has been described as the 'empirical-normative method' (p. 71).
4. **Research-informed law reform:** based on the findings from prior stages, the research could use normative theory to help recommend changes in line with TJ's goals.

Vols' method bears some parallels to Stobbs' therapeutic imperative with the first two criteria alike to the command and the third to the duty. It also explicitly links with Cooper's iteration of Wexler's metaphor since Vols' first criterion refers to the design of the law (bottle) and the second to the application of the law (wine). As such, whilst the methodology handbook offers a collection of chapters with differing perspectives on how to expedite TJ methodology, many of these concatenate and dovetail to form holistic guidance for this task.

Although the scaling systems in this paper consider measurement of wine from Wexler (2015) Vols (2019) and Cooper's (2019) articulations of methodology and primarily enriches the command aspect of Stobbs' (2019) therapeutic imperative, I am not suggesting that other aspects of applied research should be ignored. Rather, the scales should be incorporated into a broader design, as per the guidance of these scholars. Wexler (2019: 8) proposes that TJ researchers craft Amicus Justitia briefs as 'a different type of legal writing', for instance: 'bullets, suggested scripts, visual aids, crisply stated best practice'. The scales serve to coalesce this goal by providing a snappy document for dissemination.

Overall, the recent handbook (Stobbs et al., 2019a) makes important inroads for TJ methodology. However, there remains gaps, and no endorsed instrument to measure the behaviours and interactions of judicial benches during applied research studying problem-solving courts. As such, the paper assumes that TJ can have application as a methodology and seeks to respond to this by statistically validating a tool for measuring "wine". That is not to say that only one tool should be used exclusively in all relevant studies henceforth. The purpose is rather to create clarity for empiricists interested in international replicability, validity, and comparative analysis. The tool is imperfect, and critiques are offered in upcoming sections. However, the aim is to ignite a process, and extend an invitation to future TJ researchers to adapt, trial, and experiment with its properties and significance. As such, this paper sits in line with broader goals for clearer articulation of TJ but marries more specifically to the topic of TJ as a methodology.

3. Methodology and Methods

3.1. Research question

- How can TJ empirical researchers measure interactional and behavioural styles of problem-solving court judges?

3.2. The Study and Setting

This paper is part of a broader study carried out between 2015 and 2018. The study sought to investigate the functioning of a problem-solving court, Manchester Review Court ("MRC"), located the United Kingdom. MRC brings addicted criminal offenders back for regular court review of recovery and law compliance on the Drug Rehabilitation Requirement under section 210 of the Criminal Justice Act (2003). During review, progress is overseen, monitored, and supported by magistrates (BLIND, 2020: *in press*). In many ways, MRC bears semblance to the classic drug court model, but lacks too many of the key components to be classed as the same (Ashcroft et al., 2004; BLIND, 2020). One stark similarity is pertinence of the judicial role, which must align with TJ principles to increase client engagement and recovery pathways in both examples (BLIND, 2020; Justice Innovation Centre, 2018; Ashcroft et al., 2004).

The primary research objective was to explore the therapeutic quality of magistrates' interactional and behavioural styles at MRC (wine). Other aspects of the project concerned bottle level questions and analysis, whilst the discussion posed normative suppositions recommending changes to policy, practice and legislation. As such, the approach to research held a similar guise to the Vols and Stobbs (2019) recommendation even before the release of the methodology handbook (Stobbs et al., 2019a); this corroborates the intuitive value of their frameworks. However, this paper is not concerned with peripheral bottle level questions or normative considerations; rather it looks at how to carry out empirical measurement of the wine.

A tool was developed to measure the wine at MRC during two data sweeps. Firstly, it was used as a standardised observation protocol to measure the judicial interactional and behavioural styles at MRC from the perspective of the researcher. Secondly, the wording was changed to first person onto a questionnaire design using the survey method to gather insight of service-users' experience at MRC from their perspective. Questionnaires were gathered through opportunistic sampling as service-users left their court hearing. The questionnaire posed thirty questions; twelve concerned broader court operation (bottle questions) and the remaining eighteen explored participant's perceptions of magistrates' interactional styles (wine

questions). In both phases, a Likert scale of 1-5[1] was used to collect data for each variable, and the results from both phases were triangulated latterly. During the planning and preparation, designing variables to measure TJ wine was made tricky because there was no rectified instrument for this purpose. To overcome this, I created four *a priori,* theoretical scales modelled upon the manual for training judges in problem-solving techniques offered by the Canadian Institute, authored by Goldberg (2011). Goldberg's (2011) bench book offered a comprehensive and authentic approach to training therapeutic judicial benches, which has not yet been empirically tested, unlike many of the other related TJ projects.

Goldberg (2011) also made empricial measurement more viable than some TJ literature sources because she discusses various constituents (sub-skills) necessary to practice certain therapeutic styles; for instance, for empathy she suggests employing six subskills (see Goldberg for more information). The process of designing the tool involved translating eighteen of Goldberg's (2011) sub-skills into observation protocol or questionnaire items (or variables) to comprise four theoretical measurement scales. Based on Goldberg (2011), these were originally entitled: 'empathy' (p. 29), 'respect' (p. 34), 'positive focus' (p. 38), and 'active listening' (p. 35) (see table A). Some of the meanings of the subskills were inferred from Goldberg's general descriptors, and they were adapted based on preliminary fieldwork scoping relevance of the site and the variables. The purpose of measuring eighteen facets on four scales was underpinned by a social science rationale, whose research concerns human-invented concepts, such as: intelligence, emotion, humour (Bauer, 2017) or judicial interactive skills. This subjective nature makes measurement difficult, which is often offset by developing tools that measure a series of attributes that characterise the construct, rather than a vast, unrefined and vague concept like empathy (Field, 2013; Stobbs, 2019). As such, eighteen variables were arranged on four respective scales to better quantify and characterise each social science construct.

I sought to keep variables as problem-solving court generic and TJ specific as possible with some natural variations caused by the idiosyncrasies of MRC. Some of the Goldberg (2011) skills were also not easily measurable and therefore abandoned after piloting and finetuning of the tool. As such, the variables were based, but not modelled wholesale, on Goldberg's (2011) handbook. It was also important to keep the wine measurements concise for standardised court observations and during surveys to enhance validity – these methods are broad approaches that

---

[1] The options were: "strongly agree", "agree" "no opinion", "disagree", "strongly disagree".

advocate bigger masses of data over richness (Flick, 2011). As such, the variables needed to be quickly implied and readily measurable, which the Goldberg (2011) manual could facilitate. Piloting comprised of roughly three weeks' worth of court observations totalling six hours before the final tool was rolled out as observation guidance and into the surveys.

Empathy	Respect	Positive Focus	Active Listening
Variable 1: Magistrates are interested in and compassionate about my life circumstances	**Variable 5:** Magistrates speak to me without pity or disdain	**Variable 10:** Magistrates have hope and faith that I will make progress	**Variable 15:** Magistrates listen to my point of view
Variable 2: Magistrates understand what it's like to have drug and/or alcohol problems	**Variable 6:** Magistrates reiterate my goals so that they are clear	**Variable 11:** Magistrates praise me when I am doing well	**Variable 16:** Magistrates give me a voice
Variable 3: Magistrates are personable	**Variable 7:** Magistrates speak to me slowly, clearly, and loudly	**Variable 12:** Magistrates motivate me	**Variable 17:** Magistrates are attentive when I am speaking
Variable 4: Magistrates are realistic when we set my goals for next review	**Variable 8:** Magistrates are sincere and honest when they speak to me	**Variable 13:** Magistrates help me build upon my strengths	**Variable 18:** I can ask magistrates questions when I need to
	Variable 9: Magistrates do not rush or interrupt me when I am speaking	**Variable 14:** Magistrates make me feel positive about my future	

Table A: The Original Tool Prior to Statistical Validation

During the data collection stage, "empathy", "respect", "positive focus", and "active listening" were theoretical names as the variables were raw, untested, and their construct validity had not yet been statistically Goldberg (2011) also made empricial validated. Although the chances of the variables accurately measuring the labelled constructs were slim, this could be verified using statistical systems later. The forthcoming analyses explore whether the makeup of the scales was truly reflective of the theorised constructs to compound validated TJ measurement scales.

I considered employing standardised and pre-tested measurement tools, such as an empathy scale from the psychology domain (Batson, 2009; Dymond, 1949; Jolliffe & Farrington, 2006). However, I abandoned this idea to retain values more authentically cited as TJ, which a generic scale from another subject might not capture. For instance, the tested empathy scale in Batson (2009 cited in Neumann, Chan, Boyle, Wang & Westbury, 2015) from socio-psycho research measures empathy by way of the following dimensions: (a) knowing another's emotional and cognitive state; (b) matching the posture or neural response of another; (c) feeling the same as another; (d) projecting oneself into another's situation; (e) imagining how another is feeling and thinking; (f) imagining how one would think and feel in another's situation; (g) feeling distress for the suffering of another; and (h) feeling for another person who is suffering. Clearly some of these would be relevant for measuring therapeutic judging, such as: f, g, h. Indeed, these resembled some of the Goldberg (2011) skills employed as part of the empathy scale in my design: such as, "imagining how one would think and feel in another's situation" (f) is alike to 'magistrates understand what it's like to have drug and/or alcohol problems' (p. 29); "feeling distress for the suffering of another" (g) is alike to 'Magistrates are realistic when we set my goals for next review' (p. 29); and "feeling for another person who is suffering" (h) is alike to 'magistrates are personable ' (p.29). However, some standardised variables were irrelevant to therapeutic judging, such as b and c.

Existing instruments would overlook important problem-solving techniques incorporated into the tool such as: "magistrates are realistic when we set my goals for next review" because this is distinctive to this context. Similarly, whilst the "magistrates understand what it's like to have drug and/or alcohol problems" might resemble Batson (2009) dimension f, it is again "drug and/or alcohol problems" that makes the variable unique to specialist court practice. As such, an empathy scale was too generic for this project and pre-existing scaling systems would

confound the objective of measurement. Thus, I extrapolated values from a TJ manual to legitimise and mobilise TJ and problem-solving court values during measurement, and I sought to design an authentic paradigmatic TJ instrument. As such, designing a new tool was germane, and whilst TJ has broad application, I assumed that it could be invoked as a methodology and its facets were observable.

To carry out analyses to validate the instrument, I ascribed to a pragmatist ontology (Creswell & Plano Clark, 2011; Pound, 1909) by taking a practical approach to judging, courts, and MRC, by assuming that each of these things, and the law more broadly, can yield transformative results to people and on the human condition (Stobbs, 2019). Notably, this ontological approach is traditional for TJ research as the movement is founded in American Legal Realism (Leiter, 2002), which fuses humanism and to the practical effects of the law (Stobbs, 2019). The epistemological approach to knowledge was constructivist, which theorises that belief-systems, human intentions, and interpretation, govern epistemological reality (Stobbs, 2019; Hodson & Hodson, 1998). I interpreted the results of the analyses and their commonalties and themes to devise relevant names for the scales; however, I also realised that there were multiple realities to be discovered (Stobbs, 2019). As such, I note where other researchers may have come to different conclusions, and address that the scales and variable names are open for rebranding, retagging, and reinterpretation. Moreover, the paper ascribes the final view of Vols' articulation of TJ theory, by adopting a social constructivist approach to TJ rooted in a sociological genre (Hodson & Hodson, 1998; Minick, 1987). It advocates both the descriptive and normative aspects by designing a tool to observe the practice of problem-solving courts that may help to advance normative change.

Although PCA and Cronbach's alpha are fundamentally quantitative methods, I sought to extrapolate their qualitative qualities to achieve the aims of the paper. For instance, I extracted the descriptive components of the analysis and did not make inferential (and positivist) claims using the p value (Field, 2013). Most often, PCA is used as preparatory tool to explore datasets before analysts carry out more advanced statistical tests. As such, PCA is (relatively speaking) an uncomplicated and descriptive statistical method that usually leads to more complex analysis afterwards (Dancey & Reidy, 2007). Thus, whilst the proceeding analyses may appear to be statistically dense, the techniques lie within a qualitative genre. As such, quantitative methods were merely a vehicle for achieving qualitative objectives (Stobbs, 2019), namely, to

investigate how therapeutic variables link together to devise new scales through interpretive analysis and epistemological constructivism.

3.3. Principal Component Analysis ("PCA"): Rationale

By calculating how strongly variables load onto components, PCA tests the validity of variables of empirical scaling systems devised on theoretical terms (rather than when employing standardised instruments) (Field, 2013). PCA can therefore help to design and ratify research instruments by auditing the inter-correlations of variables on scales (Costello & Osbourne, 2005). This paper rearranges subvariants from theoretical scales as presented in table D onto a statistically validated tool; the journey and process of doing so will be described in upcoming sections.

Exploratory factor analysis includes both factor analysis and PCA (Costello & Osbourne, 2005; Field, 2013). Both methods explore underlying dimensions of datasets by confirming or overthrowing clusters of variables comprising measurement scales (Brown, 2009; Field, 2013). Whilst there are trends in operationalisation, guidelines are not absolute lending itself to the exploratory and descriptive nature of the techniques. Specialists advise that decisions are tailored to individual datasets for optimal results and these should be based upon analysts' own theoretical choices (Costello & Osborne, 2005). As such, for analysts seeking to carry out more complex inferential statistical analyses afterwards (and ascribing to a positivist philosophy therein), stricter rather than interpretivist analysis of the results is appropriate (Field, 2013).

There are many key differences between factor analysis and PCA. Factor analysis can say more than PCA about underlying variable structures by looking beyond the bounds of the standardisation to reveal latent variables (Costello & Osborne, 2005). Factor analysis is thus better suited to projects that roll out pre-ratified tools, rather than those that are theoretically engineered, because researchers are likely to already know the correlation between variables on scales. PCA was most appropriate in this study because although one could hypothesise how the variables correlated based upon non-empirical literature (Goldberg, 2011), the scales were only theoretical. As there was no guarantee that the variables inter-correlated empirically or that they measured the intended constructs, PCA could help validate the theoretical instrument by confirming and/or overthrowing the proposed combinations of variables (Dancey & Reidy, 2007). Whilst PCA was operationalised to ratify the levied scaling systems

in this study, and to compound TJ measurement scales, I recommend that a future study gathers new data, using the proposed scales, and retests them using confirmatory Factor analysis.

3.4. Component Extraction

Eighteen variables were arranged upon four independent scales: 'empathy' (p. 29), 'respect' (p. 34), 'positive focus' (p. 38), and 'active listening' (p. 35) (Goldberg, 2011); however, their compositions were extrapolated from a non-empirical judicial manual. Once the data from MRC had been collected, it was inputted into Statistical Package for Social Sciences ("SPSS"), and PCA was run on the eighteen variables to explore their inter-correlations. The first step was extraction: a process of deciding which (and how many) principal components should be retained (Field, 2013; Costello & Osborne, 2005). Each principal component demonstrates underlying structures in the data by demonstrating how variables weigh onto linear combinations (Field, 2013; Costello & Osborne, 2005) (demonstrated in more detail in table B later). The scree plot determines this by mapping eigenvalues (variable) has on each eigenvector (component). Identified by the point of inflexion (i.e., the bend point), retainable components include any points above (but not including) the break (Field, 2013; Costello & Osborne, 2005).
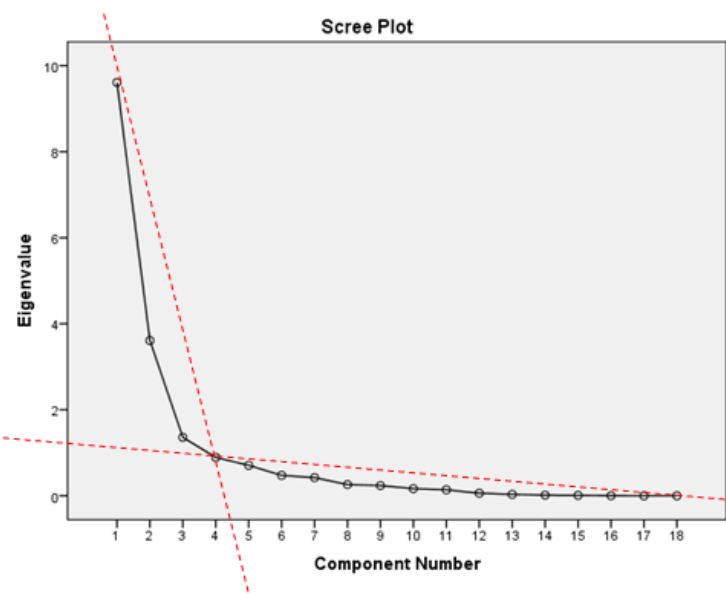

Figure A: Scree Plot Mapping Eigenvalues for Each Eigenvector

As plot bends at point four, this suggests there are three principal components. By default, SPSS uses Kaiser's (1974) standard by retaining eigenvalues greater than 1; however, this has been critiqued for being too strict and overestimating components (Jolliffe, 1972; Costello & Osborne, 2004). For clarity, the scree plot was rerun using a lower cut-off, but the results were identical to Figure A. For further clarity, the plot was rerun by manually setting component extraction at two, four, and five (Costello & Osborne, 2005) but the scree plots were again identical. As such, there was reason to believe that three principal components should be extracted overthrowing the hypothesis that there were four underlying constructs (empathy, respect, positive focus, and active listening) based upon Goldberg (2011). Table B was developed to understand percentages of variance explained by each of the three Principal Components.

	Percentage of Variance
1	53.391
2	20.055
3	7.548

Table B: Demonstrating the Importance of the Components in Percentages

Table B suggests that the three components together explain eighty percent of the total variance and that the final twenty percent is explained by the remaining less prominent components. It is not uncommon to consider solutions accounting for sixty percent of the total variance (and in some instances even less) (Williams, Brown & Onsman, 2012). Therefore, table B explains high amount of variance, which confirms the eminence of three principal components.

3.5. Component Rotation

Following extraction of three components, rotation was processed, defined as: 'any of several methods… by which the researcher attempts to relate the calculated factors to theoretical entities' (Vogt, 1993: 91). The pattern matrix is used to calculate the weight each eigenvalue (variable) has on each eigenvector (component); a larger loading indicates greater importance to that component (Field, 2013). Rotation thus retains variables with larger eigenvalues whilst ignoring smaller ones, although conclusions can be arbitrary (Field, 2013). The purpose of

rotation should be to achieve a simplified restructure in which components have few high loadings with the rest close to zero (Field, 2013).

There are two types of rotation methods: orthogonal and oblique (Vogt, 1993; Field, 2013; Brown, 2009). The overarching difference is that orthogonal rotation assumes components are uncorrelated whereas oblique rotation allows components to correlate (Costello & Osborne, 2005). It has been argued that orthogonal rotation methods should be disregarded for any data involving people because human and social constructs are inevitably highly interrelated (Costello & Osborne, 2005). As such, I rationalised that an oblique rotation method (Direct Oblimin) was appropriate because variables referred to the same social phenomenon: judicial interactional and behavioural skills. The degree to which the factors could correlate was left at default Delta Zero (Field, 2013).

3.6. Loadings

The pattern matrix delineates each variable's loading on each component following rotation (Costello & Osborne, 2005); higher loadings indicate a stronger relationship to the given component (Field, 2013). Negative loadings also indicate a strong relationship in the opposite direction (Dancey & Reidy, 2007). The pattern matrix was generated; loadings of below .32 (in either positive or negative direction) were suppressed as they possessed too little meaning to any component (Costello & Osborne, 2005).

4. Analysis

4.1 Pattern Matrix: Findings

Pattern Matrix			
	Component		
	1	2	3
Q6 (E) Magistrates are interested in interested and compassionate about my life circumstances	.907 (c1)		

Q25 (E) Magistrates understand what it's like to have drug and/or alcohol problems	**.889 (c1)**		
Q11 (PF) Magistrates have hope and faith that I will make progress	**.863 (c1)**		
Q9 (AL) Magistrates listen to my point of view	**.858 (c1)**		
Q8 (PF) Magistrates praise me when I am doing well	**.843 (c1)**		
Q26 (E) Magistrates are personable	**.788 (c1)**	.531	
Q4 (R) Magistrates speak to me without pity or disdain	-.650	**.506 (c2)**	
Q21 (R) Magistrates reiterate my goals so that they are clear	**.581 (c1)**		-.523
Q5 (PF) Magistrates motivate me	**.578 (c1)**		-.344
Q24 (AL) Magistrates give me a voice		**.975 (c2)**	
Q27 (R) Magistrates speak to me slowly, clearly, and loudly		**.948 (c2)**	
Q14 (R) Magistrates are sincere and honest when they speak to me		**.688 (c2)**	
Q28 (R) Magistrates do not rush or interrupt me when I am speaking		**.656 (c2)**	-.538
Q23 (AL) Magistrates are attentive when I am speaking		**.614 (c2)**	
Q17 (AL) I can ask Magistrates questions where I need to			**-.915 (c3)**
Q19 (E) Magistrates are realistic when we set my goals for next review			**-.827 (c3)**

Q20 (PF) Magistrates help me build upon my strengths			**-.732 (c3)**
Q12 (PF) Magistrates make me feel positive about my future	.406	.327	**-.473 (c3)**
Extraction Method: Principal Component Analysis. Rotation Method: Oblimin with Kaiser Normalization.			
a. Rotation converged in 11 iterations.			

Table C: Summary of Principal Component Analysis Results for Interactional Items on Questionnaire.

The pattern matrix in table C observes the following results:

- The analysis displays prominence of three Principal Components; this subsumes the *a priori* theoretical assumption that there were four based on Goldberg's (2011) judicial training manual.
- The original tags ("E", "R", "AL" and "PF" to represent Empathy, Respect, Active Listening and Positive Focus) were left in place disclose the scales upon which sub-variates were originally affiliated with. These were scattered relatively evenly across the three principal components.
- There are six variables that cross-load (4, 21, 26, 5, 28, and 12) onto more than one component. In doing so, they explain something about each component.
- The first two components load positively, and the third component loads negatively.
- Component one has the most variables (8), decreasing for component two (to 6) then further for component three (to 4). This is to be expected as the prominence of components reduce (Field, 2013).
- Question 4 cross-loads most problematically, loading highest (but negatively at -.650) on component one, but grouping better with component two (for a consistent positive signage within a single component). As such, it was tagged as component two for consistency of signage amongst variables.

- Question 12 loaded onto all three components; it was tagged as part of component three due to a higher loading in view of this. Similarly, the other cross-loading items (26, 21, 5, and 28) were grouped with the component on which it loaded highest.

- -.473 is the overall lowest loading, as part of component three; this is typical of the last item on the final component (Field, 2013).

- Component numbers "C1", "C2", and "C3"[2] were added to in brackets each variable to delineate which component that they were best tagged as.

4.2. Correlations between Variables

After analysing the patterns matrix above, components were further considered in the correlation matrix (Appendix A). Solutions have no real meaning if they are not high enough; if items are measuring the same underlying dimension, one would expect them to correlate with one-another (Field, 2013; Costello & Osbourne, 2005).[3] Field posits that analysts should determine this by scanning the correlation matrix and excluding variables with 'lots of' correlations of below 'about 0.3', whilst critiquing his method for being loose, arbitrary, and subjective (Field, 2013: 685). However, this critique lends itself to the fundamental nature of PCA as an exploratory, descriptive, and, qualitative tool, which does not claim to offer clear-cut answers (Field, 2013; Costello & Osbourne, 2005). The correlation matrix (Appendix A) revealed Question 4 potentially correlated too low, although a stricter analyst may also exclude 23, 24, and 27. Question 4 had already been flagged as problematic by the pattern matrix (table C) for cross-loading highly onto two components. The implication is this that stricter positivist analyses may justify their exclusion. However, given the aims and efforts of this paper, and its social constructivist rationalise, they were retained as part of the analysis.

4.3. Data Strength: Loadings and Communalities

There are no strict rules on sample sizes for PCA; rather the appropriateness rests on other factors (Costello & Osborne, 2005). Again, this is due to the exploratory nature of PCA (Field, 2013) where accepted guidelines depend upon philosophical positioning, the subject under investigation, and the purpose of the analysis (Field, 2013). Field (2014) suggests three hundred

---

[2] Tags were added as Component One, Two, and Three.
[3] High correlations are without issue in PCA and so were not considered here (Field, 2014).

cases could be appropriate whilst Dancey & Reidy (2007) recommend one hundred. According to Stevens (2002), larger samples (of one thousand plus) mean that small component loadings (.162) are statistically meaningful; however, smaller samples (of 50), necessitate larger loadings (of .722). Despite this, sample size is deemed secondary to the importance of data strength; this is determined by analysing two factors, component loadings and communalities, following extraction (Field, 2013; Costello & Osborne, 2005). Loadings and communalities will be considered in the upcoming sections, where it will be argued that strength of the current dataset was strong, allowing the tests to be run rigorously.

Like sample size, different studies and specialists posit different methods for decoding data strength (Field, 2013). Whilst noting its rare achievement, Costello & Osborne (2005) define strong data as satisfying three criteria. Firstly, there must be uniformly high communalities above .4 that do not cross-load above .5. Data in this study had consistently high communalities above .4 (see Appendix B); however, some items cross-loaded above .5 (Questions 4, 21, 28) (table C). Secondly, several variables must load strongly on each component (at .32 minimum) (Tabachnick & Fidell, 2001); in this study, every variable loaded strongly onto its component. Thirdly, components must have no fewer than three items; all components had at least three items. Whilst cross-loadings within criterion one weakened the data strength, this is considered unavoidable given the heavily connected nature of the variables (Field, 2013). As such, the dataset passed the Costello threshold, wherein a loose fit is acceptable (Costello & Osborne, 2005).

Field's (2014) guidelines for data strength are stricter than Costello & Osborne (2005); he rationalises that as the number of variables per component increases, loading sizes can decrease (Field, 2013). He posits that any component with four or more loadings above .6 are reliable regardless of sample size, a test that this study would pass if component three was altogether removed (Guadagnoli & Velicier, 1987). On the other hand, components with a few low loadings should not be interpreted unless the sample size is above 300, which the dataset accomplished despite the sample size. As communalities become lower, importance of sample size increases; if all communalities are above .6 small samples are adequate (Field, 2013), which this dataset passed but for Question 23. When dropping Question 23 only (extracted at .476 in Appendix B), communalities in this dataset are strong enough to pass the Field (2014) test. Nevertheless, it passes the Costello & Osborne (2005) test with Question 23 included as it loads highly onto component two (at .614) thereby minimising complications.

This section has discussed requisite sample size for PCA, and its relationship to data strength (based upon loadings and communalities). Discussions have emphasised that whilst larger samples may have strengthened findings, data in this study was strong enough to carry out PCA. The imprecise nature of the statistical guidelines reflects PCA's purpose as an exploratory tool and non-inferential statistical method rather than a strict data-testing mechanism. Standards should be interpreted loosely, and how strict analysts choose rests on theoretical choices. Lending itself to the more qualitative positioning of this paper and the malleable nature of TJ itself, this paper could duly accept the more subjective test of Costello demonstrating that this dataset is strong. For Costello & Osbourne (2005), the main statistical problems would be three cross-loading items above .5, which may undermine data strength (Questions 4, 21, 28). However, cross-loadings were inevitable given the interrelatedness of the constructs within the tested social phenomenon. These variables were retained but flagged as potentially dubious; retention helped to uphold the nuance and integrity of the qualitative dataset. As such, the theoretical, substantive, and methodological frameworks justified inclusion of all items.

4.4.   Proposing New Scales in Therapeutic Jurisprudence

Based on the PCA findings, items were recalibrated from Goldberg's (2011) original four scales (empathy, respect, active listening, and positive focus) onto three new scales and renamed based on their shared underpinning idea. For the rationales posed within the upcoming subheadings, they were named: "Harnessing Therapeutic Support", "Engaging Therapeutic Dialogue", and "Inspiring Therapeutic Change". Their titles were justified using the TJ court-craft series authored by problem-solving court magistrates on Spencer's TJ (https://mainstreamtj.wordpress.com/about/) blog. The verbs, to "harness", "engage", and "inspire", were added to capture the proactivity inherent within TJ courtroom problem-solving (Winick & Wexler, 2003). This process oversaw the proposal of statistically validated, empirical scales to measure judicial panels' interactional and behavioural skills in problem-solving courts. This original methodological contribution to TJ is thought to have wide impact and ambitions given the international calls for development of such instruments (Cooper, 2019; Stobbs, 2019). Although positing these scales breaks new ground for TJ methodology, I do not claim they are perfect, and will offer critique in the forthcoming sections.

Harnessing Therapeutic Support	Engaging Therapeutic Dialogue	*(Debilitating)* Inspiring Therapeutic Change
Q6 (E) Magistrates are interested in and compassionate towards my life circumstances	Q24 (AL) Magistrates give me a voice	Q17 (AL) I can*no*t ask questions when I need to
Q25 (E) Magistrates understand what it's like to have drug and/or alcohol problems	Q27 (R) Magistrates speak to me slowly, clearly, and loudly	Q19 (E) Magistrates **are not** realistic when we set my goals for next review
Q11 (PF) Magistrates have hope and faith that I will make progress	Q28 (R) Magistrates do not rush or interrupt me when I am speaking	Q20 (PF) Magistrates **do not** help me build upon my strengths
Q9 (AL) Magistrates listen to my point of view	Q14 (R) Magistrates are sincere and honest when they speak to me	Q12 (PF) Magistrates **do not** make me feel positive about my future
Q8 (PF) Magistrates praise me when I am doing well	Q4 (R) Magistrates speak to me without pity or disdain	
Q26 (E) Magistrates are personable	Q23 (AL) Magistrates are attentive when I am speaking	
Q21 (R) Magistrates reiterate my goals so that they are clear		
Q5 (PF) Magistrates motivate me		

Table D: A Tool for Measuring TJ Values in Problem-Solving Court Settings

4.5. Component One: Harnessing Therapeutic Support

King (2016a) discusses therapeutic support in TJ: 'supporting is acknowledging and identifying with a person's situation'. Support is theorised to possess the following facets, which bear striking resemblance to variables on component one. King's (2016a) examples are

provided for in the footnotes to demonstrate semblances to the variables on component one (tables C and D).

- 'Empathising'[4]. Further, Berkley University (2020) recently posited that 'emotion researchers generally define empathy as the ability to sense other people's emotions, coupled with the ability to imagine what someone else might be thinking or feeling'. This mirrors the following variables: "magistrates are interested in and compassionate about my life circumstances"; "magistrates are personable" and; "magistrates understand what it's like to have drug and/or alcohol problems".

- 'Agreement'[5]: "magistrates listen to my point of view".

- 'Offers to help'[6]: "magistrates reiterate my goals so that they are clear".

- 'Praise'[7]: "magistrates praise me when I am doing well".

- 'Reassurance'[8] : "magistrates have hope and faith that I will make progress" and; "magistrates motivate me".

The links between King's (2016a) work and the correlations of variables suggest that Component One can be appropriately entitled: "Therapeutic Support". Interestingly, the variables from the 'empathising' dimension within the first bullet were all originally empathy variables from the *a priori* theoretical scale; this buttressed the restructure following PCA and implies empathy is an underlying dimension of 'support' (King, 2016a). Perhaps the weakest fit is 'magistrates motivate me' as a sub-component of reassurance; although the example is inherently motivational – "I think you will achieve your goal of staying off drugs" – I have argued elsewhere that motivation was its own standalone dimension at MRC (BLIND, 2020). By loading unproblematically onto component one (Table B), motivation may also be a key ingredient for therapeutic support and could be added to King's 2016a bench book as an original contribution to practice. Previously, Question 21, "magistrates reiterate my goals so that they are clear", was flagged as undermining data strength; this might reflect poor wording (Field, 2013) and, thus, other researchers might wish change it to something clearer. Examples for

---

[4] 'I understand why you would be upset about using'.
[5] 'You did the right thing by telling me you have relapsed'.
[6] 'How about we reschedule your next court appearance so you can attend your son's graduation?'.
[7] 'You did a fantastic job in staying off drugs for three months!'.
[8] 'Given your enthusiasm, commitment and excellent relapse prevention plan, I think you will achieve your goal of staying off drugs'.

rewording might include: "magistrates make sure my goals for the next court session are clear" or "I come away from reviews clear on my goals for next time".

Overall, as variables from component one imbricated King's (2016a) characterisation of therapeutic 'support', it was recalibrated to: "Harnessing Therapeutic Support". However, I encourage rebranding (if and as seen fit) by future and fellow TJ colleagues. I emphasise that this is an open-minded process, whereby alternative suggestions from peers are welcome.

4.6.   Component Two: Engaging Therapeutic Dialogue

Component two is interesting because variables each tap into the notion of speech. This affirms the rigour of PCA by validating that variables loading onto component two measured the same underlying construct. This justified naming component two as: "Therapeutic Dialogue'.

According to King (2016b), problem-solving judging involves a unique form of communication: 'a dialogue between judicial officers and participants that aims to promote particular common goals.' This has four main dimensions, of which one is turn-taking: giving space, encouragement, and supporting service-users' autonomy during decision-making (King, 2016b). Component two is strongly indicative of a turn-taking dialogue: "magistrates give me a voice"; "magistrates do not rush or interrupt me when I am speaking". Turn-taking also involves well-developed judicial listening skills (King, 2016b), elucidated by the following variable: "magistrates are attentive when I am speaking". Two further aspects of effective communication are commitment to the person and commitment to the message (Adler & Proctor, 2007) The former is 'caring about what one says and being sincere' (King, 2016b), which links to existing variables: "magistrates are sincere and honest when they speak to me" and "magistrates speak to me without pity or disdain ". The latter includes 'use of language that makes sense to the other person' (King, 2016b), exemplified by the variable: "magistrates speak to me slowly, clearly, and loudly". Interestingly, four of the six items were part of the original hypothesised "respect" scale (Goldberg, 2011), which could suggest that respect is fundamental for administering therapeutic dialogue.

Although this component undoubtedly taps into key ideas within King's (2016b) conceptualisation of therapeutic dialogue, the scale does not comprehensively cover its full meaning. I therefore suggest further studies add variables to the scale based upon King (2016b),

followed by revalidation of component structures through confirmatory factor analysis. One variable open to critique is "magistrates speak to me without pity or disdain" (Question 4), which fits less clearly with King's abstraction. This item has already been identified within previous discussion as problematically cross-loading; loading highest (but negatively at -650) on component one, whilst grouping better with component two (due to consistent positive signage on the latter). Moreover, Question 28 also cross-loads onto component three: "magistrates do not rush or interrupt me when I am speaking". Lack of rigour within items 4 and 28 could be symptomatic of poor wording; "without" and "do not" operate negative phraseology, unlike the remaining sixteen variables, which may have misled respondents.

Future researchers might reconsider rewording; Question 4 could become: "magistrates are sympathetic when they speak to me" or "magistrates are compassionate when they speak to me". However, this suggestion is perhaps too like: "Magistrates are interested in and compassionate about my life circumstances" from component one. This similarity could explain why Question 4 also loaded negatively onto component one indicating that "speaking *with* pity or disdain" (i.e., to the effect of sympathy or compassion) is an inherent part of the "Harnessing Therapeutic Support" (component one). This variable also had low reliability according to Cronbach's Alpha (in forthcoming analyses). As such, future empirical researchers in TJ can decide on how they operate this variable; it could be retagged as component one, modified into clearer language, or altogether removed; again, this should be ultimately based on philosophical choices. For the purposes of this paper, it will be retained as part of component two as it promotes a methodology rooted in a qualitative framework.

4.7.   Component Three: Inspiring Therapeutic Change

On the third scale, items loaded negatively, hence its negative name given in table D: "***Debilitating***" Therapeutic Change. Negative lettering (in bold and italics) was added to each variable to reflect negative meaning to that construct. For the proposed scale, the Likert system was reversed to a positive subject: Inspiring Therapeutic Change, and the bold and italics wording of each variable could then be removed to reflect notions operating in the positive. When reversed, the items implied that magistrates were forward-focused within their interaction, attempting to promote positive self-development and therapeutic change during court conversations with service-users.

Recalibration of the scale led to the following name: "Inspiring Therapeutic Change". This was rationalised as follows; variables "setting goals for the next review" and "building upon strengths" resembled King's (2017) notion: 'research suggests achievement is promoted through the setting of goals'. Both Spencer and King put forth strategies for setting positive goals during problem-solving court craft; Spencer (2018) suggests a staircase diagram where reaching an overall goal is broken down into smaller steps, and King (2017) suggests a written protocol in which service users develop a forward-focused change plan alongside the judiciary. Spencer comments that shame and low self-worth can be overcome by setting goals, which is captured by the current variable: "feeling positive about the future". Spencer (2018) further suggests that 'asking questions' refers to individuals' active participation in their goal setting, linking to the variable: "I can ask questions when I need to". However, this could be accentuated through the wording of future questionnaires. Overall, the close fit with the expertise of the magistrates indicated that component three had measured the construct: "Therapeutic Change".

5. Reliability Analysis

Once the scale compositions had been validated through PCA, and restructured and renamed accordingly, reliability testing on the same could be performed. Within empirical analysis, the term reliability broadly relates to consistency of measurement during later retests (Field, 2013). However, scales themselves are reliable when their results are internally consistent with the overall questionnaire; if items behave similarly across the data, the scale is reliable (Field, 2013). Thus, theoretically, the eight items on the support scale should behave similarly; the same applies to the further two scales. Even if three components had been incorrectly relabelled, reliability tests can still disclose whether they consistently reflect the same measured construct, whatever that construct might be.

Reliability testing examines the consistency of measurements through inter-scale reliability checks (Field, 2013). Cronbach's Alpha is the most widely used test; by splitting scales into two random halves, scales are reliable if scores from both halves correlate highly across several data entries (Field, 2013). Since there are many ways scales could be randomly split, Cronbach (1951) created an average value for every possible split, equating to Alpha (α) (Field, 2013). Whilst statisticians debate acceptable cut-off points to determine Alpha, most posit a threshold

value of between .7 and .8, although this may vary according to the project at hand (Kline 2009).

Alpha was performed on each of the three scales (or principal components) and the results are displayed in the following tables. For each output, the "Corrected Item Total Correlation" score refers to the extent that the item correlates with the overall scale; if this is below .3, discarding the item is recommended (Field, 2013). The "If Item Deleted" column reflects the change to Alpha if the item is dropped; as such, an item scoring greater than the overall Alpha decreases reliability of the scale (Field, 2013).

5.1.   Cronbach's Alpha: Results

Scale 1

Cronbach's Alpha	N of Items
.957	8

Table E: Displaying results of Cronbach's Alpha from the Harnessing Therapeutic Support Scale

	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Q6 Magistrates are interested and compassionate in and about my life circumstances	.870	.948
Q5 Magistrates motivate me	.763	.958
Q8 Magistrates praise me when I am doing well	.923	.945
Q9 Magistrates listen to my point of view	.943	.944

Q11 Magistrates have hope and faith that I will make progress	.877	.948
Q21 Magistrates reiterate my goals so that they are clear	.833	.951
Q25 Magistrates understand what it's like to have drug and alcohol problems	.877	.948
Q26 Magistrates are personable	.647	.961

Table F: Displaying results of Cronbach's Alpha from the Harnessing Therapeutic Support Scale broken down into items

5.2.   Reliability Findings: Harnessing Therapeutic Support

The overall Alpha of .957 indicates a highly reliable scale (Field, 2013). This could be criticised for being potentially too high, indicating redundancy of items as variates measure the same thing as the construct itself. However, as numbers of scale items increase, Alpha tends to increase without changing internal consistency or undermining reliability (Field, 2013). Therefore, high Alpha for this scale could reflect the larger number of variables comprising this scale, compared to the remaining two (Field, 2013). The "Corrected Item" indicates how much each item correlates with the overall questionnaire score; less than .3 indicates that the item may not belong on that scale. As these were not below .3, items correlated strongly with one another to indicate high internal consistency, and, thus, reliability. As Question 26 is above Alpha, this decreases overall reliability of the scale. Future researchers could remove that item.

Scale 2

Cronbach's Alpha	N of Items
.865	6

Table G: Displaying results of Cronbach's Alpha from the Engaging Therapeutic Dialogue Scale

	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Q24 Magistrates give me a voice	.825	.812
Q27 Magistrates speak to me slowly, clearly, and loudly	.924	.791
Q28 Magistrates do not rush or interrupt me when I am speaking	.747	.827
Q14 Magistrates are sincere and honest when they speak to me	.747	.827
Q4 Magistrates speak to me without pity or disdain	.227	.905
Q23 Magistrates are attentive when I am speaking	.529	.869

Table H: Displaying results of Cronbach's Alpha from the Engaging Therapeutic Dialogue Scale broken down into items

5.3. Reliability Findings: Engaging Therapeutic Dialogue

The overall Alpha of .865 indicates a highly reliable scale (Field, 2013). "Corrected Item" scores were not below .3, apart from Question 4; this shows strong internal consistency overall. Questions 4 and 23 are above Alpha, decreasing reliability of the scale. Both this analysis and former PCA show that Questions 4 and 23 decrease validity and reliability of this scale; this should be accounted for within further study; a change of wording could be beneficial.

Scale 3

Cronbach's Alpha	N of Items
.895	4

Table I: Displaying results of Cronbach's Alpha from the Inspiring Therapeutic Change Scale

	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Q17 I can ask magistrates questions where I need to	.817	.875
Q19 Magistrates are realistic when setting my goals for next review	.816	.847
Q20 Magistrates help me build upon my strengths	.767	.866
Q12 Magistrates make me feel positive about my future	.772	.870

Table J: Displaying results of Cronbach's Alpha from the Inspiring Therapeutic Change Scale broken down into items

5.4. Reliability Findings: Inspiring Therapeutic Change

The overall Alpha of .895 indicates a highly reliable scale (Field, 2013). As no "Corrected Item" scores were below .3, items correlated strongly with one another. No questions are above the overall reliability score, suggesting that there are no reliability problems with the therapeutic change scale.

5.5. Summary of Alpha

Overall, Cronbach's Alpha shows that the three scales extracted from PCA were largely reliable, and any flaws have been disclosed alongside recommendations for future study. Whilst some variables undermined reliability, deletion only marginally increased reliability, and Alpha would still be within the reliable range when included (Nunnally, 1978); future researchers can decide whether to retain these items depending on their philosophical positioning. By way of comparison, Alpha was carried out on the previous four theoretical

scales, where reliability was found to be significantly lower. Therefore, verifying scaling structures through PCA increased both validity and reliability.

6. Summary Discussion

PCA can be employed as a technique to help to design and ratify research instruments by auditing the inter-correlations of variables on scales (Costello & Osbourne, 2005). This paper rearranged subvariants from four theoretical scales onto a statistically validated tool. Cronbach's alpha then demonstrated that the reliability of the scales was high.

However, there are acknowledged critiques for doing this. Perhaps shoehorning TJ into empirical measurement systems belies its purpose. TJ is a broad philosophy where no tool could adequately capture its full scope (Stobbs, 2019). As shown within earlier sections of this paper, TJ's amorphous nature, and wide range of values, means that it has potential to be malleable into various ontological forms and into many fields. Therefore, any empirical scales proposing to measure it will be insufficient. When studying problem-solving courts, complimentary analyses are therefore recommended, and full guidance can be found in Stobbs et al., (2019a).

There is also far more substance to therapeutic courts than just the judicial interaction (Ashworth et al., 2004); they face a whole host of other problems when offenders approach them, for instance: socio-economic, mental health, housing, broken relationships, physical health. As such, TJ analysis 'need(s) to take into account the big picture, as well as snapshots at various points in time' (Cooper, 2019: 292). This suggests that therapeutic responses by way of these factors cannot be captured through crude measurements, like those proposed. Whilst I accept these critiques, the following point must be emphasised: if the purpose of TJ is to accentuate the therapeutic and/or countertherapeutic aspects of law processes and legal practice, the tool did (and can) achieve exactly that; it provided insight into the therapeutic quality of the interactional and behavioural styles of MRC's judiciary (BLIND, 2020). In this sense, the instrument fulfilled TJ's purpose at MRC.

Notably, the tool could be used for other legal contexts where there is an interaction between practitioner and individual – perhaps solicitor and client, or during police interviewing. The problem-solving court model is simply a prototype, and further advancement and expansion is

recommended in line with TJ goals for broad standardisation (Cooper, 2019). Here the phraseology could be changed from "magistrates" to "solicitors", or even from third to first person for practitioners seeking to carry out self-assessment. Similarly, questions can be adapted; in this study, they were used for standardised observations, then surveys, but also latterly as a deductive theory-driven map to help analyse interview data using NVivo. As such, it could have use in qualitative and quantitative studies. I have already rolled the tool out to study another English problem-solving court using the standardised court observation method. As well as endeavouring to report the results in future research, I seek to confirm the component structures through confirmatory factor analysis.

It is important to acknowledge the small sample size in this paper. The setting in this study, MRC, had issues with service-user attendance, which made collecting large amounts of survey data difficult. In terms of the impact upon these analyses, discussions in previous sections demonstrated that the data was strong enough despite this, and the small sample is unlikely to have had huge impacts on the validity of results. With that being said, although the analysis strongly indicated three components, structures displayed in Table B might only represent the sample population, and not a broader population; equally, a larger sample might confirm and verify structures. There could also be some variations linked to subjectivism if future studies use the tool because researchers and participants might interpret the variables differently to those from this study. Moreover, different jurisdictions might vary in results due to the idiosyncrasy of the contexts, legal systems, and processes. These factors should be born in mind. The task now is now reimplementation of the tool, generating new study and results, and reflecting on any differences. Whilst this might be a process that requires further thought, it is thoughtful process that would not be possible without standardisation.

7. Conclusion

This paper sought to explore the empirical measurability of a therapeutic interactional and behavioural style, a widely asked and engaged question, currently leading TJ discussion. I collected original data from an English problem-solving court using theoretical scaling systems. I was aware that in bypassing ratified and standardised empirical scales, and devising an instrument modelled upon non-empirical literature, the scales may fail to accurately capture the intended constructs. PCA could test this, although as an exploratory tool, it does not pretend to provide clear-cut rules or exact answers.

The scree plot revealed that the eighteen variables measured three principal components; this overthrew the original assumption that there were four constructs based upon Goldberg (2011). The components were extracted, and an oblique method was used to rotate the variables because they corresponded to the same subjective phenomenon. The pattern matrix facilitated the interpretation of loadings; Question 4 was flagged as undermining the data strength by cross-loading onto components one and two. Although it loaded higher on component one, it was tagged as part of component two for consistent positive signage. However, like most decisions taken when carrying out PCA, this choice was arbitrary, and other analysts may have tagged it differently. There were six variables that cross-loaded onto more than one component, thus explaining something important about each. Whilst it is good to be critical and remember that cross-loadings undermine the quality of findings, it is also important to note that loadings are never going to be clear-cut when items broadly refer to the same intangible and subjective social science phenomena (Field, 2013).

Following validity checks, reliability of three new scales were considered using Cronbach's Alpha; overall, reliability of each scale was reported as high. Questions 26, 4, and 23 decreased reliability of their respective scales by increasing Alpha when deleted. However, deletion would increase reliability only marginally, leaving Alpha still within the reliable range (Nunnally, 1978). They were therefore retained although whether future researchers make similar decisions depends upon their own theoretical choices; positivist analyses might exclude them, whilst constructivist coding structures might justify inclusion.

The original scales were based upon a manual for training Judges in TJ communication styles. For this reason, all variables were entitled: "judicial interactional and behavioural skills", which courtroom specialist, King (2006: 92), defines as: the 'ability to listen and communicate effectively with an ethic of care and the ability to motivate others to consider positive change'. These skills are a fundamental tenet of TJ and are consistently and broadly discussed throughout the TJ literature. However, perhaps they better reflect therapeutic alliance, interpersonal styles, sources of a therapeutic relationship, or something else: an uncertainty that stems from the intangibility and subjective nature of social science variables. Whatever that "something" is, the analyses have indicated high validity and reliability of the measurement systems used to measure it.

This paper represents the first validated attempt to develop empirical scales for measuring interactional and behavioural styles of problem-solving court judges. This is a positive step for TJ, as it represents the doctrine infiltrating new terrains, and responds to some of the outstanding critiques that it is undertheorized. Although the paper breaks new ground in this area, I have stressed that this is an open-minded process that welcomes suggestions by future and fellow TJ researchers. The scales are now due for adaption, modification, and development within further empirical pilots.

Acknowledgements

---

[9] 'In the creating of BLIND analytic scales, BLIND crafted three categories: "Harnessing Therapeutic Support", "Engaging Therapeutic Dialogue", and "Inspiring Therapeutic Change". I can say this without fear of contradiction, as someone who has been studying problem-solving courts for twenty years: These are categories that should be used, going forward, by all scholars who analyze these courts. I don't think anyone else could do better. As she focuses on these, she stresses the significance of empathy (as an underlying dimension of support), I wrote in the margins when I read this first, "This is so valuable!", and on second reading, I have not changed my mind at all!'.

	Q4	Q5	Q6	Q8	Q9	Q11	Q12	Q14	Q17	Q19	Q20	Q21	Q23	Q24	Q25	Q26	Q27	Q28
Q4	1.000	-.414	-.595	-.519	.566	-.453	-.199	.104	-.227	-.363	-.184	-.505	-.015	.406	-.461	-.208	.417	.104
Q5	-.414	1	.685	.796	.787	.557	.765	.521	.600	.632	.682	.684	.359	.219	.708	.515	.246	.412
Q6	-.595	.685	1	.772	.918	.898	.590	.479	.345	.552	.491	.822	.123	-.0552	.740	.485	-.059	.216
Q8	-.519	.796	.772	1	.917	.793	.754	.465	.602	.575	.736	.739	.332	.202	.853	.779	.180	.465
Q9	-.566	.787	.918	.917	1	.870	.709	.469	.523	.642	.652	.818	.259	.045	.816	.564	.015	.339
Q11	-.453	.557	.898	.793	.870	1	.632	.451	.465	.580	.597	.856	.136	.023	.845	.602	.008	.314

Q12	-.199	.765	.590	.754	.709	.632	1.0000	.620	.616	.817	.700	.639	.593	.361	.627	.550	.388	.620
Q14	.104	.521	.479	.465	.469	.451	.620	1	.466	.514	.481	.440	.537	.633	.309	.510	.783	.742
Q17	-.227	.600	.345	.602	.523	.465	.616	.466	1	.740	.897	.683	.250	.178	.511	.327	.365	.706
Q19	-.363	.632	.552	.575	.642	.580	.817	.514	.740	1	.669	.698	.400	.061	.465	.236	.216	.514
Q20	-.184	.682	.491	.736	.652	.597	.700	.481	.897	.669	1	.730	.285	.156	.654	.480	.307	.642
Q21	-.505	.684	.822	.739	.818	.856	.639	.440	.683	.698	.730	1	.130	-.142	.747	.438	.000	.440
Q23	-.015	.359	.123	.332	.259	.136	.593	.537	.250	.400	.285	.130	1	.535	.226	.352	.540	.537
Q24	.406	.219	-.0552	.202	.045	.023	.361	.633	.178	.061	.156	-.142	.535	1	.180	.484	.887	.633

Q25	-.4661	.708	.740	.853	.816	.845	.627	.309	.511	.465	.654	.747	.226	.180	1	.693	.059	.309
Q26	-.2088	.515	.485	.779	.564	.602	.550	.510	.327	.236	.480	.438	.352	.484	.693	1	.467	.510
Q27	.4177	.246	-.0559	.180	.015	.008	.388	.783	.365	.216	.307	.000	.540	.887	.059	.467	1	.783

Appendix B: Communality Table

	Extraction
Q4 Magistrates speak to me without pity or disdain	.623
Q5 Magistrates motivate me	.711
Q6 Magistrates are interested in and compassionate about my life circumstances	.858
Q8 Magistrates praise me when I am doing well	.910
Q9 Magistrates listen to my point of view	.923
Q11 Magistrates have hope and faith that I will make progress	.839
Q12 Magistrates make me feel positive about my future	.791
Q14 Magistrates are sincere and honest when they speak to me	.755
Q17 I can ask magistrates questions where I need to	.871
Q19 Magistrates are realistic when setting my goals for next review	.812
Q20 Magistrates help me build upon my own strengths	.811
Q21 Magistrates reiterate my goals so that they are clear	.873
Q23 Magistrates are attentive when I am speaking	*.476*
Q24 Magistrates give me a voice	.906

Q25 Magistrates understand what it's like to have drug and/or alcohol problems	.827
Q26 Magistrates are personable	.802
Q27 Magistrates speak to me slowly, clearly, and loudly	.944
Q28 Magistrates do not rush or interrupt me when I am speaking	.848

Reference list

Adler, R. B., Proctor II, R. R., and Towne, N. (2007). *Looking Out, Looking In.* Nelson Education.

Ashcroft, J., Daniels D.J, & Herraiz, D.S. (2004). *Defining drug Courts: The Key Components*. U.S. Department of Justice, National Association of Drug Court Professionals. Retrieved from: https://www.ncjrs.gov/pdffiles1/bja/205621.pdf

Batson, C. D. (2009) These things called empathy: Eight related but distinct phenomena. In J. Decety & W. Ickes (Eds.), *Social neuroscience. The social neuroscience of empathy* (p. 3–15). MIT Press. https://doi.org/10.7551/mitpress/9780262012973.003.0002

Bauer, D. J. (2017). A more general model for testing measurement invariance and differential item functioning. *Psychological Methods, 22(3),* 507.

Berkley University, The Greater Good Science Center. (2020). What is Empathy? Retrieved from: https://greatergood.berkeley.edu/topic/empathy/definition

Birgden, A. & Ward, T. (2003). Pragmatic Psychology Through A Therapeutic Jurisprudence Lens Psycholegal Soft Spots in the Crimical Justice System *Psychology, Public Policy, and Law 2003, Vol. 9, No. 3/4, 334–360*

BLIND (2020) Problem-Solving Courts, Criminal Justice and the International Gold Standard: A therapeutic jurisprudence perspective. London, Routledge.

Brown, J. D. (2009). Principal components analysis and exploratory factor analysis - definitions, differences, and choices. *Statistics, 13(1),* January 2009, 26 - 30

Clark, M.D. (2002) A Change-Focused Approach for Judges. In D.B Wexler & Winick, B. J. *Judging in a therapeutic key: Therapeutic jurisprudence and the courts* (pp. 137-147) (2003). Durham, N.C: Carolina Academic Press.

Cooper, C. (2017, December 22). Re: TJ approach measure [online discussion group]. Retrieved from: https://tinyurl.com/y2bos4v5

Cooper, C. (2019) Evaluating the application of TJ principles: lessons from the drug court experience. In N. Stobbs, L. Bartels, & M Vols (eds), *The Methodology and Practice of Therapeutic Jurisprudence* (pp. 287-204). Durham, N.C: Carolina Academic Press.

Costello, A. B., and Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation, 10(7),* 1-9.

Cramer, R. & Vols, M. (2016). Hoarding disorder and the legal system: A comparative analysis of South African and Dutch law. *International Journal of Law and Psychiatry Volume 49, Part A, November–December 2016, Pages 114-123.*
Creswell, J.W. and Plano Clark, V.L. (2011). *Designing and Conducting Mixed Methods Research.* 2nd edition, Sage, Thousand Oaks, CA.

Criminal Justice Act 2003, Section 210.
Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16(3),* 297-334.

Daicoff, S. & Wexler, D.B (2003) Therapeutic jurisprudence. In A. M. Goldstein (Ed.) *Handbook of psychology: Forensic psychology*, Vol. 11 (p. 561–580). John Wiley & Sons Inc.
Dancey, C. P., and Reidy, J. (2007). *Statistics without Maths for Psychology*. Pearson Education.

Dymond, R.F. (1949) A scale for the measurement of empathic ability. *J Consult Psychol. 1949 Apr;13(2):127-33.* DOI: 10.1037/h0061728

Field, A. (2013). *Discovering statistics using IBM SPSS statistics.* Sage.

Flick, U., (2011). I*ntroducing research methodology: A beginner's guide to doing a research project.* London: Sage.

Gal, T. & Wexler, D.B. (2015). Synergizing therapeutic jurisprudence and positive criminology. In N. Ronel and D. Segev (Eds), *Positive Criminology* (pp. 138-151). London: Routledge.

Glover, M. (2012). The Therapeutic Function of Testamentary Formality. *61 U. Kan. L. Rev. 139.*

Goldberg, S. (2011). *Problem-solving in Canada's courtrooms: A guide to therapeutic justice.* National Judicial Institute for Canada.

Guadagnoli, E., and Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin, 103(2),* 265 -275. doi: 10.1037/0033-2909.103.2.265

Harrison J.L. Class, (1994). personality, contract, and unconscionability. *William and Mary Law Review, 35. 445-501.*

Harwin, J., Alrouh, B., Ryan, M., & Tunnard, J. (2014). *Changing lifestyles, keeping children safe: an evaluation of the first Family Drug and Alcohol Court (FDAC) in care proceedings*. **London: Brunel University.**

Hodson, D. & Hodson, J. (1998). From Constructivism to Social Constructivism: A Vygotskian Perspective on Teaching and Learning Science. *School Science Review*, v79 n289 p33-41

Hopkins, A. & Bartels, L. (2019). Paying attention to the person: compassion, equality and therapeutic jurisprudence. In N. Stobbs, L. Bartels, & M Vols (eds), *The Methodology and Practice of Therapeutic Jurisprudence* (pp. 107-129). Durham, N.C: Carolina Academic Press.

Hora, F. P. (2002). a dozen years of drug treatment courts: Uncovering our theoretical foundation and the construction of a mainstream paradigm. *Substance use and Misuse, 37(12-13),* 1469-1488. doi:10.1081/JA-120014419

Hora, P. F. (2011). Courting new solutions using problem-solving justice: Key components, guiding principles, strategies, responses, models, approaches, blueprints and tool kits. *Chapman Journal of Criminal Justice 2 (1),* 7-52

Hora, P., and Schma, W.Rosenthal., J. (1999). Therapeutic jurisprudence and the drug treatment court movement: Revolutionizing the criminal justice System's response to drug abuse and crime in America. *Notre Dame Law Review, 74(2),* 439-527.

Hudson, C. (2017, September 18). Re: working on a way to measure legal actor contributions [online discussion group]. Retrieved from: <https://tinyurl.com/y3jc6ha6>

Joliffe, D. & Farrington, D.P. (2006) Development and validation of the Basic Empathy Scale. *J Adolesc. 2006 Aug;29(4):589-611*. DOI: 10.1016/j.adolescence.2005.08.010

Jolliffe, I. T. (1972). Discarding variables in a principal component analysis. *I: Artificial data. Applied Statistics, 21 (2),* 160-173

Justice Innovation Centre (2018) *Problem solving courts in Manchester.* Retrieved from: https://justiceinnovation.org/articles/problem-solving-courts-manchester
Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39(1),* 31-36.

Kerr, J., Tompkins, C., Tomaszewski, W., Dickens, S., Grimshaw, R., Wright, N. and Barnard, M. (2011). The dedicated drug courts pilot evaluation process study. *Ministry of Justice Research Series, 1 (11).* London: Ministry of Justice.

King, M. (2006). The therapeutic dimension of judging: The example of sentencing. *Journal of Judicial Administration, 16,* 92- 105.

King, M. (2016a, July 16). Better judicial conversations [Blog post]. Retrieved from https://mainstreamtj.wordpress.com/2016/07/16/better-judicial-conversations-tj-court-craft-series-4/

King, M. (2016b, June 3). *Enhancing judicial communication* [Blog post]. Retrieved from https://mainstreamtj.wordpress.com/2016/06/03/enhancing-judicial-communication-tj-court-craft-series-1/

King, M. (2017, February 28). *A judicial officer assists offenders to set rehabilitation goals and strategies* [Blog post]. Retrieved from https://mainstreamtj.wordpress.com/2017/02/28/a-judicial-officer-assists-offenders-to-set-rehabilitation-goals-strategies-tj-court-craft-series-8/

King, M., Freiberg, A. M., Batagol, B. M., & Hyams, R. L. (2014). Non-Adversarial Justice. (2nd ed.) Annandale NSW: Federationa Press.

Kline, P. (1999). *The Handbook of Psychological Testing*. Routledge, London.

Koch, H. & Diesen, C. (2016) Therapeutic jurisprudence-continuous quality management and positivity in the civil justice system in Sweden and UK. *MOJ Public Health;5(1):37–39. DOI: 10.15406/mojph.2017.05.00118*

KPMG Consulting. (2014). *Evaluation of the drug court of Victoria*. Government Advisory Services. Final Report. Magistrates' Court of Victoria, 18.

La Prairie, C., Gliksman, L., Erickson, P.G., Wall, R., Newton-Taylor, B. (2002) Drug treatment courts--a viable option for Canada? Sentencing issues and preliminary findings from the Toronto court. *Subst Use Misuse. 2002;37(12-13):1529-66.*

Leiter, B. (2002). American Legal Realism. *U of Texas Law, Public Law Research Paper No. 42.* DOI: http://dx.doi.org/10.2139/ssrn.339562

Lens, V., Katz, C. C., & Suarez, K. S. (2016). Case Workers in Family Court: A Therapeutic Jurisprudence Analysis. Children and Youth Services Review, 68, 107-114. https://doi.org/10.1016/j.childyouth.2016.06.027

Lynch, A., and Perlin, M. L. (2016). 'Life's Hurried Tangled Road': A Therapeutic Jurisprudence Analysis of Why Dedicated Counsel Must Be Assigned to Represent Persons with Mental Disabilities in Community Settings. NYLS Legal Studies Research Paper No. 2833365. Doi: http://dx.doi.org/10.2139/ssrn.2833365

Marson, J. Ferris, K. & Kawalek, A. (2019a) A Vineyard in a Law Clinic: The Practical Application of a Therapeutic Jurisprudence Philosophy in a UK Law Clinic *Singapore Journal of Legal Studies, Mar 19, 124-146.*

Marson, J. Ferris, K. & Kawalek, A. (2019b) Applying Refugee Family Reunion Law Therapeutically. *Beijing Law Review, 10(5), 1172-1196*

McIvor, G. (2009). Therapeutic jurisprudence and procedural justice in Scottish drug courts. *Criminology and Criminal Justice, 9(1),* 29-49.

Minick, N. (1987). The development of Vygotsky's thought: An introduction. In R. W. Rieber and A. S. Carton (Eds.), The collected works of L. S. Vygotsky. Vol. 1. (N. Minick, Trans.). New York: Plenum

National Institute of Justice. (2012, May 1). Overview of Drug Courts. <https://nij.ojp.gov/topics/articles/overview-drug-courts>

Neumann, D. L., Chan, R. C. K., Boyle, G. J., Wang, Y., & Westbury, H. R. (2015). Measures of empathy: Self-report, behavioral, and neuroscientific approaches. In G. J. Boyle, D. H. Saklofske, & G. Matthews (Eds.), *Measures of personality and social psychological constructs* (p. 257–289). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-386915-9.00010-3

Nunnally, J. (1978). *Psychometric methods*. McGraw-Hill.

Perlin, M. & Lynch, A (2016). Had to be held down by big police: A therapeutic jurisprudence perspective on interactions between police and persons with mental disabilities. *Fordham Urban Law Journal, 43*, 685.

Perlin, M. L. (2018, July 11). Re: Blog post re measuring therapeutic contribution of legal actors in a courtroom [online discussion group]. Retrieved from: https://tinyurl.com/y5bj5roe

Perlin, M. L. (2012). 'Justice's Beautiful Face': Bob Sadoff and the Redemptive Promise of Therapeutic Jurisprudence. *Journal of Psychiatry and Law, 265.*

Perlin, M. L. (2013a). *A prescription for dignity: Rethinking criminal justice and mental disability law.* Routledge.

Perlin, P. (2013b). The judge, he cast his robe aside: Mental health courts, dignity and due process. *Mental Health Law and Policy Journal, 3 (1). 1–29*

Perlin, M. L. (2013c). 'There are no trials inside the gates of Eden': Mental health courts, the convention on the rights of persons with disabilities, dignity, and the promise of therapeutic jurisprudence (pp. 193). In B McSherry, and I Freckelton. (Eds) *Coercive care: Law and policy.* Abingdon, Oxon: Routledge.

Perlin, M. L. (2017). 'Have You Seen Dignity?': The Story of the Development of Therapeutic Jurisprudence. *U.N.Z. Law Review, 1135, 27.*

Perlin, M. L. (2019) Dignity and Therapeutic Jurisprudence: How We Can Best End Shame and Humiliation. In C Chowdhury, M. Britto & Hartling, L. *human dignity: practices, discourses, and transformations* (113). Lake Oswego: Human Dignity Press. Available at SSRN: https://ssrn.com/abstract=3502587 or http://dx.doi.org/10.2139/ssrn.3502587

Petrucci, C. (2002). Respect as a component in the judge-defendant interaction in a specialized domestic violence court that utilizes therapeutic jurisprudence. *Criminal Law Bulletin, 28 (2).* 263-295. doi:https://ssrn.com/abstract=2263908

Pound, R. (1909). Liberty of Contract. *Yale Law Journal. (18)7, 454-487*

Richardson, L. (2019) social control theory in mental health courts; reflections on writing a "TJ" PhD. In N. Stobbs, L. Bartels, & M Vols (eds), *The Methodology and Practice of Therapeutic Jurisprudence* (pp. 305-326). Durham, N.C: Carolina Academic Press.

Roderick, D. & Krumholz, S. (2006) Much Ado About Nothing? A Critical Examination of Therapeutic Jurisprudence *University of Massachusetts Law Review: Vol. 1: Iss.1 Article 9.*
Shuman, D.W. (1992). Therapeutic Jurisprudence and Tort Law: A Limited Subjective Standard of Care. *46 S.M.U. L. Rev. 409.*

Slobogin, C. (1995). Therapeutic jurisprudence: Five dilemmas to ponder. **Psychology, Public Policy, and Law, 1***(1), 193–219.* DOI: https://doi.org/10.1037/1076-8971.1.1.193

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Hillsdale, NS: Erlbaum.

Stobbs, N. (2019). Therapeutic Jurisprudence as theoretical and applied research. In N. Stobbs, L. Bartels, & M Vols (eds), *The Methodology and Practice of Therapeutic Jurisprudence* (pp. 29-58). Durham, N.C: Carolina Academic Press.

Stobbs, N., Bartels, L., & Vols, M. (2019) *The Methodology and Practice of Therapeutic Jurisprudence*. Durham, N.C: Carolina Academic Press.

Stobbs, N., Bartels, L., & Vols, M. (2019) Therapeutic Jurisprudence – a strong community and maturing discipline. In N. Stobbs, L. Bartels, & M Vols (eds), *The Methodology and Practice of Therapeutic Jurisprudence* (pp. 15-28). Durham, N.C: Carolina Academic Press.

Tabachnick, B. G., and Fidell, L. S. (2001). *Computer-assisted research design and analysis*. Allyn and Bacon Boston.

Vogt, W. (1993). *Dictionary of statistics and methodology (vol. 1).*

Vols, M. (2019) theory and methodology of therapeutic jurisprudence. In N. Stobbs, L. Bartels, & M Vols (eds), *The Methodology and Practice of Therapeutic Jurisprudence* (pp. 59-82). Durham, N.C: Carolina Academic Press.

Warren, R.K. (2001) Public Trust and Procedural Justice. In D.B Wexler & Winick, B. J. *Judging in a therapeutic key: Therapeutic jurisprudence and the courts* (pp. 132-136) (2003). Durham, N.C: Carolina Academic Press

Wexler, D. B., and Winick, B. J. (1991). Therapeutic jurisprudence as a new approach to mental health law policy analysis and research. *University of Miami Law Review, 45(5),* 979-1004.

Wexler, D.B. (1992). Putting mental health into mental health law. *Law Hum Behav 16, 27–38* https://doi.org/10.1007/BF02351047

Wexler, D.B. (1993) New Directions in Therapeutic Jurisprudence: Breaking the Bounds of Conventional Mental Health Law Scholarship. *N.Y.L. SCH. J. RTs. 759*

Wexler, D. B. (1995). Reflections on the scope of therapeutic jurisprudence. *Psychology, Public Policy, and Law, 1(1*), 220.

Wexler, D. B., and Winick, B. J. (1999). *Law in a therapeutic key: Developments in therapeutic jurisprudence.* Carolina Academic Press, Durham, North Carolina. doi: http://www.austlii.edu.au/au/journals/CICrimJust/1997/9.pdf

Wexler, D. B. (2015). Wine and bottles: A metaphor and a methodology for mainstreaming TJ. *Arizona Legal Studies Discussion Paper No. 15 (05*). doi: https://ssrn.com/abstract=2553868

Wexler, D. B. (2017, December 11). Re: TJ approach measure---maybe expand inquiry to court observations and recommendations for improved judicial practice and better TAL Translation of Spanish instrument to English. Retrieved from: https://tinyurl.com/yyk9jfpg

William, B., Brown, T. & Onsman, A. (2012) Exploratory Factor Analysis: A Five-Step Guide for Novices. *Australasian Journal of Paramedicine, 8.*

Winick, B. (2005) *Civil Commitment: A Therapeutic Jurisprudence Model*. Carolina Academic Press.

Winick, B. J., & Wexler, D. B. (2003). *Judging in a therapeutic key: Therapeutic jurisprudence and the courts*. Durham, N.C: Carolina Academic Press.

Winick, B.J. (2003) Therapeutic jurisprudence and problem solving court. *Fordham Urban Law Journal, 30 532-569*