An age-old issue: Evaluating the applicability of adult criminal risk assessment tools for use with youth offenders

By

Anna M. Ferguson

A thesis
submitted to Victoria University of Wellington
in fulfilment of the requirements for the degree of
Master of Science
in Forensic Psychology

Victoria University of Wellington
2015

Abstract

Research has consistently recognized that youth and adults share risk factors for crime, although whether certain factors are of increased importance during adolescence is debated. The present research evaluated the extent to which two risk assessment tools could predict criminal and breach reconviction in a matched sample of youth (aged 17-19) and adult (aged 20-60) community-supervised offenders: The Dynamic Risk Assessment for Offender Re-entry (DRAOR) and the static Risk of re-Conviction X Risk of re-Imprisonment (RoC*RoI). Cox regression and AUC analyses revealed initial DRAOR scores had mixed predictive validity for both groups, while proximal risk scores showed comparably moderate to high accuracy for youth and adults. Protective scores were consistently poor predictors for adults. The proximal assessment predicted reconviction better than the initial assessment, and decreases in risk scores between assessments were associated with a reduction in the likelihood of reconviction, showing the value in monitoring risk and updating assessment. The RoC*RoI predicted criminal reconviction for adults but did not predict either reconviction outcome for youth. These findings support the use of the DRAOR for identifying which youth and adults are likely to reoffend, and suggest that dynamic factors might be more useful predictors than static for assessing and monitoring youth offenders.

Acknowledgments

I would first like to thank my primary supervisor Dr Clare-Ann Fortune and secondary supervisor Professor Devon Polaschek for their encouragement, guidance, and invaluable feedback. Thank you for giving me the opportunity to conduct research in an area that I am really interested in. I feel I have learned a lot over the years I have spent in contact with you both, and through being a part of the criminal justice lab. Thanks also to the lab for their support throughout the year, ensuring I set goals and achieved them. I am also grateful to the Department of Corrections for supplying the data for my thesis.

Thank you to Gabriel Braniff and Jessica Scanlan for thought-provoking discussions and friendship throughout the past two years. Thank you to Damien Huynen for your constant moral support and encouragement both with my research and in life. Thank you Julie, Bernie and Bianca for Monday nights and for listening to me talk about my research. And big thanks to my family: Mom, Dad, Ian and James, who have always been there for me, encouraged me to aim high, and who I love and appreciate so much.

Finally, a big thank you to everyone who allowed me to discuss, explain, and (at times) complain about my research; I could not have done it without you.

Table of Contents

Chapter 4

Chapter 5

List of Tables

List of Figures

List of Appendices

Chapter 1

Introduction

Consider the case of Chris who is serving a community supervision sentence after committing burglary. He has 10 previous convictions. While serving his sentence, he continues to associate with antisocial friends. Despite suffering from occasional bouts of irritability, he has high expectations for his future and shows a good understanding of the costs his offending has had for himself and others. How likely is Chris to reoffend? If you assessed Chris using a standard risk assessment tool, do you think the likelihood of his reoffending would change depending on whether he was 17 or 30 years old? Research suggests that it might.

Despite youth and adult offenders sharing similar major risk factors for reoffending (e.g., antisocial attitudes, history of antisocial behaviour; Andrews & Bonta, 2010), certain risk factors may be better predictors during different time periods in an individual's life (Hoge, Vincent, & Guy, 2012). For example, during adolescence, associating with antisocial peers assumes greater importance as a predictor of antisocial behaviour (Borum, 2003). Meanwhile factors such as drug and alcohol abuse decrease as predictors, due to the increased prevalence of these behaviours among all adolescents, not simply those engaging in crime (Hoge et al., 2012). Further complicating risk assessment are the conflicting opinions about when adolescence ends and adulthood begins (see Blakemore, Burnett, & Dahl, 2010; Males, 2009; Moshman, 2011; Scott & Steinberg, 2008).

To address the issue that certain risk factors may have a stronger, or weaker, relationship with reoffending depending on age, separate assessment tools for youth and adults have been developed and are currently used. Presently however, the limited research on risk assessment during the period of late adolescence has yielded mixed results when it comes to which tools are most useful during this time. The present research aims to bring some clarity to this issue through an investigation of whether two different criminal risk assessment tools designed and validated on adult offenders can predict reconviction just as effectively for older youth offenders (aged 17-19), compared to adult offenders (aged 20-60), serving community supervision sentences.

Defining 'Youth'

"The notion that a single line can be drawn between adolescence and adulthood for different purposes under the law is at odds with developmental science."

— Steinberg, Cauffman, Woolard, Graham, & Banich, 2009, p. 583

Age is one of the most reliable correlates of crime (Indig, Frewen, & Moore, 2014). Studies have consistently shown a greater number of youth are engaging in antisocial behaviour during adolescence. The prevalence of crime increases during this period, peaking around age 17 and then declining until about age 25; a phenomenon known as the age-crime curve (Moffitt, 1993). Youth offenders are not only more likely to reoffend, but also have a shorter time frame between their offences (Indig et al., 2014).

In the literature, the term *youth* is used somewhat liberally to refer to a broad range of ages; generally clustered around the period of adolescence.[1] Even the age criterion set by the United Nations is broad, defining a youth as anyone between the ages of 15-24. Achieving a universal definition is hampered by differing legal definitions of youth set by various jurisdictions/institutions as well as the overall lack of consensus about the age at which adolescent development ends and adulthood begins.

Legal Definition of Youth

For the purposes of sentencing and punishment, the line must be drawn somewhere. The New Zealand Government's Children, Young Persons and Their Families Act 1989 defines a young person as "a boy or girl of or over the age of 14 years but under 17 years…" (section 2). Thus, young offenders up to 16 years of age fall under the jurisdiction of the youth justice system. Once an offender turns 17, they are transferred to the adult justice system, and their sentences are managed by the New Zealand Department of Corrections. Yet, the Department of Corrections categorises any offenders under the age of 20 as youth, despite treating them "in the same manner as adults" (Ministry of Justice, n.d., definition - child and youth). Offenders aged between 17 and 19 fall into a grey area

[1] As such, I endeavour to use each author's own words wherever possible when describing the age range used in studies. Therefore, youth aged 12-25 are described collectively as youth, with those aged 16 and under usually described as younger youth, and those aged upwards of 17 usually described as older youth, young adults, or as being in the period of late adolescence or early adulthood.

in the overlapping definitions used by both systems.

But despite being treated as such, age is far from being a fixed and consistent maturity marker for everyone. One youth aged 18 can have a very different grasp of the consequences of their actions than another of the same age, and youth are not only shaped by the quantity of time spent alive, but also the quality of that time: early socialisation experiences, and the environment they live in. Thus, the late teenage and early adulthood years could involve a “delayed outgrowing of adolescent behaviours” (Loeber & Farrington, 2012, p. 5). Some researchers (see Scott & Steinberg, 2008) have even argued that, rather than view youth and adults as two distinct categories, there should be a separate legal category for adolescents. This category would include youth in their late teenage and early adulthood years who do not yet possess the ability to think and consider consequences like adults, but have passed the point at which they are able to have their crimes pardoned on account of their age.

Developmental Definition of Youth

Arguments supporting the idea that youth are distinct from adults largely centre on the idea of immature judgment. That is, whether youth possess the capacity to reason and comprehend their behaviour and its long-term consequences, and, if so, this necessarily translates into them using these abilities constructively (Scott & Steinberg, 2008). Developmental psychology has mostly focused on child development while, to date, comparatively less attention has been given to adolescence: the ill defined “period of physical and psychological development between childhood and adulthood” (Blakemore et al., 2010, p. 926). Research attempting to better define this period has identified three common themes that make adolescence distinct from both childhood and adulthood: it is a time of increased brain maturation, particular social pressures, and unstable personality traits.

Brain maturation. While one side of the debate supports the notion that youth have different brain functioning to adults, opponents of this view believe adolescents show similar reasoning capabilities to adults and there is “simply no empirical basis” on which the two groups should be categorised separately (Moshman, 2011, p. 171). These opponents criticise neurological research into adolescence as too vague, alleging it is particularly hard to pinpoint exactly when and how changes in the brain occur (Moshman, 2009).

Indeed, from a developmental standpoint, there is no clear-cut point at which the transition from youth to adulthood is made (Farrington, Loeber, & Howell, 2012). Research into neurological development has suggested that during adolescence certain brain regions crucial to behavioural control, in particular the frontal lobe (believed to be responsible for decision-making processes and inhibiting impulses) are not yet fully developed (Blakemore & Choudhury, 2006). Furthermore, the brain does not mature completely until an individual is in their mid-twenties (Scott & Steinberg, 2008). Even into late adolescence brain regions responsible for crucial functions such as the ability to weigh up the likely costs and benefits of one's actions continue to develop (Scott & Steinberg, 2008). Farrington et al. (2012) have suggested older youth aged 18-24 have similar cognitive functioning to youth aged 15-17. Despite having reached intellectual maturity, older youth may still lack psychosocial maturity (required for such tasks as resisting pressure from peers) and self-regulation capability (required to make prosocial decisions) when faced with the prospect of engaging in antisocial behaviour (Cauffman & Steinberg, 2000; Steinberg et al., 2009).

Social pressures. During adolescence and early adulthood, individuals are particularly vulnerable to social pressures. Adolescence is a time during which many young people start to distance themselves from their parents, and approval and acceptance by peers assumes greater importance (Hoge et al., 2012). Peer relationships can provide a source of prosocial support for youth, or they can support and reinforce antisocial behaviours (Dahlberg, 1998). In order to please their peers, and because they have trouble anticipating the consequences of their actions, youth may place greater importance on the immediate rewards gleaned from engaging in risky and antisocial behaviour, more so than most adults (Scott & Steinberg, 2008).

Unstable personality traits. While many youth engage in antisocial and/or criminal activities to feel part of a group or assert their independence, adults are more inclined to do so due to their personal preferences (Scott & Steinberg, 2008). This idea implies an element of fluidity to most antisocial behaviour in youth, painting it as opportunistic and perhaps a means of asserting their independence. While extroversion and openness to experience peak in the late teenage years, increased development of traits like conscientiousness and emotional stability do not occur until the early to mid twenties (Roberts, Walton, & Viechtbauer, 2006). In recognition of the fact that personality is unstable during adolescence, clinicians are hesitant to label evidence of callousness-unemotional traits in youth as being indicative of psychopathy (Hoge et al., 2012). Furthermore, a diagnosis of antisocial personality disorder (ASPD) requires that an individual be at least 18 years old (American Psychiatric Association, 2013). It follows that if personality traits in youth are still developing, there is still opportunity for change, meaning traits that might make youth more willing to engage in antisocial behaviour will not necessarily persist or manifest in the same way in adulthood.

Summary. Research suggests that even older youth are fundamentally different to adults because the interaction between their underdeveloped brains, unstable personality traits and social pressures relating to the period of life they are in puts them at a greater likelihood of making poor decisions. Taken together, these three features of adolescence suggest youth are more susceptible to engaging in antisocial behaviour than most adults.

Do these recognised cognitive and social differences between youth and adults necessarily translate into different risk factors for crime? That depends. While most factors are shared, some operate at varying strengths for youth. I now discuss shared and unique risk factors for youth offending within the context of adult risk assessment.

Youth and Adult Risk Factors: Similarities and Differences

Andrews and Bonta (2010) have identified the *Central Eight* risk factors, theorised to be universal predictors of antisocial behaviour. The Central Eight comprises four of the strongest correlates of crime (history of antisocial behaviour, antisocial associates, antisocial attitudes and antisocial personality), and four moderate correlates (family/marital relationships, education/employment, leisure activities and substance abuse). An additional shared risk factor for crime is gender. Being male is associated with a higher likelihood of criminal conduct compared to being female (Andrews & Bonta, 2010).

Although the Central Eight risk factors are theorised to apply across age and gender, the strength of certain risk factors varies by developmental stage. Having a history of antisocial behaviour is a risk factor relevant for all ages, while individual characteristics (e.g., temperament), family and school factors are important predictors for younger youth, and education/employment, peer associations and substance use are stronger predictors for older youth (Borum, 2003). Because the focus of the present research is older youth, following a discussion of the risk factor of history of antisocial behaviour, I will focus on two risk factors pertinent to older youth in further detail: antisocial peers, and substance use/abuse.

History of Antisocial Behaviour

One of the best methods to identify risk factors for persistent offending is to use longitudinal studies, which track the same groups of people over time. This method shows how the propensity for crime develops over the course of an offender's life and what risk factors might be central to the onset, maintenance and persistence of antisocial behaviour. Assessing the same participants regularly and thoroughly can help explain why some people offend and others do not.

One such study is the *From a Boy to a Man* study conducted in Finland (Sourander et al., 2006). Researchers tracked a cohort of boys born in 1981, assessing them on a number of measures at age eight (to measure conduct problems, family structure, education, health problems, etc.) and then through to young adulthood in order to investigate early predictors of later negative outcomes (e.g., antisocial behaviour). One study using this data examined offences committed by youth between the ages of 17-20 (Sourander et al., 2006). A plurality of risk factors characterised the childhoods of those who committed five or more crimes during these years. In particular, 43% of these offenders came from broken homes and 82% had parents with a low level of education, with many offenders themselves also achieving poorly at school. However, the strongest predictors of both number of offences and all crime types were childhood conduct and hyperactivity problems; the more conduct disordered symptoms a child showed at age eight the greater number of crimes committed between the ages of 17-20 (Sourander et al., 2006).

In research generated from a similar longitudinal study, the *Christchurch Health and Development* study (where a cohort of individuals born in Christchurch, New Zealand was followed from birth), conduct problems in childhood corresponded not only with higher rates of criminal convictions and imprisonments but also with mental health issues, substance use and risky sexual behaviour in early adulthood (defined as ages 21-25; Fergusson, Horwood, & Ridder, 2005). As with the *From a Boy to a Man* study (Sourander et al., 2006), Fergusson et al. (2005) also found a dose-response relationship, where the severity of conduct problems in childhood predicted more adverse outcomes in early adulthood.

Both longitudinal studies provide evidence that it is possible to identify those at risk of future offending by examining the extent of childhood conduct problems. Both adults and youth who continue to offend are likely to present with similar antisocial backgrounds stretching back into childhood (Andrews & Bonta, 2010). The earlier the onset of the antisocial behaviour, and the greater the scope of problems, the more likely an individual is to persistently offend (Loeber & Farrington, 2000). This is partly due to the accumulating consequences of antisocial behaviour: early negative developmental experiences (e.g., difficult temperament, poor parenting) may preclude opportunities to learn and practice prosocial behaviour, thereby preventing a youth from recognising and seizing opportunities for change if and when they arise (Farrington & Loeber, 2002; Moffitt, 1993).

Antisocial Peer Associations

Although antisocial peers are a risk factor for both adults and youth, it has been suggested that youth are more susceptible to the influences of their peers who, as mentioned previously, have increased importance during adolescence (Stouthamer-Loeber, Wei, Loeber, & Masten, 2004).

Scott & Steinberg (2008) explain that social comparison and conformity indirectly and directly underpin peer influence respectively. Having moved away from the influence of their parents, youth look to their peers as models, and are likely to behave in a way that fits with what their peers are doing, eager not to be the odd one out. Youth who were antisocial children, and grew up with inadequate parenting and supervision, might finally find a sense of belonging in an antisocial peer group or gang full of similar others that share their positive beliefs about crime. Youth whose antisocial behaviour originated in, and is limited to, adolescence may look to antisocial peers as role models, viewing their rule-breaking, risky behaviour as a means of asserting independence (Andrews & Bonta, 2010).

Antisocial peers can both directly encourage criminal conduct, but also may moderate the expression of other risk factors. For example, peers may support and reinforce antisociality, viewing antisocial attitudes and behaviour as desirable traits (Stouthamer-Loeber et al., 2004). As mentioned, antisocial attitudes (that is, positive attitudes about criminal conduct) are one of the Central Eight risk factors for crime.

Although youth and adults may show similar patterns of risk taking when alone, youth take comparatively more risks than adults while in the presence of friends and peers (Scott & Steinberg, 2008). Therefore, the risk factor of antisocial peers is likely to be a stronger predictor of crime during adolescence compared to adulthood.

Substance (Ab)use

The term substance *use* is often used interchangeably with substance *abuse* in the literature, but the general consensus is that persistent, problematic use makes it a risk factor for crime (Andrews & Bonta, 2010). However, although present at higher levels in offender populations than in the general population, substance abuse tends to be a poor predictor, on its own, of who is likely to reoffend. Substance abuse is instead likely to operate as a factor that serves to increase the potential for violence or antisocial behaviour via its effect on other predictors (e.g., by exacerbating levels of arousal and/or impulsivity, leading to violence). Pullmann (2010) found older youth (aged 16-25) who had been involved with both the mental health and justice system, and had been diagnosed with a substance use disorder, were more likely to have a subsequent criminal conviction for violence or drugs (but not property offences), suggesting that substance use increases aggression and propensity for violent behaviour.

Some authors (see Hoge et al., 2012) have even argued that substance abuse actually decreases as a predictor of crime during adolescence, when using substances is somewhat normative and peers encourage experimentation. For example, because alcohol is widely used during this period, alcohol abuse is a weaker predictor of youth offending than using illegal drugs (Indig et al., 2014).

Therefore, substance (ab)use is a risk factor that, although often co-occurring with other risk factors for both youth and adults, arguably has a more complex relationship with youth than adult offending. Because adolescence affords greater opportunity for experimentation with substances, substance abuse may not be directly related to youth offending but instead serve to aggravate other risk factors.

Interaction Between Factors

As alluded to previously, it is hard to tease out the effect that individual factors might be having on the likelihood of youth offending independent of others, as factors appear to be best able to explain offending when considered in combination. As an example, poor parenting and negative early socialisation experiences might lead a youth to then fall in with a peer group that fosters antisocial attitudes and encourages frequent drug use. Peer pressure may cause a youth to drop out of school, leaving plenty of free time to instead spend in disadvantaged neighbourhoods where acts of violence are commonly witnessed and copied.

The method by which risk factors are routinely assessed often takes this into account, providing a cumulative evaluation, rather than simply focusing on the individual effort of single predictors. I turn now to a discussion of risk assessment, its evolution over time, and current practice.

Generations, Factors, and Current Practice in Risk Assessment

Risk assessment involves the identification of historical, situational, and individual factors that increase an individual's likelihood of reoffending (recidivism). Risk assessment is based on the premise that the more risk factors an individual has, and the greater their severity, the greater the individual's likelihood of engaging in criminal activity (de Ruiter & Nicholls, 2011). The word likelihood is used, as it is possible to have an abundance of risk factors, but not offend. Albeit challenging, risk assessment is crucial to criminal justice systems, informing multiple decisions such as community release, and the level of rehabilitation an offender receives (Olver, Stockdale, & Wong, 2012). Accurate risk assessment is important, as an inaccurate assessment not only jeopardises public safety but also potentially the rights of an offender.

Risk assessment has evolved over the years in response to the need for a more rigorous and empirically based approach. Four generations of risk assessment have been identified to date. The earliest assessment methods, termed the *first generation*, involved an unstructured assessment of risk made by a clinician, with the expectation that a clinician's own expertise and experience working with offenders might give them some success in distinguishing reoffenders from desisters. This intuitive approach was prone to errors of overconfidence, hindsight bias, and the failure to account for the base rate of certain crimes; meaning unreliable assessments were made (Perrault, Paiva-Sailsbury, & Vincent, 2012). After it was accepted that structure was needed, actuarial risk assessment tools were developed. Actuarial tools involve assigning numerical values to predictors and using a scoring system to give some indication of whether an offender is at a low, medium or high risk of reoffending

(Douglas & Skeem, 2005). These tools comprise the second and third generations of risk assessment. They consist of scales typically including factors empirically derived from studies of criminal recidivism. Three broadly used and widely accepted factor types are discussed: static, dynamic, and protective factors.

Static Risk

Second generation actuarial tools include solely static risk factors—historical or permanent variables that cannot be changed as a result of treatment or intervention—such as number of previous convictions, or age at first offence (Douglas & Skeem, 2005). The appeal in using them is that they are often consistently strong predictors of recidivism, as past criminal behaviour is a good predictor of future criminal behaviour (Caudy, Durso, & Taxman, 2013). Static factors are usually easily obtained (e.g., from an offender's case file) and do not involve clinical judgment. However, static factors fail to take into account improvements in risk level over time or factors that may be operating to reduce the likelihood of recidivism. Their inability/resistance to change, and their absolute nature has earned them the title of tombstone predictors (Becroft, 2009).

Dynamic Risk

Third generation tools resulted from recognition that assessment needed to be informed not only by historical risk factors but also more immediate individual and situational risk factors. Andrews, Bonta, and Hoge (1990) changed the way scholars conceptualise risk and risk factors with their Risk-Need-Responsivity (RNR) principle. The main premise of the RNR is that to effectively rehabilitate offenders, those who pose the highest *risk* to society should be given more intensive services, while low risk offenders should have little to no rehabilitation. Further, criminogenic *needs*—that is, changeable factors empirically related to recidivism—should be the primary factors assessed and the exclusive targets of treatment. Finally, treatment should be *responsive,* that is, appropriate to an offender's level of comprehension and abilities.

Therefore, concentrating on the need principle, in order to more accurately determine the risk an offender poses to society, only factors shown to reliably predict criminal behaviour should be assessed and treated (Andrews & Bonta, 2010). Dynamic risk factors are amenable to change, and can be further subdivided into *stable* factors that are expected to change slowly (e.g., peer associations or impulse control), and *acute* factors that are expected to change more rapidly (e.g., mood or employment status; Douglas & Skeem, 2005). A benefit of assessing dynamic factors is that their sensitivity to change allows them to serve a unique dual-purpose: to measure and monitor risk on a regular basis, and as treatment targets (Douglas & Skeem, 2005). However, dynamic factors are unable to account for criminal history—mentioned previously as a consistent predictor of criminal behaviour—and without any historical point of reference, using dynamic factors alone may under or overestimate likelihood of recidivism.

Change over time. The value in measuring dynamic risk factors is that they are expected to change over time. It then follows that changes in risk should be associated with an increase or decrease in the likelihood of recidivism, depending on the direction of the change (Olver, Wong, Nicholaichuk, & Gordon, 2007). Studies have examined the predictive validity of dynamic risk factors, using scores taken from a single assessment, offering purely a snapshot of dynamic risk (Douglas & Skeem, 2005). In this way, risk measured is a static measure of dynamic risk and there is little way of knowing whether the factors change over time; in other words, whether they are truly dynamic. Studies examining whether risk factors are dynamic must use scores from assessments at multiple time points to examine whether any change made between assessments increases or decreases an offender's likelihood of recidivism. Research assessing dynamic change using youth offenders is virtually nonexistent. Much of the contemporary research on dynamic change using multiple time points comes from studies of adult violent and sex offenders. Research has found both evidence for and against the idea that dynamic risk factors change in adults, and that this change is able to predict recidivism. The results of several adult studies measuring change using multiple time points are now discussed.

Schlager and Pacheco (2011) found that primarily violent parolees' total scores on the Level of Service Inventory–Revised (LSI–R; Andrews & Bonta, 1995), a tool with both static and dynamic factors, decreased over time (likely) in response to treatment that targeted the dynamic risk factors identified in assessment. The authors suggested LSI–R scores should not be "viewed in a vacuum" but instead as one part of a series of assessments (Schlager & Pacheco, 2011, p. 550). In other words, it is important individual risk assessment scores are neither viewed nor treated as standalone predictors, but rather are considered within the context of the overall pattern of dynamic risk over time, and whether marked changes have been observed.

Olver et al. (2007) found evidence for the changeability of dynamic items in their study, which assessed primarily adult sex offenders who had undergone high intensity treatment. When controlling for static risk, change in risk scores on the Violence Risk Scale - Sexual Offender Version (VRS-SO, a dynamic tool designed to assess changes in risk for sex offenders; Wong, Olver, Nicholaichuk, & Gordon, 2003), in the direction of a reduction in risk, was associated with a reduction in the likelihood of sexual recidivism for high-risk, but not low-risk offenders. Given this result, it appears particularly important to monitor changes in dynamic risk for high risk offenders, who have more room to make change. In a similar vein, Lewis, Olver, and Wong (2013) also found that, after controlling for pretreatment risk level, change made between pre- and post-treatment scores on the Violence Risk Scale (VRS, a dynamic tool designed to assess change in risk for violence; Wong & Gordon, 2006) in the direction of a reduction in risk, corresponded with a reduction in the likelihood of reconviction over a five year follow-up period. Taken together, these two studies suggest factors on instruments specifically designed to be administered at multiple time points are useful for monitoring purposes not only because they are dynamic and able to measure imminent risk, but because changes occur that are associated with reductions in recidivism.

In contrast, Hanson, Harris, Scott, and Helmus (2007) found that while stable and acute factors on two dynamic tools (measured at different time points) were able to predict recidivism in a sample of adult sex offenders on community supervision, change on these factors was not. Over a six-month period, offenders showed little change on stable items, suggesting stable factors might not need to be updated regularly unless an assessor notices marked changes. Contrary to expectation, Hanson et al. (2007) also found acute factors, expected to be the most changeable by nature, might actually be capturing enduring personality characteristics rather than the immediate risk they are intended to. Although the acute assessment closest to reoffence was a good predictor of recidivism, an average of six months of assessments prior to reoffence was a better predictor. Averaging scores captures consistency in behaviour (e.g., chronic irritability), rather than what could be exclusive features of a single assessment (e.g., displaying heightened anger right before an offence). Hanson et al. (2007)

found that the highest predictive accuracy came from using both static and dynamic tools in combination, with scores taken from more recent dynamic assessments performing better than ones earlier in time to the offence, showing the value in updating risk assessments regularly.

Combined, these studies illustrate that dynamic factors are best examined using multiple time points. Although results are mixed as to whether dynamic factors actually change, and whether change alters the likelihood of recidivism, updated assessments appear to perform better than those conducted earlier in time, and predictive accuracy improves when dynamic factors are used in addition to static.

Protective Factors and Desistance

Given that dynamic change over time has been linked with change in the likelihood of recidivism, perhaps this indicates that more than simply being correlates of crime, certain dynamic risk factors are causal; they can be manipulated through treatment leading to reductions in risk. However, focusing solely on risk factors as causal mechanisms negates the perhaps mediational “influence of normative life events” during early adulthood (Hoge et al., 2012, p. 175). Experiences at work or school, changes in the quality of relationships between youth and their parents and/or peers, contact with the correctional system and even internal sources of resilience can all affect the likelihood of recidivism (Hoge et al., 2012).

Protective factors are static or dynamic strengths an offender possesses which promote resilience and may make recidivism less likely (Serin, Mailloux, & Wilson, 2012). Like risk factors, protective factors have also been theorised to operate via a dose-response relationship; the greater the volume of protective factors, the lesser the likelihood of offending (Lösel & Farrington, 2012; Rennie & Dolan, 2010). Factors have been proposed in the individual (e.g., intelligence), family (e.g., close relationship to a parent), school (e.g., positive school climate), neighbourhood (e.g., nonviolent neighbourhood), and peer domains (e.g., nondeviant friends; Lösel & Farrington, 2012). Protective factors can operate in one of two ways: either directly, producing a lower likelihood of reoffending by virtue of being present (e.g., being female); or indirectly, by reducing the effect risk factors might have (Lösel & Farrington, 2012). For example, a romantic relationship might make an individual more inclined to hold down a steady job that provides a legitimate source of income, limiting the need to steal. In this way, a good quality relationship reduces the effect unemployment might have on likelihood of criminal activity, thus being indirectly protective.

Burgeoning research done in the area of protective factors recognises the importance of not simply focusing on an offender's risk, but also on their strengths and potential for desistance. Desistance refers to the cessation of criminal activity (Stouthamer-Loeber et al., 2004). Desistance has been suggested to occur for a variety of reasons ranging from increases in maturity that come with age, to greater opportunities to form prosocial bonds with peers and the community, to an individual's rational choice to disengage from crime, once they are developmentally able to do so (Stouthamer-Loeber et al., 2004). The issue with measuring desistance is that it is a process with no clear end point. For example, a previously chronic offender might not completely desist but his crime rate may drop significantly over time. Is this offender likely to desist completely? Furthermore, is it realistic to expect that he might, given his chronic history of offending? The way risk factors are commonly studied—using short-term follow-ups, and examining one outcome—give limited information about which factors actually predict long-term desistance. Thus, determining whether an offender has completely desisted from crime is difficult, and often unrealistic. Studies measuring desistance often try to have long follow-up lengths, or employ a longitudinal approach in recognition of this difficulty.

Youth desistance. Regardless of age of onset for antisocial behaviour, the majority of desistance occurs in late adolescence and early adulthood, while those who continue to offend past adolescence tend to be mostly life-course persistent individuals (Farrington et al., 2012; Moffitt, 1993; Stouthamer-Loeber et al., 2004). Sampson and Laub (2005) examined the criminal activity of a group of offenders over their entire life course to examine why some adolescents appeared to age-out of crime while others persisted. Sampson and Laub (2005) have theorised that persistence with and desistance from crime depend largely on factors of social control. That is, crime is most likely when one has a weak bond to society. The majority of offenders, therefore, desist due to changes in their social connections – and thus social control – at various stages throughout their lives. Transitions from one life stage to another provide turning points; that is, opportunities to desist from crime (Sampson & Laub, 2005). Being a period of transition, adolescence affords greater opportunities for turning points and greater chances to exercise social control. Sampson and Laub (2005) asked a group of offenders to retrospectively identify turning points in their lives, and found the most common turning points were also associated with transitions in age, a change in routine, and increased opportunity to shift or develop a new identity (e.g., marriage, military service, employment; Sampson & Laub, 2005). These turning points promote desistance “by default” through encouraging long-term conformity, giving individuals some societal investment (Sampson & Laub, 2005, p. 37). Turning points are not only one-time significant events, but can also be events that repeat themselves throughout life, making them dynamic in nature. Sampson and Laub (2005) give the example of marriage, which is not a fixed event but rather can be entered into and out of multiple times, and thus its effect on offending can vary depending on its current quality.

Stouthamer-Loeber et al. (2004) investigated factors theorised to promote desistance from adolescence into early adulthood in a longitudinal sample of individuals ranging in age from 13-25 (as part of the *Pittsburgh Youth Study*). The authors measured self-reported predictors of both persistence and desistance in three age groups, 13-16, 17-19 and 20-25 (Stouthamer-Loeber et al., 2004). They found over half of their sample desisted partially (kept offending at a lower level) or totally between the ages of 19-20. Factors at age 17-19 associated with desistance were low nonphysical aggression, the belief that one is likely to be caught, possessing employable skills, low level of substance abuse among peers, and having positive interactions with interviewers (perhaps indicative of good social skills and the ability to get along with others). Many of these factors also predicted desistance in early adulthood, supporting the idea that 17-19 and 20-25 year olds share protective factors. Once all variables were entered into a regression equation to account for the influence of each other, only one predicted desistance in adulthood: possessing employable skills. Furthermore, once significant predictors of desistance from ages 13-16 were entered into the regression equation, along with significant risk factors from ages 17-19, predictive accuracy improved, suggesting that early adulthood delinquency and desistance can be anticipated using predictors from adolescence (Stouthamer-Loeber et al., 2004). The authors concluded it is a combination of factors from adolescence, as well as coexisting adult risk factors, that help to explain persistence and desistance.

Protective factors, and a focus on how to best promote desistance, have become increasingly more important over the years as evidence has mounted in support of sentences that include a treatment/rehabilitation component, with a focus on building strengths, rather than being purely punitive (Andrews & Bonta, 2010). Risk assessment has also become increasingly strengths-focused in an effort to find alternative explanations as to why two offenders with similar risk factors may not have the same likelihood of reoffending. I turn now to a discussion of current practice in risk assessment.

Current Practice in Risk Assessment

Utilising structured tools containing empirically derived static, dynamic, and protective factors helps to reduce judgment biases that result from assessing factors that have little to no direct relationship with recidivism (e.g., ethnicity). Clinical judgment still has its place in risk assessment however, as tools containing dynamic risk and protective factors are scored by clinicians and probation officers and used to inform their risk judgments. One way this is done is by using actuarial tools and scoring factors as simply present or absent, or scoring factors according to severity. The resulting final score guides what action is taken (Vincent, Perrault, Guy, & Gershenson, 2012). Another approach, known as a structured professional judgment, is similar but the clinician instead observes the risk factors and has the discretion to make the final judgment (Perrault et al., 2012). These approaches are still dependent on "competent exercise of clinical judgments and skills," but the idea is that having a structured tool will give clinicians and probation officers a set of assessment guidelines (Harris, 2006, p. 12). Assessment then becomes dependent on whether the assessor can capably and consistently identify risk factors, and score tools accordingly.

Fourth generation tools make up the current generation in risk assessment. These tools go a step further, combining risk assessment with a case-management plan. This approach integrates a variety of information in order to assess risk for the purposes of monitoring an offender and determining how they might get the most benefit out of treatment (Andrews & Bonta, 2010). A fourth generation assessment might also address other issues that, although are not risk factors per se, might inhibit responsiveness to treatment if left unaddressed (e.g., anxiety and stress; Andrews & Bonta, 2010).

Summary

Over time, the concept of risk has shifted focus from 'risk status' to 'risk state'. Risk status involves identifying an offender's risk level by relative comparison to other offenders. Static approaches are used here, as an offender is given a statistical probability representing their likelihood of recidivism relative to other offenders (Douglas & Skeem, 2005). However, this assumes the majority of the worst offenders will not change, and does little to identify those who have made changes over time. Risk state, in contrast, focuses on current, dynamic risk, acknowledging that offenders do have the capacity for change and that risk factors fluctuate at an individual level (Douglas & Skeem, 2005). Therefore, while static approaches tend to be useful for identifying the highest risk individuals, based on their past behaviour; information provided by dynamic factors should be considered when deciding how to manage and prevent criminal behaviour from occurring.

Because, as mentioned, most desistance occurs during late adolescence, accurate youth risk assessment proves an arguably greater challenge than adult assessment. This challenge stems from two ideas specific to adolescence: first, not all youth offenders persist into adulthood, and second, despite many not persisting into adulthood they may indeed persist throughout the course of their youth and early adulthood years. Therefore, it is important that youth assessment focuses on factors that change in the short-term in order to identify youth who may be in the process of aging-out of crime. The following section details how youth risk assessment is approached, and describes several commonly used tools.

Why Might Risk Assessment of Youth Differ from that of Adults?

"Whatever the age, youth represent "moving targets" from the point of view of risk assessment specialists."

— Vincent, Guy, & Grisso, 2012, p. 21

A primary concern when conducting a risk assessment with youth is striking the right balance between community safety and the rights of the offender. Adolescence is a period of "vast change and development" thus making accurate risk assessment particularly challenging (Vincent, Perrault, Guy & Gershenson, 2012, p. 365). Youth might score highly on certain risk factors such as impulsivity and lack of empathy "due to their transient, developmental immaturity" as opposed to factors that will remain stable in the long-term (Vincent, Perrault, Guy, & Gershenson, 2012, p. 368). Although a disproportionate number of youth begin offending, and may continue to do so during their teenage years, not all persist into adulthood, suggesting they have greater opportunity and room for change than adults. An inaccurate risk assessment could affect what punishment or rehabilitation a youth receives, and inadvertently encourage a realistically low risk youth further down the path to a life of crime (Scott & Steinberg, 2008).

Borum (2003) identifies five reasons why youth risk assessment might differ from adult risk assessment: 1) youth and adults have differing base rates for crime (e.g., youth have higher base rates for violent crime), 2) different risk factors, 3) different behavioural norms, 4) psychosocial maturation is ongoing in youth, and 5) individual factors may be less stable in youth. Therefore, youth risk assessment must not only consider factors such as history of antisocial behaviour that tend to be good long-term predictors of recidivism, but also ongoing dynamic factors that provide information about short-term risk (Borum, 2003). In order to get the fullest picture possible, information should ideally be collected from multiple sources (e.g., the offender, their parents, law enforcement; Olver, Stockdale, & Wormith, 2009).

Youth risk assessment approaches have been largely modeled on adult approaches (Vincent, Perrault, Guy, & Gershenson, 2012). However, separate tools for youth and adults are a response to the recognised need for "developmentally informed" risk assessment (Borum, 2003, p. 117). The majority of youth risk assessment tools have upper limits of 17 or 18, while tools designed for offenders aged 18 and older are considered adult risk assessment tools (Hoge et al., 2012). Empirically valid risk assessment tools refer to those that are consistently capable of distinguishing between individuals who will reoffend and those who will not. Both adult and youth risk assessment tools vary widely from jurisdiction to jurisdiction; with some correctional facilities using internationally developed and validated tools and others using locally developed and validated tools (Hoge et al., 2012). Nevertheless, some tools have had little to no validation done, yet are still used (Hoge et al., 2012).

Researchers have advised caution when using risk assessment tools on groups that are not widely represented in the samples they were developed and/or validated on (see Austin, 2006; Farrington & Loeber, 2002; Singh, Grann, & Fazel, 2011). It is vital to establish the validity of risk assessment tools on different samples of offenders to ensure their effectiveness is not simply due to "chance relationships" specific to the group they were developed on (Putniņš, 2005, p. 326). This is important not only for specific populations (e.g., youth, ethnic minorities) but also community-sentenced offender samples as opposed to parole samples. Unless tools are tested on "multiple populations under varying conditions" they cannot be expected to perform equally well for all populations (Austin, 2006, p. 59).

Risk Assessment Tools with Youth

Risk assessment tools have been traditionally developed and validated on adult male offenders, who make up the majority of the prison and parole populations. Although adult tools might work well enough for youth because of shared risk factors, there is some evidence to suggest tools that have been specifically designed or adapted to assess risk in certain populations (e.g., youth) have higher predictive validity and accuracy overall, when compared to those designed for general offender populations (Singh et al., 2011). For example, the Youth Level of Service/Case Management Inventory (YLS/CMI; Hoge & Andrews, 2006), now in its second version, is an adapted version of its adult-equivalent, the Level of Service/Case Management Inventory (LS/CMI; Andrews, Bonta, & Wormith, 2004). The YLS/CMI incorporates static, dynamic, and protective factors that are scored as present or absent to assess risk of general recidivism. The YLS/CMI comprises many adult risk factors (e.g., three or more prior convictions, few positive friends), but includes a focus on areas of particular risk for youth, in the family and school domains (e.g., inappropriate discipline by a family member, disruptive classroom behaviour). It has predictive validity across a wide range of demographically diverse youth (Olver et al., 2012).

Another youth risk tool with good predictive validity (Catchpole & Gretton, 2003; Singh et al., 2011) is the Structured Assessment of Violence Risk in Youth, used to assess the likelihood of future violent offending (SAVRY; Borum, Bartel, & Forth, 2006). It was created by drawing on the youth recidivism literature and assesses static, dynamic, and protective factors across a range of domains as part of a structured clinical judgment (Lodewijks, de Ruiter, & Doreleijers, 2010). Some of the risk and protective factors assessed are youth-specific (e.g., poor parental management, parental/caregiver criminality, poor school achievement), while many are shared with adults (e.g., history of violence, substance use difficulties, risk taking/impulsivity; Borum, 2003; Lodewijks et al., 2010).

The factors on these two well-established youth tools are geared, for the most part, towards younger youth, many of whom still have ties to family and school. Although adult and youth tools have been shown to have roughly equivalent predictive and construct validity when used on their intended populations, factors vary somewhat by tool, and a tool designed for older youth should "include variables identified as having particular relevance for the early adult period" (Hoge et al., 2012, p. 176).

Little research has been done, however, on the applicability of adult risk assessment tools for use with youth. Research examining the predictive validity of adult tools tends to include 17-19 year olds as part of the adult age group, making it hard to determine if the tools are working differentially for older youth. Only a few studies have investigated how adult risk tools perform when used to assess youth (of any age); two recent studies will be discussed here.

Hoge et al. (2012) reported an unpublished review by several of the authors in which they retrospectively examined studies with relatively long follow-up lengths that had used either youth or adult developed risk assessment tools to predict recidivism outcomes in early adulthood (ages 18-29). Limited information was given about the range of tools that were included in the study and what kinds of factors (static, dynamic, or protective) were primarily assessed. Findings nevertheless indicated that the youth tools that predicted reoffending with reasonable accuracy prior to age 18 maintained their predictive accuracy into early adulthood. Adult tools that predicted recidivism well in early adulthood continued to predict well in later adulthood. This research provides some evidence of overlap in the ability of youth and adult tools to predict early adulthood offending.

Ralston and Epperson (2013) investigated the extent to which different age-specific risk assessment tools could be used for both adults and youth. Their study focused on younger youth (aged 11-17) who had perpetrated sexual offences, and were retrospectively assessed on two adult-developed tools: the Minnesota Sex Offender Screening Tool-Revised (MnSOST-R; Epperson et al., 2004) and the Static-99 (Hanson & Thornton, 2000), and two youth-specific tools: the Juvenile-Sex Offender Assessment Protocol-II (J-SOAP-II; Prentky & Righthand, 2003) and the Juvenile Risk Assessment Scale (JRAS; New Jersey Attorney General's Office, 2006). Recidivism data were collected for two time points: prior to age 18 (youth offending) and aged 18 onwards (adulthood offending). Results indicated that the primarily static adult sex offender tools were predicting youth sexual and violent recidivism just as well as the youth-designed sex offender tools that included dynamic factors. All tools were better at predicting recidivism in the short term, when perpetrators were still youth. Long-term predictions of recidivism were less accurate, and as the youth grew into adulthood, the predictive accuracy that the tools had had in the short-term waned. The authors concluded that this was evidence for the idea that risk of adolescent offending does not necessarily translate into risk of adult offending and thus is challenging to predict. However, the two adult-developed tools (MnSOST-R and Static-99) used in the study consisted of primarily static variables (the MnSOST-R does include dynamic factors but as the study was conducted retrospectively, these could not be scored), while the two juvenile sex offender tools (JRAS and J-SOAP-II) included dynamic subscales in addition to static. Ralston and Epperson (2013) acknowledged that the two adult tools used in their study had age related items (e.g., an age difference of five years or more between the perpetrator and victim) and the manuals themselves explicitly stated the tools were not designed for use with youth. They also noted that, as most risk factors in the adult tools they examined were static, their study could not determine how well adult-designed dynamic tools might work for youth (Ralston & Epperson, 2013). The present study aims to investigate this.

Introduction to the Present Study

> "Unless a tool is validated in a local sanctioning system—and then periodically revalidated—there is little assurance that it works."
>
> *— Monahan & Skeem, 2014, p. 162*

There is mixed evidence around whether adult risk assessment tools can be used with youth. Adult-designed risk tools might be expected to work for youth because they measure shared risk factors, however these tools might work less well for youth compared to adults because factors have been suggested to predict reoffending at varying strengths during adolescence compared to adulthood. Therefore, research that directly compares youth and adults is needed to determine whether adult-designed tools are equally accurate at predicting recidivism for both groups. Researchers have identified the need to examine the applicability of adult risk assessment tools (specifically those containing dynamic factors) for use with older youth, because this age group is on the cusp of the cut-offs for many risk assessment tools (Ralston & Epperson, 2013; Vincent, Perrault, Guy, & Gershenson, 2012).

The present study examines both adult and youth[2] offenders serving community-based sentences with a supervision component. Similar to parolees, community-supervised offenders are monitored and may have sanctions or special conditions imposed on them (e.g., employment or peer association restrictions, or required attendance of a treatment programme). Offenders most likely to be sentenced to supervision are those at low risk of reoffending, that the court deems will benefit most from community-based rehabilitation. Community-based supervision sentences were chosen for the current study because they are the most common of all sentence types for youth offenders (Department of Corrections, 2013). In 2013, just under 5,000 youth were serving a community-based sentence (Department of Corrections, 2013). Community-based sentences last anywhere between six months to two years (Department of Corrections, n.d., supervision). Probation officers regularly meet with offenders during their sentence and offenders report to their probation officer as required. Non-compliance could mean the offender receives another sentence (including prison) or is fined (Department of Corrections, n.d., supervision).

Offenders are assessed at the beginning of, and during, their supervision sentence using two adult risk assessment tools (the RoC*RoI and DRAOR; described in detail in the method). The aim of the present research is to determine whether these tools can predict reconviction equally well for both youth and adult offenders.

More specifically, I aim to answer the following research questions:

1. To what extent does the DRAOR predict reconviction outcomes for youth and adults?
2. To what extent do DRAOR scores change over time and does change predict reconviction for youth and adults?
3. Is the (dynamic) DRAOR a better predictor of reconviction than the (static) RoC*RoI for youth and adults?

[2] Thus far I have called youth in the 17-19 year age bracket ‘older youth’ so as to better identify them as a separate population when comparing them to younger youth or adults, but in keeping with the Department of Corrections’ definition, and for ease of reading, they are referred to in the present study as youth.

Chapter 2

Method

Data

Archival data were provided by the New Zealand Department of Corrections and included a random sample of offenders who had served a community supervision sentence between 1 January 2011 and 31 December 2013. The initial sample (N = 1004) consisted of 455 male youth (aged 17-19) and 549 male adult offenders (aged 20-60) who had been assessed with the DRAOR and RoC*RoI.[3]

Measures

The Dynamic Risk Assessment for Offender Re-entry (DRAOR). The DRAOR (Serin, 2007, Serin et al., 2012) is an actuarial risk assessment tool designed for use with adult offenders (Hanby, 2013). The DRAOR is administered multiple times to offenders in the community (including those who are on parole and also those who are serving a community-based sentence) in order to monitor their risk of reoffending, and track any change in risk that occurs over time. Probation officers complete the assessment via an interview with the offender and using third-party information (e.g., police intelligence and talking to an offender's family members). The DRAOR has face validity as its 19 items are theoretically derived from factors associated with risk of reoffending (see Table 1). The items make up three subscales: stable dynamic risk, acute dynamic risk and protective factors. Each item is scored on a three-point scale (0, 1, 2). For stable and acute items, a score of 0 represents no problem, 1 a slight problem, and 2 a definite problem (Serin et al., 2012). For the protective items, a score of 0 represents an item that is not an asset, 1 a slight asset, and 2 a definite asset (Serin et al., 2012). DRAOR total scores are calculated by adding the acute and stable dynamic risk factors and subtracting the protective factors, creating a measure that takes all subscales into account. Although it has been contested in the literature as to whether protective factors directly remove the effect of risk factors, a total score is the best way to examine both risk and protective factors together, and is an approach used in previous research (see de Vries Robbé, de Vogel, Douglas, & Nijman, 2015)

[3] Although female offender data were also supplied, female offenders were not included in the current study as they have been shown to perpetrate crime at different rates to males and may have different pathways into and out of crime (Indig et al., 2014).

including research using the DRAOR (Hanby, 2013; Yesberg & Polaschek, 2015). The maximum possible total score someone can obtain on the DRAOR is 26 (scoring a 2 for all risk factors and a 0 for all protective factors), and the lowest possible total score is -12 (scoring a 2 for all protective factors and a 0 for all risk factors).

Although the DRAOR has yet to be tested on community supervision samples, a pilot study found the DRAOR had good predictive validity and reliability on a sample of New Zealand adult parolees (Tamatea & Wilson, 2009). It has since been shown to predict recidivism for various further offender samples including (primarily) adult male (Hanby, 2013; Yesberg & Polaschek, 2015) and adult female offenders (Yesberg, Scanlan, Polaschek, Hanby, & Serin, 2015). The only previous research into its use with youth is an unpublished validation study (Fortune, Ferguson, Hanby, & Serin, in preparation) that found it was unable to predict recidivism for a sample of 72 youth parolees.

Table 1

DRAOR Items by Subscale

Dynamic Risk Assessment for Offender Re-entry (DRAOR)		
Acute	Stable	Protective
Substance abuse	Peer associations	Responsive to advice
Anger/hostility	Attitudes towards authority	Prosocial identity
Opportunity/access to victims	Impulse control	High expectations
Negative mood	Problem-solving	Costs/benefits
Employment	Sense of entitlement	Social support
Interpersonal relationships	Attachment with others	Social control
Living situation		

Risk of re-Conviction X Risk of re-Imprisonment (RoC*RoI). Developed by Bakker, Riley, and O'Malley (1999), the RoC*RoI is a computer-generated algorithm that weights 16 static items (including age, gender, total estimated years spent in prison) according to their relationship with reoffending. It calculates a risk score (ranging from 0 - 1) interpreted as the likelihood of an offender committing an offence resulting in a sentence of imprisonment within a five-year period. For example, a score on the RoC*RoI of 0.7 is interpreted as indicating that an offender has a 70% likelihood of being imprisoned within the next five years, giving them a classification of high risk by the Department of Corrections. A low risk offender would likely get a score of less than 0.3, and a medium risk offender would get a score between 0.3 and 0.7. Early research has found the RoC*RoI has high predictive validity (AUC = .76; Bakker et al., 1999), a finding that has since been replicated in further samples of offenders (e.g., Nadesu, 2007).

Data Preparation and Exclusion/Inclusion Criteria

Date of birth and age. The age supplied in the dataset had been rounded to the nearest whole number (e.g., 17.5 year olds had their age recorded as 18, etc.). Because age is used to distinguish youth from adults in the current study, it was important that it be as accurate as possible. All birthdates provided had been reset to the first of the month in order to protect the identity of offenders. All birthdates were subsequently changed to the 15th of the month, so they would only be incorrect by a potential 15 days either side of the month, rather than a possible 30. A more exact age was calculated by subtracting this new date from the date of sentence commencement.

DRAOR assessments. The dataset included the dates of all recorded DRAOR assessments and scores on each of the items. As the DRAOR is intended to be a dynamic instrument, regular assessments are essential in order for scores to be as up-to-date and accurate as possible. Youth and adults were excluded from this research if they had a time period greater than fourteen weeks between the start of their sentence and first DRAOR assessment, or if there was a gap of this length between any of the assessments throughout the course of the sentence. Because change over time is examined in this research, offenders were also removed if they had fewer than five recorded DRAOR assessments.

For the current study, several DRAOR scores were of interest. I extracted scores on each of the three subscales and the total score for each offender's third DRAOR assessment during the period of their sentence (the *initial* score), as well as the most recent DRAOR assessment prior to reconviction for reoffenders or the last DRAOR assessment on sentence for those who were not reconvicted during their time at risk (the *proximal* score). The scores were selected for their expected variability, and the different information they might give about the DRAOR's ability to predict risk.

The third DRAOR assessment on sentence was chosen as the initial score instead of the first assessment because it can take three to four assessments for a probation officer to become familiar enough with an offender that they can accurately assess their risk level (a precedent set by Hanby, 2013). In some cases, an offender's regular probation officer may not have been the one who completed the first assessment. Familiarity with an offender is important because the DRAOR's interview-based format requires a probation officer to have detailed knowledge of the offender's present situation. The proximal score was chosen because it is closer in time to reconviction and might be able to provide more pertinent information about risk.

RoC*RoI assessments. Two RoC*RoI scores were included in the dataset: the RoC*RoI score calculated closest to the start of an offender's sentence, and the most recent RoC*RoI score that had been calculated during the term of the offender's sentence. In order to be included in the study, offenders had to have a RoC*RoI score that was calculated near the start of their supervision sentence and before any reconvictions had occurred.[4]

Exclusion based on the above criteria resulted in 274 youth and 446 adults remaining in the available pool of choices for propensity score matching. Their demographic variables are shown in Appendix A, and discussed below.

Procedure

Propensity Score Matching (PSM). Youth were matched to adults using a PSM procedure, in order to control for differences in variables that might predict reconviction such as ethnicity, number of previous convictions, violent convictions, and imprisonments, index offence,[5] sentence length, and static risk (RoC*RoI score). The procedure involved entering these variables simultaneously as predictors into a logistic regression, with age group (youth, adult) as the dependent variable. The regression model was significant, $\chi^2(11, n = 720) = 420.94$, $p < .001$, and was able to explain between 44.3% (Cox and Snell R square) and 60.2% (Nagelkerke R squared) of the variance in predicting age group. The regression model correctly classified 83.2% of cases, indicating the groups were sufficiently different on the predictor variables and that they could be matched. The

[4] This is because, as a static tool, the RoC*RoI score increases with the number of new convictions. Therefore, any score calculated after a reconviction might be accounting for that offence and thus represent a level of risk that was inaccurate at the start of an offender's sentence.

[5] Index offences included: Non-violent index offences (e.g., burglary, driving, and drug offences), violent or sexual offences (e.g., assault, threatening to kill; the majority of these offences were violent, but sexual offences were included with them due to the invasive nature of the offence – two adults had committed an indecent assault, and two youth an unlawful sexual connection), justice/administrative offences (e.g., breaching sentence conditions, resisting police), or unknown offences.

regression process generated a probability for each offender ranging between 0 and 1, interpreted as the likelihood that they belonged either to the adult or youth group based on these variables. Using these probabilities, each youth was then manually matched to an adult, using the nearest neighbour method, whereby the closest possible adult match for each youth was chosen, with a range of ± 0.02.[6] The offenders were matched without replacement; each chosen offender was used only once. This procedure resulted in 122 youth being matched to 122 adults, giving a final sample of 244.

Demographic Information

Sample characteristics. The age of the youth group (recorded at start of supervision) ranged from 17.0 to 19.8 years, with a mean age of 18.6 years (*SD* = 0.61). The age of the adult group ranged from 20.0 to 60.3 years, with a mean age of 26.4 years (*SD* = 8.0). Because offenders were matched on key demographic and criminal history variables, as expected, there were no other significant differences between the two groups.

Table 2 reports sample characteristics by group. To summarise, the overall sample was primarily European (43.9%, *n* = 107) and Māori (39.3%, *n* = 96). Offenders had an average of 10.1 previous convictions (*SD* = 6.68), one of which was for a violent offence (*M* = 0.90, *SD* = 1.08). Very few had been previously imprisoned (*M* = 0.41, *SD* = 0.97). The majority of the sample had committed a non-violent index offence (62.3%, *n* = 152), followed by a violent offence (28.7%, *n* = 70). The mean sentence length was 268.38 days (*SD* = 64.62). Offenders had a mean RoC*RoI score of .37 (*SD* = .17), placing them, on average, at medium risk of imprisonment within five years.

Appendix A shows a summary table of youth and adult group demographics before the matching process. The table reports pre-matching demographics for all offenders including those who were matched. Youth who were not matched had slightly shorter sentence lengths and fewer previous convictions, but higher RoC*RoI scores than their matched counterparts. On average, adults who were not matched had slightly longer sentence lengths, more previous convictions, but slightly lower RoC*RoI scores. A major difference is that adults who could not be matched had, on average, 27.15

[6] This criterion was chosen at my discretion, as the purpose of PSM is to match the sample as closely as possible while still maintaining enough cases in the sample that statistical power is not compromised.

(*SD* = 24.81) previous convictions, over double the amount of the adults in the matched sample (*M* = 10.27, *SD* = 5.68).

Table 2

Summary of Youth and Adult Group Demographic and Offence-related Variables

Variable	Group			t-test statistic
	Youth *M* (*SD*) *n* = 122	Adult *M* (*SD*) *n* = 122		
Sentence length (in days)	270.25 (67.09)	266.50 (62.27)		$t(242) = 0.45$, $p = .651$, d[7] $= 0.06$, $M_{diff} = 3.75$, 95% CI [-12.57, 20.08].
Number of previous convictions	9.87 (7.56)	10.27 (5.68)		$t(224.59) = -0.47$, $p = .639$, $d = -0.06$, $M_{diff} = -0.40$, 95% CI [-2.09, 1.29].
Number of previous violent convictions	0.93 (1.10)	0.87 (1.06)		$t(242) = 0.42$, $p = .679$, $d = 0.06$, $M_{diff} = 0.06$, 95% CI [-0.22, 0.33].
Number of previous imprisonments	0.39 (1.09)	0.44 (0.83)		$t(242) = -0.46$, $p = .644$, $d = -0.05$, $M_{diff} = 0.05$, 95% CI [-0.30, 0.19].
RoC*RoI score	.37 (.16)	.37 (.17)		$t(242) = -0.11$, $p = .914$, $d = 0$, $M_{diff} = 0$, 95% CI [-0.04, 0.04].
	Number of youth	Number of adults	Total	Chi-square statistic
Ethnicity				
Māori	52	44	96	$\chi^2(3, N = 244) = 1.48$, $p = .686$. Φ[8] $= 0.08$.
European	49	58	107	
Pacific Peoples	12	12	24	
Other/Unknown	9	8	17	
Index offence				
Non-violent	76	76	152	$\chi^2(3, N = 244) = 2.86$, $p = .414$, $\Phi = 0.11$.
Violent/sexual	36	34	70	
Justice/admin	8	12	20	
Unknown	2	0	2	

[7] See interpretation of results section below for interpretation of Cohen's *d*
[8] See interpretation of results section below for interpretation of *phi* (*Φ*)

In other words, adults with the fewest convictions were the best matches for youth. It is no surprise that adults were more likely than youth to have a greater number of convictions, having had more years in which to acquire them. However, due to the nature of this research as a comparative study between youth and adults, matching to control for criminal history variables was essential.

Normality. To examine the distribution of the matching variables across the two groups in greater depth, *z* scores were calculated. The skewness and kurtosis values of each predictor (for each group) were divided by their respective standard error. By convention, for large samples, if *z* scores fall within the range of ± 3.29 (two-tailed probability of .001), the predictor can be considered likely to have been drawn from a normally distributed population (Allen & Bennett, 2010). Sentence length for both groups, and adult RoC*RoI score had both skewness and kurtosis fall within the normal distribution. Despite youth RoC*RoI score and number of previous convictions of all types violating the assumption of normality for both groups, it is worth noting that the population these data are drawn from is a medium risk community supervision sample of matched offenders; extreme values in predictors are expected. The total sample size (n = 244) was also large enough to accommodate heterogeneity in the variance (i.e. over 100 cases; Tabachnick & Fidell, 2013).

Prior to analysis, all DRAOR scores used were also checked for normality (this includes initial and proximal DRAOR scores, and change scores; described in detail later). Several of the scores deviated from normality but, again, the sample size was large enough to accommodate these deviations.

Offence-related Variables

Recidivism. Two recidivism outcomes were used in this research: any new criminal reconvictions excluding breaches (hereafter referred to as *criminal reconviction*) and any breaches of sentence conditions (hereafter referred to as *breach reconviction*). Criminal reconviction was defined as the first new conviction for any criminal offence that occurred during the duration of the offender's sentence and up until data extraction (i.e. their *time at risk*, see below). Recidivism data were extracted for all offenders on 13 June 2014. Breach reconvictions by definition could not occur outside of an offender's sentence, and so were defined as the first new breach reconviction that occurred during the supervision sentence. New dichotomous variables (yes, no) were created for both outcomes based on whether a reconviction occurred during an offender's time at risk (see below).

Time at risk. For the criminal reconviction outcome, subtracting the start date of each offender's sentence from the data extraction date gave the total number of days they were at risk of reconviction. The minimum time at risk for youth was 382 days and the maximum was 1044 (M = 808.71, SD = 175.77, median = 849.50). The minimum number of days at risk for adults was 374 and the maximum was 1208 (M = 812.36, SD = 222.28, median = 815.50). There was no statistical difference between the groups $t(229.77) = -0.14$, $p = .89$, $d = 0.02$, $M_{diff} = -3.65$, 95% CI [-54.20, 46.90].

Time at risk for the breach reconviction outcome was the duration of an offender's sentence, with the minimum time at risk for youth being 181 days and the maximum being 539 (M = 270.25, SD = 67.09, median = 273.00). The minimum number of days at risk for adults was 180 and the maximum was 365 (M = 266.50, SD = 62.27, median = 273.00). There was no statistical difference between the groups $t(242) = 0.45$, $p = .651$, $d = 0.06$, $M_{diff} = 3.75$, 95% CI [-12.57, 20.08].

Survival days. The length of survival differed by offender. For criminal reconvictions, it was the length of time from the start of an offender's sentence until their reconviction. For offenders who were not reconvicted of a criminal offence, it was the length of time from the start of their sentence to the data extraction date. For breach reconvictions the period for offenders who were reconvicted included the start of sentence until their first breach offence while on sentence, or start of their sentence until end of their sentence for offenders who were not reconvicted.[9]

Analysis

Statistical methods used in the present study are described in detail below. All data were analysed using IBM SPSS Statistics version 22.

Kaplan-Meier survival analysis. The Kaplan-Meier survival analysis method uses reconviction outcomes and survival days to compare the true rate of reconviction between groups. The analysis censors data points to represent individuals who had not been reconvicted by the end of

---

[9] As mentioned previously, once an offender had completed their supervision sentence they no longer had the opportunity to breach the conditions of that sentence during the follow-up period.

their follow-up period (Norušis, 2004). Thus, cumulative survival (i.e. the proportion of offenders who had not been reconvicted at any given time) is estimated while controlling for varying times at risk.

There are several significance tests that can be used in survival analysis. For the current study, the Tarone-Ware statistic was used to test significance, as it takes into account both the start and end points of the analysis (start of supervision, and reconviction or data extraction respectively), rather than weighting one point more heavily than the other (Norušis, 2004). This weighting was ideal for the current analysis, to examine the overall pattern of reconviction throughout the period of interest.

Cox regression. Cox regression allows an examination of the nature and strength of the relationship of both continuous and categorical variables in predicting an outcome for censored data. One or more predictor variables are entered into the regression equation, in order to assess whether the model with the predictors is better than one operating at a level of chance and if so, whether any of the predictors are driving the model's predictive ability. A multivariate regression model does this by assessing the unique contribution each predictor makes while accounting simultaneously for the contribution of the other variables. In other words, as with all regression models, Cox regression assesses whether one or more predictor variables are able to explain variance in an outcome that other predictors cannot explain (Allen & Bennett, 2010). The Cox regression method also has the benefit of taking into account variables that are time-dependent (Garson, 2013), which was essential for the current study given the need to evaluate how much of a contribution the tools were making towards predicting recidivism outcomes while considering offenders' differing survival lengths. The hazard ratio indicates the extent to which an increase or decrease in the predictor variable affects the outcome; in the case of the current study, to what extent a one-unit increase in risk scores is associated with an increased/decreased likelihood of reconviction. Because it gives a measure of the rate at which a predictor variable affects the dependent variable, the hazard ratio can be interpreted as a measure of effect size. A hazard ratio greater than 1 signifies that a one-unit increase in the predictor is associated with an increase in reconviction. A hazard ratio of less than 1 signifies that a one-unit increase in the predictor is associated with a decrease in reconviction. If confidence intervals surrounding the hazard ratio cross 1, the predictor is unlikely to be affecting the likelihood of reconviction.

Area Under the receiver operating Curve analysis (AUC). AUC analysis tests the accuracy of a regression model. This analysis uses the X * Beta probabilities generated from Cox regression to give each offender a score from 0 to 1.[10] The AUC value generated from the analysis is interpreted as the model's ability to discriminate true positive from false positive outcomes (represented as the 'area under the curve' when plotted on a graph; Wilson & Rolleston, 2004). In the current study, the closer to 1 a value is, the more accurately the regression model is able to distinguish between offenders who were reconvicted and offenders who were not. Therefore, an AUC value of .50 would mean the predictor(s) contained within the model were operating at a level no better than chance (Zhou, Obuchowski, & McClish, 2002). AUC values between .56 and .63 suggest the predictor(s) are operating at a level of low accuracy; values between .64 and .71 indicate moderate accuracy; and values above .71 indicate high accuracy (Rice & Harris, 2005). An advantage of AUC analysis in forensic psychology research is that, unlike Cox regression, it accounts for differences in base rates, which allows for direct comparisons to be made between groups that have low or unequal ratios of reconvictions (Miller, 2015). When comparing groups using this analysis, overlap in the confidence intervals for the AUC values suggests that accuracy is comparable for both groups, as the true mean for both groups could plausibly be contained within the overlap.

Interpretation of Results

I have endeavoured to interpret the results of my analyses in terms of effect size wherever possible. Geoff Cummings' (2014) New Statistics approach recommends researchers move away from the traditional dichotomous significant/non-significant approach of null hypothesis significance testing (NHST) and instead look at effect size to infer meaning from their results. This involves examining the size of point estimates (i.e. means and mean differences) and their confidence intervals and is particularly useful for making comparisons between groups, which is frequently done in the present research. Confidence intervals express the "extent of uncertainty" around the point estimate,

[10] The X * Beta probabilities are described as the "best linear combination of two or more predictor variables [from the regression] including time-dependent variables" (Brown, St. Amand, & Zamble, 2009, p. 33)

and thus the precision of the analysis (Cumming, 2014, p. 13). If there are no overlapping values in the confidence intervals between groups, one can be fairly sure there is a statistical difference between the groups. Highly overlapping confidence intervals suggest it is most likely there is no difference in point estimates between groups. I also report Cohen's *d*, a standardised effect size used to compare the extent of the difference between two means. For independent and paired samples t-tests comparing differences between means, a small effect is represented by a *d* value of .20, a medium effect has a *d* value of .50 and a large effect has a *d* value of .80 or greater (Cohen, 1992). Chi-square effect sizes are represented by the *phi* (Φ) value, where a small effect is represented by a Φ value of .10, a medium effect has a Φ value of .30 and a large effect has a Φ value of .50 or greater (Cohen, 1992). For correlations, a small/weak effect is represented by an *r* coefficient of .10, a medium/moderate effect has an *r* coefficient of .30 and a large/strong effect has an *r* coefficient of .50 or greater (Cohen, 1992). In addition to reporting mean differences, confidence intervals, and effect sizes, *p*-values have also been included as this is how results are routinely interpreted and reported in the literature. An alpha level of .05 was used for all significance tests.

Chapter 3

Results

Rate of Criminal Reconviction

More youth than adults had incurred criminal reconvictions; of the 122 youth, 93 (76.2%) had a reconviction for any new criminal offence compared to 47 (38.5%) of the 122 adults.

Figures 1 and 2 show the survival curve graphs for the first new criminal and breach reconviction outcomes. The x-axis represents the time to event; that is the number of days from the start of an offender's sentence until their first reconviction, or data extraction. The y-axis represents the proportion of 'survivors', that is, offenders who had not been reconvicted at any given time.

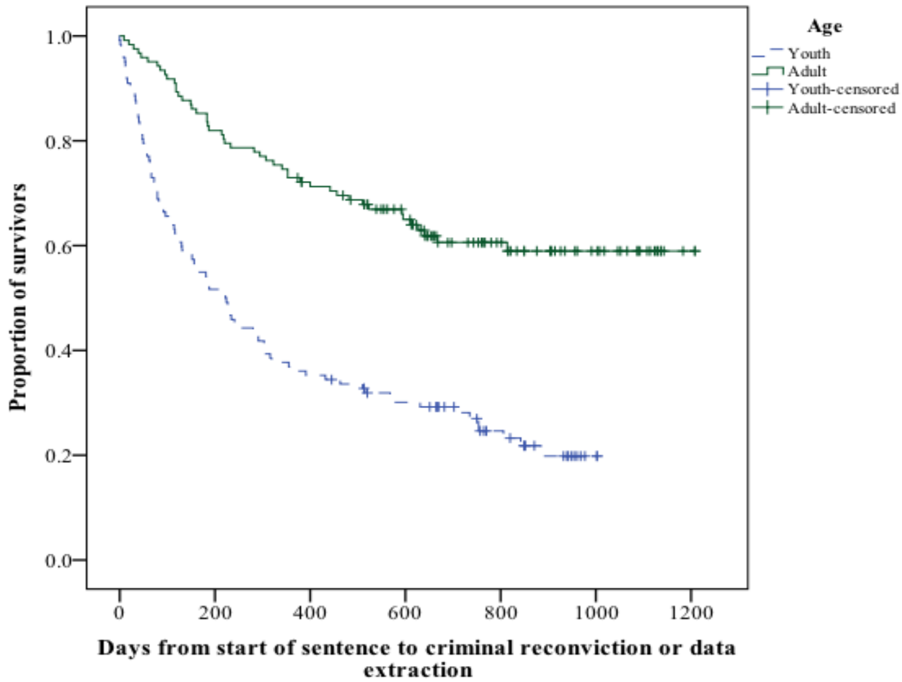

Figure 1. Survival curve for youth and adult criminal reconviction

Figure 1 shows the survival curve for criminal reconviction for the youth and adult groups. Although the median is traditionally reported in survival analysis, as the adult proportion of survivors did not drop below 0.5 (and in this case there is no median) the mean is reported for both groups in order to keep the reported indices of central tendency consistent. Overall mean survival time for youth was 386.54 days (*SE* = 34.60), 95% CI [318.71, 454.36]. The adult mean survival time was 837.75 days (*SE* = 43.07), 95% CI [753.32, 922.18].

An examination of the slopes of the survival curves, and accompanying statistics, reveals that overall more youth were reconvicted than adults, and at a faster rate: Tarone-Ware statistic (*df* = 1) χ^2 = 41.33, p < .001.

Rate of Breach Reconviction

Similarly, more youth than adults had also incurred breach reconvictions; of the 122 youth, 61 (50.0%) had a reconviction for any new breach offence compared to 26 (21.3%) of the 122 adults.

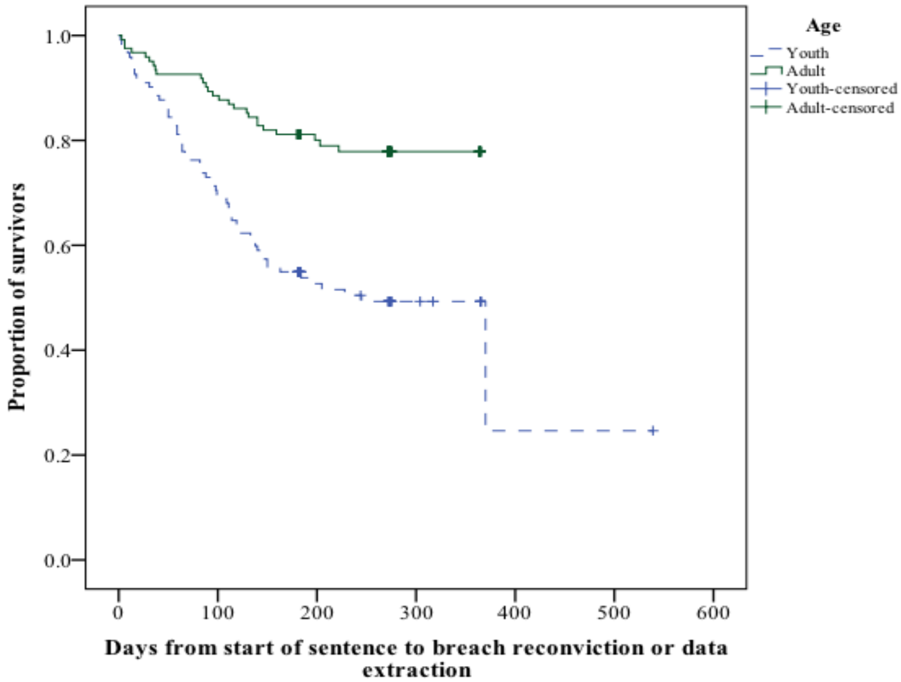

Figure 2. Survival curve for youth and adult breach reconviction

Figure 2 shows the survival curve for breach reconviction for the youth and adult groups. Overall mean survival time for youth was 269.77 days (*SE* = 34.12), 95% CI [202.90, 336.64]. Mean survival time for adults was 305.85 days (*SE* = 10.53), 95% CI [285.22, 326.48]. These times are shorter on average than those for criminal reconviction as recall that breach reconvictions could only occur during an offender's sentence, and for those who were not reconvicted the end of the follow-up period was their sentence end date. Again, more youth were reconvicted than adults, and at a faster rate: Tarone-Ware statistic (*df* = 1) χ^2 = 20.93, $p < .001$.

Reporting Analyses for Both Reconviction Outcomes

The three research questions will be answered for each reconviction outcome separately. The remainder of this chapter will be devoted to criminal reconviction and the next chapter will concentrate exclusively on breach reconviction.

Criminal Reconviction

To what extent does the DRAOR predict criminal reconviction outcomes for youth and adults?

I first undertook a series of analyses to investigate whether DRAOR scores had any relationship with criminal reconviction. More specifically, could these scores predict which youth and adults were likely to have been reconvicted of a criminal offence? The three DRAOR subscale scores (acute, stable, protective) and the composite total score were first examined individually and then the subscales were examined in combination to explore how well they could predict any new criminal reconviction, and if so, whether their accuracy differed between the age groups. By examining the subscales in combination, I could account for overlapping variance in the prediction of reconviction to see whether any of the subscales might make a unique contribution towards the DRAOR's predictive validity. Evidence for predictive validity would suggest the DRAOR is adequately assessing factors related to criminal reconviction.

As the DRAOR is a dynamic tool and scores are expected to change in response to risk, it was vital that I took care to use it only to measure an offender's risk before they had incurred any reconvictions. Therefore, offenders were only included in the following analyses if their reconviction occurred after their initial DRAOR assessment. Offenders also had to have their initial and proximal assessments be different assessments, so any offenders whose initial DRAOR assessment was also the one closest to their reconviction were not included. All offenders excluded based on these criteria also had their matched pair excluded, so the sample would remain matched. This resulted in a slightly reduced sample of 100 youth and 100 adults to be used in subsequent criminal reconviction analyses.

To what extent does the initial DRAOR score predict criminal reconviction for both groups? As previously mentioned (see method), the initial DRAOR score serves as a measure of an offender's dynamic risk near the start of his sentence.

The mean initial DRAOR scores for both groups of offenders are shown in Table 3. Compared to adults, youth had higher mean scores on the acute subscale, $t(198) = 2.04$, $p = .043$, $d = 0.29$, $M_{diff} = 0.74$, 95% CI [0.02, 1.46], lower mean scores on the protective subscale, $t(198) = -2.15$, $p = .033$, $d = -0.30$, $M_{diff} = -0.67$, 95% CI [-1.28, -0.06], and higher total scores, $t(198) = 2.22$, $p = .028$, $d = 0.31$, $M_{diff} = 1.73$, 95% CI [0.19, 3.27], while the small difference between groups on the stable subscale did not reach statistical significance, $t(198) = 1.11$, $p = .270$, $d = 0.16$, $M_{diff} = 0.32$, 95% CI [-0.25, 0.89].

Even the statistically significant differences are small according to the effect size guidelines for t-tests given by Cohen (1992, see method for interpretation). For example, recall that scores on the acute subscale can range from 0 to 14, so a 0.74 difference in scores is small. However, even small mean differences in the direction of youth scoring higher on risk and lower on protective factors are promising, as recall considerably more youth than adults were reconvicted of a criminal offence.

Table 3

Mean Initial DRAOR Scores for all Subscales and Total Score: Criminal Reconviction

Score	Youth *M* (*SD*)	Adult *M* (*SD*)
Acute subscale	5.06 (2.63)*	4.32 (2.51)
Stable subscale	5.82 (2.22)	5.50 (1.86)
Protective subscale[11]	6.19 (2.38)	6.86 (2.01)*
Total score	4.69 (5.81)*	2.96 (5.21)

* $p < .05$

Preliminary Pearson bivariate correlation analyses were undertaken to check the linear relationship between the three DRAOR subscales and total score. Checking the relationship of the variables in this way would provide convergent validity; with moderate correlations between the subscales providing justification for investigating their predictive validity further using Cox regression. The correlation coefficient *r* was interpreted as an effect size using Cohen's (1992) guidelines (see method). For both groups all the initial DRAOR subscale scores were correlated with each other in the expected directions at a level ranging from moderate to strong (the risk subscales being positively correlated with each other, and the protective subscale being negatively correlated with the risk subscales, see Table 4). These correlations are a promising result that suggests the subscales are measuring a related concept (ideally, likelihood of reconviction). For youth, and particularly for adults, the stable and protective subscales had the strongest correlation with one another; meaning offenders who scored higher on the stable subscale tended to also score lower on the protective subscale. As expected, being a measure that included contributions from all three subscales,

[11] Recall, higher protective scores mean the offender has a greater number of positive assets.

the total score was strongly correlated with each of them.

Table 4

Correlations Between Initial DRAOR Subscale Scores and Total Score for Youth and Adults: Criminal Reconviction

	Acute	Stable	Protective	Total score
Youth				
Acute	1			
Stable	.46**	1		
Protective	-.42**	-.55**	1	
Total score	.80**	.81**	-.81**	1
Adult				
Acute	1			
Stable	.47**	1		
Protective	-.42**	-.63**	1	
Total score	.81**	.83**	-.81**	1

** $p < .01$

However, strong correlations suggest issues with multicollinearity, a problem that can affect the precision and power of multivariate regression models. When many predictors share variance, it can cause imprecise estimates around the amount of variance accounted for by each predictor, and thus reduce the statistical power of the regression. As a result, it becomes hard to see the effects individual predictors might be having over the variance accounted for by other predictors. Larger sample sizes can reduce these effects, however due to the nature of this research—using archival data and matching offenders— it was not possible to increase the sample size.

To test the extent of multicollinearity among the predictors used in the present research I ran multiple regressions to calculate the *Variance Inflation Factor (VIF)* of each predictor entered together in the regression equation. The VIF provides a measure of how much greater the variances are above what would be expected if the predictors were not correlated (Lin, 2008). It has been suggested that multicollinearity is only a serious problem if VIF values are greater than 10 (Lin, 2008). VIF values were calculated for all variables that were to be entered into multiple regressions in this research, with all analyses yielding values ranging between 1 and 5. Thus, multicollinearity is unlikely to affect the validity of the present research.

Univariate predictive validity. Although related to each other, to what extent are scores on the subscales and total score related to reconviction? In other words, would offenders who had received higher initial DRAOR scores be at an increased likelihood of reconviction (or a decreased likelihood, in the case of the protective subscale)? In order to investigate how well initial scores were able to predict youth and adult criminal reconviction, the three subscale scores and the total score were each entered as a covariate in the first block of four separate univariate Cox regressions. The time variable entered was survival days (days to criminal reconviction or data extraction for those who were not reconvicted), and the status variable was the presence or absence of a criminal reconviction.

Table 5 shows the results of these regressions investigating the predictive validity of the scores on all subscales without taking shared variance into account (for ease of reading and comparison these have been combined into one table). For youth, all initial subscale scores were able to predict criminal reconviction, acute: $\chi^2(1, n = 100) = 6.09$, $p = .014$; stable: $\chi^2(1, n = 100) = 5.60$, $p = .018$; protective: $\chi^2(1, n = 100) = 8.15$, $p = .004$, as was the total score, $\chi^2(1, n = 100) = 10.53$, $p = .001$. For adults, only the initial acute subscale score, $\chi^2(1, n = 100) = 5.74$, $p = .017$, and total score, $\chi^2(1, n = 100) = 4.03$, $p = .045$, were able to predict criminal reconviction. The stable, $\chi^2(1, n = 100) = 1.71$, $p = .191$, and protective, $\chi^2(1, n = 100) = 1.17$, $p = .280$, subscale scores were not.

The *hazard ratio (HR)* represents the strength of the DRAOR scores as predictors. It shows the effect a one-unit increase in scores would have on the likelihood of reconviction. As anticipated, the acute and stable subscales and total score have HRs higher than 1 (higher risk scores correspond with an increased likelihood of reconviction), and the protective subscale has a HR lower than 1 (higher protective scores correspond with a reduced likelihood of reconviction). As an example, with every one-unit increase in stable subscale scores, the likelihood of a youth being reconvicted of a criminal offence increased by 14%, HR = 1.14, 95% CI [1.02, 1.26]. For protective factors, this interpretation is reversed. For youth, a one-unit increase in protective scores corresponds to a 15% decrease in the likelihood of criminal reconviction, HR = 0.85, 95% CI [0.76, 0.95] (this is calculated by subtracting the value of the protective HR from 1 to get the percentage decrease). The highest adult HR came from the acute subscale, with every one-unit increase in acute scores increasing an adult's likelihood of criminal reconviction by 15%, HR = 1.15, 95% CI [1.02, 1.29]. Highly overlapping confidence intervals around the subscale HRs for both groups support previous correlation findings that suggested the subscales shared variance.

Table 5

Univariate Regression Models Containing Initial DRAOR Subscale Scores and Total Score Predicting Criminal Reconviction

Model for initial DRAOR scores	*β* (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]	AUC [95% CI]
Youth					
Acute score	.12 (.05)	6.04*	1.12	[1.02, 1.24]	.60 [.48, .71]
Stable score	.13 (.05)	5.56*	1.14	[1.02, 1.26]	.61 [.49, .73]
Protective score	-.16 (.06)	8.18**	0.85	[0.76, 0.95]	.60 [.48, .71]
Total score	.07 (.02)	10.45**	1.08	[1.03, 1.12]	.64* [.52, .75]
Adult					
Acute score	.14 (.06)	5.64*	1.15	[1.02, 1.29]	.63* [.52, .75]
Stable score	.12 (.09)	1.71	1.12	[0.94, 1.34]	.56 [.45, .68]
Protective score	-.08 (.08)	1.18	0.92	[0.79, 1.07]	.57 [.46, .69]
Total score	.06 (.03)	4.07*	1.06	[1.00, 1.12]	.62* [.51, .74]

** $p < .01$, * $p < .05$

The DRAOR's accuracy at distinguishing between offenders who were reconvicted and offenders who were not reconvicted, taking into account base rates, is represented by the *area under the curve (AUC)* value. Recall that AUC values of .50 mean a predictor is operating at chance level (refer to method for further detail). The highest youth AUC value came from the total score, AUC = .64, 95% CI [.52, .75], meaning that if a youth who was reconvicted and a youth who was not reconvicted were chosen at random, there is a 64% likelihood that the reconvicted youth would have a higher total score than the youth who was not reconvicted. For adults the highest AUC values came from the acute subscale, AUC = .63, 95% CI [.52, .75], and total score, AUC = .62, 95% CI [.51, .74].

As previously mentioned, the total score represents the contribution of all three subscales. Despite the total score having a slightly a higher AUC value for youth, highly overlapping confidence intervals between the youth and adult groups for the total score AUC value indicates when taking into account base rates of criminal reconviction, accuracy plausibly does not differ between groups (see

Cumming, 2014).

An inspection of the confidence intervals around the youth AUC values revealed that the acute, stable and protective subscale values spanned .50, meaning each of these subscale models were plausibly no better (or even worse) than chance at distinguishing between youth who were and were not reconvicted based on their scores. Although all three of these subscales reached a level of statistical significance for youth in their univariate Cox regressions, their AUC values suggest that when accounting for the base rate of criminal reconviction, their predictive ability was not accurate. For adults, the stable and protective subscale AUC values supported the non-significant findings of their Cox regressions; it is likely that the accuracy of these subscale scores at correctly distinguishing between offenders who were reconvicted and those who were not is no better than chance.

Summary. Results from Cox regression analyses suggest that scores on all three subscales recorded at the initial DRAOR assessment can predict youth criminal reconviction, with the total score having the greatest statistical support as a predictor: operating with moderate accuracy. That is, on average, higher scores on all risk subscales are associated with a greater likelihood of reconviction, while higher scores on the protective subscale are associated with a lesser likelihood of reconviction. For adults, only the acute and total score regression models were able to predict criminal reconviction with above chance accuracy. Despite more predictors reaching a level of statistical significance for youth in the Cox regressions, when accounting for criminal reconviction base rates, accuracy appears comparable between youth and adults.

Multivariate predictive validity. Because DRAOR subscale scores were moderately correlated with each other, it is reasonable to expect that they share variance. The next step was to examine the subscales in combination, to determine whether any of them could account for more of the variance in reconviction than the others. In other words, is any one subscale able to give more information about who is likely to be reconvicted than the other subscales? All initial DRAOR subscale scores (acute, stable, protective) were entered as covariates into the first block of a Cox regression (see Table 6).

Table 6

Multivariate Regression Models Containing Initial DRAOR Subscale Scores Predicting Criminal Reconviction

Model for initial DRAOR scores	β (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]	AUC [95% CI]
Youth					
Acute score	.07 (.06)	1.38	1.07	[0.96, 1.20]	.63* [.53, .74]
Stable score	.04 (.07)	0.37	1.04	[0.91, 1.19]	
Protective score	-.12 (.06)	3.44	0.89	[0.79, 1.01]	
Adult					
Acute score	.14 (.08)	3.35	1.15	[0.99, 1.33]	.63* [.52, .75]
Stable score	.02 (.12)	0.04	1.02	[0.80, 1.30]	
Protective score	.02 (.10)	0.04	1.02	[0.84, 1.24]	

* $p < .05$

For youth, the multivariate regression model with all predictors was significant, $\chi^2(3, n = 100) = 10.91$, $p = .012$, which is expected given that all the univariate regressions were significant. The regression model revealed that all subscales in combination were contributing to predicting criminal reconviction, and no one subscale in particular was better than the others at distinguishing youth who were reconvicted from youth who were not (see Table 6). However, the protective subscale had the largest HR and was trending towards significance as a unique predictor ($p = .064$), with the upper limit of the confidence interval for its HR only crossing 1, HR = 0.89, 95% CI [0.79, 1.01], suggesting that although statistically it was not incrementally predictive, it might be capturing something different to the other two subscales. Given that protective factors are conceptually different to risk factors, this result makes sense. The AUC value suggests that a randomly selected reconvicted youth was 63% more likely than a non-reconvicted youth to have higher risk scores and lower protective scores.

For adults, the multivariate regression model did not reach a level of statistical significance, $\chi^2(3, n = 100) = 5.77$, $p = .124$. Recall that in the previous univariate regression analyses the acute subscale was a significant predictor of adult reconviction its own, but here, in combination with the other two subscales, it became a non-significant predictor. This indicates the acute subscale could not predict at a level above the other subscales due to shared variance and that perhaps additional unexplained variance introduced by the stable and protective subscales was hindering the acute subscale's individual contribution (see Table 6).

Summary. For youth, the subscale scores taken from the initial DRAOR assessment are comparable in their ability to predict criminal reconviction. In terms of effect size, the protective subscale might be of increased importance for youth as it is trending towards being a unique predictor. Despite the multivariate regression model containing all subscales being non-significant for adults—possibly due to the stable and protective scores failing to reach significance individually—the AUC values for the multivariate models suggest comparable accuracy between groups. That is, the initial DRAOR subscales might be operating at varying strengths when predicting youth and adult reconviction, but overall accuracy is just as good for both groups.

To what extent does the proximal DRAOR assessment predict criminal reconviction for both groups? Next, scores taken from the proximal DRAOR assessment were examined. As previously mentioned, the proximal DRAOR score serves as a measure of an offender's dynamic risk at the closest possible time to reconviction, or to data extraction for offenders who were not reconvicted. I expected these scores to be good predictors of reconviction because they should, in theory, be capturing dynamic risk that may have been relevant at the time of reconviction.

Table 7

Mean Proximal DRAOR Scores for all Subscales and Total Score: Criminal Reconviction

Score	Youth *M* (*SD*)	Adult *M* (*SD*)
Acute subscale	4.11 (2.47)**	3.20 (2.22)
Stable subscale	4.82 (2.58)	4.55 (2.35)
Protective subscale	6.80 (2.86)	7.85 (2.28)**
Total score	2.13 (6.86)*	-0.10 (5.94)

** $p < .01$, * $p < .05$

Table 7 displays the mean proximal DRAOR scores for youth and adult offenders. As with the initial scores, compared to adults, youth also had higher acute, $t(198) = 2.74$, $p = .007$, $d = 0.39$, $M_{diff} = 0.91$, 95% CI [0.26, 1.56], and total scores, $t(198) = 2.46$, $p = .015$, $d = 0.35$, $M_{diff} = 2.23$, 95% CI [0.44, 4.02], and lower protective scores, $t(198) = -2.87$, $p = .005$, $d = -0.41$, $M_{diff} = -1.05$, 95% CI [-1.77, -0.33]. The difference between groups was greatest for the protective subscale, which had the highest Cohen's d value ($d = -0.41$). However, again, these differences are small (Cohen, 1992). There was no statistical difference between scores on the stable subscale, $t(198) = 0.78$, $p = .439$, $d =$

0.11, M_{diff} = 0.27, 95% CI [-0.42, 0.96]. Of interest, the adult total score was negative; indicating that, on average, adult offenders scored higher on the protective subscale than on the acute and stable risk subscales at their proximal assessment. This result is promising, as most adults were not reconvicted.

For both youth and adults the proximal DRAOR subscales and total score were correlated with each other in the expected directions, again implying convergent validity (see Table 8). Correlations between proximal scores on the three subscales and total score were larger than those for the initial score for both groups, suggesting stronger relationships between proximal subscale and total scores.

Table 8

Correlations Between Proximal DRAOR Subscale Scores and Total Score for Youth and Adults: Criminal Reconviction

	Acute	Stable	Protective	Total score
Youth				
Acute	1			
Stable	.59**	1		
Protective	-.59**	-.71**	1	
Total score	.83**	.88**	-.89**	1
Adult				
Acute	1			
Stable	.61**	1		
Protective	-.61**	-.67**	1	
Total score	.85**	.88**	-.88**	1

** $p < .01$

Univariate predictive validity. The four proximal DRAOR predictors (the three subscale scores and the total score) were each entered as a covariate in four separate univariate Cox regressions to test their ability to predict criminal reconviction without taking shared variance into account. The results of these analyses are shown in Table 9.

For youth, all proximal subscale scores were able to predict criminal reconviction, acute: $\chi^2(1, n = 100) = 13.17, p < .001$; stable: $\chi^2(1, n = 100) = 24.46, p < .001$; protective: $\chi^2(1, n = 100) = 12.50, p < .001$; as was the total score: $\chi^2(1, n = 100) = 21.80, p < .001$. For adults, the acute, $\chi^2(1, n = 100) = 9.98, p = .002$; stable, $\chi^2(1, n = 100) = 8.20, p = .004$; and total score, $\chi^2(1, n = 100) = 8.98, p = .003$ were able to predict criminal reconviction, and protective, $\chi^2(1, n = 100) = 3.03, p = .082$ was not.

Table 9

Univariate Regression Models Containing Proximal DRAOR Subscale Scores and Total Score Predicting Criminal Reconviction

Model for proximal DRAOR scores	β (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]	AUC [95% CI]
Youth					
Acute score	.17 (.05)	12.75**	1.19	[1.08, 1.31]	.68** [.56, .79]
Stable score	.23 (.05)	23.20**	1.26	[1.15, 1.39]	.74** [.64, .84]
Protective score	-.15 (.04)	12.27**	0.86	[0.79, 0.93]	.65* [.54, .76]
Total score	.08 (.02)	21.01**	1.09	[1.05, 1.12]	.74** [.63, .84]
Adult					
Acute score	.22 (.07)	9.59**	1.25	[1.09, 1.44]	.63* [.51, .75]
Stable score	.22 (.08)	8.02**	1.25	[1.07, 1.45]	.64* [.53, .76]
Protective score	-.12 (.07)	3.02	0.88	[0.77, 1.02]	.59 [.47, .70]
Total score	.08 (.03)	8.89**	1.09	[1.03, 1.15]	.65* [.54, .76]

** $p < .01$, * $p < .05$

For youth, the stable subscale had the highest HR with every one-unit increase in stable scores placing an offender at a 26% greater likelihood of reconviction. The stable subscale performed with high accuracy, AUC = .74, 95% CI [.64, .84].

For adults, the HR for both the stable and acute subscales suggests that a one-unit increase in risk scores corresponds to a 25% increased likelihood of reconviction, each score operating with low to moderate accuracy. As with initial protective scores, proximal protective scores did not predict adult criminal reconviction at a level of statistical significance, and the AUC value suggested the protective subscale was operating at a level no better than chance.

Comparing accuracy between groups, youth have slightly higher AUC values than adults for all subscales and the total score. This is particularly true for the stable subscale, which has the largest between-group difference in AUC values (.10). However, when comparing the confidence intervals around these values, substantial overlap indicates differences in accuracy between groups are not likely to be great, or may not exist.

Multivariate predictive validity. Table 10 displays the multivariate regression models for youth and adults showing the contribution of each proximal DRAOR subscale when accounting for shared variance. For youth, the multivariate model containing all predictors was significant, $\chi^2(3, n = 100) = 25.60, p < .001$. The stable score was the only unique predictor in the model for youth, operating at a level above the acute and protective subscales. This means that, despite being significant predictors on their own, the acute and protective subscales could not tell anything more about who would or would not be reconvicted than stable subscale could. This suggests most of the predictive ability of the proximal DRAOR total score is coming from the stable subscale.

Table 10

Multivariate Regression Models Containing Proximal DRAOR Subscale Scores Predicting Criminal Reconviction

Model for proximal DRAOR scores	β (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]	AUC [95% CI]
Youth					
Acute score	.06 (.06)	0.92	1.06	[0.94, 1.20]	
Stable score	.21 (.07)	8.10**	1.23	[1.07, 1.43]	.74** [.64, .85]
Protective score	.01 (.07)	0.05	1.01	[0.89, 1.16]	
Adult					
Acute score	.17 (.10)	3.20	1.19	[0.98, 1.44]	
Stable score	.17 (.10)	2.62	1.18	[0.97, 1.45]	.66** [.55, .78]
Protective score	.08 (.10)	0.59	1.08	[0.89, 1.31]	

** $p < .01$

For adults the multivariate regression model with all predictors was also significant, $\chi^2(3, n = 100) = 12.30, p = .006$. No one subscale was able to predict over and above the others, suggesting overlapping variance between the acute and stable subscales, which were significant in their respective univariate models. However, the lower bound of the confidence intervals around the HR border very close to 1 for both risk subscales suggesting a difference in variance accounted-for could exist, and might be detected with greater statistical power. Again, substantially overlapping confidence intervals indicate there is likely no difference in the accuracy of the multivariate models between groups, despite the AUC value being higher for youth.

Summary. For youth, higher proximal scores on all risk subscales are associated with a greater likelihood of reconviction, and higher proximal scores on the protective subscale are associated with a lesser likelihood of reconviction. The stable subscale has the largest HR and appears to be predicting criminal reconviction at a level above the acute and protective subscales. For adults, proximal scores on the risk subscales are comparable in their ability to predict criminal reconviction, while scores on the protective subscale are not good predictors.

In contrast to the previous youth and adult multivariate models containing the initial subscale scores, the multivariate models containing the proximal scores did show a difference in the AUC values between youth and adults. However, overlapping confidence intervals suggest that plausibly there is no difference.

Which DRAOR score is a better predictor: initial or proximal? I expected the proximal DRAOR scores to be better predictors of reconviction than the initial DRAOR scores as, in theory, dynamic assessments should improve in accuracy with regular updates. It follows that offenders who were reconvicted should have higher DRAOR risk scores for the assessment closest to reconviction and lower protective scores, while those who were not reconvicted should have lower risk scores and higher protective scores. Evidence from Cox regressions thus far suggests proximal scores may indeed be stronger predictors than the initial DRAOR scores, as evidenced by slightly higher HRs and AUC values. In order to answer this question, the two scores must be directly compared in a Cox regression.

Initial and proximal scores for all subscales were highly positively correlated with each other (see Table 11). This is unsurprising as they are meant to be measuring a related concept. Offenders who had higher initial scores also tended to have higher proximal scores. Correlations between the different initial and proximal scores for both groups ranged from moderate to strong, while correlations between scores on the same subscales for the two assessments (in bold) were strong.

Table 11

Correlations Between Initial and Proximal DRAOR Subscale Scores and Total Score for Youth and Adults: Criminal Reconviction

	Initial Acute	Initial Stable	Initial Protective	Initial Total score
Youth				
Proximal Acute	**.64****	.37**	-.39**	.59**
Proximal Stable	.40**	**.63****	-.43**	.60**
Proximal Protective	-.47**	-.51**	**.76****	-.72**
Proximal Total score	.58**	.58**	-.62**	**.74****
Adult				
Proximal Acute	**.66****	.37**	-.44**	.62**
Proximal Stable	.42**	**.61****	-.49**	.61**
Proximal Protective	-.33**	-.40**	**.71****	-.57**
Proximal Total score	.54**	.53**	-.63**	**.69****

** $p < .01$

Incremental validity: Initial and proximal DRAOR scores. The initial and proximal scores for the four DRAOR predictors (the three subscale scores and the total score) were entered together as covariates in the first block of four separate multivariate Cox regressions. In the cases where the initial score did not predict reconviction, and the proximal score did, the proximal was the better predictor. Adults' initial and proximal protective, and initial stable subscale scores did not predict criminal reconviction. Therefore, their stable and protective subscales were not included here but can be found in Appendix B. Regressions are reported here if both the initial and proximal scores were found to be predicting reconviction in univariate regressions.

For youth, all Cox regression models containing both initial and proximal subscale scores predicted criminal reconviction: acute, $\chi^2(2, n = 100) = 13.54$, $p = .001$; stable, $\chi^2(2, n = 100) = 24.52$, $p < .001$; protective, $\chi^2(2, n = 100) = 13.25$, $p = .001$; as did the total score model, $\chi^2(2, n = 100) = 22.00$, $p < .001$. For adults, both the acute, $\chi^2(2, n = 100) = 10.42$, $p = .005$; and total score, $\chi^2(2, n = 100) = 9.00$, $p = .011$, models were significant. These results are expected, as all subscales were predictive in their univariate models.

Table 12

Multivariate Regression Models Containing Initial and Proximal DRAOR Subscale Scores and Total Score Predicting Criminal Reconviction

Multivariate model for initial and proximal DRAOR scores	β (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]
Youth				
Acute score				
Initial	.04 (.06)	0.37	1.04	[0.92, 1.16]
Proximal	.15 (.06)	6.80**	1.17	[1.04, 1.31]
Stable score				
Initial	-.06 (.08)	0.65	0.94	[0.81, 1.09]
Proximal	.27 (.07)	14.73**	1.31	[1.14, 1.51]
Protective score				
Initial	-.06 (.07)	0.63	0.94	[0.82, 1.09]
Proximal	-.12 (.06)	4.53*	0.88	[0.79, 0.99]
Total score				
Initial	.01 (.03)	0.03	1.01	[0.95, 1.07]
Proximal	.08 (.03)	9.12**	1.08	[1.03, 1.14]
Adult				
Acute score				
Initial	.04 (.08)	0.22	1.04	[0.89, 1.22]
Proximal	.19 (.10)	3.93*	1.21	[1.00, 1.47]
Total score				
Initial	-.01 (.04)	0.02	1.00	[0.92, 1.08]
Proximal	.09 (.04)	4.58*	1.09	[1.01, 1.18]

** $p < .01$, * $p < .05$

An inspection of the HRs in Table 12 shows that, for youth, the proximal score was accounting for more of the variance in predicting criminal reconviction than the initial score for all subscales and the total score. Therefore, proximal scores appear to be stronger predictors of youth criminal reconviction than initial scores in all cases. This difference is particularly evident in the case of proximal stable scores for youth. The proximal stable score not only has the highest HR, with a one-unit increase in proximal scores corresponding with a 31% increase in the likelihood of criminal reconviction when accounting for initial scores, but its confidence interval does not overlap with that of the initial score, indicating a statistical difference in predictive ability between the two assessments.

For adults the proximal acute and total scores also accounted for more of the variance in reconviction than the initial scores, suggesting proximal scores are better predictors than initial scores.

Summary. Multivariate Cox regressions containing both initial and proximal scores as predictors show that for both groups, proximal scores are able to give more information than initial scores about which offenders are likely to be reconvicted for all subscales and the total score. The adult protective subscale is the exception to this, with scores taken from both initial and proximal assessments being equally poor predictors of reconviction.

To what extent do DRAOR scores change over time and does change predict reconviction?

If the DRAOR is a truly dynamic measure of risk, it follows that scores should change over time in response to risk level. A paired samples t-test was undertaken to determine whether there were significant differences between the initial and proximal scores that might indicate change had occurred between the two assessments. As seen in Table 13, proximal scores were significantly lower than initial DRAOR scores on both risk subscales and the total score for both youth: acute, $t(99)$ = 4.39, $p < .001$, d = 0.44, M_{diff} = 0.95, 95% CI [0.52, 1.38]; stable, $t(99)$ = 4.81, $p < .001$, d = 0.48, M_{diff} = 1.00, 95% CI [0.59, 1.41]; total score, $t(99)$ = 5.45, $p < .001$, d = 0.55, M_{diff} = 2.56, 95% CI [1.63, 3.49]; and adults: acute, $t(99)$ = 5.71, $p < .001$, d = 0.57, M_{diff} = 1.12, 95% CI [0.73, 1.51]; stable, $t(99)$ = 4.98, $p < .001$, d = 0.50, M_{diff} = 0.95, 95% CI [0.57, 1.33]; total score, $t(99)$ = 6.91, $p < .001$, d = 0.69, M_{diff} = 3.06, 95% CI [2.18, 3.94]. Proximal protective scores were higher than initial protective scores for both groups: youth protective, $t(99)$ = -3.24, p = .002, d = -0.32, M_{diff} = -0.61, 95% CI [-0.98, -0.24]; adult protective, $t(99)$ = -5.96, $p < .001$, d = -0.60, M_{diff} = -0.99, 95% CI [-1.32, -0.66]. These findings suggest that, on average, risk had decreased between the two assessments, and protective factors had increased. The Cohen's *d* effect sizes of all subscale scores were larger for adults, indicating the magnitude of change within the adult group from the initial to proximal assessment was greater than that within the youth group.

To what extent do DRAOR scores change over time? Using the initial and proximal scores, a change score was calculated for each of the DRAOR subscales and the total score. Subtracting the proximal DRAOR scores from the initial DRAOR scores gave a number representing the average direction and magnitude of change between the two assessments.[12] A positive number meant that risk had decreased between the two assessments, a negative number meant risk had increased, and zero represented no difference between assessment scores (for protective factors, a negative number meant protective factors had increased and a positive number meant protective factors had decreased).

Table 13

Mean Change Made Between Initial and Proximal DRAOR Assessments: Criminal Reconviction

	Mean initial score (*SD*)	Mean proximal score (*SD*)	Mean change score (*SD*)	Range of change	No change (%)
Youth					
Acute	5.06 (2.63)**	4.11 (2.47)	0.95 (2.16)	[-4, 7]	25
Stable	5.82 (2.22)**	4.82 (2.58)	1.00 (2.08)	[-3, 8]	50
Protective	6.19 (2.38)	6.80 (2.86)**	-.61 (1.89)	[-6, 6]	62
Total Score	4.69 (5.81)**	2.13 (6.86)	2.56 (4.70)	[-8, 16]	25
Adult					
Acute	4.32 (2.51)**	3.20 (2.22)	1.12 (1.96)	[-3, 7]	31
Stable	5.50 (1.86)**	4.55 (2.35)	0.95 (1.91)	[-4, 8]	50
Protective	6.86 (2.01)	7.85 (2.28)**	-0.99 (1.66)	[-6, 3]	47
Total Score	2.96 (5.21)**	-0.10 (5.94)	3.06 (4.43)	[-7, 17]	20

** $p < .01$

Table 13 shows the mean change scores by group, as well as the range of change (i.e. the limits within which change occurred). To illustrate, for the acute subscale predicting criminal reconviction, the most any youth increased in acute risk between the initial and proximal assessments was four units, the most any youth decreased in acute risk was seven units.[13] The most any adult increased in acute risk was three units, while the most any adult decreased in risk was seven units. For the protective subscale, the most any youth increased in protective factors was six units (more assets), and the most any youth decreased in protective factors was also six units (fewer assets). The most any adult increased in protective factors (more assets) was six units, and the most any adult decreased in protective factors (fewer assets) was three units.

[12] Change' refers to mean change; as of course scores could be the same at both assessments, suggesting no change, when in fact change could have occurred in the assessments in between.

[13] The acute subscale has slightly greater room for change compared to the stable and protective subscales as it contains one extra item (7 items compared to the protective and stable subscales' 6). Being a composite of the 3 subscales, the total scores also had greater room for change, so total score change appears larger.

Although the Cohen's *d* values for paired samples t-tests between initial and proximal scores suggested the difference between the two scores was larger for adults, independent samples t-tests found no statistical difference between youth and adults' change scores for any of the subscales or total score, acute: $t(198) = -0.58$, $p = .561$, $d = -0.08$, $M_{diff} = -0.17$, 95% CI [-0.75, 0.41]; stable: $t(198) = 0.18$, $p = .860$, $d = 0.03$, $M_{diff} = 0.05$, 95% CI [0.28, -0.51]; protective: $t(198) = 1.51$, $p = .132$, $d = 0.21$, $M_{diff} = 0.38$, 95% CI [-0.12, 0.88]; and total score: $t(198) = -0.77$, $p = .440$, $d = -0.11$, $M_{diff} = -0.50$, 95% CI [-1.77, 0.77].

Table 13 also shows the proportion of youth and adults whose scores did not change between assessments. Half of the youth and adults (50%) made no change on the stable subscale. The acute subscale had the smallest number of youth making no change (25%) compared to the stable (50%) and protective subscales (62%), which is promising because the acute subscale is theorised to be the most changeable dynamic subscale. Despite a plurality of youth making no change on the protective subscale, recall both initial and proximal scores were good predictors. The protective subscale has not been found thus far to predict adult criminal reconviction at either assessment; perhaps this is because almost half of the adults (47%) made no change in their scores on this subscale.

Does change predict criminal reconviction? In order to investigate the extent to which changes in DRAOR scores from the initial to proximal assessment were able to predict youth and adult criminal reconviction, it was first necessary to investigate the relationship between initial scores and change scores using correlation analysis. In brief, all change scores had weak to moderate positive correlations with their initial scores. However, youth protective and total score correlations failed to reach a level of statistical significance, despite still being correlated in a positive direction. For the risk subscales and total score, a positive correlation indicated that greater decreases in risk were associated with higher initial risk scores. For the protective subscale, a positive correlation indicated that greater increases in protective factors were associated with lower initial protective scores. Put simply, as expected, the amount of change made on the risk and protective subscales was limited by an offender's initial score (e.g., an offender with a score of 2 on the acute subscale could only decrease in risk by a potential 2 points, and an offender with a score of 14 could decrease in risk by a potential, albeit unlikely, 14 points). These correlation tables can be found in Appendix C.

Next, the four change scores for the DRAOR subscales and total score were each entered as a covariate in the second block of four separate univariate Cox regressions. The first block contained the respective initial DRAOR score for each subscale/total score, because, as mentioned, an offender's initial assessment limited the amount of change able to be made, and the purpose was to determine whether change would predict reconviction at a level above an offender's initial risk. In other words, after controlling for baseline dynamic risk, can change made on the DRAOR predict criminal reconviction?

Table 14

Multivariate Regression Models Containing Change Scores Controlling for the Initial DRAOR Subscale Scores and Total Score Predicting Criminal Reconviction

Model for change scores	*β* (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]	AUC [95% CI]
Youth					
Initial acute score	.19 (.06)	11.78**	1.21	[1.08, 1.35]	.67** [.56, .79]
Acute change score	-.15 (.06)	6.80**	0.86	[0.77, 0.96]	
Initial stable score	.21 (.06)	14.34**	1.23	[1.11, 1.38]	.74** [.64, .85]
Stable change score	-.27 (.07)	14.73**	0.76	[0.66, 0.88]	
Initial protective score	-.18 (.06)	10.67**	0.83	[0.75, 0.93]	.65* [.53, .76]
Protective change score	.12 (.06)	4.53*	1.13	[1.01, 1.27]	
Initial total score	.08 (.02)	14.82**	1.09	[1.04, 1.13]	.73** [.63, .84]
Total score change score	-.08 (.03)	9.12**	0.92	[0.88, 0.97]	
Adult					
Initial acute score	.23 (.07)	9.95**	1.26	[1.09, 1.45]	.63* [.51, .75]
Acute change score	-.19 (.10)	3.93*	0.82	[0.68, 1.00]	
Initial stable score	.19 (.09)	4.17*	1.21	[1.01, 1.45]	65* [.54, .76]
Stable change score	-.26 (.11)	6.27*	0.77	[0.63, 0.95]	
Initial protective score	-.12 (.08)	2.03	0.89	[0.76, 1.04]	.59 [.47, .70]
Protective change score	.14 (.11)	1.78	1.15	[0.94, 1.42]	
Initial total score	.08 (.03)	7.28**	1.09	[1.02, 1.15]	.65* [.54, .76]
Total score change score	-.09 (.04)	4.58*	0.92	[0.85, 0.99]	

** $p < .01$, * $p < .05$

For youth, all Cox regression models containing both initial DRAOR scores and change scores predicted criminal reconviction: acute, $\chi^2(2, n = 100) = 13.54$, $p = .001$; stable, $\chi^2(2, n = 100) = 24.52$, $p < .001$; protective, $\chi^2(2, n = 100) = 13.25$, $p = .001$; as did the total score model, $\chi^2(2, n = 100) = 22.00$, $p < .001$. Recall that for the acute, stable, and total scores, a positive change score represents a decrease in risk between the two assessments. Therefore, a hazard ratio of less than 1 means that for every one unit of change an offender made in the direction of a decrease in risk, the likelihood of criminal reconviction also decreased (when taking into account the variation in initial DRAOR scores). To illustrate, for youth, stable change has the highest hazard ratio; where one unit of change in the direction of a decrease in stable risk (e.g., from 1 to 2) is associated with a 24% reduction in the likelihood of criminal reconviction, HR = 0.76, 95% CI [0.66, 0.88], (see Table 14). For the protective subscale, a positive number represents a decrease in protective scores between the two assessments. So, every one unit of change made on the protective subscale in the direction of a decrease in protective factors (e.g., from -2 to -1) is associated with an increase in the likelihood of criminal reconviction. To illustrate, for youth, one unit of change in the direction of a decrease in protective factors between initial and proximal assessments is associated with a 13% increase in the likelihood of criminal reconviction, HR = 1.13, 95% CI [1.01, 1.27].

For adults, the acute, $\chi^2(2, n = 100) = 10.42$, $p = .005$; stable, $\chi^2(2, n = 100) = 8.32$, $p = .016$; and total score, $\chi^2(2, n = 100) = 9.00$, $p = .011$ models predicted criminal reconviction but the protective model did not, $\chi^2(2, n = 100) = 3.06$, $p = .217$. As with youth, stable subscale change also had the highest HR for adults, with one unit of change in the direction of a decrease in stable risk associated with a 23% reduction in the likelihood of criminal reconviction, HR = 0.77, 95% CI [0.63, 0.95].[14]

For this analysis, AUC values are interpreted differently also. For example, when considering the stable subscale, there is a 74% likelihood that a randomly selected reconvicted youth would have a lower change score than a randomly selected youth who was not reconvicted—and thus have made less change in the direction of a decrease in stable risk—when taking into account the variation in initial stable scores. Although stable and total score change both appear to be trending towards higher

[14] It is worth mentioning that for adults, initial stable DRAOR scores could not predict criminal reconviction on their own but are now adding something unique to the multivariate model. This could be due to the initial score being an integral part of the stable change score, and suggests that change scores possibly mediate the relationship between initial stable scores and criminal reconviction. A fuller discussion of this is beyond the scope of this research.

accuracy for youth when compared with adults, substantial overlap in confidence intervals around the youth and adult AUC values suggests that accuracy is comparable between groups for these subscales.

Summary. Multivariate Cox regressions examining change scores while controlling for initial scores show that for youth, change made on all DRAOR subscales predicts criminal reconviction. This is an interesting result, as recall 50% of youth made no change on the stable subscale and 62% made no change on the protective subscale. For adults, change on all subscales predicted reconviction with the exception of the protective subscale. For both groups, change made on the stable subscale looked to have the largest effect on the likelihood of criminal reconviction, with greater decreases in risk corresponding to greater decreases in the likelihood of reconviction. Despite the noticeable difference in the accuracy of stable and total change scores at predicting criminal reconviction—with both scores performing with higher accuracy for youth— accuracy was comparable across the youth and adult groups.

Is the DRAOR a better predictor of criminal reconviction than the RoC*RoI?

As change was found to be predicting criminal reconviction for both groups (with the exception of adult protective change) the next step was to determine whether the DRAOR could operate over and above a measure of static risk. This analysis used offenders' RoC*RoI scores in addition to their initial and proximal DRAOR scores on each of the subscales and total score. The RoC*RoI changes only in response to age and further convictions, meaning that scores are more likely to increase than decrease over time. In contrast, previous analyses have shown that on average, DRAOR scores decreased over time, suggesting that probation officers were responding to reductions in dynamic risk that they witnessed. To what extent would a tool measuring solely static factors be able to predict criminal reconviction? Furthermore, how would it perform when directly compared to a tool containing dynamic subscales, the scores on which have been shown to change over time?

Does the RoC*RoI predict criminal reconviction? First it was necessary to establish whether the RoC*RoI predicted reconviction on its own. Recall, the groups were matched on their RoC*RoI scores, so these did not significantly differ between groups even after excluding offenders (and their matched pairs) who were reconvicted before their initial DRAOR assessment. For the sample of 100 youth and 100 adults, the youth mean score was .36 (*SD* = .16, range: .10 to .82) and the adult mean score was .37 (*SD* = .16, range: .02 to .71).

RoC*RoI score was entered as a covariate into the first block of a Cox regression. As with previous analyses, the time variable entered was survival days (days to criminal reconviction or data extraction for those who were not reconvicted), and the status variable was the presence or absence of a criminal reconviction.

The RoC*RoI did not predict criminal reconviction for youth, $\chi^2(1, n = 100) = 1.41, p = .235$ (see Table 15). Taken together, the HR from the Cox regression and corresponding AUC value suggest that the RoC*RoI's ability to distinguish between youth who were reconvicted and those who were not is no better, or may even be worse than chance. The RoC*RoI did predict criminal reconviction for adults, however, $\chi^2(1, n = 100) = 5.89, p = .015$, with a hazard ratio of 12.89, meaning a 12 unit increase in RoC*RoI scores is associated with an 89% increase in the likelihood of criminal reconviction. As an example, a 12 unit increase would be an offender's score of .37 increasing to .49.

Table 15

*Univariate Regression Models Containing RoC*RoI Scores Predicting Criminal Reconviction*

Model for RoC*RoI scores	β (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]	AUC [95% CI]
Youth					
RoC*RoI	0.86 (.72)	1.41	2.36	[0.57, 9.70]	.57 [.44, .69]
Adult					
RoC*RoI	2.56 (1.06)	5.79*	12.89	[1.61, 103.34]	.62* [.51, .73]

* $p < .05$

Therefore, the RoC*RoI is able to predict adult but not youth criminal reconviction. However, the confidence intervals around the HR for adults are large suggesting the precision of the analysis is low, and the adult AUC value represents a level of relatively low accuracy so these results must be interpreted with those factors in mind.

Does the DRAOR add incremental validity over the RoC*RoI? The next step was to determine how the RoC*RoI would operate, when accounting for variance shared with the DRAOR.

This analysis could only be done for the adult group, as incremental validity is contingent on the RoC*RoI being a significant predictor individually. See Appendix D for the youth results.

First, correlation analyses were performed to check how related RoC*RoI scores were to both the initial and proximal scores for all DRAOR subscales and the total score (see Table 16). Adults who had high RoC*RoI scores also tended to have high initial and proximal acute scores and initial total scores, although the size of these correlations were small. All other correlations did not reach a level of statistical significance, suggesting the RoC*RoI is measuring something different to both the initial and proximal DRAOR scores.

Table 16

*Correlations Between RoC*RoI Scores and Initial (I) and Proximal (P) DRAOR Subscale Scores and Total Score for Adults: Criminal Reconviction*

	Acute		Stable		Protective		Total score	
	I	P	I	P	I	P	I	P
Adult								
RoC*RoI	.27**	.27**	.12	.11	-.16	-.07	.24*	.17

** $p < .01$, * $p < .05$

All DRAOR subscale scores and RoC*RoI scores were entered individually as covariates in the first block of a Cox regression. Entering the scores from both DRAOR assessments would control for baseline dynamic risk as well as including a more up-to-date assessment (which, overall, was found to be the better predictor).

All adult regression models containing initial and proximal DRAOR scores and the RoC*RoI score were significant, acute, $\chi^2(3, n = 100) = 14.59$, $p = .002$; stable, $\chi^2(3, n = 100) = 14.62$, $p = .002$; protective, $\chi^2(3, n = 100) = 9.82$, $p = .020$; as was the total score model, $\chi^2(3, n = 100) = 14.82$, $p = .002$. Although the RoC*RoI looks to be a better predictor than the acute and protective subscales, in these cases the confidence intervals around the proximal DRAOR's HR for both subscales only just cross 1, suggesting greater statistical power might yield a statistically significant result (see Table 17). Proximal stable subscale scores and RoC*RoI scores were both significant predictors in their model, suggesting that, after taking shared variance into account, they were each able to give unique information about the likelihood of criminal reconviction. The proximal total scores also showed this pattern, arguably being driven primarily by the stable subscale. In order to test whether the stable and total DRAOR scores added incremental validity over the RoC*RoI because they were simply measuring dynamic factors or because they were measuring up-to-date dynamic factors, regressions using just the initial scores and RoC*RoI scores were run. Although not reported here in detail, these analyses showed the RoC*RoI was a better predictor than the initial stable and total scores, suggesting that the incremental validity of the DRAOR, over the RoC*RoI, comes from the fact that DRAOR scores are up-to-date dynamic predictors rather than simply because they are measures of dynamic factors.

Table 17

*Multivariate Regression Model Containing RoC*RoI Scores and DRAOR Subscale Scores and Total Score Predicting Adult Criminal Reconviction*

Adult models for incremental predictive validity	β (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]	AUC [95% CI]
Acute model					
Initial acute score	.04 (.08)	0.24	1.04	[0.89, 1.22]	
Proximal acute score	.18 (.09)	3.50	1.19	[0.99, 1.44]	.65* [.54, .76]
RoC*RoI score	2.36 (1.10)	4.60*	10.55	[1.23, 90.74]	
Stable model					
Initial stable score	-.09 (.12)	0.57	0.92	[0.73, 1.15]	
Proximal stable score	.30 (.11)	7.86**	1.34	[1.09, 1.65]	.66** [.55, .77]
RoC*RoI score	3.19 (1.17)	7.39**	24.33	[2.44, 242.97]	
Protective model					
Initial protective score	.06 (.12)	0.29	1.07	[0.85, 1.34]	
Proximal protective score	-.18 (.10)	2.91	0.84	[0.69, 1.03]	.63* [.52, .74]
RoC*RoI score	2.83 (1.10)	6.59*	16.96	[1.95, 147.22]	
Total score model					
Initial total score	-.01 (.04)	0.09	0.99	[0.91, 1.07]	
Proximal total score	.09 (.04)	5.41*	1.10	[1.02, 1.19]	.66* [.55, .77]
RoC*RoI score	2.78 (1.12)	6.18*	16.04	[1.80, 143.07]	

** $p < .01$, * $p < .05$

Summary. All proximal DRAOR scores are better predictors of youth criminal reconviction after accounting for any variance measured by the RoC*RoI and initial DRAOR scores. Proximal stable scores and proximal total scores give different information than the RoC*RoI about which adults are likely to be reconvicted, and are thus equally useful predictors. RoC*RoI scores appear to be better predictors of adult criminal reconviction than the initial and proximal acute and protective subscale scores.

Chapter 4

Breach Reconviction

To what extent does the DRAOR predict breach reconviction outcomes for youth and adults?

In addition to criminal reconviction, I also wanted to investigate whether the DRAOR was able to predict breach reconviction. The previous chapter established that the DRAOR was able to predict criminal reconviction, and indeed the DRAOR is designed to predict this type of reconviction. However, it is not specifically designed to predict other types of reconviction. Investigating the DRAOR's ability to predict breach reconviction will provide some insight as to whether DRAOR scores are measuring risk information relevant to an offender's breach of sentence conditions, which could have clinical utility.

As done previously, offenders were only included in the following analyses if their breach reconviction occurred after their initial DRAOR assessment, and their initial and proximal scores were from different assessments. After also excluding the matched pairs of offenders who did not meet these criteria, a sample of 101 youth and 101 adults remained to be used in subsequent breach reconviction analyses.

To what extent does the initial DRAOR score predict breach reconviction outcomes for both groups? Table 18 shows the mean initial DRAOR scores for both groups. As different people were excluded for each outcome, these numbers are slightly different to those used in the previous chapter.

Table 18

Mean Initial DRAOR Scores for all Subscales and Total Score: Breach Reconviction

Score	Youth *M* (*SD*)	Adult *M* (*SD*)
Acute subscale	5.11 (2.65)	4.49 (2.57)
Stable subscale	5.90 (2.24)	5.64 (1.92)
Protective subscale	6.08 (2.42)	6.61 (2.02)
Total score	4.93 (6.00)	3.51 (5.20)

Although the pattern of scores was similar to the mean initial scores for criminal reconviction (with youth having higher mean risk scores and lower protective scores), Cohen's *d* values indicated effect sizes were small, and there was not a statistical difference between the groups for any of the scores: acute, $t(200) = 1.70$, $p = .091$, $d = 0.24$, $M_{diff} = 0.62$, 95% CI [-0.10, 1.35]; stable, $t(200) = 0.88$, $p = .382$, $d = 0.12$, $M_{diff} = 0.26$, 95% CI [-0.32, 0.84]; protective, $t(200) = -1.71$, $p = .090$, $d = -0.24$, $M_{diff} = -.53$, 95% CI [-1.15, 0.08]; and total score, $t(200) = 1.79$, $p = .075$, $d = 0.25$, $M_{diff} = 1.42$, 95% CI [-0.14, 2.97].

Subscale scores were highly correlated with each other for both groups, with the acute and protective subscales showing moderate correlation with each other (see Table 19). Offenders who scored higher on the acute subscale also tended to score higher on the stable subscale and offenders who scored higher on the risk subscales tended to also score lower on the protective subscale. This high level of association suggests the subscales and total score are measuring a similar concept. Despite this high relatedness, VIF analyses indicated no significant issues with multicollinearity (refer back to the first research question in the previous chapter for information about VIF analysis).

Table 19

Correlations Between Initial DRAOR Subscale scores and Total Score for Youth and Adults: Breach Reconviction

	Acute	Stable	Protective	Total score
Youth				
Acute	1			
Stable	.51**	1		
Protective	-.46**	-.56**	1	
Total score	.82**	.83**	-.82**	1
Adult				
Acute	1			
Stable	.50**	1		
Protective	-.34**	-.55**	1	
Total score	.81**	.83**	-.76**	1

** $p < .01$

Univariate predictive validity. Would offenders who had received higher initial DRAOR scores be at an increased likelihood of breach reconviction (or a decreased likelihood, in the case of the protective subscale)? The three subscale scores and the total score were each entered as a covariate in the first block of four separate univariate Cox regressions to test their predictive validity. The time variable entered was survival days (days to breach reconviction or end of sentence for those who were not reconvicted), and the status variable was the presence or absence of a breach reconviction.

Table 20 shows the results of these regressions investigating the predictive validity of all subscale scores without taking shared variance into account (combined into one table). For youth, none of the initial subscale scores, nor the total score were able to predict breach reconvictions, acute: $\chi^2(1, n = 101) = 2.00, p = .158$; stable: $\chi^2(1, n = 101) = 0.76, p = .385$; protective: $\chi^2(1, n = 101) = 0.14, p = .709$; total score: $\chi^2(1, n = 101) = 1.25, p = .264$. For adults, initial scores on the acute subscale were the only ones that predicted breach reconviction, $\chi^2(1, n = 101) = 7.84, p = .005$, while scores on the stable, $\chi^2(1, n = 101) = 1.18, p = .278$; and protective subscales, $\chi^2(1, n = 101) = 0.26, p = .613$; and total scores, $\chi^2(1, n = 101) = 3.76, p = .052$, did not.

Table 20

Univariate Regression Models Containing Initial DRAOR Subscale Scores and Total Score Predicting Breach Reconviction

Model for initial DRAOR scores	β (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]	AUC [95% CI]
Youth					
Acute score	.08 (.06)	1.98	1.09	[0.97, 1.22]	.60 [.48, .72]
Stable score	.06 (.07)	0.76	1.07	[0.92, 1.23]	.55 [.44 .67]
Protective score	-.03 (.07)	0.14	0.98	[0.85, 1.12]	.52 [.40, .64]
Total score	.03 (.03)	1.25	1.03	[0.98, 1.09]	.57 [.45, .69]
Adult					
Acute score	.23 (.09)	7.41**	1.26	[1.07, 1.49]	.71** [.56, .86]
Stable score	.15 (.14)	1.18	1.16	[0.89, 1.52]	.57 [.41, .73]
Protective score	-.06 (.12)	0.26	0.94	[0.75, 1.19]	.57 [.42, .73]
Total score	.08 (.04)	3.81	1.08	[1.00, 1.17]	.66* [.51, .80]

** $p < .01$, * $p < .05$

A one-unit increase in adult acute scores corresponds to a 26% increase in the likelihood of a breach reconviction, HR = 1.26, 95% CI [1.07, 1.49]. However, the confidence intervals surrounding the HR are quite large for this subscale, indicating low precision and that this increased likelihood could plausibly be as low as 7% or as high as 49%. Although the adult total score did not reach a level of statistical significance as a predictor, its HR confidence interval only just includes 1, which taken together with the corresponding AUC value, suggests it is predicting at a level above chance. Because the total score is made up of the three subscales, the acute subscale might be driving this result as it was a significant predictor in its own univariate regression.

The acute score for adults is the only score—both within and across groups—that predicts breach reconviction with high accuracy, AUC = .71, 95% CI [.56, .86]. However, because the confidence intervals of the acute AUC values overlap for youth and adults, it is plausible that accuracy does not differ between groups.

Summary. Results from Cox regression analyses suggest initial DRAOR scores are poor predictors of breach reconviction for youth; all three subscales and the total score were unable to predict youth breach reconviction at a level above chance. For adults, only scores on the acute subscale and the total score were able to predict breach reconviction with accuracy above chance level. Despite the relationship between initial acute scores and breach reconviction appearing stronger for adults compared to youth, when accounting for reconviction base rates, accuracy is comparable between groups.

Multivariate predictive validity. Next, the subscales were examined in combination, to determine whether any of them could account for more of the variance in reconviction than the others. All initial DRAOR subscale scores (acute, stable, protective) were entered as covariates into the first block of a Cox regression (see Table 21).

For youth the multivariate regression model with all predictors remained non-significant, $\chi^2(3, n = 101) = 2.13$, $p = .545$, an expected result given that all univariate regressions were non-significant. Notice that when taking into account variance shared by all three subscales, the ability of the protective subscale worsens: a one-unit increase in protective scores (which in theory should insulate against reconviction) corresponds to a 3% increase in the likelihood of breach reconviction.

In contrast, for adults the multivariate regression model was statistically significant, $\chi^2(3, n = 101) = 8.18$, $p = .042$. Recall that in the previous univariate regression analyses the acute subscale was the only significant predictor, and here, in combination with the other two subscales, it maintains its significance, HR = 1.34, 95% CI [1.06, 1.70]. When taking shared variance into account, the acute subscale captures unique variance and predicts breach reconviction at a level above the other two subscales. However the confidence intervals around the HR are very large, suggesting low precision and that the true value of the HR could be as low as 1.06 or as high as 1.70. Interestingly, once all subscales are entered into the model together, the HRs of the stable and protective subscales reverse direction. Here, an increase in stable scores is associated with a decrease in likelihood of breach reconviction, HR = 0.97, 95% CI [0.67, 1.40] and an increase in protective scores is associated with an increase in the likelihood of reconviction, HR = 1.11, 95% CI [0.82, 1.49]. This finding is contrary to how the risk subscales are designed to operate and suggests that for adults, when taking into account shared variance, scores on the acute subscale are such strong predictors that they render scores on the other subscales poorer predictors. The adult model containing all subscales has a higher AUC value than the youth model; there is a 70% likelihood that a randomly chosen reconvicted adult will have a higher acute score than a randomly chosen adult who was not reconvicted, when taking into account variance shared with the other subscales. However overlapping confidence intervals suggest plausibly there is no difference in accuracy between groups.

Table 21

Multivariate Regression Models Containing Initial DRAOR Subscale Scores Predicting Breach Reconviction

Model for initial DRAOR scores	β (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]	AUC [95% CI]
Youth					
Acute score	.08 (.07)	1.30	1.08	[0.94, 1.24]	.62 [.50, .73]
Stable score	.03 (.09)	0.08	1.03	[0.86, 1.23]	
Protective score	.03 (.08)	0.09	1.03	[0.87, 1.21]	
Adult					
Acute score	.29 (.12)	6.11*	1.34	[1.06, 1.70]	.70* [.55, .86]
Stable score	-.03 (.19)	0.03	0.97	[0.67, 1.40]	
Protective score	.10 (.15)	0.46	1.11	[0.82, 1.49]	

* $p < .05$

Summary. For youth, initial DRAOR scores are poor predictors of breach reconviction, with accuracy plausibly no better than chance. For adults, the acute subscale appears to be the best predictor of breach reconviction. However, despite this apparent difference between youth and adults, AUC values for the multivariate models suggest when taking base rates into account, accuracy is likely comparable between the two groups.

To what extent does the proximal DRAOR assessment predict breach reconviction for both groups? Next, scores taken from the proximal DRAOR assessment were examined. Table 22 displays the mean proximal DRAOR scores for youth and adult offenders.

Table 22

Mean Proximal DRAOR Scores for all Subscales and Total Score: Breach Reconviction

Score	Youth *M* (*SD*)	Adult *M* (*SD*)
Acute subscale	4.12 (2.49)*	3.37 (2.26)
Stable subscale	4.94 (2.71)	4.80 (2.36)
Protective subscale	6.85 (2.81)	7.66 (2.20)*
Total score	2.21 (7.07)	0.50 (5.66)

* $p < .05$

Youth had statistically higher acute, $t(200) = 2.25$, $p = .025$, $d = 0.32$, $M_{diff} = 0.75$, 95% CI [0.09, 1.41], and lower protective scores, $t(200) = -2.29$, $p = .023$, $d = -0.32$, $M_{diff} = -0.81$, 95% CI [-1.51, -0.11], than adults. However, these differences are small (Cohen, 1992). There was no statistical difference between the two groups' scores on the stable subscale, $t(200) = 0.39$, $p = .699$, $d = 0.06$, $M_{diff} = 0.14$, 95% CI [-0.57, 0.84], or total scores $t(190.89) = 1.89$, $p = .060$, $d = 0.27$, $M_{diff} = 1.70$, 95% CI [-0.08, 3.48].

Table 23

Correlations Between Proximal DRAOR Subscale Scores and Total Score for Youth and Adults: Breach Reconviction

	Acute	Stable	Protective	Total score
Youth				
Acute	1			
Stable	.63**	1		
Protective	-.63**	-.74**	1	
Total score	.84**	.90**	-.90**	1
Adult				
Acute	1			
Stable	.58**	1		
Protective	-.45**	-.56**	1	
Total score	.82**	.87**	-.80**	1

** $p < .01$

For both youth and adults the proximal DRAOR subscales and total score were moderately to highly correlated with each other in the expected directions (positively for the risk subscales, negatively for risk and protective subscales), again implying convergent validity (see Table 23).

Correlations between proximal scores on the three subscales and total score were stronger than those for the initial score for both groups, suggesting stronger relationships between the proximal subscale scores and total scores.

Univariate predictive validity. The four proximal DRAOR predictors (the three subscale scores and the total score) were each entered as a covariate in four separate univariate Cox regressions to test their ability to predict breach reconviction without taking shared variance into account.

For youth, all proximal subscale scores predicted breach reconviction, acute: $\chi^2(1, n = 101) = 12.97, p < .001$; stable: $\chi^2(1, n = 101) = 10.14, p = .001$; protective: $\chi^2(1, n = 101) =6.44, p = .011$; as did the total score: $\chi^2(1, n = 101) = 12.39, p < .001$. For adults, the acute, $\chi^2(1, n = 101) = 22.54, p < .001$; stable, $\chi^2(1, n = 101) = 7.86, p = .005$; and total score, $\chi^2(1, n = 101) = 11.62, p = .001$, predicted breach reconviction, while the protective did not, $\chi^2(1, n = 101) = 1.32, p = .250$.

Table 24

Univariate Regression Models Containing Proximal DRAOR Subscale Scores and Total Score Predicting Breach Reconviction

Model for proximal DRAOR scores	β (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]	AUC [95% CI]
Youth					
Acute score	.20 (.06)	12.68**	1.22	[1.09, 1.35]	.70** [.59, .81]
Stable score	.18 (.06)	9.65**	1.20	[1.07, 1.35]	.72** [.62, .81]
Protective score	-.15 (.06)	6.33*	0.86	[0.76, 0.97]	.65* [.55, .76]
Total score	.08 (.02)	12.23**	1.08	[1.03, 1.13]	.71** [.60, .81]
Adult					
Acute score	.43 (.10)	18.74**	1.53	[1.26, 1.86]	.77** [.63, .91]
Stable score	.33 (.12)	7.49**	1.38	[1.10, 1.75]	.68* [.52, .85]
Protective score	-.13 (.11)	1.32	0.88	[0.71, 1.09]	.59 [.46, .71]
Total score	.14 (.04)	11.12**	1.15	[1.06, 1.25]	.74** [.61, .87]

** $p < .01$, * $p < .05$

For youth, the acute subscale had the highest HR, 1.22, 95% CI [1.09, 1.35] —with every one-unit increase in acute scores placing an offender at a 22% greater likelihood of reconviction—and performed with moderate accuracy, AUC = .70, 95% CI [.59, .81], (see Table 24). The stable subscale also had a high HR, 1.20, 95% CI [1.07, 1.35], and performed with high accuracy, AUC = .72, 95% CI [.62, .81].

For adults, the HR for the acute subscale suggests that every one-unit increase in acute scores corresponds to a 53% increased likelihood of reconviction, with high accuracy, AUC = .77, 95% CI [.63, .91]. As with the initial score, the proximal protective subscale could not predict adult breach reconviction at a level of statistical significance, and its AUC value suggested it was operating at a level no better than chance.

Despite adults having higher HRs than youth for scores on all subscales except protective, when comparing the confidence intervals around the AUC values, substantial overlap indicates there is plausibly no difference in predictive accuracy between adults and youth.

Multivariate predictive validity. Table 25 displays the multivariate regression models for youth and adults showing the contribution of each proximal DRAOR subscale when accounting for shared variance. For youth, the multivariate model containing all subscales was significant, $\chi^2(3, n = 101) = 14.68$, $p = .002$. Although predicting breach reconviction in combination, none of the subscales reached a level of statistical significance individually in the model. However, the confidence interval of the acute subscale only just crossed 1, suggesting perhaps the acute subscale was capturing slightly more of the variance in predicting breach reconviction.

Table 25

Multivariate Regression Models Containing Proximal DRAOR Subscale Scores Predicting Breach Reconviction

Model for proximal DRAOR scores	β (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]	AUC [95% CI]
Youth					
Acute score	.14 (.08)	3.20	1.15	[0.99, 1.34]	
Stable score	.12 (.10)	1.57	1.13	[0.93, 1.36]	.73** [.62, .83]
Protective score	.03 (.10)	0.10	1.03	[0.85, 1.26]	
Adult					
Acute score	.43 (.13)	11.60**	1.53	[1.20, 1.96]	
Stable score	.12 (.15)	0.63	1.13	[0.84, 1.51]	.76** [.61, .91]
Protective score	.17 (.14)	1.45	1.18	[0.90, 1.54]	

** $p < .01$

For adults the multivariate regression model with all predictors was also significant, $\chi^2(3, n =$

101) = 24.82, $p < .001$. For adults, the acute score was the only significant predictor in the model, operating at a level above the stable and protective subscales and suggesting that most of the predictive ability of the proximal DRAOR scores in predicting breach reconviction is coming from the acute subscale.

Again, despite the AUC value being slightly higher for adults, substantially overlapping confidence intervals indicate no difference in the accuracy of the multivariate models between groups.

***Summary*.** For youth, it appears subscale scores taken from the proximal DRAOR assessment are comparable in their ability to predict breach reconviction. That is, on average, higher scores on all risk subscales are associated with a greater likelihood of reconviction and higher protective scores are associated with a lesser likelihood of reconviction. This finding is in contrast to initial scores, which did not predict breach reconviction for youth. For adults, the acute subscale appears to be the best predictor of breach reconviction after accounting for variance shared with the other subscales. The protective subscale does not appear to predict adult breach reconviction well. However, overlapping confidence intervals still suggest that plausibly there is no difference in accuracy between groups.

Which DRAOR score is a better predictor: initial or proximal? In the previous chapter, all proximal DRAOR scores were better predictors of criminal reconviction than the initial DRAOR scores (with the exception of adult protective scores). Evidence from Cox regressions examining scores predicting breach reconviction so far suggests proximal scores may be stronger predictors than initial DRAOR scores, particularly for youth, as evidenced by proximal scores having higher HRs and AUC values.

Initial and proximal scores for all subscales were highly positively correlated with each other (see Table 26). Correlations between the different initial and proximal scores for both groups ranged from moderate to strong (with the exception of adult proximal acute and initial stable scores, and proximal acute and initial protective scores, which were weakly correlated with each other). Correlations between scores on the same subscales for the two assessments (in bold) were strong.

Table 26

Correlations Between Initial and Proximal DRAOR Subscale Scores and Total Score for Youth and Adults: Breach Reconviction

	Initial Acute	Initial Stable	Initial Protective	Initial Total score
Youth				
Proximal Acute	**.62****	.33**	-.38**	.55**
Proximal Stable	.44**	**.61****	-.48**	.62**
Proximal Protective	-.51**	-.49**	**.71****	-.70**
Proximal Total score	.59**	.54**	-.60**	**.71****
Adult				
Proximal Acute	**.63****	.24*	-.29**	.51**
Proximal Stable	.43**	**.56****	-.41**	.58**
Proximal Protective	-.31**	-.34**	**.70****	-.55**
Proximal Total score	.55**	.46**	-.56**	**.66****

** $p < .01$, * $p < .05$

Incremental validity: Initial and proximal DRAOR scores. The initial and proximal scores for the four DRAOR predictors were entered together as covariates in the first block of four separate multivariate Cox regressions. The regression using adult scores from the acute subscale is reported here because it was the only subscale where both the initial and proximal scores were found to be predicting reconviction in univariate regressions. In the cases where the initial score did not predict reconviction, and the proximal score did, the proximal was the better predictor. All youth initial scores did not predict breach reconviction, and adult initial stable, protective, and total scores, and proximal protective scores did not predict breach reconviction. These scores were not included here but can be found in Appendix E.

The adult Cox regression model containing both initial and proximal acute subscale scores predicted breach reconviction, $\chi^2(2, n = 101) = 22.57$, $p < .001$. The proximal score was the only significant predictor in the model, suggesting it is capturing more information related to reconviction than the initial score. The proximal score's HR has increased while the initial score's HR has lowered, suggesting that when accounting for overlapping variance, the proximal score is the better predictor of adult breach reconviction. A one-unit increase in proximal acute scores corresponds to a 65% increased likelihood of breach reconviction (see Table 27).

Table 27

Multivariate Regression Model Containing Initial and Proximal DRAOR Acute Subscale Scores Predicting Adult Breach Reconviction

Multivariate model for initial and proximal DRAOR scores	β (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]
Adult				
Acute score				
Initial	-.08 (.14)	0.32	0.92	[0.70, 1.22]
Proximal	.50 (.16)	9.47**	1.65	[1.20, 2.26]

** $p < .01$

Summary. Multivariate Cox regressions containing both initial and proximal scores show that for both groups, proximal scores are better predictors of breach reconviction than initial scores (with the exception of the adult protective scores, which did not predict reconviction at either assessment).

To what extent do DRAOR scores change over time and does change predict reconviction?

Previous analyses indicate it is possible DRAOR scores had changed between assessments, as they had become better predictors of reconviction, particularly for youth. A paired samples t-test revealed proximal scores were significantly lower than initial DRAOR scores on both risk subscales and the total score for youth: acute, $t(100) = 4.43$, $p < .001$, $d = 0.44$, $M_{diff} = 0.99$, 95% CI [0.55, 1.43]; stable, $t(100) = 4.34$, $p < .001$, $d = 0.43$, $M_{diff} = 0.96$, 95% CI [0.52, 1.40]; total score $t(100) = 5.37$, $p < .001$, $d = 0.53$, $M_{diff} = 2.72$, 95% CI [1.72, 3.73]; and adults: acute, $t(100) = 5.38$, $p < .001$, $d = 0.54$, $M_{diff} = 1.12$, 95% CI [0.71, 1.53]; stable, $t(100) = 4.15$, $p < .001$, $d = 0.41$, $M_{diff} = 0.84$, 95% CI [0.44, 1.24]; total score, $t(100) = 6.73$, $p < .001$, $d = 0.67$, $M_{diff} = 3.01$, 95% CI [2.12, 3.90]. Proximal protective scores were higher than the initial scores for both groups, youth: $t(100) = -3.86$, $p < .001$, $d = -0.38$, $M_{diff} = -0.77$, 95% CI [-1.17, -0.38]; and adults: $t(100) = -6.43$, $p < .001$, $d = -0.64$, $M_{diff} = -1.05$, 95% CI [-1.37, -0.73]. These findings suggest that, on average, risk had decreased between the two assessments (see Table 28). The Cohen's *d* effect sizes of all scores were larger for adults except in the case of the stable subscale, indicating that the magnitude of change within the adult group from the initial to proximal assessment was greater than that within the youth group for the acute, protective and total scores.

To what extent do DRAOR scores change over time? Change scores were calculated for each of the DRAOR subscales and the total score by subtracting the proximal DRAOR scores from the initial DRAOR scores.[15] As in the previous chapter, a positive number meant risk had decreased between the two assessments, a negative number meant risk had increased, and zero represented no difference between assessment scores (for protective factors, a negative number meant protective factors had increased and a positive number meant protective factors had decreased).

Table 28 shows the mean change scores for both groups. Although the Cohen's *d* values for paired samples t-tests between initial and proximal scores suggested the magnitude of change was greater for adults, the independent samples t-test found no statistical difference between youth and adults' change scores for any of the subscales or total score, acute: $t(200) = -0.42$, $p = .674$, $d = -0.06$, $M_{diff} = -0.13$, 95% CI [-0.73, 0.47]; stable: $t(200) = 0.40$, $p = .693$, $d = 0.06$, $M_{diff} = 0.12$, 95% CI [-0.47, 0.71]; protective: $t(200) = 1.07$, $p = .284$, $d = 0.15$, $M_{diff} = 0.28$, 95% CI [-0.23, 0.79]; and total score: $t(200) = -0.43$, $p = .672$, $d = -0.06$, $M_{diff} = -0.29$, 95% CI [-1.62, 1.05].

Table 28

Mean Change Made Between Initial and Proximal DRAOR Assessments: Breach Reconviction

	Mean initial score (*SD*)	Mean proximal score (*SD*)	Mean change score (*SD*)	Range of change	No change (%)
Youth					
Acute	5.11 (2.65)**	4.12 (2.49)	0.99 (2.25)	[-4, 9]	24.8
Stable	5.90 (2.24)**	4.94 (2.71)	0.96 (2.23)	[-4, 8]	49.5
Protective	6.08 (2.42)	6.85 (2.81)**	-.77 (2.01)	[-6, 6]	54.5
Total Score	4.93 (6.00)**	2.21 (7.07)	2.72 (5.10)	[-8, 18]	24.8
Adult					
Acute	4.49 (2.57)**	3.37 (2.26)	1.12 (2.09)	[-4, 7]	28.7
Stable	5.64 (1.92)**	4.80 (2.36)	0.84 (2.04)	[-4, 8]	38.6
Protective	6.61 (2.02)	7.66 (2.20)**	-1.05 (1.64)	[-6, 3]	45.5
Total Score	3.51 (5.20)**	0.50 (5.66)	3.01 (4.50)	[-8, 17]	16.8

** $p < .01$

[15] Again, 'change' refers to mean change; as scores could be the same at both assessments, suggesting no change, when in fact change could have occurred in the assessments in between.

Table 28 also shows the limits within which change occurred (range of change). Results were similar to those found in the previous chapter with criminal reconviction. To illustrate, for the acute subscale, the most any youth increased in acute risk between the initial and proximal assessments was four units, the most any youth decreased in acute risk was nine units. The most any adult increased in acute risk was four units, while the most any adult decreased in risk was seven units. For the protective subscale, the most any youth increased in protective factors was six units (more assets), and the most any youth decreased in protective factors was also six units (fewer assets). The most any adult increased in protective factors (more assets) was six units, and the most any adult decreased in protective factors (fewer assets) was three units. Within the youth and adult groups, the protective subscale was the one on which the most offenders made no change between assessments, with 54.5% of youth and 45.5% of adults making no change. Roughly half of the youth (49.5%) also made no change on the stable subscale.

Does change predict breach reconviction? Any regressions using change scores would also have to contain initial scores, because, as mentioned previously, the amount of change is dependent on and limited by an offender's initial score. Therefore, the relationship between initial scores and change scores was first examined. All change scores had weak to moderate positive correlations with their initial scores, with acute change scores having strong positive correlations with initial scores for both groups. These strong correlations indicated for both youth and adults, greater decreases in acute risk were associated with higher initial acute risk scores. This is likely due, in part, to the extra item the acute subscale contains – and thus, greater scope for change – but recall also that for both groups, fewer offenders made no change on the acute subscale compared to the stable and protective, suggesting perhaps it the most changeable subscale. These correlations can be found in Appendix F.

Next, the four change scores for the DRAOR subscales and total score were each entered as a covariate in the second block of four separate univariate Cox regressions. The first block contained the respective initial DRAOR score for each subscale/total score, to determine whether change would predict reconviction at a level above an offender's initial risk.

For youth, all Cox regression models containing both initial DRAOR scores and change scores predicted breach reconviction: acute, $\chi^2(2, n = 101) = 13.41$, $p = .001$; stable, $\chi^2(2, n = 101) =$

11.00, p = .004; protective, $\chi^2(2, n = 101) = 9.81$, p = .007; as did the total score model, $\chi^2(2, n = 101)$ =14.53, p = .001.

Table 29

Multivariate Regression Models Containing Change Scores Controlling for the Initial DRAOR Subscale Scores and Total Score Predicting Breach Reconviction

Model for change scores	β (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]	AUC [95% CI]
Youth					
Initial acute score	.18 (.06)	7.71**	1.19	[1.05, 1.35]	.71** [.60, .82]
Acute change score	-.23 (.07)	9.98**	0.80	[0.69, 0.92]	
Initial stable score	.13 (.08)	2.69	1.13	[0.98, 1.32]	.72** [.63, .82]
Stable change score	-.24 (.08)	9.23**	0.79	[0.68, 0.92]	
Initial protective score	-.08 (.07)	1.24	0.92	[0.80, 1.06]	.69** [.59, .79]
Protective change score	.26 (.08)	10.41**	1.30	[1.11, 1.53]	
Initial total score	.04 (.03)	2.67	1.05	[0.99, 1.10]	.73** [.64, .83]
Total score change score	-.13 (.04)	11.29**	0.88	[0.82, 0.95]	
Adult					
Initial acute score	.42 (.10)	17.29**	1.52	[1.25, 1.85]	.76** [.62, .91]
Acute change score	-.50 (.16)	9.47**	0.61	[0.44, 0.83]	
Initial stable score	.26 (.15)	3.07	1.29	[0.97, 1.72]	70* [.54, .87]
Stable change score	-.41 (.16)	6.38*	0.66	[0.48, 0.91]	
Initial protective score	-.10 (.13)	0.66	0.90	[0.71, 1.15]	.58 [.46, .71]
Protective change score	.18 (.16)	1.19	1.19	[0.87, 1.63]	
Initial total score	.13 (.05)	7.50**	1.14	[1.04, 1.25]	.74** [.60, .88]
Total score change score	-.19 (.07)	6.80**	0.83	[0.72, 0.95]	

** p < .01, * p < .05

Recall that for the acute, stable and total scores, a positive change score represents a decrease in risk between the two assessments. A hazard ratio of less than 1 means that for every one unit of change an offender made in the direction of a decrease in risk, the likelihood of breach reconviction also decreased (when taking into account the variation in initial DRAOR scores). For youth, as with criminal reconviction, stable change had the highest hazard ratio for a risk subscale, where one unit of change in the direction of a decrease in stable risk (e.g., from 1 to 2) is associated with a 21% reduction in the likelihood of breach reconviction (see Table 29). For the protective subscale, a positive number represents a decrease in protective scores between the two assessments. So, every one unit of change on the protective subscale made in the direction of a decrease in protective factors (e.g., from -2 to -1) is associated with an increase in the likelihood of breach reconviction. For youth, one unit of change in the direction of a decrease in protective factors between the initial and proximal assessments is associated with a 30% increase in the likelihood of breach reconviction.

For adults, the acute, $\chi^2(2, n = 101) = 22.57, p < .001$; stable, $\chi^2(2, n = 101) = 8.08, p = .018$; and total score, $\chi^2(2, n = 101) = 11.78, p = .003$, models predicted breach reconviction but the protective model did not, $\chi^2(2, n = 101) = 1.48, p = .476$. Acute change had the highest HR, with one unit of change in the direction of a decrease in acute risk associated with a 39% reduction in the likelihood of breach reconviction.[16]

Risk change scores show high accuracy (all AUCs range from .70 to .76), and appear comparable between youth and adults, while protective change appears to be trending towards higher accuracy for youth. Taking into account the variation in initial DRAOR scores, a randomly selected reconvicted youth is 69% more likely to have made less change in the direction of an increase in protective factors than a randomly chosen youth who was not reconvicted, AUC = .69, 95% CI [.59, .79]. However, substantial overlap in confidence intervals around the adult and youth AUC values suggests that, when taking into account base rates, the protective subscale's accuracy is plausibly no different between groups.

Summary. Multivariate Cox regressions examining change scores while controlling for initial scores show that for youth, change made on all DRAOR subscales predicts breach reconviction, with the protective subscale having the largest effect on likelihood of breach reconviction; performing with moderate accuracy. That is, for youth, greater decreases in protective factors corresponded to greater increases in the likelihood of reconviction. For adults, change on all subscales predicted reconviction with the exception of the protective subscale. The largest effect came from the acute subscale, where greater change in the direction of a decrease in risk was associated with a decrease in the likelihood of reconviction. However, overlapping confidence intervals suggest that the accuracy of change scores at predicting reconviction is comparable across the youth and adult groups.

[16] Despite all initial DRAOR scores being poor predictors for youth, the initial acute DRAOR score is now adding something unique to the multivariate model. The initial total score has also become a significant predictor for adults, when alone it was not. These results again suggest that, in these cases, change scores might possibly mediate the relationship between initial scores and criminal reconviction, a fuller discussion of which is beyond the scope of this research.

Is the DRAOR a better predictor of breach reconviction than the RoC*RoI?

This analysis used offenders' RoC*RoI scores and their initial and proximal DRAOR scores on each of the subscales and total score to determine whether the DRAOR could predict breach reconviction at a level over and above the RoC*RoI. Recall in the previous chapter, the RoC*RoI did not predict youth criminal reconviction, and was a better predictor for adults; predicting reconviction at a level above the acute and protective subscales, while having comparable predictive ability to the stable subscale and total score. Might this pattern of results continue for breach reconviction?

Does the RoC*RoI predict breach reconviction? Recall, the groups were matched on their RoC*RoI scores, so these did not significantly differ between groups. For the sample of 101 youth and 101 adults, the youth mean score was .37 (*SD* = .17, range: .10 to .82) and the adult mean score was .37 (*SD* = .17, range: .02 to .74).

RoC*RoI scores were entered as a covariate into the first block of a Cox regression. The time variable entered was survival days (days to breach reconviction or end of sentence for those who were not reconvicted), and the status variable was the presence or absence of a breach reconviction.

The RoC*RoI did not predict breach reconviction for youth, $\chi^2(1, n = 101) = 0.01, p = .923$ (see Table 30). Taken together, the HR from the Cox regression and the AUC value suggest that the RoC*RoI's ability to predict breach reconviction is likely worse than chance level. The RoC*RoI also did not predict breach reconviction for adults, $\chi^2(1, n = 101) = 1.02, p = .312$. Therefore, the RoC*RoI is a poor predictor of which adults and youth are likely to be reconvicted of a breach offence.

Table 30

*Univariate Regression Models Containing RoC*RoI Scores Predicting Breach Reconviction*

Model for RoC*RoI scores	*β* (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]	AUC [95% CI]
Youth					
RoC*RoI	-0.09 (.97)	0.01	0.91	[0.14, 6.07]	.49 [.37, .60]
Adult					
RoC*RoI	1.51 (1.50)	1.02	4.51	[0.24, 84.64]	.56 [.40, .72]

Does the DRAOR add incremental validity over the RoC*RoI? Incremental validity is contingent on the RoC*RoI being a significant predictor individually. Since the RoC*RoI did not predict breach reconviction for either group, and all proximal DRAOR scores did (with the exception of protective subscale scores for adults), proximal scores are better predictors of breach reconviction for both youth and adults. The exception to this is adult protective subscale scores, which, as mentioned, were poor predictors individually. Therefore, correlations and incremental Cox regressions taking into account variance shared by DRAOR and RoC*RoI scores are not reported here, but instead in Appendix G.

Summary. Proximal DRAOR scores on all risk subscales were able to account for more of the variance in predicting breach reconviction than both the initial DRAOR scores and RoC*RoI scores for both groups. Therefore, for both youth and adults, proximal DRAOR scores (with the exception of adult protective scores) are better predictors of breach reconviction.

Chapter 5

Discussion

The primary aim of the present research was to compare the ability of two adult risk assessment tools —one static and one dynamic—to predict reconviction in a matched sample of community-supervised youth and adult offenders. In short, the tools' ability to predict reconviction for both groups depends on which factors they assess, and how close in time to the reconviction the assessment was made. While an up-to-date dynamic risk assessment (the DRAOR) predicted criminal and breach reconviction well for both youth and adults, a static risk assessment (the RoC*RoI) was only able to predict adult criminal reconviction, and neither reconviction outcome for youth.

Comparing Youth and Adult Risk Assessment Using the DRAOR

Specifically, the first research question examined whether scores on the DRAOR, taken from different assessments: one near the start of an offender's sentence (initial) and one closer to reconviction or sentence end (proximal), were able to distinguish between youth and adults who were reconvicted, with equal accuracy. I found that while initial scores yielded mixed results across reconviction outcomes, proximal scores always emerged as better predictors of both criminal and breach reconviction for both groups. That is, youth and adults who had received higher risk scores at the proximal assessment were more likely to be reconvicted than youth and adults with lower risk scores. Youth who had higher protective scores at the proximal assessment, and thus greater assets, were less likely to be reconvicted than youth who had lower protective scores. Adult protective scores proved an exception to this, with protective scores from both assessments being poor predictors across both reconviction outcomes.

Similarities and differences in how the DRAOR works for youth and adults. Previous literature has suggested that despite developmental and social pressures acting upon youth, they largely share risk factors with adults. However, certain dynamic factors have been hypothesised to have a stronger or weaker relationship with reoffending during the period of adolescence (Borum, 2003). The little previous research that has been done examining the predictive validity of adult risk assessment tools for use with youth, suggests that adult tools have good predictive validity for youth, at least in the short-term (see: Hoge et al., 2012; Ralston & Epperson, 2013). The present research also provides support for the short-term use of dynamic adult risk tools with youth aged 17-19. Although the DRAOR was developed by drawing on the extant adult risk assessment literature, it worked equally as well—if not slightly better—for the youth in the present study. Although no individual item analyses were done, speculation based on the literature might suggest that the DRAOR's efficacy at predicting youth reconviction is due to its inclusion of several pertinent youth risk factors (e.g., peer associations and substance abuse), and several important protective factors (e.g., social support and social control). The DRAOR might also be working well for youth because it adheres to several best principles for youth risk assessment (e.g., collecting information from a range of sources, including dynamic factors, and regular assessments; Olver et al., 2009).

Surprisingly, for the most part, the DRAOR trended towards performing with higher accuracy for youth compared to adults. However, despite the apparent differences in effect size and relative contribution of the DRAOR subscales and total score between groups; there was no statistical difference in the level of accuracy shown by the subscales and total score when predicting reconviction for youth and adults. Despite the similarity in accuracy, there were a few differences in how certain subscales of the DRAOR were working across assessments and groups. These will now be discussed.

Important DRAOR subscales for predicting criminal reconviction. It appears that, when assessed close in time to the reconviction, both the acute and stable subscales predict adult criminal reconviction equally well. In contrast, for youth, it appears that much of the DRAOR's ability to predict criminal reconviction comes from the stable subscale, which emerged as the best proximal predictor of criminal reconviction when also considering contributions from the acute and protective subscales. In other words, youth scores on the stable subscale were able to give more information relevant to predicting youth reconviction than information provided by the other two subscales. This result makes sense, because the stable subscale includes items such as peer associations, attachment with others, and impulse control; factors that have been theorised to become greater sources of risk during adolescence when relationships with peers assume greater importance, and the brain is still developing the capacity to regulate behaviour (Hoge et al., 2012; Scott & Steinberg, 2008). Further research might examine the DRAOR at the item level, to investigate whether youth-pertinent items such as peer associations, impulse control, or even substance abuse are among those contributing most to the DRAOR's predictive ability.

Important DRAOR subscales for predicting breach reconviction. When predicting youth breach reconviction, all three of the DRAOR subscales (acute, stable, and protective) work equally well when assessed close in time to the reconviction. In contrast, for adults, the acute subscale emerged as an important predictor of breach reconviction at the initial and particularly the proximal assessment. In other words, adult acute scores were able to consistently give more information relevant to breach reconviction than stable and protective scores. Theoretically this makes sense, because the acute subscale is intended to pick up on more proximal sources of risk, which might make an offender more likely to breach their sentence conditions. However, it is also possible that acute factors are good predictors because they represent ongoing struggles, not simply imminent risk factors that are unique to a single assessment (Hanson et al., 2007).

The DRAOR's ability to predict breach reconviction is an interesting finding, because it is not designed to predict this kind of reconviction. In an early pilot study of the DRAOR, Tamatea and Wilson (2009) found that neither initial nor proximal DRAOR scores predicted breach reconvictions for a group of primarily adult offenders on probation, and concluded that the risk and protective factors contained within the DRAOR were not "directly related to sentence compliance" (p. 5). A previous study investigating the use of the DRAOR with youth examined solely initial scores, finding they were unable to predict reconviction for a sample of 17-19 year old parolees whose reconvictions consisted primarily of breaches (Fortune et al., in preparation). This finding was replicated in the present study for youth initial scores, but not proximal scores for either group. All proximal subscale scores (with the exception of protective for adults) and total score predicted breach reconviction, suggesting that updating risk assessment is particularly useful for monitoring both youth and adult offenders who may be at risk of breaching their sentence conditions.

The contribution of protective factors. The most consistent difference in how the DRAOR was operating for youth and adults came from the protective subscale. Previous research has found that youth and adults share protective factors (Stouthamer-Loeber et al., 2004). However, while the present study found higher protective scores to be associated with a reduction in the likelihood of youth reconviction, protective scores were consistently poor predictors for adults. Adults in the present study scored higher than youth on protective factors for both initial and proximal assessments, which is expected, given they were reconvicted less. However, despite being higher, protective scores did not appear to significantly reduce the likelihood of reconviction for the adult group. Recall that, on average, risk scores changed between assessments, with scores improving in predictive accuracy over time. But despite significant change also being made on the protective subscale, this change did not correspond to a reduction in the likelihood of reconviction for adults. The protective subscale of the DRAOR has previously been shown to independently predict breach convictions and any new reconvictions in a sample of high-risk adult offenders on parole (but not incrementally, when considered alongside the risk subscales; Yesberg & Polaschek, 2015). However, in validating the DRAOR on a primarily adult community-supervised sample of offenders in the U.S., Chadwick (2014) found similar results to the present research in that the protective subscale did not predict any form of reconviction; corroborating evidence that the protective subscale might not be best serving its intended purpose.

The protective subscale might have been predicting reconviction for youth but not for adults for several reasons. First, the majority of contemporary research on protective factors concentrates on youth samples, presumably because they are in a period of life characterised by transitions, where personality and social factors are not yet set in stone, ideally affording them more opportunities for ‘turning points’ than adults (Sampson & Laub, 2005). Much of the limited research that has examined adult protective factors has focused on a specific type of offender; sex offenders (and more recently, violent offenders; see de Vries Robbé, Mann, Maruna, & Thornton, 2014; Ullrich & Coid, 2011). Very few previous studies have looked at the effect of protective factors on the likelihood of more general forms of criminal recidivism using adult samples. Although the protective subscale was measuring factors that clearly promoted youth resilience, perhaps the factors on the protective subscale were simply not assets for this group of adult offenders, and thus did not serve to insulate them from reconviction.

However it is possible that the DRAOR’s protective factors *are* related to a reduction in reconviction, but they are harder to score for adults, compared to youth. Adults did, on average, increase in protective factors over time, suggesting that probation officers were noting differences in protective factors and recording these changes. Probation officers may simply have not had enough information about offenders' protective assets to make an informed judgment. Most of the probation officer and offender interview session might be devoted to discussing risk, rather than strengths (i.e. either risk is easier to discuss, or is are seen as more important to discuss). It might also be harder for adult offenders, who are used to being seen as a risk to society, to talk about their strengths. In addition, protective items such as 'social control' or 'responsive to advice' may be more subjective than risk items such as 'peer associations' and 'substance abuse'. While it might be easy to tell whether an offender spends time with antisocial associates, whether they are responsive to advice is more open to interpretation, and requires a deeper understanding of the offender's motivations. For example, what would an offender score if he is responsive to a prosocial family member's advice, but ignores the advice of his probation officer? In other words, scores could depend on the quality of the relationship between the probation officer and the offender (Bonta, Rugge, Scott, Bourgon, & Yessine, 2008).

Furthermore, the optimum way to examine protective factors might not be to separate them out and assess their contribution individually, but to instead include them as part of a total risk score. Vincent, Perrault, Guy, and Gershenson (2012) found that protective factors included in the SAVRY (Borum et al., 2006), did not improve the tool's prediction of recidivism over and above its risk factors, and suggested that protective factors should instead "interact with risk" rather than be treated as separate entities (Vincent, Perrault, Guy, & Gershenson, 2012, p. 378). Indeed, in the present study protective factors did function as part of the total score – which was a good predictor of adult reconviction – but did not predict when examined separately. Perhaps protective factors are less likely to function separately for adults compared to youth, and instead are more likely to moderate the expression of risk. Recall that fewer adults in the present study were reconvicted than youth. This suggests, for this group of adults, there were indeed factors operating to reduce their likelihood of reconviction, albeit factors not captured exclusively by the protective subscale of the DRAOR. It is vitally important that factors related to desistance are accurately measured so as not to overestimate adult risk. Further research is needed into why higher scores on the protective subscale of the DRAOR were not related to reductions in adult reconviction in the present study.

Investigating the DRAOR's Utility as a Dynamic Risk Assessment Tool

The second research question examined the DRAOR's utility as a dynamic risk assessment tool. Presently, there is mixed evidence in the literature about whether dynamic scores change and whether that change can be linked to recidivism (see Hanson et al., 2007; Olver et al., 2012). Research that uses multiple time points to examine offenders' change over time on dynamic factors is sparse, with even less research having been done using youth samples with a matched comparison sample of adults. I wanted to investigate change in scores across two separate assessments because if change over time occurred and predicted reconviction then not only could I conclude that the DRAOR is operating as it is designed to, but I could also link change to an outcome: a reduction in the likelihood of reconviction. I used a mean change score—operationalised as the difference between initial and proximal assessment scores—to see whether DRAOR scores changed over time, and whether I could link this change to changes in risk level by using it to predict reconviction. I found that, on average, dynamic risk (as measured by all three subscales of the DRAOR and the total score) decreased over time. That is, more youth and adults were scored lower on risk factors and higher on protective factors at the proximal assessment than at the initial one. Not only did DRAOR scores change over time, but that change predicted both criminal and breach reconviction for youth and adults. In other words, the more youth and adults decreased in risk, the less likely they were to be reconvicted. Despite the protective subscale changing for adults, this change was neither linked to criminal nor breach reconviction, meaning that there was no clear pattern associated with higher or lower protective scores that was able to distinguish adults who were reconvicted from adults who were not.

The implications of change between assessments. Reassessment of dynamic factors has been proposed to be particularly crucial during adolescence, given that it is a period of significant change (Vincent, Guy, & Grisso, 2012). For this research I compared two scores. An initial score: measuring an offender's dynamic risk near the start of his sentence, and a proximal score: recorded closer in time to reconviction, providing a more up to date assessment of dynamic risk. I found that in almost all cases, the proximal assessment was better, suggesting dynamic factors work best when used for their intended purpose, that is, to provide a measure of imminent risk. This finding supports the DRAOR's utility as a risk assessment and monitoring tool because it shows the DRAOR is sensitive to changes in risk. On average, scores decreased over time, and in turn reductions in risk were associated with decreases in the likelihood of reconviction. Average initial DRAOR risk scores in the present study (mean scores ranging from 4.32 to 5.90) suggested that when considering the minimum and maximum score possible for each subscale (12 for stable and 14 for acute), there was actually more room for change in the direction of an increase in risk than a decrease, so it is encouraging that despite this, risk tended to decrease overall. Initial scores, despite being reasonable predictors of reconviction in some cases, did not predict reconviction as well or as consistently as proximal scores. Proximal scores were not only lower, but also more accurate, showing the value in updating risk assessments. It appears that the closer in time to the offence a score is taken, the more useful information it contains (and thus, the more accurate it is) relevant to predicting reconviction. These results have clinical utility because they show that risk factors can and do change and presumably, if they are targeted in treatment, greater reductions in risk will produce lower rates of reconviction.

One possible reason why initial scores were not as good as proximal scores at predicting reconviction is that over time any treatment or intervention an offender was receiving while serving their sentence may have been operating to reduce their risk level. Thus, lower proximal scores were a true reflection of a reduction in risk over time. Another possibility is that increased contact with a probation officer leads to more accurate assessments of risk. Although youth did tend to score, on average, slightly higher on the risk subscales and lower on the protective subscale than adults across assessments and outcomes, there were several occasions where these differences were neither large nor statistically significant. Similar scores suggested that dynamic risk was similar for both groups, but because more youth were reconvicted than adults, in theory, youth scores should have reflected this: with youth consistently being scored higher than adults on dynamic risk factors, particularly at the proximal assessment. One reason youth did not always score significantly higher than adults could be that initially, probation officers might adopt a conservative stance, scoring higher by default, so as not to underestimate risk due to lack of information. Over time a probation officer gets to know an offender better, and so it follows that assessments farther away in time from initial meetings are more accurate. Despite being a structured actuarial tool, the DRAOR requires the competent scoring of items by a probation officer. Probation officers will also have access to an offender's RoC*RoI score and case file containing historical information. This could be inflating their judgments of dynamic risk, particularly initially, when a probation officer might not have all the information necessary to make a fully informed judgment. In a study by Hilton, Harris, Rawson, and Beach (2005) when clinicians were given descriptive information about static risk alongside a score representing an offender's probability of reoffending relative to others, their risk perceptions tended to inflate, resulting in more inaccurate assessments compared to when only the probability score was provided. Although the Hilton et al., (2005) study used static descriptive information, the same principle may apply here, and the descriptive information provided by the DRAOR's dynamic subscales, when considered alongside the RoC*RoI score (expressed as a probability), might also have initially increased probation officers' estimates of risk.

The present study was retrospective, and it is impossible for probation officers to know at the time which assessment will be an offender's most proximal. Results from the present study suggest that by paying attention to dynamic change on the DRAOR, probation officers will be better able to identify likely precursors to reoffending and perhaps act to prevent reconvictions from occurring. Rather than rely solely on single points of dynamic risk when deciding whether an offender will commit an offence, a possible way to keep track of changes in risk over time would be to plot scores on a graph. That way if there is a noticeable increase, rather than decrease, in risk over time, it can be addressed. Further research could examine DRAOR scores using multiple, rather than dual, time points in order to see whether offenders who show a consistent decrease in risk over time have a lower likelihood of reconviction than offenders whose risk level fluctuates over the course of their sentence.

It must be mentioned, however, that despite change predicting reconviction in the present research, the way change was measured (a mean change score) means although overall offenders decreased in risk and increased in protective factors over time, I cannot draw any conclusions about how risk might have fluctuated in the assessments between an offender's initial and proximal assessment. As I did not look at the number of days between assessments, I am also unable to determine whether greater time between assessments was associated with greater decreases in risk (one might expect that greater time in which to change would lead to greater mean change). Furthermore, it may be true that one or more items on a given subscale were exclusively responsible for its predictive ability but as no item-level analyses were performed, I can only conclude that aggregate change on factors contained within the subscales predicted reconviction. In other words, I was not able to tell whether, for example, peer associations was the only item changing on the stable subscale and thus responsible for stable change predicting youth criminal reconviction.

Evaluating the Relative Contributions of the DRAOR and RoC*RoI

For the final research question, I wanted to know whether the DRAOR could give more information than the RoC*RoI about which youth and adults were likely to be reconvicted. While, in the literature, both static and dynamic factors are consistently found to be strong predictors of recidivism for adults, dynamic factors have been theorised to be particularly important for youth, because they pick up on more imminent sources of risk that might have a greater association with reconviction during the period of adolescence (Borum, 2003). I found that, indeed, the DRAOR was a better predictor of both breach and criminal reconviction for youth, with RoC*RoI scores having no relationship with reconviction; failing to distinguish between youth who were reconvicted and youth who were not. The RoC*RoI was a better predictor of criminal reconviction for adults than two subscales of the DRAOR: the acute and protective subscales, while it predicted equally as well as the proximal stable subscale score and proximal total score. All proximal DRAOR scores, except the protective subscale score, outperformed the RoC*RoI as predictors of adult breach reconviction.

Comparing static and dynamic risk assessment for youth and adults. Although static tools are useful for identifying the “risk status” of adults who pose the greatest risk of criminal reconviction, the present research illustrates that assessing youth using a tool containing solely static factors might yield unreliable results. It has been argued that dynamic and static tools should be considered in combination, as part of an overall risk assessment for youth, and not separately as this might “dissipate the predictive acumen of [adolescent] risk assessment tools” (Vincent, Perrault, Guy, & Gershenson, 2012, p. 380). However, the present study suggests caution when including static factors alongside dynamic for youth assessment, finding instead that the RoC*RoI did not appear to add anything to assessment over and above the ability of the DRAOR. That is, risk assessment accuracy did not improve when static factors were examined alongside dynamic.

The finding that the static risk tool used in the present study did not predict reconviction well for youth is inconsistent with Ralston and Epperson's (2013) research finding that adult static risk tools predicted recidivism well for a group of younger youth. One reason for this discrepancy in findings may be that Ralston and Epperson (2013) used a sample of youth sex offenders, while youth who had perpetrated sexual offences were a minority in the present research. While static tools might better predict risk of reconviction for crimes with low base rates, like sex offending, dynamic tools are clearly shown in the present research to better predict more general forms of reconviction, including breach of sentence conditions.

Recall that youth and adults in the present study were matched on RoC*RoI scores, but that youth reoffended at a higher rate, suggesting that perhaps the RoC*RoI was under-classifying youth. On average, youth offenders in the present study had a 37% likelihood of being sentenced to prison within 5 years as estimated by the RoC*RoI. Already a large number of them had reoffended after being at risk for an average time of just over 2 years, (76.2% of youth for criminal reconviction and 50% of youth for breach reconviction, compared to 38.5% and 21.3% of adults respectively). Nevertheless, although used in the present study as a proxy for static risk, the RoC*RoI is designed to predict more serious reconviction outcomes that result in a prison sentence. Data on reimprisonment was not analysed in the present research, and the criminal and breach reconviction outcomes used did not necessarily result in imprisonment for most youth in the sample; therefore, these results should be interpreted conservatively. It cannot be ruled out that perhaps the RoC*RoI might have been a good predictor of reconvictions that led to imprisonment for youth.

The results of the present research raise the important question, however, of whether a static tool designed to predict reimprisonment is relevant for youth assessment. RoC*RoI scores are not generated as frequently as DRAOR scores, and are seldom updated. For youth, dynamic factors outperformed static, showing that ongoing, current sources of risk are more relevant to their reconviction than a single assessment of criminal history. The full extent of youth offenders' history of antisocial behaviour—that might have lead to higher (and perhaps more accurate) estimates of static risk—was not captured by the criminal history items measured by the RoC*RoI. As mentioned, childhood conduct disordered behaviour is generally a good predictor of persistent crime in young adulthood, however few static risk tools provide a way to take this, often unrecorded, behaviour into account. The RoC*RoI measures past antisocial behaviour via variables that youth may not have had the time or experience to acquire (i.e. frequency of convictions, or number of years spent in prison). More useful information for youth might be the number of appearances they have had in the youth court, or frequency of contact with law enforcement prior to age 17. This information is not included in a RoC*RoI assessment, because it is unlikely to have been officially recorded, due in part for a need to avoid giving youth a premature criminal record (Becroft, 2009).

If youth have a lack of criminal history information available, and few continue into adulthood to perpetrate serious crime, how useful is a static score as a baseline risk estimate if it does not accommodate ongoing adolescent development or provide an accurate picture of risk level? When one considers that the aim of youth risk assessment is to give a realistic estimate of risk of future offending while also bearing in mind the importance of deterring a youth from a life of crime; static risk factors "do very little in terms of guiding case planning or informing juvenile justice personnel about how to decrease a specific youth's risk for reoffending" (Vincent, Chapman, & Cook, 2011, p. 58). Therefore, the importance of placing greater weight on dynamic factors when assessing youth offenders is underscored.

Limitations/Future Directions

Every possible effort was made to conduct the present research using approaches that were methodologically sound. However, there are still limitations to this research that provide avenues for future study.

The matching process. Matching youth and adults strengthened the present research by controlling for any between-group differences in variables related to reconviction that might make it more likely that some offenders would be reconvicted, independent of age. However, one limitation with this is that even though both groups were matched on static risk as measured by the RoC*RoI, youth were clearly the higher risk group. Consistent with the age-crime curve (Moffitt, 1993), youth were reconvicted at a faster rate and higher volume than the adult group; a result reflected by DRAOR scores in most cases, but not RoC*RoI scores.

Difficulties with matching highlight the fact that youth and adults inherently differ on several important criminal history variables—explicitly due to age—making it hard to compare age directly. The average age of the adult group was 26, so roughly half of the offenders were still in the period of late adolescence, and the aging-out period for crime (recall, in theory, late adolescence extends to age 25; Farrington et al., 2012; Moffitt, 1993). Perhaps greater between-group differences in the DRAOR's predictive ability would have been found had the average age of the adult sample been older. One of the issues with conducting research comparing youth and adults is that adults have had more opportunity for crime and perhaps greater experience embedded in an antisocial lifestyle than have youth. Thus, it is no surprise that younger adults were the ones who could be matched to youth. Less than half the original number of youth could be matched to adults, and those who could were often the ones with more previous convictions. Youth that could be matched to adults had more previous convictions than youth who could not be matched, while adults who could be matched had on average 17 fewer convictions than adults who could not be matched, leading to an arguably higher risk sample of youth and lower risk sample of adults than might be expected among community-supervised offenders. In this case, results of the present research might have limited generalisability to other community-supervised samples, as it is possible the groups obtained via the matching process are not representative samples of youth and adults serving community supervision sentences in New Zealand. Furthermore, it is plausible that a youth who already has a large number of previous convictions at age 18 might not realistically have a similar risk level to a 30 year old adult with the same number of previous convictions.

Future research might match youth with adults who had the same number of convictions between the ages of 17-19, information that was unavailable for the present study. For example, a 17 year old with two previous convictions could be matched with a 30-year-old adult who, at the age of 17 had two convictions, despite potentially having a larger number at age 30. This could mean, however that static risk would be higher for the adult sample (due to the possibility that adults would have more convictions by virtue of the passage of time), but might be a good way to examine the predictive validity of the DRAOR using samples that are potentially more comparable. However, an additional issue with that is that if most youth age out of crime, many adults who one might expect to have fewer previous convictions at age 17 may no longer be in the criminal justice system.

Overlapping confidence intervals and limited utility of factors. The present research showed that certain subscales of the DRAOR operated at varying strengths for youth and adults. However, due to the overlap in confidence intervals for all analyses, I am unable to conclude that there a statistical difference in the DRAOR's ability to predict reconviction for youth and adults. Perhaps a larger sample would increase statistical power and yield smaller confidence intervals surrounding point estimates (i.e. greater precision). However, it is also likely that the comparable accuracy found between groups is due to the individual DRAOR items having a reachable limit to their predictive ability. It has been noted by researchers that the process of using factors to assess risk "makes recidivism impossible to predict beyond a certain level of accuracy" (Monahan & Skeem, 2014, p. 162). In other words, tools may be unable to predict beyond a certain level simply due to the fact that their capacity for prediction is limited by the factors selected, or there are other unmeasurable factors that the tool cannot account for. Therefore, a major limitation of any tool is what it cannot predict: the interaction between the immediate environment and human agency; guaranteeing that no prediction method will capture all random variation in recidivism and yield 100% accuracy in prediction. Further, although the DRAOR predicted youth criminal and breach reconviction with a level of accuracy that ranged from moderate to high, the present research cannot rule out that placing greater weight on DRAOR items important during adolescence (e.g., antisocial peers and substance abuse) might increase its predictive ability with youth.

However, encouragingly, research suggests that provided a risk assessment tool is assessing enough recidivism-related factors, it might not matter which tool is used. Yang, Wong and Coid (2010) conducted a meta-analysis of nine static and/or dynamic risk assessment tools and their subscales, concluding that their ability to predict violence was about the same in all cases. All the tools they examined performed moderately, at a level above what would be expected by chance, but none of them outperformed the others. The authors conclude that selection of a risk assessment tool should therefore not be based on how it compares to other tools, but rather which tool will best fit the purpose it is being used for (e.g., management of offenders; Yang et al., 2010). When using factors for the purposes of risk prediction—as in the present study—it has been argued that it does not matter so much how they are working, but rather whether they work as predictors at all (Monahan & Skeem, 2014). The present study was able to show that the DRAOR subscales were measuring items related to reconviction but not which particular items or *how* they were related. Since the DRAOR is used for monitoring rather than prediction purposes, understanding how factors might interact or what specific DRAOR items are most related to reconviction would be a useful avenue for future research.

Conclusion

At the start of this research I asked you to imagine an offender, Chris, who was serving a community supervision sentence. The information provided to you about Chris included several dynamic and static risk factors that are scored in a DRAOR and RoC*RoI assessment respectively. Chris's age was deliberately ambiguous because the Department of Corrections does not assess risk factors using separate age-specific tools. You were asked to consider whether Chris's likelihood of reoffending would change, depending on his age, if he were assessed using a standard risk assessment tool. The present study found one such tool, the DRAOR, to be able to assess the likelihood of reconviction with comparable accuracy for both youth and adult offenders in the community. When scored regularly, DRAOR subscales give useful information pertaining to an offender's likelihood of reconviction. Higher risk scores indicate that both youth and adult offenders are at a greater likelihood of reconviction, while higher protective scores indicate that a youth is less likely to be reconvicted. The RoC*RoI, however, was unable to predict youth and adult breach reconviction or youth criminal reconviction, suggesting it should be used with caution when formulating a risk judgment, particularly for youth. In effect, these findings suggest that, although the RoC*RoI (and arguably static factors) may predict criminal reconviction better for adults compared to youth; the DRAOR is an equally useful tool to assess and monitor community-sentenced youth and adults' risk of both criminal and breach reconviction.

Despite these encouraging results, continued validation of the DRAOR on further samples of both youth and adult offenders is crucial, to ensure these results are not simply due to unique characteristics of the current matched sample, and because, over time, certain factors may lose their predictive ability (Monahan & Skeem, 2014). Presently, however, when using the DRAOR to regularly assess an offender like Chris, a probation officer can be confident that they are assessing factors related to reconviction, whether Chris is 17 or 30 years old.

References

Allen, P., & Bennett, K. (2010). *PASW statistics by SPSS: A practical guide, version 18.0*. (1st ed.). Australia: Cengage Learning.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5*. Washington, D.C.: American Psychiatric Association.

Andrews, D. A., & Bonta, J. (1995). *The Level of Service Inventory–Revised*. Toronto, Canada: Multi-Health Systems.

Andrews, D. A., & Bonta, J. (2010). *The psychology of criminal conduct*. (5th ed.). New Providence, NJ: Anderson Publishing.

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, *17*, 19-52. doi: 10.1177/0093854890017001004

Andrews, D. A., Bonta, J., & Wormith, J. S. (2004). *The Level of Service/Case Management Inventory (LS:CMI): User's Manual.* Toronto: Multi-Health Systems.

Austin, J. (2006). How much risk can we take? The misuse of risk assessment in corrections. *Federal Probation, 70*(2), 58-63. Retrieved from http://search.proquest.com/docview/621493446?accountid=14782

Bakker, L. W., Riley, D., & O'Malley, J. (1999). *ROC, risk of reconviction: Statistical models predicting four types of re-offending*. Wellington, NZ: Department of Corrections.

Becroft, A. (2009). How to turn a child offender into an adult criminal - in 10 easy steps. Paper presented at Children and the Law International Conference, Prato, Italy. Retrieved from http://www.rethinking.org.nz/assets/young_people_and_crime/10_Easy_Steps.pdf

Blakemore, S-J., Burnett, S., & Dahl, R. E. (2010). The role of puberty in the developing adolescent brain. *Human Brain Mapping, 31*, 926-933. doi: 10.1002/hbm.21052

Blakemore, S-J., & Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry, 47*, 296-312. doi: 10.1111/j.1469-7610.2006.01611.x

Bonta, J., Rugge, T., Scott, T-L., Bourgon, G., & Yessine, A. K. (2008). Exploring the black box of community supervision. *Journal of Offender Rehabilitation, 47*, 248-270. doi: 10.1080/10509670802134085

Borum, R. (2003). Managing at-risk juvenile offenders in the community: Putting evidence-based principles into practice. *Journal of Contemporary Criminal Justice, 19*, 114-137. doi: 10.1177/1043986202239745

Borum, R., Bartel, P., & Forth, A. (2006). *SAVRY: Structured Assessment of Violence Risk in Youth: Professional Manual*. PAR.

Brown, S. L., St. Amand, M. D., & Zamble, E. (2009). The dynamic prediction of criminal recidivism: A three-wave prospective study. *Law and Human Behavior, 33*, 25-45. doi: 10.1007/s10979-008-9139-7

Catchpole, R. E. H., & Gretton, H. M. (2003). The predictive validity of risk assessment with violent young offenders: A 1-year examination of criminal outcome. *Criminal Justice and Behavior, 30*, 688-708. doi: 10.1177/0093854803256455

Caudy, M. S., Durso, J. M., & Taxman, F. S. (2013). How well do dynamic needs predict recidivism? Implications for risk assessment and risk reduction. *Journal of Criminal Justice, 41*, 458-466. doi: 10.1016/j.jcrimjus.2013.08.004

Cauffman, E., & Steinberg, L. (2000). (Im)maturity of judgment in adolescence: why adolescents may be less culpable than adults. *Behavioral Sciences and the Law, 18*, 741-760. doi: 10.1002/bsl.416

Chadwick, N. (2014). Validating the Dynamic Risk Assessment for Offender Re-entry (DRAOR) in a sample of U.S. probationers and parolees. Unpublished master's thesis, Carleton University, Ottawa, Canada.

Children, Young Persons, and Their Families Act, New Zealand Statutes. (1989). Retrieved from http://www.legislation.govt.nz/act/public/1989/0024/latest/DLM147094.html ?search=sw_096be8ed80e154b8_17_25_se&p=1&sr=1

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155-159. doi: 10.1037/0033-2909.112.1.155

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*, 7-29. doi: 10.1177/0956797613504966

Dahlberg, L. L. (1998). Youth violence in the United States: Major trends, risk factors, and prevention approaches. *American Journal of Preventative Medicine*, *14*, 259-272. doi: 10.1016/S0749-3797(98)00009-9

Department of Corrections. (2013). *Trends in the offender population* (2013 calendar year). New Zealand. Retrieved from http://www.corrections.govt.nz/__data/assets/pdf_file/0012/738939/Trends_in_the_Offender_Population_2013_updated.pdf

Department of Corrections. (n.d.). *Sentences and orders*. New Zealand. Retrieved from http://www.corrections.govt.nz/working_with_offenders/community_sentences/sentences_and_orders.html

de Ruiter, C., & Nicholls, T. L. (2011). Protective factors in forensic mental health: A new frontier. *International Journal of Forensic Mental Health*, *10*, 160-170. doi: 10.1080/14999013.2011.600602

de Vries Robbé, M., de Vogel, V., Douglas, K. S., & Nijman, H. L. I. (2015). Changes in dynamic risk and protective factors for violence during inpatient psychiatric treatment: Predicting reductions in postdischarge community recidivism. *Law and Human Behavior*, *39*, 53-61. doi: 10.1037/lhb0000089

de Vries Robbé, M., Mann, R. E., Maruna, S., & Thornton, D. (2014). An exploration of protective factors supporting desistance from sexual offending. *Sexual Abuse: A Journal of Research and Treatment, 27*, 16-33. doi: 10.1177/1079063214547582

Douglas, K. S., & Skeem, J. L. (2005). Violence risk assessment: Getting specific about being dynamic. *Psychology, Public Policy, and Law*, *11*, 347-383. doi: 10.1037/1076-8971.11.3.347

Epperson, D. L., Kaul, J. D., Goldman, R., Huot, S., Hesselton, D., & Alexander, W. (2004). *Minnesota Sex Offender Screening Tool–Revised (MnSOST–R)*. St. Paul, MN: Minnesota Department of Corrections.

Farrington, D. P., & Loeber, R. (2002). Serious and violent juvenile offenders. In M. K. Rosenheim, F. E. Zimring, D. S. Tanenhaus & B. Dohrn (Eds.), *A century of juvenile justice*. Chicago, USA: The University of Chicago Press.

Farrington, D. P., Loeber, R., & Howell, J. C. (2012). Young adult offenders: The need for more effective legislative options and justice processing. *Criminology & Public Policy, 11*, 729-750. doi: 10.1111/j.1745-9133.2012.00842.x

Fergusson, D. M., Horwood, L. J., & Ridder, E. M. (2005). Show me the child at seven: the consequences of conduct problems in childhood for psychosocial functioning in adulthood. *Journal of Child Psychology and Psychiatry, 46*, 837-849. doi: 10.1111/j.1469-7610.2004.00387.x

Fortune, C-A., Ferguson, A. M., Hanby, L., & Serin, R. C. (in preparation). *Static and dynamic risk and protective factors in parolees: Comparing the validity of two risk assessment tools with youth and adult offenders.* Manuscript in preparation.

Garson, D. (2013). *Cox Regression*. Asheboro, NC: Statistical Associates Publishing.

Hanby, L. (2013). *A longitudinal study of dynamic risk, protective factors and criminal recidivism: Change over time and the impact of assessment timing.* Unpublished doctoral thesis, Carleton University, Ottawa, Canada. Retrieved from https://curve.carleton.ca/system/files/theses /27525.pdf

Hanson, R. K., Harris, A. J., Scott, T-L., & Helmus, L. (2007). *Assessing the risk of sexual offenders on community supervision: The dynamic supervision project* (Research Report No. 2007-05). Ottawa, Canada: Public Safety Canada. Retrieved from http://www.publicsafety.gc.ca/cnt/ rsrcs/pblctns/ssssng-rsk-sxl-ffndrs/ssssng-rsk-sxl-ffndrs-eng.pdf

Hanson, R. K., & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. *Law and Human Behavior, 24,* 119–136. doi: 10.1023/ A:1005482921333

Harris, P. M. (2006). What community supervision officers need to know about actuarial risk assessment and clinical judgment. *Federal Probation, 70*(2), 8-14. Retrieved from http:// search.proquest.com/docview/9809738?accountid=14782

Hilton, N. Z., Harris, G. T., Rawson, K., & Beach, C. A. (2005). Communicating violence risk information to forensic decision makers. *Criminal Justice and Behavior*, *32*, 97-116. doi: 10.1177/0093854804270630

Hoge, R. D., & Andrews, D. A. (2006). *Youth Level of Service/Case Management Inventory (YLS/CMI): User's Manual*. Multi-Health Systems.

Hoge, R. D., Vincent, G. M., & Guy, L .S. (2012). Prediction and risk/needs assessment. In R. Loeber, & D. P. Farrington (Eds.), *From juvenile delinquency to adult crime: criminal careers, justice policy, and prevention*. New York: Oxford University Press.

Indig, D., Frewen, A., & Moore, E. (2014). Predictors and correlates of re-incarceration among Australian young people in custody. *Australian & New Zealand Journal of Criminology, 0*, 1-17. doi: 10.1177/0004865814550534

Lewis, K., Olver, M. E., & Wong, S. C. P. (2013). The violence risk scale: Predictive validity and linking changes in risk with violent recidivism in a sample of high-risk offenders with psychopathic traits. *Assessment, 20*, 150-164. doi: 10.1177/1073191112441242

Lin, F-J. (2008). Solving multicollinearity in the process of fitting regression model using the nested estimate procedure. *Quality and Quantity, 42*, 417-426. doi: 10.1007/s11135-006-9055-1

Lodewijks, H. P. B., de Ruiter, C., & Doreleijers, T. A. H. (2010). The impact of protective factors in desistance from violent reoffending: A study in three samples of adolescent offenders. *Journal of Interpersonal Violence*, *25*, 568-587. doi: 10.1177/0886260509334403

Loeber, R., & Farrington, D. P. (2012). Introduction. In R. Loeber, & D. P. Farrington (Eds.), *From juvenile delinquency to adult crime: criminal careers, justice policy, and prevention*. New York: Oxford University Press.

Loeber, R., & Farrington, D. P. (2000). Young children who commit crime: Epidemiology, developmental origins, risk factors, early interventions, and policy implications. *Development and Psychopathology, 12*, 737-762. doi: 10.1017/s0954579400004107

Lösel, F., & Farrington, D. P. (2012). Direct protective and buffering protective factors in the development of youth violence. *American Journal of Preventive Medicine*, *43*, S8-S23. doi: 10.1016/j.amepre.2012.04.029

Males, M. (2009). Does the adolescent brain make risk taking inevitable?: A skeptical appraisal. *Journal of Adolescent Research, 24*, 3-20. doi: 10.1177/0743558408326913

Miller, T. W. (2015). *Modeling techniques in predictive analytics: A guide to data science.* Python edition. New Jersey, USA: Pearson Education, Inc.

Ministry of Justice. (n.d.). *Policy development: Child offending and youth justice processes*. Retrieved from http://www.justice.govt.nz/policy/crime-prevention/youth-justice/child-offending-and-youth-justice-processes

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*, 674-701. doi: 10.1037/0033-295x.100.4.674

Monahan, J., & Skeem, J. L. (2014). Risk redux: The resurgence of risk assessment in criminal sanctioning. *Federal Sentencing Reporter, 26,* 158-166. doi: 10.1525/fsr.2014.26.3.158

Moshman, D. (2009). Identity, morality, and adolescent development: Commentary on Proulx and Chandler. *Human Development, 52*, 287-290. doi: 10.1159/000233260

Moshman, D. (2011). Adolescents are young adults, not immature brains. *Applied Developmental Science, 15*, 171-174. doi: 10.1080/10888691.2011.618098

Nadesu, A. (2007). *Reconviction patterns of released prisoners: A 36-months follow-up analysis*. Wellington: Department of Corrections.

New Jersey Attorney General's Office. (2006). *Juvenile Risk Assessment Scale manual*. Retrieved from http://www.state.nj.us/lps/dcj/megan/jras-manual-scale-606.pdf

Norušis, M. J. (2004). SPSS 13.0 Advanced statistical procedures companion. Upper Saddle River, NJ: Prentice Hall.

Olver, M. E., Stockdale, K. C., & Wong, S. C. (2012). Short and long-term prediction of recidivism using the youth level of service/case management inventory in a sample of serious young offenders. *Law and Human Behavior*, *36*, 331-344. doi: 10.1037/h0093927

Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2009). Risk assessment with young offenders: A meta-analysis of three assessment measures. *Criminal Justice and Behavior*, *36*, 329-353. doi: 10.1177/0093854809331457

Olver, M. E., Wong, S. C. P., Nicholaichuk, T., & Gordon, A. (2007). The validity and reliability of the Violence Risk Scale—Sexual Offender Version: Assessing sex offender risk and evaluating therapeutic change. *Psychological Assessment, 19*, 318-329. doi: 10.1037/1040-3590.19.3.318

Perrault, R. T., Paiva-Salisbury, M., & Vincent, G. M. (2012). Probation officers' perceptions of youths' risk of reoffending and use of risk assessment in case management. *Behavioral Sciences and the Law*, *30*, 487-505. doi: 10.1002/bsl.2015

Prentky, R. A., & Righthand, S. (2003). *Juvenile Sex Offender Assessment Protocol: Manual*. Bridgewater, MA: Justice Resource Institute.

Pullmann, M. D. (2010). Predictors of criminal charges for youth in public mental health during the transition to adulthood. *Journal of Child and Family Studies, 19*, 483-491. doi: 10.1007/s10826-009-9320-8

Putniņš, A. L. (2005). Assessing recidivism risk among young offenders. *The Australian and New Zealand Journal of Criminology, 38*, 324-339. doi: 10.1375/000486505774824785

Ralston, C.A., & Epperson, D. L. (2013). Predictive validity of adult risk assessment tools with juveniles who offended sexually. *Psychological Assessment, 25*, 905-916. doi: 10.1037/a0032683

Rennie, C. E., & Dolan, M. C. (2010). The significance of protective factors in the assessment of risk. *Criminal Behaviour and Mental Health*, *20*, 8-22. doi: 10.1002/cbm.750

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC area, Cohen's d, and r. *Law and Human Behavior, 29*, 615-620. doi: 10.1007/s10979-005-6832-7

Roberts, B.W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132*, 1-25. doi: 10.1037/0033-2909.132.1.1

Sampson, R. J., & Laub, J. H. (2005). A life-course view of the development of crime. *Annals of the American Academy of Political and Social Science, 602*, 12-45. doi: 10.1177/ 0002716205280075

Schlager, M. D., & Pacheco, D. (2011). An examination of changes in LSI-R scores over time: Making the case for needs-based case management. *Criminal Justice and Behaviour, 38*, 541-553. doi: 10.1177/0093854811402300

Scott, E. S., & Steinberg, L. (2008). *Rethinking juvenile justice*. Cambridge, MA: Harvard University Press.

Serin, R. C. (2007). The Dynamic Risk Assessment of Offender Re-entry (DRAOR). Unpublished user manual.

Serin, R. C., Mailloux, D., & Wilson, N. J. (2012). *Practice manual for use with Dynamic Risk Assessment For Offender Re-Entry (DRAOR) Scale (New Zealand Revised Version Three - 2012)*: Carleton University, Canada & Department of Corrections, New Zealand.

Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review, 31*, 499-513. doi: 10.1016/j.cpr.2010.11.009

Sourander, A., Elonheimo, H., Niemelä, S., Nuutila, A-M., Helenius, H., Sillanmäki, L., Piha, J., Tamminen, T., Kumpulainen, K., Moilanen, I., & Almqvist, F. (2006). Childhood predictors of male criminality: A prospective population-based follow-up study from age 8 to late adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry, 45*, 578-586. doi: 10.1097/01.chi0000205699.58626.b5

Steinberg, L., Cauffman, E., Woolard, J., Graham, S., & Banich, M. (2009). Are adolescents less mature than adults?: Minors' access to abortion, the juvenile death penalty, and the alleged APA "flip-flop". *American Psychologist, 64*, 583-594. doi: 10.1037/a0014763

Stouthamer-Loeber, M., Wei, E., Loeber, R., & Masten, A. S. (2004). Desistance from persistent serious delinquency in the transition to adulthood. *Development and Psychopathology, 16*, 897-918. doi: 10.1017/s0954579404040064

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics (6th ed.).* USA: Pearson Education Inc.

Tamatea, A., & Wilson, N. (2009). *Dynamic risk assessment for offender re-entry (DRAOR): A pilot study*. New Zealand: Department of Corrections. Unpublished government report.

Ullrich, S., & Coid, J. (2011). Protective factors for violence among released prisoners—effects over time and interactions with static risk. *Journal of Consulting and Clinical Psychology, 79*, 381-390. doi: 10.1037/a0023613

Vincent, G. M., Chapman, J., & Cook, N. E. (2011). Risk-needs assessment in juvenile justice: Predictive validity of the SAVRY, racial differences, and the contribution of needs factors. *Criminal Justice and Behavior, 38*, 42-62. doi: 10.1177/0093854810386000

Vincent, G. M., Guy, L. S., & Grisso, T. (2012). Risk assessment in juvenile justice: A guidebook for implementation. *National Youth Screening & Assessment Project.* MacArthur Foundation.

Vincent, G. M., Perrault, R. T., Guy, L. S., & Gershenson, B. G. (2012). Developmental issues in risk assessment: Implications for juvenile justice. *Victims and Offenders, 7*, 364-384. doi: 10.1080/15564886.2012.713900

Wilson, D. N., & Rolleston, R. (2004). A risk-need profile using four measures for youth offenders incarcerated in young offender units. *New Zealand: New Zealand Department of Corrections*.

Wong, S. C. P., & Gordon, A. (2006). The validity and reliability of the violence risk scale: A treatment-friendly violence risk assessment tool. *Psychology, Public Policy, and Law*, *12*, 279-309. doi: 10.1037/1076-8971.12.3.279

Wong, S. C. P., Olver, M. E., Nicholaichuk, T. P., & Gordon, A. (2003). The violence risk scale: Sexual offender version (VRS-SO). *Regional Psychiatric Centre and University of Saskatchewan, Saskatoon, Saskatchewan, Canada*.

Yang, M., Wong, S. C. P., & Coid J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin*, *136*, 740-767. doi: 10.1037/a0020473

Yesberg, J. A., & Polaschek, D. L. L. (2015). Assessing dynamic risk and protective factors in the community: Examining the validity of the Dynamic Risk Assessment for Offender Re-entry. *Psychology, Crime, & Law, 21,* 80-99. doi: 10.1080/1068316X.2014.935775

Yesberg, J. A., Scanlan, J. M., Polaschek, D. L. L., Hanby, L. J., & Serin, R. C. (2015). Predicting women's recidivism: Validating a dynamic community-based 'gender-neutral' tool. *Probation Journal, 62*, 33-48. doi: 10.1177/0264550514562851

Zhou, X. H., Obuchowski, N. A., & McClish, D. K. (2002). Statistical methods in diagnostic medicine. Wiley & Sons: New York.

Appendices

Appendix A

Summary of Youth and Adult Group Demographic and Offence-related Variables Pre-matching

Variable	Group	
	Youth *M* (*SD*) *n* =274	Adult *M* (*SD*) *n* = 446
Sentence length (in days)	264.25 (67.14)	273.97 (70.69)
Number of previous convictions	8.63 (6.03)	22.54 (22.64)
Number of previous violent convictions	0.72 (0.97)	2.10 (2.52)
Number of previous imprisonments	1.69 (1.34)	3.18 (3.26)
RoC*RoI score	.42 (.15)	.34 (.20)

	Number of youth	Number of adults	Total
Ethnicity			
Māori	117	200	317
European	112	179	291
Pacific Peoples	31	45	76
Other/Unknown	14	22	36
Index offence			
Non-violent	175	256	431
Violent/sexual	72	150	222
Justice/admin	25	34	59
Unknown	2	6	8

Appendix B

Multivariate Regression Models Containing Initial and Proximal DRAOR Subscale Scores and Total Score Predicting Adult Criminal Reconviction (for Non-significant Univariate Predictors)

Multivariate model for initial and proximal DRAOR scores	β (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]
Adult				
Stable score[a]				
Initial	-.07 (.12)	0.37	0.93	[0.74, 1.17]
Proximal	.26 (.11)	6.27*	1.30	[1.06, 1.60]
Protective score[b]				
Initial	.03 (.12)	0.05	1.03	[0.82, 1.29]
Proximal	-.14 (.11)	1.78	0.87	[0.71, 1.07]

* $p < .05$

[a]Model $\chi^2(2, n = 100) = 8.32, p = .016$
[b]Model $\chi^2(2, n = 100) = 3.06, p = .217$

Appendix C

Correlations Between Change Scores and Initial DRAOR Subscale Scores and Total Score for Youth and Adults: Criminal Reconviction

	Initial Acute	Initial Stable	Initial Protective	Initial Total score
Youth				
Acute change	.48**			
Stable change		.28**		
Protective change			.12	
Total score change				.16
Adult				
Acute change	.53**			
Stable change		.23*		
Protective change			.24*	
Total score change				.25*

** $p < .01$, * $p < .05$

Appendix D

*D1. Correlations Between RoC*RoI Scores and Initial (I) and Proximal (P) DRAOR Subscale Scores and Total Score for Youth: Criminal Reconviction*

	Acute		Stable		Protective		Total score	
	I	P	I	P	I	P	I	P
Youth								
RoC*RoI	.08	.08	.20	.08	-.16	-.25*	.18	.16

* $p < .05$

*D2. Multivariate Regression Models Containing RoC*RoI Scores and DRAOR Subscale Scores and Total Score Predicting Youth Criminal Reconviction*

Youth models for incremental predictive validity	β (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]	AUC [95% CI]
Acute model[a]					
Initial acute score	.03 (.06)	0.31	1.03	[0.92, 1.16]	
Proximal acute score	.15 (.06)	6.57*	1.16	[1.04, 1.31]	.68** [.56, .79]
RoC*RoI score	.34 (.75)	0.20	1.40	[0.32, 6.09]	
Stable model[b]					
Initial stable score	-.07 (.08)	0.74	0.94	[0.80, 1.09]	
Proximal stable score	.27 (.07)	14.55**	1.31	[1.14, 1.51]	.75** [.65, .85]
RoC*RoI score	.32 (.75)	0.19	1.38	[0.32, 6.02]	
Protective model[c]					
Initial protective score	-.06 (.07)	0.64	0.94	[0.82, 1.09]	
Proximal protective score	-.12 (.06)	3.91*	0.89	[0.78, 1.00]	.65* [.54, .76]
RoC*RoI score	.07 (.79)	0.01	1.07	[0.23, 5.06]	
Total score[d]					
Initial total score	.01 (.03)	0.03	1.01	[0.95, 1.07]	
Proximal total score	.08 (.03)	9.04**	1.08	[1.03 1.14]	.73** [.63 .84]
RoC*RoI score	-.05 (.76)	0.00	0.95	[0.22, 4.21]	

** $p < .01$, * $p < .05$

[a]Model $\chi^2(3, n = 100) = 13.64, p = .003$
[b]Model $\chi^2(3, n = 100) = 24.60, p < .001$
[c]Model $\chi^2(3, n = 100) = 13.25, p = .004$
[d]Model $\chi^2(3, n = 100) = 22.07, p < .001$

Appendix E

Multivariate Regression Models Containing Initial and Proximal DRAOR Subscale Scores and Total Score Predicting Youth and Adult Breach Reconviction (for Non-significant Univariate Predictors)

Multivariate model for initial and proximal DRAOR scores	β (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]
Youth				
Acute score[a]				
Initial	-.05 (.08)	0.44	0.95	[0.82, 1.11]
Proximal	.23 (.07)	9.98**	1.25	[1.09, 1.44]
Stable score[b]				
Initial	-.11 (.10)	1.23	0.90	[0.74, 1.09]
Proximal	.24 (.08)	9.23**	1.27	[1.09, 1.48]
Protective score[c]				
Initial	.18 (.09)	3.82	1.20	[1.00, 1.44]
Proximal	-.26 (.08)	10.41**	0.77	[0.66, 0.90]
Total score[d]				
Initial	-.08 (.05)	3.02	0.92	[0.84, 1.01]
Proximal	.13 (.04)	11.29**	1.14	[1.05, 1.22]
Adult				
Stable score[e]				
Initial	-.16 (.19)	0.65	0.86	[0.59, 1.25]
Proximal	.41 (.16)	6.38*	1.51	[1.10, 2.08]
Protective score[f]				
Initial	.07 (.18)	0.18	1.08	[0.77, 1.52]
Proximal	-.18 (.16)	1.19	0.84	[0.61, 1.15]
Total score[g]				
Initial	-.06 (.07)	0.63	0.94	[0.82, 1.09]
Proximal	.19 (.07)	6.80**	1.21	[1.05, 1.39]

** $p < .01$, * $p < .05$

Model[a] $\chi^2(2, n = 101) = 13.42, p = .001$
Model[b] $\chi^2(2, n = 101) = 11.00, p = .004$
Model[c] $\chi^2(2, n = 101) = 9.81, p = .007$
Model[d] $\chi^2(2, n = 101) = 14.53, p = .001$
Model[e] $\chi^2(2, n = 101) = 8.08, p = .018$
Model[f] $\chi^2(2, n = 101) = 1.48, p = .476$
Model[g] $\chi^2(2, n = 101) = 11.78, p = .003$

Appendix F

Correlations Between Change Scores and Initial DRAOR Subscale Scores and Total Score for Youth and Adults: Breach Reconviction

	Initial Acute	Initial Stable	Initial Protective	Initial Total score
Youth				
Acute change	.50**			
Stable change		.27**		
Protective change			.21*	
Total score change				.20*
Adult				
Acute change	.55**			
Stable change		.29**		
Protective change			.29**	
Total score change				.33*

** $p < .01$, * $p < .05$

Appendix G

*G1. Correlations Between RoC*RoI Scores and Initial (I) and Proximal (P) DRAOR Subscale Scores and Total Score for Youth and Adults: Breach Reconviction*

	Acute		Stable		Protective		Total score	
	I	P	I	P	I	P	I	P
Youth								
RoC*RoI	.18	.11	.17	.06	-.18	-.26**	.22*	.17
Adult								
RoC*RoI	.24*	.28**	.09	.22*	-.12	-.07	.20*	.23**

** p < .01, * p < .05

*G2. Multivariate Regression Model Containing RoC*RoI Scores and DRAOR Subscale Scores and Total Score Predicting Youth Breach Reconviction*

Youth models for incremental predictive validity	β (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]	AUC [95% CI]
Acute model[a]					
Initial acute score	-.05 (.08)	0.37	0.95	[0.82, 1.11]	
Proximal acute score	.23 (.07)	9.77**	1.25	[1.09, 1.44]	.71** [.60, .82]
RoC*RoI score	-.20 (1.07)	0.04	0.82	[0.10, 6.65]	
Stable model[b]					
Initial stable score	-.11 (.10)	1.22	0.89	[0.73, 1.09]	
Proximal stable score	.24 (.08)	9.14**	1.27	[1.09, 1.48]	.72** [.62, .82]
RoC*RoI score	.10 (0.99)	0.01	1.10	[0.16, 7.66]	
Protective model[c]					
Initial protective score[17]	.20 (.10)	4.08*	1.22	[1.01, 1.49]	
Proximal protective score	-.30 (.09)	10.55**	0.74	[0.62, 0.89]	.68** [.58, .79]
RoC*RoI score	-1.02 (0.99)	1.07	0.36	[0.05, 2.49]	
Total score model[d]					
Initial total score	-.08 (.05)	2.80	0.92	[0.84, 1.01]	
Proximal total score	.13 (.04)	11.10**	1.13	[1.05 1.22]	.73** [.63 .83]
RoC*RoI score	-.13 (1.02)	0.02	0.88	[0.12, 6.53]	

** $p < .01$, * $p < .05$

Model[a] $\chi^2(3, n = 101) = 13.51, p = .004$
Model[b] $\chi^2(3, n = 101) = 11.00, p = .012$
Model[c] $\chi^2(3, n = 101) = 10.57, p = .014$
Model[d] $\chi^2(3, n = 101) = 14.69, p = .002$

[17] The initial protective score becomes inversely predictive for youth once entered into a regression with the proximal score and RoC*RoI score. That is, a one unit increase in protective scores corresponds to a 22% likelihood of breach reconviction. This finding is not consistent with how protective factors are intended to operate (they should promote resilience, not reconviction).

*G3. Multivariate Regression Model Containing RoC*RoI Scores and DRAOR Subscale Scores and Total Score Predicting Adult Breach Reconviction*

Adult models for incremental predictive validity	β (*SE*)	Wald	Hazard ratio	95% CI [Lower, Upper]	AUC [95% CI]
Acute model[a]					
Initial acute score	-.08 (.14)	0.35	0.92	[0.70, 1.21]	
Proximal acute score	.51 (.17)	9.23**	1.67	[1.20, 2.32]	.76** [.62, .91]
RoC*RoI score	-.41 (1.52)	0.07	0.66	[0.03, 12.95]	
Stable model[b]					
Initial stable score	-.15 (.20)	0.58	0.86	[0.58, 1.27]	
Proximal stable score	.40 (.18)	5.27*	1.50	[1.06, 2.11]	.70* [.54, .86]
RoC*RoI score	.17 (1.59)	0.01	1.19	[0.05, 26.89]	
Protective model[c]					
Initial protective score	.08 (.18)	0.19	1.08	[0.76, 1.53]	
Proximal protective score	-.17 (.16)	1.19	0.84	[0.62, 1.14]	.62 [.48, .77]
RoC*RoI score	1.44 (1.50)	0.93	4.23	[0.23, 79.11]	
Total score model[d]					
Initial total score	-.06 (.08)	0.59	0.94	[0.82, 1.09]	
Proximal total score	.19 (.07)	6.17*	1.20	[1.04, 1.39]	74** [.60, .88]
RoC*RoI score	.20 (1.50)	0.02	1.22	[0.07, 23.07]	

** $p < .01$, * $p < .05$

Model[a] $\chi^2(3, n = 101) = 22.61, p < .001$
Model[b] $\chi^2(3, n = 101) = 8.25, p = .041$
Model[c] $\chi^2(3, n = 101) = 2.52, p = .471$
Model[d] $\chi^2(3, n = 101) = 11.92, p = .008$