Comparing different versions of Road to Mental Readiness to determine optimal content

Testing instruction type, homework, and intelligence effects at two timepoints

Deniz Fikretoglu, Ph.D.
Erin Beatty, Ph.D.
Aihua, Liu, Ph.D.

DRDC - Toronto Research Centre

Defence Research and Development Canada

Comparing different versions of Road to Mental Readiness to determine optimal content

Testing instruction type, homework, and intelligence effects at two timepoints

Deniz Fikretoglu, Ph.D.
Erin Beatty, Ph.D.
Aihua, Liu, Ph.D.

DRDC - Toronto Research Centre

Defence Research and Development Canada

Abstract

Observations of Road To Mental Readiness (R2MR) mental health training sessions at the Basic Military Qualification (BMQ) over a two-year period from 2012 to 2014, and a series of studies conducted during the same time period revealed problems with the delivery, the receipt, and the enactment of R2MR concepts and skills. The primary purpose of this study was to compare the two different versions (Versions 5 and 6) of R2MR that were recently developed to identify the version that may lead to better receipt and enactment of key R2MR concepts, especially stress management (and Cognitive Restructuring) skills. A secondary objective was to examine the effects of providing supplemental homework booklets for practicing R2MR concepts/Big 4 skills outside of the 160-minute classroom session. Given the statistically significant and robust effects of intelligence on the uptake and application of R2MR skills observed in a previous study, the current study controlled for the effects of intelligence in looking at the effects of version and homework. The study was a mixed methods design with two between subjects variables (Homework versus No Homework and Version 5 versus Version 6) and one within subjects variable (Time 1 versus Time 2). It was conducted over a 16-day period which included data collection at two time points. Four platoons (approximately 200 recruits) participated in the study. Findings show that Version 5, which was shorter and less technical, consistently outperformed Version 6. Contrary to our hypotheses, supplemental homework did not show beneficial effects on learning. These findings suggest that keeping R2MR material simple and straightforward may lead to better retention of R2MR concepts and also that there may be a natural limit to how much learning can be expected to take place in a single exposure of R2MR.The findings also highlight the importance of empirically testing the effects of any supplemental learning aids that may be developed in the future, as even learning aids with robust face validity may not always produce the intended effects.

Significance to defence and security

This report summarizes the findings from a study aimed at comparing different versions of R2MR to determine optimal content. Our findings suggest keeping R2MR material simple and straightforward may lead to better retention of R2MR concepts within the recruit population and that there may be a natural limit to how much learning can be expected to take place in a single exposure of R2MR.

Résumé

Les observations sur les cours de formation en matière de santé mentale *En route vers la préparation mentale* (RVPM) qui ont été données dans le cadre de la Qualification militaire de base (QMB) au cours d'une période de deux ans, de 2012 à 2014, en plus d'une série d'études réalisées en même temps, ont révélé des problèmes avec l'enseignement, l'assimilation et la mise en pratique des concepts et des compétences des cours RVPM. La présente étude avait pour but principal de comparer deux versions différentes (Version 5 et Version 6) du RVPM que l'on avait récemment développées pour déterminer la version qui assurerait une meilleure assimilation et mise en pratique des concepts du cours RVPM, notamment les compétences de gestion du stress (et de restructuration cognitive). Cette étude avait aussi comme autre but d'examiner les effets d'un cahier de devoirs supplémentaires sur les concepts et les 4 compétences de base du cours RVPM pour s'exercer à l'extérieur des séances de 160 minutes en classe. Étant donné les effets statistiquement significatifs et robustes tirés des observations sur l'assimilation et l'application des compétences du cours RVPM qui avait été constatés dans une étude précédente, dans la présente étude, nous avons contrôlé les effets observés dans l'examen des versions et des devoirs. Nous avons employé une méthode mixte, dans laquelle nous avons employé deux variables inter-sujets (avec devoir / sans devoir et Version 5 / Version 6) et une variable intra-sujets (Temps 1/Temps 2). Cette étude a été menée sur une période de 16 jours. Elle comprenait la cueillette de données à deux intervalles précis dans le temps. Un total de quatre pelotons (approximativement 200 recrues) y ont participé. Les constatations démontrent que la Version 5, plus courte et moins technique que la Version 6, produisait uniformément de meilleurs résultats que la Version 6. Contrairement à nos hypothèses, les devoirs supplémentaires n'ont pas eu d'effets bénéfiques sur l'apprentissage. Ces constatations suggèrent qu'un contenu de cours simple et facile pourrait mener à une meilleure rétention des concepts du cours RVPM, et qu'il existerait une limite naturelle à l'apprentissage auquel on peut s'attendre dans le cadre d'une exposition unique au cours RVPM. Ces constatations mettent aussi en évidence l'importance d'évaluer de façon empirique les effets du matériel d'apprentissage supplémentaire qui pourrait être conçu à l'avenir, puisque ce matériel, même s'il montre une validité claire de prime abord, ne produit pas nécessairement les effets escomptés.

Importance pour la défense et la sécurité

Ce rapport résume les constatations d'une étude qui visait à comparer différentes versions du cours RVPM afin de déterminer un contenu de cours optimal. Nos constatations suggèrent qu'un contenu simple et facile pourrait mener à une meilleure rétention des concepts du cours RVPM au sein de la population des recrues, et qu'il existerait une limite naturelle à l'apprentissage auquel on peut s'attendre dans le cadre d'une exposition unique au cours RVPM.

Table of contents

List of tables

Acknowledgements

This page intentionally left blank.

1   Introduction and purpose of the study

The Road to Mental Readiness (R2MR) program is a mental health intervention that is currently being delivered throughout the Canadian Armed Forces (CAF) career cycle, beginning at the recruit level. At the recruit level, the R2MR material is delivered during Basic Military Qualification (BMQ) in a PowerPoint presentation, during a single 160-minute session. The three key objectives of R2MR, as delivered to the recruits, are: 1) to teach recruits basic mental health literacy concepts (e.g., definition of good and poor mental health), 2) to teach recruits stress management skills they can use to reduce psychological distress and improve performance, and 3) to change recruits' attitudes towards mental health problems and mental health service use.

Defence Research and Development Canada (DRDC) Toronto has undertaken a program of research, the primary purpose of which is to test the efficacy of R2MR in achieving these three key objectives through a group randomized control trial (gRCT). In clinical trials of mental health interventions, "adequate levels of independent treatment components (delivery, receipt, and enactment) *are prerequisite to asserting whether a valid clinical trial has been conducted*" (emphasis added) (Lichstein, Riedel, & Grieve, 1994, p. 1). Delivery refers to the extent to which the intervention has been delivered in the intended (standardized) manner; receipt refers to the comprehension/uptake of the key active ingredients/concepts in the intervention by the target audience, and enactment refers to the extent to which the key active ingredients or concepts i) can be and ii) have been applied by the target audience. Observations of R2MR sessions at the BMQ over a two-year period from 2012 to 2014, and a series of studies conducted during the same time period, revealed problems with all three components (the delivery, the receipt, and the enactment) of R2MR.

More specifically, early observations revealed that there was significant variation in the extent to which R2MR instructors adhered to the standardized material (i.e., delivery). Instructors were observed skipping definitions and practical exercises built into the PowerPoint presentation, especially in the module focusing on the Big 4 skills (i.e., Tactical (Diaphragmatic) Breathing, Goal Setting, Visualization, and Self-Talk (Positive Mantras and Cognitive Restructuring) and in the module focusing on the practical application of concepts and skills through hypothetical scenarios. The largest deviations from the standard material were seen in the delivery of the Cognitive Restructuring part of Self-Talk (Fikretoglu, Lam, & Beatty, 2013).

Cognitive Restructuring is a skill in which different types of maladaptive or irrational thoughts are identified and then challenged by considering the evidence behind them. Maladaptive thoughts are theorized to increase physiological arousal in response to stress (which is in turn theorized to adversely impact performance and psychological well-being). The idea is to challenge and replace maladaptive, negative thoughts with more positive ones, thus reducing physiological arousal and improving psychological well-being and performance. With Cognitive Restructuring, R2MR instructors were observed skipping the definition of each type of maladaptive thought (or defining it poorly) and failing to demonstrate how each type of maladaptive thought can be challenged and replaced with a more positive thought by considering the evidence. Follow-up discussions with R2MR stakeholders revealed that R2MR instructors received limited training in teaching the Big 4 skills; this is problematic considering many R2MR instructors are not mental health professionals and lack clinical training in teaching stress management skills such as the Big 4.

Following these early observations and discussions with R2MR stakeholders, a 2012 study on the effects of sleep and fatigue on the uptake of R2MR showed that a day after taking the BMQ R2MR course, recruits failed to remember and describe a number of key concepts (i.e., receipt): Among these concepts, the most noteworthy problems were detected around the Cognitive Restructuring part of Self-Talk (Fikretoglu et al., 2013). Following discussions with R2MR stakeholders, in addition to creating an Intervention/Treatment Fidelity Checklist to improve the extent to which instructors adhered to standardized R2MR material (i.e., delivery), all stress management skills slides of the PowerPoint presentation were jointly revised by R2MR program developers and DRDC Toronto researchers to improve the uptake and application of the Big 4 stress management skills (i.e., receipt and enactment). For the Self-Talk slides, two versions were created: Version 4 and Version 5. Version 4, with a total of 13 slides, described each type of maladaptive thought (i.e., Overgeneralizing, Catastrophizing, Emotional Reasoning, and Shoulds), gave a BMQ example that the recruits could relate to, and then provided a "remedy" for restructuring and modifying that maladaptive thought. In contrast, Version 5, with a total of nine slides, did not describe each type of maladaptive thought but rather focused on describing the three-step process by which any maladaptive thought can be challenged and restructured (1-Pay attention to what you are saying to yourself, 2-Challenge your negative thinking, 3-Tell yourself helpful things). Essentially, Version 5 was a simpler way of teaching Self-Talk that moved away from technical language, and distilled the key messages around the process of cognitive restructuring. A 2013 study compared Version 4 and Version 5 and found – even after controlling for the effects of intelligence and instructor type (mental health professional versus peer educator versus joint) – that Version 5 outperformed Version 4: recruits who received Version 5 scored significantly higher on the R2MR quiz assessing the receipt and enactment of R2MR skills (Fikretoglu, Beatty, Liu, & Smith, 2014).

However, the 2013 study revealed that there was still room for improvement in the receipt and enactment of R2MR concepts, especially for Self-Talk: even under ideal conditions (i.e., Version 5 being taught by a well-trained mental health professional who fully adhered to the standard material), recruits received a score of 75.32 out of 100 on the Quiz, they scored 3.12 on a four-point item on the Quiz assessing the uptake/receipt of all Big 4 skills, and perhaps most noteworthy, scored 1.67 on a four-point item assessing both receipt and enactment of the Cognitive Restructuring part of Self-Talk (Fikretoglu et al., 2014). A careful review of recruits' responses to the Self-Talk items on the Quiz revealed that many of them confused the two component skills of Self-Talk (i.e., Positive Mantras and Cognitive Restructuring). In order to better differentiate between Positive Mantras and Cognitive Restructuring, Version 6 of the PowerPoint presentation was created, again, jointly by R2MR program developers and DRDC researchers. Version 6 teaches Self-Talk with 12 slides; it explicitly makes a distinction between Positive Mantras and Cognitive Restructuring and describes how the first skill is best used in the moment in high-stakes performance situations and the second skill is best used outside of performance situations with maladaptive thoughts that tend to occur frequently. In addition to the changes made to Self-Talk, both Version 5 and Version 6 were modified in the following manner, with an eye towards increasing and improving the practicing (enactment) of all the Big 4 skills in the 160-minute session: the material on the application of Goal-Setting was expanded, the material on the application of Tactical Breathing was expanded, and the order in which the Big 4 skills are presented was changed so that an easy skill would be followed by a difficult skill (Tactical Breathing (easy), Goal Setting (difficult), Visualization (easy), and Self-Talk (difficult). Finally, recognizing that there was a limit to how much additional practical application could be

built around the Big 4 skills in the 160-minute classroom session, a homework booklet was developed to encourage recruits to practice the skills outside of the classroom.

The primary purpose of this study was to compare the two different versions (Versions 5 and 6) of R2MR that were recently developed to identify the version that may lead to better receipt and enactment of key R2MR concepts, especially stress management (and Cognitive Restructuring) skills. A secondary objective was to examine the effects of providing supplemental homework booklets for practicing of R2MR concepts / Big 4 skills outside of the 160-minute classroom session. Given the statistically significant and robust effects of intelligence on the uptake and application of R2MR skills observed in the previous 2013 study (Fikretoglu et al., 2014), the current study controlled for the effects of intelligence in looking at the effects of version and homework. We used a slightly modified version of the R2MR Quiz that had been developed in previous research. Given that the possible beneficial effects of homework were expected to be detected not immediately after R2MR exposure, but over the ensuing days and weeks, we administered the Quiz twice: first, a day after exposure to R2MR (Time 1, around Week 2 of the BMQ), and second, approximately two weeks later (Time 2, around Week 4 of the BMQ).

1.1　Hypotheses

We hypothesized that Version 6 may lead to better receipt and enactment of R2MR concepts and skills; we further hypothesized that homework may lead to better receipt and enactment of R2MR concepts and skills, especially at Time 2; We did not specify interaction effects as homework was expected to have uniform effects across Versions 5 and 6.

2   Methodology

2.1   Selection of human subjects

All participants in this study were CAF recruits at the Canadian Forces Leadership and Recruit School (CFLRS), in St. Jean-Sur-Richelieu, Quebec, going through the BMQ, who attended their assigned platoon's R2MR session. There were 235 participants who consented to participate in the study; Of those original 235 participants, 194 completed sessions one and two, 19 participants completed session one only, and 22 participants were removed from the analyses because they were recoursed into the platoon between Time 1 and Time 2 and therefore did not receive the R2MR session being evaluated in the current study.

2.2   Design

This study is a mixed methods design with two between subjects variables (Homework versus No Homework and Version 5 versus Version 6) and one within subjects variable (Time 1 versus Time 2). A schematic of the study design is presented below.

Table 1: *Schema of the study design.*

Educational Material	Educational supplement			
	Homework		No Homework	
Version 5	Time 1	Time 2	Time 1	Time 2
Version 6	Time 1	Time 2	Time 1	Time 2

2.3   Materials

2.3.1   Assessment instruments

2.3.1.1   R2MR retention quiz

Receipt and enactment of R2MR material were assessed through a non-validated Quiz composed of 14 multiple-choice and short-answer questions (Annex B). Slight modifications were made to an earlier version of the Quiz and a few additional items were added for the current study. It is important to note that with this very brief Quiz, we could assess the degree of receipt and enactment of R2MR concepts and skills among recruits only in a cursory way. The Quiz items, which can be seen in the annex, assessed a) recruits' understanding of basic mental health concepts, b) recruits' ability to apply the mental health continuum model, c) recruits' ability to remember and describe fully the Big 4 (Tactical Breathing, Visualisation, Self Talk, and Goal Setting), and d) recruits' ability to apply Cognitive Restructuring to various maladaptive thoughts.

In assessing recruits' ability to remember and describe fully the Big 4 skills (Tactical Breathing, Visualisation, Self Talk, and Goal Setting), responses were scored 0–2 per skill (0 = no response

or incorrect response, 1 = provided the name of the skill but only a partial description, 2 = provided the name and a full description of the skill). This original scoring of the Big 4 item on the Quiz thus ranged from a low of zero to a high of eight (this variable was labelled Big4Summary). A second way of scoring this item looked at whether the recruits could remember and describe the item well enough to be able to use the skill or not; 0s and 1s on the original item were collapsed into 0s (i.e., did not describe the skill in way that suggests he/she can use the skill) and 2s became 1s (i.e., described the skill in way that suggests he/she can use the skill). This second, binary, and stricter way of scoring the Big 4 item was labelled Big4SummaryV2. We took a similar approach to the scoring of Self-Talk (Cognitive Restructuring) items on the Quiz. The Self-talk items (10 a & b and 11 a & b) were first scored on a scale of 0–2, with 0s reflecting no or incorrect responses, 1's reflecting partial responses, and 2s reflecting perfect responses (labelled Self-Talk Summary). A second, binary, stricter scoring was also computed, collapsing 0s and 1s from the original variables into 0s, and 2s from the original scoring into 1s (labelled Self-Talk Summary V2). Both versions of the Big4Summary and Self-Talk Summary variables were included in the analyses for a complete picture of recruits' understanding of R2MR concepts and skills.

2.3.1.2   Intelligence / Shipley 2

Based on previous study results, there was a need to control for the impact of intelligence on the receipt and enactment of R2MR concepts and skills. Intelligence is commonly broken down to two types: Crystalized and fluid. Crystalized intelligence relies on prior knowledge and experience, whereas fluid intelligence requires the use of abstract reasoning that allows an individual to solve problems and perceive connections on novel problems without relying on prior knowledge.

The Shipley 2 scale was administered as a test of intelligence. The Shipley 2 has three subscales: verbal (crystalized intelligence), block patterns (fluid intelligence), and abstraction (fluid intelligence) (Shipley, 1940; Shipley, Gruber, Martin, & Klein, 2009). The verbal subscale asks participants to match the target word with another which has the same meaning, choosing from four possible options. There are 40 items which increase in difficulty. For example, large, have the options of red, big, silent and wet. The block pattern subscale presents 12 items of increasing difficulty where participants are presented with two copies of a two dimensional black-and-white block pattern. The second copy has missing pieces and the participants must select the correct options to complete the missing pieces so that the two patterns are identical from four or six options. The task requires participants to mentally rotate the blocks and hold an increasing number of pieces of information in mind in order to solve the problems correctly. The abstraction scale is a measure of problem solving ability. Participants solve 25 problems of increasing complexity. Each problem contains a series of numbers, letters or words that are related by an underlying rule. The participants must determine what the rule is in order to complete the sequence and supply the next item in the series. For example, big little, high low, cold, _ _ _ (answer: hot). Each subscale produces a standardized score. In order to produce an overall intelligence score, the verbal score is combined with one of the measures of fluid intelligence. In the current study, the two subscales combined for the age adjusted overall intelligence score were verbal and abstraction (labelled SSCompA). The combination of verbal and abstraction (problem solving) was selected as it is very common for intelligence assessments to use some combination of verbal skills and some type of problem solving and by choosing this assessment combination, the results from the current study are highly compatible with existing literature. The Shipley 2 has

acceptable levels of internal reliability (split half .91) and test-retest reliability (correlation range from .74–.94). Concurrent validity has been demonstrated between the Shipley 2 and various measures of intelligence including the Wechsler Adult Intelligence Scale III (WAIS III, Wechsler, 1997) and the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999) (Shipley, Gruber, Martin, & Klein, 2009). The Shipley 2 composite score shows a high correlation (0.86) with the WAIS III full scale intelligence quotient. The Shipley 2 is a valid measure of intelligence that can be administered quickly and in a group setting and thus was ideal for the current study.

2.3.1.3   Mental health service use attitudes[1]

Mental health service use attitudes, subjective norms, perceived behavioral control, and intentions (Ajzen, 1991) and the specific beliefs behind these four constructs were measured by a 74-item self-report measure developed specifically among CAF recruits by DRDC Toronto researchers. The CAF Recruit Mental Health Service Use Questionnaire (CAF Recruit-MHSUQ) was developed through a program of research using mixed methods, and has shown promising psychometrics (i.e., internal consistency reliability and factorial validity) in seven separate studies (Fikretoglu, Blais, & Lam, under review). The CAF-MHSUQ was included in the current study in an opportunistic fashion to collect additional data for reliability and validity and was not used in the analyses described below.

2.3.1.4   Marlowe-Crowne social desirability scale[1]

The Marlowe-Crowne Social Desirability Scale (MC-SDS; Crowne & Marlowe, 1960) was designed to measure social desirability independent of psychopathology. It assesses whether respondents are responding truthfully or are misrepresenting themselves. There are 33 items using a true/false response format. Sample items include "I never hesitate to go out of my way to help someone in trouble" and "It is sometimes hard for me to go on with my work if I am not encouraged". The MC-SD was also included in the current study in an opportunistic fashion to collect data on whether responses to the CAF-MHSUQ are affected by the tendency to respond in a socially desirable manner. The MC-SD was therefore not included in the analyses below.

2.4   Procedures

This study was conducted over a 16-day period which included data collection at two time points. The R2MR Quiz, The Shipley, and the CAF Recruit-MHSUQ were administered for the first time the day after R2MR was delivered (Time 1). The R2MR Quiz, the CAF Recruit-MHSUQ, and MC-SD were administered 16 days after R2MR was delivered (Time 2).

Recruits were told at the beginning and end of the R2MR class that there would be an exam on the material they were learning; this is expected to increase the attentional effort recruits expend to uptake the R2MR material, and is now a standard statement that all R2MR instructors are expected to make at the beginning and end of the R2MR session.

At the conclusion of the data collection for Time 1, that is, after completing the Quiz, the Shipley, and the CAF Recruit MHSUQ, participants were made aware of the research study and were

---

[1] Note that the initial findings of the CAF Recruit - MHSUQ and the MC-SDS scales are not included in this report as they are the focus of additional research questions that will be summarized in a future report.

asked to give their informed consent to participate in the study (i.e., have their quiz results, CAF Recruit MHSUQ scores and Shipley scores used for research purposes). Consent was requested after the administration of the study materials as we wanted to have the R2MR Quiz resemble as closely as possible other exams the recruits complete as part of their recruit training. That is, we were interested in testing recruits' receipt and enactment of R2MR materials under conditions that closely mirror all other recruit training exams at CFLRS (not under research study conditions). Completing the Quiz knowing that it is part of a research study may have led to biased estimates (underestimates) of recruits' understanding and ability to apply R2MR concepts and skills. Data from only the participants who gave consent were used for analyses; if a participant did not consent for their data to be used for research purposes, they retained their study materials and did not give them back to the researchers. At Time 2, the assessment session began with seeking consent to link the data from Time 2 to Time 1 for research purposes; following this, the Quiz, the CAF Recruit MHSUQ, and the MC-SD were administered in that order. At Time 2, there was no method by which we could hide from the recruits that fact that completion of the Quiz was part of a research study; the recruits knew that they were completing the Quiz as part of a research study, simply as a result of having participated in the study at Time 1 and having met the research assistant doing the data collection at Time 1.

2.5   Data analyses

Descriptive statistics including minimum value, maximum value, median, lower quartile, upper quartile, mean, and standard deviation (SD) were used to describe the distributions of the three quiz outcomes including: 1) SelfTalkSummary, 2) Big4Summary, and 3) Quiz overall scores at both Time 1 and Time 2, using both versions of scoring. Paired t-test was used to test if there is significant decrease in the outcome scores from Time 1 to Time 2. Given that there was only one platoon per cell in our study design, we also wanted to rule out the possibility of something unusual happening to a specific platoon, thus affecting that platoon's outcome scores (and skewing results). To do this, we used one-way ANOVA (analysis of variance) to test if there was a difference in the mean values of the outcomes among the four platoons.

Given that the participants for this study were recruited in platoons and the intervention (R2MR) was delivered at the platoon level, the resulting data were clustered. We therefore used multivariable mixed linear model to model the outcome variables at Time 1 and Time 2, separately, with platoons treated as random effects to take into account the within-platoon correlation (Gueorguieva, & Krystal, 2004) . For Time 2 outcomes, we constructed two versions of the models. One version did not include the corresponding Time 1 scores in the model; this essentially modelled the absolute value of the outcomes at Time 2. The second version included the corresponding Time 1 scores in the model. This second model essentially modelled changes in the outcomes from Time 1 to Time 2.

Two variables for describing the interventions were created. The first variable indicated the version of R2MR (Version 5 or Version 6) received, and the second variable indicated if the homework booklet was issued or not. To test 1) if the two versions of R2MR lead to different Quiz outcome scores; and 2) to see if homework booklet helps improve the Quiz outcomes, we assessed the fixed effects of these variables from the mixed models. In the presence of results showing a difference between the two versions of R2MR or between using the homework booklet and not using the homework booklet, we calculated Cohen's d, a measure of effect size to

quantify the difference. Cohen's d was computed as the difference in the mean scores of two samples divided by the pooled standard deviation (Rosnow & Rosenthal, 1996). Based on the results of the previous 2013 study (Fikretoglu et al., 2014), the intelligence summary variable (SSCompA) from the Shipley scale was *a priori* considered an important predictor of all the outcomes, and thus was included in all the models.

All analyses were conducted using SAS software version 9.3 (SAS Institute, Cary, North Carolina).

3   Results

Four platoons participated in the study. Platoon 1 was in the Version 5 Homework condition. Platoon 2 was in the Version 5 No Homework condition. Platoon 3 was in the Version 6 no Homework condition. Platoon 4 was in the Version 6 Homework condition. Table 1 shows the characteristics of the study platoons/participants. It can be seen that the four platoons that participated in the study had similar number of recruits. Approximately half of the recruits were in the Version 5 group. Also, approximately half of them were in the homework group.

Table 2: *Characteristics of the study participants.*

Characteristics	Frequency (%) or mean, SD
Platoon	
1 (Version 5, HW)	58 (27.23%)
2 (Version 5, No HW)	59 (27.70%)
3 (Version 6, No HW)	49 (23.00%)
4 (Version 6, HW)	47 (22.07%)
Version of R2MR	
Version 5	117 (54.93%)
Version 6	96 (45.07%)
Homework booklet	
Yes	105 (49.30%)
No	108 (50.70%)
Intelligence score	108.85, 9.93

Table 3 shows the distribution of the Quiz outcome variables at Time 1 and Time 2, separately. Small decreases from Time 1 to Time 2 were seen in the four outcome variables for Big4Summary and Quiz overall scores but not for the two outcome variables for SelfTalkSummary. Results from paired t-tests for testing Time 1 to Time 2 differences indicate that there are significant decreases from Time 1 to Time 2 for four out of the six outcomes: the two BigSummary outcomes and the two Quiz overall scores.

Table 3: *Descriptive statistics of the distribution of the quiz outcome scores.*

Variable name	N	Minimum	Maximum	Lower Quartile	Median	Upper Quartile	Mean	SD
Time 1								
Big4summaryV2	213	0.00	4.00	1.00	2.00	3.00	2.24	1.16
Big4Summary	213	0.00	8.00	5.00	6.00	7.00	5.57	1.84
SelfTalksummaryV2	213	0.00	4.00	1.00	2.00	2.00	1.57	1.10
SelfTalkSummary	213	0.00	8.00	3.00	5.00	6.00	4.60	1.92
QuizOverallOriginal	213	7.00	23.50	15.50	17.50	19.50	17.25	3.29
QuizOveralladdQ	213	9.00	25.50	17.50	19.00	21.50	19.17	3.35
Time 2								
Big4summaryV2*	194	0.00	4.00	1.00	2.00	3.00	1.87	1.25
Big4Summary*	194	0.00	8.00	4.00	5.00	7.00	4.97	2.12
SelfTalksummaryV2	194	0.00	4.00	1.00	2.00	2.00	1.59	1.08
SelfTalkSummary	194	0.00	8.00	3.00	5.00	6.00	4.58	2.00
QuizOverallOriginal *	194	5.00	23.00	14.00	16.50	19.00	16.32	3.84
QuizOveralladdQ*	194	7.00	25.00	16.00	18.50	21.00	18.25	3.92

Note: 1) Big4Summary (data range: 0–8) and Big4SummaryV2 (data range: 0–4) assess ability to remember and describe the Big 4 skills of Tactical Breathing, Goal Setting, Visualization, and Self-Talk; 2) Self-Talk Summary (data range: 0–8) and Self-Talk SummaryV2 (data range: 0–4) assess the ability to apply Cognitive Restructuring Skills to correct maladaptive, negative thoughts; 3) Quiz Overall scores range from 0–25. Quiz OveralladdQ scores range from 0–27 and include the scoring of two additional new items; 4) * scores of Big4summaryV2, Big4Summary, QuizOverallOriginal, and QuizOveralladdQ at Time 2 are significantly lower than those at Time 1 (P-values < 0.0001 from all four paired t-tests).

Table 4 shows the mean values of the Quiz outcome scores for the four platoons. It can be seen that all the four platoons had similar values for the outcomes. Results from ANOVA also indicated that there was no significant difference between any of the two platoons on any of the outcome scores.

Table 4: *Mean values and Standard deviations (SDs) of the quiz outcome scores by platoons.*

Variable	Platoon 1	Platoon 2	Platoon 3	Platoon 4
Time 1				
Big4summaryV2	2.14 (1.28)	2.17 (1.13)	2.14 (1.10)	2.57 (1.06)
Big4Summary	5.36 (1.94)	5.51 (1.83)	5.47 (1.86)	6.02 (1.69)
SelfTalksummaryV2	1.78 (1.09)	1.31 (1.00)	1.63 (1.05)	1.57 (1.25)
SelfTalkSummary	5.00 (1.77)	4.14 (1.79)	4.88 91.80)	4.40 (2.25)
QuizOverallOriginal	17.29 (3.13)	17.06 (3.22)	17.35 (3.23)	17.34 (3.71)
QuizOveralladdQ	19.15 (3.24)	18.93 (3.32)	19.34 (3.22)	19.32 (3.70)
Time 2				
Big4summaryV2	1.94 (1.26)	1.93 (1.30)	1.81 (1.17)	1.77 (1.27)
Big4Summary	5.04 (2.15)	5.22 (2.06)	4.89 (2.05)	4.62 (2.29)
SelfTalksummaryV2	1.62 (1.04)	1.51 (1.09)	1.72 (1.14)	1.51 (1.07)
SelfTalkSummary	4.53 (2.12)	4.64 (2.08)	4.87 (1.88)	4.21 (1.85)
QuizOverallOriginal	16.34 (3.76)	16.68 (3.94)	16.54 (3.73)	15.54 (3.97)
QuizOveralladdQ	18.15 (3.96)	18.64 (3.97)	18.51 (3.79)	17.51 (4.00)

Note: 1) Big4Summary (data range: 0–8) and Big4SummaryV2 (data range: 0–4) assess ability to remember and describe the Big 4 skills of Tactical Breathing, Goal Setting, Visualization, and Self-Talk; 2) Self-Talk Summary (data range: 0–8) and Self-Talk SummaryV2 (data range: 0–4) assess the ability to apply Cognitive Restructuring Skills to correct maladaptive, negative thoughts; 3) Quiz Overall scores range from 0–25. Quiz OveralladdQ scores range from 0–27 and include the scoring of two additional new items.

Table 5 shows the results from the multivariable mixed models for examining the differences in the absolute Quiz outcome scores for using Version 5 versus Version 6 of R2MR. At Time 1, no significant difference was detected between the two versions. At Time 2, two borderline significant differences were detected for the Big4Summary score and Quiz overall score, respectively, indicating an increase in the two outcome scores for those using Version 5. The effect size (Cohen's d) of the beneficial effects of using Version 5 of R2MR were 0.11 and 0.054 for the Big4Summary and the Quiz overall score, respectively, meaning that compared to using Version 6 of R2MR, using Version 5 increased the Big4Summary and Quiz overall scores by 11% and 5.4% of the population means. It is worth noting that for all the six outcomes at Time 2 and four outcomes at Time 1, the results from the mixed models showed beneficial effects of using Version 5 of R2MR compared to using Version 6 of R2MR. Thus, although the beneficial effects for using Version 5 do not reach statistical significance for all outcome variables, they are consistent across the two time points and exist for four out of the six outcomes.

Table 5: *Results from the mixed linear models for the difference between two R2MR versions in quiz outcome absolute scores in the whole sample.*

Outcome	Time 1			Time 2		
	Estimate	P-value	Cohen's d	Estimate	P-value	Cohen's d
Big4Summary_v2	-0.12	0.56	-	0.25	0.17	-
Big4Summary	-0.20	0.52	-	0.54	0.08	0.11
SelfTalkSummary_v2	0.01	0.97	-	0.03	0.87	-
SelfTalkSummary	0.06	0.93	-	0.20	0.56	-
QuizOverallOriginal	0.12	0.78	-	0.88	0.10	0.05
QuizOverallOriginalAddQ	0.00	0.99	-	0.80	0.14	-

Note: 1) Big4Summary (data range: 0–8) and Big4SummaryV2 (data range: 0–4) assess ability to remember and describe the Big 4 skills of Tactical Breathing, Goal Setting, Visualization, and Self-Talk; 2) Self-Talk Summary (data range: 0–8) and Self-Talk SummaryV2 (data range: 0–4) assess the ability to apply Cognitive Restructuring Skills to correct maladaptive, negative thoughts; 3) Quiz Overall scores range from 0–25. Quiz OveralladdQ scores range from 0–27 and include the scoring of two additional new items.

Table 6 shows the results from the mixed models for examining the differences in the Quiz outcome absolute scores between using the homework booklet and not using the homework booklet conditions. The results from the models constructed among the whole sample show opposite effects for using the homework booklet between Time 1 and Time 2. Specifically, for the whole study sample, at Time 1, homework booklet showed beneficial but non-significant effects for all of the Quiz outcomes. Contrary to this, at Time 2, the homework booklet showed harmful and non-significant effects for all the Quiz outcomes.

In order to further explore the effects of using the homework booklet, we constructed the mixed models among subjects receiving Version 5 and Version 6 of R2MR separately. The results are also presented in Table 6. In general, the results are consistent for Time 2 outcomes but are inconsistent for Time 1 outcomes between the two versions R2MR samples. Specifically, for both Version 5 and 6 samples, the results show decreases in most of the Time 2 Quiz outcome scores when using the homework booklet. Contrary to this, at Time 1, for both Version 5 and 6 samples, the results for using the homework booklet were mixed, with beneficial and harmful effects. For the Version 5 R2MR sample, the effects of using the homework booklet were beneficial and significant for the two SelfTalkSummary outcomes, with effect size (Cohen's d) of 0.32 and 0.20. For the other outcomes, the effects were mixed with beneficial but non-significant effects for the Quiz overall scores, and harmful and non-significant effects for the Big4Summary scores. For the Version 6 R2MR sample, the results of the effects of using the homework booklet were also mixed with beneficial but non-significant effects for the Big4Summary, and harmful and non-significant effects for the SelfTalkSummary and the Quiz overall scores.

In summary, for testing the effects of using a supplemental homework booklet in Quiz outcome absolute scores, the only consistent finding was that it is harmful for almost all of Quiz outcomes at Time 2. Beyond this, the results for the effect of using the homework booklet are inconsistent across time, across different outcomes, and across different versions of R2MR.

Table 6: *Results from the mixed linear models for the difference between using homework booklet and not using homework booklet in quiz outcome absolute scores.*

Outcome	Time 1			Time 2		
	Estimate	P-value	Cohen's d	Estimate	P-value	Cohen's d
Whole sample						
Big4Summary_v2	0.17	0.39	-	-0.03	0.88	-
Big4Summary	0.16	0.60	-	-0.25	0.40	-
SelfTalkSummaryV2	0.20	0.50	-	-0.05	0.78	-
SelfTalkSummary	0.18	0.81	-	-0.40	0.25	-
QuizOverallOriginal	0.08	0.85	-	-0.71	0.17	-
QuizOverallOriginalAddQ	0.06	0.89	-	-0.79	0.14	-
Version 5 R2MR sample						
Big4Summary_v2	-0.02	0.94	-	0.03	0.91	-
Big4Summary	-0.13	0.70	-	-0.17	0.68	-
SelfTalkSummaryV2	0.49	0.099	0.32	0.13	0.49	-
SelfTalkSummary	0.89	0.01	0.20	-0.07	0.84	-
QuizOverallOriginal	0.28	0.62	-	-0.28	0.69	-
QuizOverallOriginalAddQ	0.27	0.65	-	-0.42	0.56	-
Version 6 R2MR sample						
Big4Summary_v2	0.38	0.08	0.16	-0.13	0.60	-
Big4Summary	0.47	0.18	-	-0.44	0.32	-
SelfTalkSummaryV2	-0.09	0.68	-	-0.24	0.32	-
SelfTalkSummary	-0.55	0.19	-	-0.74	0.07	0.16
QuizOverallOriginal	-0.20	0.77	-	-1.29	0.10	-
QuizOverallOriginalAddQ	-0.21	0.75	-	-1.28	0.11	-

Note: 1) Big4Summary (data range: 0–8) and Big4SummaryV2 (data range: 0–4) assess ability to remember and describe the Big 4 skills of Tactical Breathing, Goal Setting, Visualization, and Self-Talk; 2) Self-Talk Summary (data range: 0–8) and Self-Talk SummaryV2 (data range: 0–4) assess the ability to apply Cognitive Restructuring Skills to correct maladaptive, negative thoughts; 3) Quiz Overall scores range from 0–25. Quiz OveralladdQ scores range from 0–27 and include the scoring of two additional new items.

Table 7 shows the results from the mixed linear models for investigating the effects of using different versions of R2MR and using the homework booklet in the changes of the Quiz outcome scores from Time 1 to Time 2. Given that the scores for Big4summary and Quiz overall score generally decreased from Time 1 to Time 2, a positive effect on these outcomes means less decrease from Time 1 to Time 2, and a negative effect means more decrease from Time 1 to Time 2. For SelfTalkSummary scores, given that the scores are similar between the two time points, a positive effect could be interpreted as an increase in the outcome score from Time 1 to Time 2 and a negative effect could be interpreted as a decrease in the outcome score from Time 1 to Time 2. The results indicate that compared to Version 6 of R2MR, Version 5 leads to less decrease from Time 1 to Time 2 in Big4Summary and Quiz overall scores. Contrary to this, using homework booklet was found to be associated with greater decrease in SelfTalkSummary score from Time 1 to Time 2.

Table 7: Results from the mixed linear models for modeling the difference in quiz outcome scores from Time 1 to Time 2.

Outcome	Effects of Version 5 R2MR			Effects of homework booklet		
	Estimate	P-value	Cohen's d	Estimate	P-value	Cohen's d
Big4Summary_v2	0.31	0.10	0.17	-0.10	0.59	-
Big4Summary	0.67	0.04	0.13	-0.28	0.39	-
SelfTalkSummary_v2	-0.00	1.00	-	-0.10	0.43	-
SelfTalkSummary	0.13	0.60	-	-0.41	0.08	0.09
QuizOverallOriginal	0.71	0.10	0.04	-0.53	0.21	-
QuizOverallOriginalAddQ	0.70	0.11	-	-0.57	0.18	-

Note: 1) Big4Summary (data range: 0–8) and Big4SummaryV2 (data range: 0–4) assess ability to remember and describe the Big 4 skills of Tactical Breathing, Goal Setting, Visualization, and Self-Talk; 2) Self-Talk Summary (data range: 0–8) and Self-Talk SummaryV2 (data range: 0–4) assess the ability to apply Cognitive Restructuring Skills to correct maladaptive, negative thoughts; 3) Quiz Overall scores range from 0–25. Quiz OveralladdQ scores range from 0–27 and include the scoring of two additional new items; 4) the effects of Version 5 R2MR was calculated in relative to Version 6 R2MR; 5) the effects of homework booklet was calculated in relative to without homework booklet.

4   Conclusion

The primary purpose of this study was to compare the two different versions (Versions 5 and 6) of R2MR that were recently developed to identify the version that may lead to better receipt and enactment of key R2MR concepts, especially stress management (and Cognitive Restructuring) skills. A secondary objective was to examine the effects of providing supplemental homework booklets for practicing of R2MR concepts / Big 4 skills outside of the 160-minute classroom session.

Looking at the pattern of results for testing to see if Version 6 led to better learning and application of R2MR concepts and skills than Version 5, we found, contrary to our hypothesis, that Version 5 consistently outperformed Version 6, although only a few of the results trended towards statistical significance, likely due to small statistical power. The effect sizes (Cohen's d) for the differences favoring Version 5 were very small. All in all, these results suggest that Version 5, which teaches Self-Talk with nine slides and distills the messages around what constitutes Cognitive Restructuring does a better job than Version 6, which teaches Self-Talk with 12 slides and explicitly makes a distinction between Positive Mantras and Cognitive Restructuring. Considering findings from this study in the context of the larger set of findings from testing different versions of R2MR, we note that Version 4 and Version 6 of R2MR both included more material than Version 5. In fact, informal reports from the R2MR instructor who delivered both versions of R2MR in the current study was that using Version 6 took significantly more time and led to a "time crunch" at the end of the 160-minute session. We therefore conclude that especially when it comes to the teaching of Self-Talk, less (material) is better than more. We recommend keeping Version 5 as the "best working version of R2MR" for the future group randomized control trial. It should also be noted that the small magnitude of the differences between Versions 5 and 6 suggest that there is likely a limit to how much the receipt and enactment of R2MR skills can be improved with modifications to the standard PowerPoint material for the 160-minute classroom session. The effect size (Cohen's d) of the differences seen in the previous study (comparing Versions 4 and 5) was about 0.3. The differences between Versions 5 and 6 in the current study are much, much smaller. Based on the set of findings from the current and the previous study, and the significant effort already invested in developing different versions of R2MR, we believe that further investments in the modification of in-class material are unlikely to lead to significant improvements in receipt or enactment of R2MR skills.

If modifications to the material for the 160-minute classroom session are no longer leading to significant improvements in uptake or enactment, then it makes sense to look at ways to enhance engagement with and practicing of R2MR skills and concepts outside of the classroom session. Looking at the pattern of findings from the current study for testing to see if supplementing learning with a homework booklet outside of the 160-min classroom session was beneficial, we found, contrary to our hypotheses, that for the most part, trying to supplement learning with a homework booklet was not beneficial or even harmful. This is surprising in that assigning homework for the practice of clinical skills such as Cognitive Restructuring outside of the clinical setting is extremely common and a core component of many evidence based therapies such as cognitive behavioral therapy (CBT) (Young and Beck, 1980); furthermore completing homework has been found to be associated with greater reductions in patient symptoms in CBT (Kazantzis, Deane, & Ronan, 2000). There could be a number of reasons why homework had harmful effects: one, it may be that assigning homework reinforced the perception that the Big4 skills are difficult

to employ in real life, this may have led to decreased perceived mastery over the material and may have discouraged further engagement with the material. Two, it may be that the recruits, at the conclusion of the quiz data collection at Time 1 wrongly assumed that the research study was over and did not hold onto the homework booklet, although this explains the absence of beneficial effects rather than the presence of harmful effects. Finally, it may be that recruits indeed followed direction and used the homework booklets between Time 1 and 2 but having poorly or partially understood concepts and skills in the classroom session, further engagement with the material led to greater confusion than before. Unfortunately, the available information from the current study does not allow for testing these potential explanations. What is clear, however, is that assigning homework booklets as supplemental learning materials did not have the expected beneficial effects in the current study and therefore should be abandoned at this time, absent a rather lengthy, multipronged research program as to what happens when recruits are assigned homework (i.e., do they feel pressured, do they do it, how often do they do it, and what are the obstacles in the larger BMQ training context to following through with homework).

There are other potential methods for increasing recruits' engagement with and application of R2MR concepts and skills. First, additional practice sessions following the 160-minute session at week 2 can be added to the BMQ training schedule, with the additional session being devoted fully to the practicing of the Big 4 skills. Unfortunately, it is difficult to add onto an already busy 13-week training period; furthermore, this requires additional human resources than currently available in the R2MR program. It is also possible to train CFLRS staff to reinforce R2MR concepts and skills throughout BMQ training; this is being planned for after the group randomized control trial. Other efforts to take learning beyond the 160-minute classroom session include the development of apps that can be used to practice the Big 4. We believe that all of these are worthwhile venues to explore, as there seems to be a natural limit to how much recruits can learn in a single classroom session, even under the best case scenario of a skillful clinician delivering the material perfectly.

There are two noteworthy incidental findings in the current study. First, the decrease in the overall Quiz scores, Big 4 summary scores, and Self-Talk summary scores from Time 1 to Time 2 was quite small in magnitude. This is somewhat reassuring in that it shows recruits retain most of what they have learned from the classroom session several weeks after exposure to R2MR.

Second, although the findings are not reported in the current report, similar to the findings from the previous study that tested the effects of intelligence (Fikretoglu et al., 2014), we found in the current study that intelligence indeed has robust, statistically significant, and consistent effects in the receipt and enactment of R2MR concepts and skills. The distribution of intelligence scores in the recruit population in both studies was nearly normal. As well, the previous study showed that the magnitude of the difference in the Quiz outcome scores between the high and low intelligence subgroups was moderate-to-large (Cohen's d=0.57-0.68). These intelligence findings, together with the findings from the two studies that show simpler, less technical versions leading to better outcomes reinforce the message that it may be prudent to keep R2MR material simple and straightforward; also, these findings highlight the fact that there may be a natural limit to how much recruits can be expected to learn in their first exposure to R2MR, especially in a brief 160-minute single classroom session and that further modifications to the in-class material may not be sufficient to overcome these natural limits. Whether the recruits' level of learning following a single, 160-minute session is sufficient to bring about the hoped for reductions in psychological

distress and improvements in training performance is an empirical question; the planned group randomized control study is intended to answer this empirical question.

It is important that a number of limitations be considered in interpreting the above findings. First, as stated in the Methods section of this report, with the Quiz used in this study, we could assess the degree of receipt and enactment of R2MR concepts and skills among recruits only in a cursory way. There are no validated measures of the Big 4 skills and the mental health continuum model that we are aware of; we therefore developed a brief measure to assess the Big 4 skills and the mental health concepts (including the continuum model) in R2MR. While the Quiz gives us a cursory look at problems that may exist with the receipt and enactment of R2MR skills and concepts, it is limited in assessing the degree or the extent of these problems.

Second, as stated in the Methods section, we could not hide the fact that the Quiz was part of a research study at Time 2. This may lead to possible underestimates of receipt and enactment of R2MR skills and concepts at Time 2.

Third, and perhaps most importantly, in this study family-wise error correction was not applied to adjust the p-values for defining statistical significance. This decision was based on two important considerations. First, p-values should be adjusted for family-wise error only when multiple independent tests are performed for one single hypothesis testing. This is not the case in our study. The significance tests performed in our study are not independent of each other. Although a total of six outcome variables are used in our study, as explained in the Methods section, there are actually three outcomes with two different ways of scoring for each of them. Thus, the two outcome variables coming from the same questions but with different scoring methods are highly correlated (correlation coefficients for the correlations between the two different scoring variables are 0.90, 0.91, and 0.997 for SelfTalkSummary, Big4Summary, and Quiz overall outcome pairs, respectively). Moreover, the outcome variables from the three different questions are also correlated. For example, the correlation between Quiz overall score and 1) Big4Summary score, 2) SelfTalkSummary score are as high as 0.73 and 0.74, respectively. Given that these outcome variables are correlated, it is highly plausible that similar results from statistical testing will be obtained for different outcomes. In this kind of situation, type I error rate is far less inflated than in the case where all the outcomes used in statistical testing are independent of each other. In other words, applying multiple correction procedure to adjust p-values for significance tests might be inappropriate when the outcomes are correlated as it has the potential of leading to under-estimation and under-reporting of study findings (Perneger, 1998; Schulz & Grimes, 2005). Second, our study is a pilot study. This means that we have low statistical power for detecting significant effects based on our sample. The pilot/exploratory nature of the study means that we would like to explore possible findings based on consistency across different outcomes and clinical meaning, and we pay less attention to statistical testing. In other words, we reason, as others do in the literature, that in small, exploratory pilot studies "it is better to tolerate findings that may later prove to be false than to prematurely discard potentially useful observations because of Type 2 errors caused by corrections for multiplicity" (Streiner & Norman, 2011, p.17). Nevertheless, we note that not applying multiple correction is an important limitation to our study. To further test the statistically significant findings from the current study, a bigger study with adequate power and with appropriate approaches for correcting multiple testing is needed.

Despite these limitations, this study adds to a body of knowledge on factors that may help optimize R2MR at BMQ. The findings point to the importance of continuing to consider ways to take R2MR learning beyond the classroom context as a way to enhance possible beneficial effects.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*, 179-211.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354.

Fikretoglu, D., Beatty, E., Liu, A., & Smith, I. (2014). Optimizing R2MR at BMQ: Looking at the effects of instructor type, instruction type, intelligence, and personality. Presentation given to the Directorate General Health Services (DGHS) (R2MR project stakeholders), Ottawa, ON, January 28, 2014.

Fikretoglu, D., Blais, A-R., Lam, Q., & McCreary, D.R. Development and initial psychometrics of a military mental health service use attitude questionnaire. Manuscript under review at *Assessment*.

Fikretoglu, D., Lam, Q., & Beatty, E. Recommendations to improve treatment fidelity and the uptake of R2MR material prior to a group randomized control trial. DRDC Toronto LR-2013-10055-1-1404HD0500 to Directorate General Health Services, Directorate of Mental Health, June 24, 2013.

Lichstein, K.L., Riedel, B.W., & Grieve, R. (1994). Fair tests of clinical trials: A treatment implementation model. *Advances in Behaviour Research and Therapy, 16,* 1-29.

Gueorguieva R., & Krystal, J.H. (2004). Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry. *Archives of General Psychiatry, 61(3),* 310-7.

Kazantzis, N., Deane, F. P., & Ronan, K. R. (2000). Homework assignments in cognitive and behavioral therapy: A meta-analysis. *Clinical Psychology: Science and Practice, 7,* 189–202.

Perneger T.V. (1998). What's wrong with Bonferroni adjustments. *British Medical Journal,* 316,11, 236-38.

Rosnow, R. L., & Rosenthal, R. (1996). Computing contrasts, effect sizes, and counternulls on other people's published data: General procedures for research consumers. *Psychological Methods*, 1, 331-340.

Schulz K.F. and Grimes D.A. (2005). Multiplicity in randomized trials I: endpoints and treatments. *Lancet,* 2005, 365, 1591-95.

Shipley, W. C. (1940). A self-administering scale for measuring intellectual impairment and deterioration. *The Journal of Psychology, 9,* 371-377.

Shipley, W. C., Gruber, C. P., Martin, T. A. & Klein, A. M. (2009). *Shipley-2 Manual.* Los Angeles: Western Psychological Services.

Streiner D.L and Norman G.R. (2011) Correction for multiple testing – is there a resolution? *Chest* 2011,140(1), 16-18.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale, Third Edition.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence.* San Antonio, TX: Psychological Corporation.

Young, J., & Beck, A. T. (1980). *Cognitive Therapy Scale: Rating manual.* (Available from the Beck Institute for Cognitive Therapy and Research, GSB Building, One Belmont Avenue, Suite 700, Bala Cynwyd, PA 19004-1610).

Annex A Electronic copy of Homework Booklet

The electronic copy of the Homework Booklet can be accessed by double clicking on the image below.
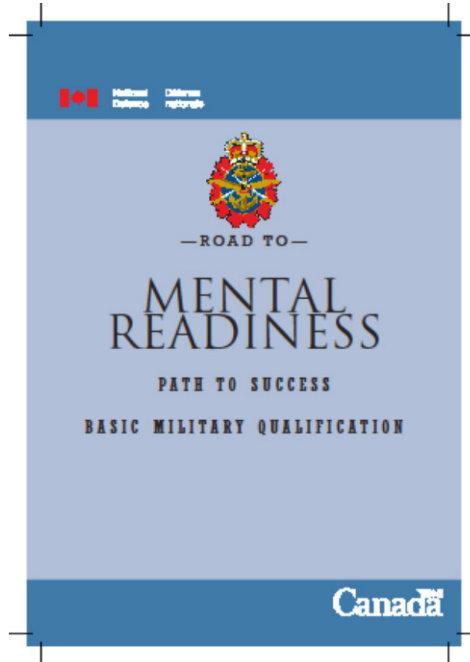
Figure A.1: Homework Booklet.

This page intentionally left blank.

Annex B   Mental Toughness and Basic Training Recall Test

Knowledge Questions:

1. All stress is negative and impacts our performance.

    a. True

    b. False

2. Mental health is an all or nothing state; you are either sick or healthy.

    a. True

    b. False

3. Please review the following case:

    Joe is a member of the CAF. He just returned from a tour of duty. Over the past three months he has begun reacting with anger to minor inconveniences. His attitude has become very negative at his job and he has trouble focussing. At night he has trouble sleeping due to his nightmares. He started online gambling, but is now losing more than he can afford. Joe has been avoiding meeting up with his friends and often fights with his wife.
    Thinking back to the **Mental Health Continuum Model**, try to remember where in the continuum this person's symptoms fall:

    a. Green (Healthy)

    b. Yellow (Reacting)

    c. Orange (Injured)

    d. Purple (Permanently sick)

4. Peer support is an important concept in the CAF. Name three helpful actions you can employ to take care of your buddy/teammate?

    a.

    b.

    c.

5. What are the four key strategies (The Big Four) in managing stress and arousal levels? List and describe below, if you do not remember what the strategy is called then just describe it. If you are having a hard time describing it, then provide an example.

a. Strategy 1: _____. Describe in your own words what it is:

b. Strategy 2: _____. Describe in your own words what it is:

c. Strategy 3: _____. Describe in your own words what it is:

d. Strategy 4: _____. Describe in your own words what it is:

6. Whether to seek help or not for a mental health problem is a personal decision; we would suggest seeking help when you are in the following range of the **Mental Health Continuum Model**:

   a. Green

   b. Red

   c. Yellow-orange

   d. Purple

   e. Never, deal with it on your own.

7. The CAF has mental health resources which are easily accessible and available to assist someone all along the **Mental Health Continuum Model**.

   a. True

   b. False

8. Donna is a member of the CAF. Donna has been posted to a new position and had to move with her husband and child to a new base in a different province. She has been feeling sad and overwhelmed. Lately she has been procrastinating over fulfilling commitments in her personal life. At work she has been putting in extra time and leaves very late every night. She now has regular headaches and has trouble sleeping. Donna has begun drinking regularly in the evenings. She has been avoiding her friends and turns down their invitations.

   Thinking back to the **Mental Health Continuum Model**, try to remember where in the continuum this person's symptoms fall:

   a. Green (Healthy)

   b. Yellow (Reacting)

   c. Orange (Injured)

    d.   Purple (Permanently sick)

9. Why do people wait to seek care? (list three reasons)

    a.

    b.

    c.

10. After failing his swim test Ted kept saying to himself that: "this always happens to me" and then he said "I am never going to pass this course".

This page intentionally left blank.

List of symbols/abbreviations/acronyms/initialisms

ANOVA	Analysis of Variance
BMQ	Basic Military Qualification
CAF	Canadian Armed Forces
CAF-MHSUQ	CAF mental Health Service Use Questionnaire
CBT	Cognitive Behavioral Therapy
CFLRS	Canadian Forces Leadership and Recruit School
DRDC	Defence Research and Development Canada
GRCT	Group Randomized Control Trial
MC-SD	Marlowe-Crowne Social Desirability
R2MR	Road To Mental Readiness
SD	Standard Deviation

This page intentionally left blank.

DOCUMENT CONTROL DATA		
(Security markings for the title, abstract and indexing annotation must be entered when the document is Classified or Designated)		
1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g., Centre sponsoring a contractor's report, or tasking agency, are entered in Section 8.) DRDC – Toronto Research Centre Defence Research and Development Canada 1133 Sheppard Avenue West Toronto, Ontario M3K 2C9 Canada	2a. SECURITY MARKING (Overall security marking of the document including special supplemental markings if applicable.) UNCLASSIFIED	
	2b. CONTROLLED GOODS (NON-CONTROLLED GOODS) DMC A REVIEW: GCEC DECEMBER 2012	
3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.) Comparing different versions of Road to Mental Readiness to determine optimal content : Testing instruction type, homework, and intelligence effects at two timepoints		
4. AUTHORS (last name, followed by initials – ranks, titles, etc., not to be used) Fikretoglu, D. Ph.D.; Beatty, E. Ph.D.; Aihua, L. Ph.D.]		
5. DATE OF PUBLICATION (Month and year of publication of document.) December 2014	6a. NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.) 43	6b. NO. OF REFS (Total cited in document.) 17
7. DESCRIPTIVE NOTES (The category of the document, e.g., technical report, technical note or memorandum. If appropriate, enter the type of report, e.g., interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.) Scientific Report		
8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.) DRDC – Toronto Research Centre Defence Research and Development Canada 1133 Sheppard Avenue West Toronto, Ontario M3K 2C9 Canada		
9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)	
10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.) DRDC-RDDC-2014-R164	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)	
11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.) Unlimited		
12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.)) Unlimited		

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

Observations of Road To Mental Readiness (R2MR) mental health training sessions at the Basic Military Qualification (BMQ) over a two-year period from 2012 to 2014, and a series of studies conducted during the same time period revealed problems with the delivery, the receipt, and the enactment of R2MR concepts and skills. The primary purpose of this study was to compare the two different versions (Versions 5 and 6) of R2MR that were recently developed to identify the version that may lead to better receipt and enactment of key R2MR concepts, especially stress management (and Cognitive Restructuring) skills. A secondary objective was to examine the effects of providing supplemental homework booklets for practicing R2MR concepts/Big 4 skills outside of the 160-minute classroom session. Given the statistically significant and robust effects of intelligence on the uptake and application of R2MR skills observed in a previous study, the current study controlled for the effects of intelligence in looking at the effects of version and homework. The study was a mixed methods design with two between subjects variables (Homework versus No Homework and Version 5 versus Version 6) and one within subjects variable (Time 1 versus Time 2). It was conducted over a 16-day period which included data collection at two time points. Four platoons (approximately 200 recruits) participated in the study. Findings show that Version 5, which was shorter and less technical, consistently outperformed Version 6. Contrary to our hypotheses, supplemental homework did not show beneficial effects on learning. These findings suggest that keeping R2MR material simple and straightforward may lead to better retention of R2MR concepts and also that there may be a natural limit to how much learning can be expected to take place in a single exposure of R2MR.The findings also highlight the importance of empirically testing the effects of any supplemental learning aids that may be developed in the future, as even learning aids with robust face validity may not always produce the intended effects.

_____-

Les observations sur les cours de formation en matière de santé mentale *En route vers la préparation mentale* (RVPM) qui ont été données dans le cadre de la Qualification militaire de base (QMB) au cours d'une période de deux ans, de 2012 à 2014, en plus d'une série d'études réalisées en même temps, ont révélé des problèmes avec l'enseignement, l'assimilation et la mise en pratique des concepts et des compétences des cours RVPM. La présente étude avait pour but principal de comparer deux versions différentes (Version 5 et Version 6) du RVPM que l'on avait récemment développées pour déterminer la version qui assurerait une meilleure assimilation et mise en pratique des concepts du cours RVPM, notamment les compétences de gestion du stress (et de restructuration cognitive). Cette étude avait aussi comme autre but d'examiner les effets d'un cahier de devoirs supplémentaires sur les concepts et les 4 compétences de base du cours RVPM pour s'exercer à l'extérieur des séances de 160 minutes en classe. Étant donné les effets statistiquement significatifs et robustes tirés des observations sur l'assimilation et l'application des compétences du cours RVPM qui avait été constatés dans une étude précédente, dans la présente étude, nous avons contrôlé les effets observés dans l'examen des versions et des devoirs. Nous avons employé une méthode mixte, dans laquelle nous avons employé deux variables inter-sujets (avec devoir / sans devoir et Version 5 / Version 6) et une variable intra-sujets (Temps 1/Temps 2). Cette étude a été menée sur une période de 16 jours. Elle comprenait la cueillette de données à deux intervalles précis dans le temps. Un total de quatre pelotons (approximativement 200 recrues) y ont participé. Les constatations

démontrent que la Version 5, plus courte et moins technique que la Version 6, produisait uniformément de meilleurs résultats que la Version 6. Contrairement à nos hypothèses, les devoirs supplémentaires n'ont pas eu d'effets bénéfiques sur l'apprentissage. Ces constatations suggèrent qu'un contenu de cours simple et facile pourrait mener à une meilleure rétention des concepts du cours RVPM, et qu'il existerait une limite naturelle à l'apprentissage auquel on peut s'attendre dans le cadre d'une exposition unique au cours RVPM. Ces constatations mettent aussi en évidence l'importance d'évaluer de façon empirique les effets du matériel d'apprentissage supplémentaire qui pourrait être conçu à l'avenir, puisque ce matériel, même s'il montre une validité claire de prime abord, ne produit pas nécessairement les effets escomptés.

---

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g., Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Road to Mental Readiness