

Case Study 1: An Evidence-Based Practice Review Report**Theme: School Based Interventions for Learning*****'Is "keepin' it REAL" and effective intervention programme for mainstream secondary aged students?'*****Summary**

With an increased emphasis on current social, emotional and mental health issues in young people, substance abuse is an area of current focus for the UK Government (HM Government, 2017b) for early intervention to reduce the risk of secondary mental health issues. With the UK Government looking to increase funding for early intervention for substance abuse (HM Government, 2017b), it is important to consider which interventions are likely to be most effective and therefore support reduction in social, emotional and mental health issues. The "keepin' it REAL" substance use intervention programme by Miller-Day and Hecht (kiR - keepin' it REAL, n.d.; Gosin, Marsiglia & Hecht, 2003), is a ten week, 45-minute, school-based, class intervention for secondary school aged students. This report systematically reviews five studies in order to consider the effectiveness of the "keepin' it REAL" intervention programme for secondary school aged, mainstream students. All studies were group-based designs, comparing pre- and post-intervention data with a control/comparison group. The effect sizes of the reduction and discontinuation of substance use were calculated and the study relevance in relation to this review was considered critically using Gough's (2007) 'Weight of Evidence Framework'. This review concludes that the "keepin' it REAL" intervention programme has limited support for reduction or discontinuation of substance use, with most effect sizes showing that the intervention had no effect and a small number showing that the intervention had a 'small' effect.

Introduction

Substance abuse is a major risk factor that can lead to additional social, emotional and mental health issues such as self-harm and suicide (HM Government, 2017a). The current government initiative, 'Five Year Forward View for Mental Health' (HM Government, 2017b), aims to address the increase in mental health issues in a number of ways, part of which aims to address substance abuse through additional funding of "£30 million for outcome-based interventions to tackle alcoholism and drug addiction through proven approaches". Sussman, Skara and Ames (2008) define adolescent substance abuse and the six key criteria differences from adult substance abuse in regards to impact on neurology, physicality and behaviour. This highlights that any substance use within adolescence can have an increased negative impact in comparison to adult substance use. This paper will therefore refer to any use of legal or illegal drug use in adolescence as 'substance use'. This heightened focus on supporting those with substance abuse problems comes from the never-ending stream of research on the impact of substance abuse on social, emotional and mental health. For example, in a review of mortality reports, Chang et al. (2010) state that substance abuse was found to be one of the highest mortality risks. Additionally, research into marijuana use in early adolescence has been shown to have detrimental effects on brain development and consequently one's social, emotional and mental health (Volkow, Baler, Compton & Weiss, 2014). It can therefore be seen, that in an attempt to reduce social, emotional and mental health issues in adolescents; the underlying aspects of physical and psychological wellbeing are key.

The “What works in enhancing social and emotional skills development during childhood and adolescence?” document by the Early Intervention Foundation (2015) outlines a number of approaches for use in school-based practice. One of these approaches recommended for use within this document is the “keepin’ it REAL” (kiR) substance use intervention programme.

“Keepin’ it REAL” is a substance use intervention programme developed by Michelle Miller-Day and Michael Hecht (kiR - keepin’ it REAL, n.d.; Gosin, Marsiglia & Hecht, 2003) throughout the 1990’s and 2000’s. Over more recent years (2005-2017), the programme has been used more extensively within the United States of America (USA), particularly with students of a Latino heritage (Kulis et al., 2005). The programme was officially designed for use with secondary aged students across multiple cultures, through utilising culturally sensitive materials to support students to use taught strategies to avoid substances, specifically alcohol, cigarettes and marijuana. These strategies aim to support students to Refuse, Explain, Avoid and Leave (REAL). The curriculum consists of ten 45-minute lessons taught over a period of ten weeks, including the use of five videos of real stories. It is an interactive intervention involving group activities and games and is implemented by a trained facilitator including trained class teachers. Table 1 below details the ten lessons.

Table 1

“Keepin’ it REAL” lessons

Lesson Number	Topic
1	Options and Choices
2	Risks
3	Communication and Conflict
4	Refuse
5	Explain
6	Avoid
7	Leave
8	Norms
9	Feelings
10	Support Network

“Keepin’ it REAL” has been adapted for cultural differences including a Mexican American version, an African American version and a European version. Additionally, other adaptations have been created through research by the “kiR” research team including a “kiR-Plus” which is adapted for use with pre-/early adolescents and “kiR-Acculturation Enhanced” (kiR-AE) which addresses the needs of students who have moved from Mexico to the USA (Elek, Wagstaff & Hecht, 2010).

The “kiR” intervention programme is grounded on social learning behavioural psychology (Miller, 2002), in that it promotes the social learning aspect of the programme through role modelling preferred behaviours and focusing on a group-

based intervention to enhance confidence in refusal of substances through peer support; making refusal appear more socially typical. Additionally the programme uses experiential/traditional learning theory (Miller, 2002) through group-based activities focused on practicing the strategies that have been taught to improve refusal, explanation and avoidance. The intervention aims to provide opportunity to practice refusal and explanation through role-play and reinforce refusal behaviours through social acceptance. It is through using this that the intervention supports students to learn social behaviours as part of a group to promote social acceptance of substance refusal and reduce social demand behaviours; i.e. behaviours that result from peer influence and pressure to avoid social exclusion.

Relevance to Educational Psychology Practice

The research has demonstrated that early adolescent substance use can be detrimental to brain development and it may be that educational provisions could play a vital role in role modelling anti-drug strategies through a taught curriculum (Sussman, Skara & Ames, 2008; Volkow, Baler, Compton & Weiss, 2014). Additionally, through providing interventions targeted at self-efficacy, education can support students to develop an internal attribution and enhance the student's ability to consider the choices they make as their own (Gulliford & Miller, 2015). Through exploring the effectiveness of substance use interventions, the educational psychologist is able to support schools to provide evidence-based, effective, universal programmes to secondary school students through training and supporting school staff to implement them. Social, emotional and mental health concerns are increasingly transferring into the school environment and with an increased pressure on Child and Adolescent Mental Health Services, this role is increasingly falling to

the school setting and Educational Psychology practice. The Early Intervention Foundation (2015) document demonstrates the increased range of practice that is occurring within the school environment that relates to Educational Psychology practice in regards to social, emotional and mental health. It is with this document in mind, and with the awareness of the “kiR” intervention programme as a supposedly effective school-based intervention, that this study aims to explore the effectiveness of the programme using a number of more recent research papers. Therefore, the review question is:

‘Is “keepin’ it REAL” an effective intervention programme for mainstream secondary aged students?’

Critical Review of the Evidence Base

Literature search

Comprehensive literature searches were carried out on 28th December 2016, 11th January 2017 and 27th January 2017 on five online electronic databases; ERIC (EBSCO), PsycInfo, British Education Index (BEI), Medline and SCOPUS. Medline and SCOPUS were chosen due to their large selection of publications; PsycInfo for it's focus on psychology orientated research; and ERIC (EBSCO) and the BEI for their specific focus on education research. Table 2 shows the search terms that were used to search within the full text of articles on these databases.

Table 2

Search Terms Used in Database Searches

Intervention	Outcome	Participants
"keepin' it real"	--	"school*"
OR		
"keepin it real"		
OR		
"KiR"		

Note. * stands for a truncation to the search term

Note. Outcome search terms were not entered as they were not required for the current search

Inclusion and exclusion criteria

From the database search, 188 studies were identified; of these 79 were excluded due to being duplicates. A title screening was conducted of the 109 remaining studies, and a further 80 excluded based on inclusion and exclusion criteria (see Table 3). These criteria were then used to exclude a further 16 studies at abstract

screening. Thirteen studies were therefore remaining for full text screening and a further four excluded based on the criteria in Table 3. The screening process is visually demonstrated in the flow diagram in Figure 1. Of the nine papers remaining, eight papers used duplicated data sets; in this instance the study most closely meeting the inclusion criteria was selected for inclusion in the review. This resulted in five papers being retained for critical analysis towards answering the review question; these are listed in Table 4 and a summary of each study is given in Appendix B.

Table 3

Inclusion and Exclusion Criteria

Criterion		Inclusion	Exclusion	Rationale
Aspect	Reference			
Publication Type	1a	Study is published in a peer-reviewed journal	Study is not published in a peer-reviewed journal	Studies in peer-reviewed journals have undergone a level of scrutiny by peers within the field, making them less likely to contain errors
	1b	Study contains primary empirical data	Study does not contain primary empirical data	Primary data is collected first hand by the authors and can therefore be systematically reviewed for effect
Participants and Setting	2a	Study participants are at secondary school level (aged between 11 and 18 years)	Study participants are not at secondary level, e.g. participants are at primary level	The “keepin’ it REAL” programme was originally designed for secondary school aged students
	2b	Study is based in a mainstream school setting	Study is not based in a mainstream schools setting,	This relates to the review question; considering the use of the programme in

			e.g. alternative school setting	mainstream classes
Intervention and Study Details	3a	Study addresses intervention outcomes of the “keepin’ it REAL” programme on substance use reduction and/or discontinuation	Study does not address intervention outcomes of the “keepin’ it REAL” programme, e.g. looks at delivery effects or carried out during creation of the intervention	The review question concerns the effectiveness of the programme on substance use therefore only studies that look at the effect and that have been carried out using the formalised model are considered
	3b	Study uses an experimental design, reporting pre- and post-quantitative data	Study does not use an experimental design, e.g. uses an exploratory or qualitative design	To allow the reviewer to consider impact of the intervention by generating effect sizes for analysis of the effectiveness
	3c	Intervention implemented is the “keepin’ it REAL” substance use programme including adapted versions “Plus” and “AE”	Intervention implemented is not the “keepin’ it REAL” substance use programme e.g. REAL groups	The programme of interest for this review is the “keepin’ it REAL” substance use programme

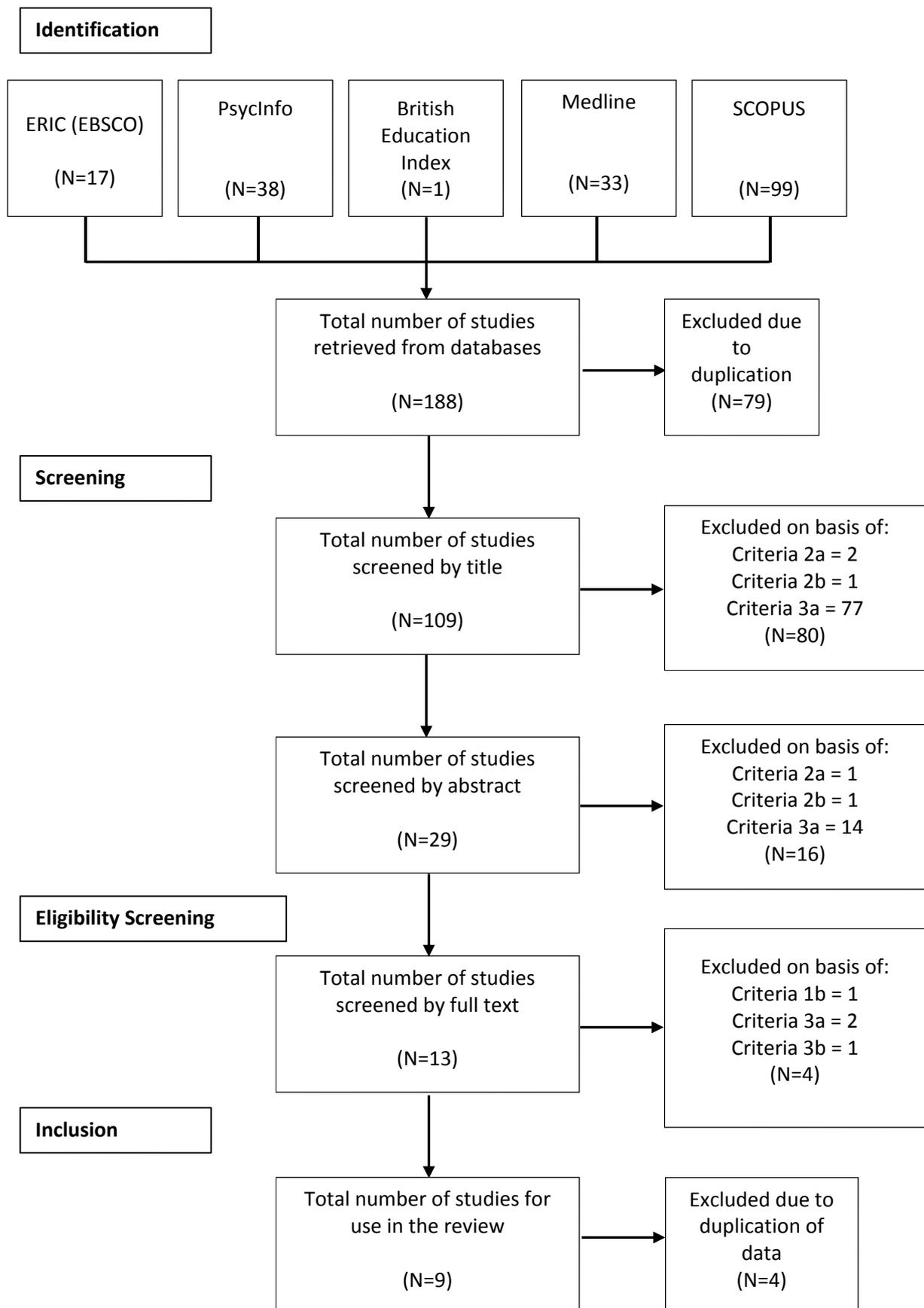


Figure 1: Flow Diagram of Literature Screening Process

Table 4

Studies Eligible for Use in the Final Review

Study Reference	Eligible Studies
1	Elek, E., Wagstaff, D. A., & Hecht, M. L. (2010). Effects of the 5 th and 7 th grade enhanced versions of the keepin' it REAL substance use prevention curriculum. <i>Journal of Drug Education</i> , 40(1), 61-79.
2	Hecht, M. L., Marsiglia, F. F., Elek, E., Wagstaff, D. A., Kulis, S., Dustman, P., & Miller-Day, M. (2003). Culturally Grounded Substance Use Prevention: An Evaluation of the keepin' it R.E.A.L. Curriculum. <i>Prevention Science</i> , 4(4), 233–248.
3	Kulis, S. S., Ayers, S. L., & Harthun, M. L. (2016). Substance Use Prevention for Urban American Indian Youth: An Efficacy Trial of the Culturally Adapted Living in 2 Worlds Program. <i>The Journal of Primary Prevention</i> , 1-22.
4	Kulis, S. S., Nieri, T., Yabiku, S., Stromwall, L. K., & Marsiglia, F. F. (2007). Promoting reduced and discontinued substance use among adolescent substance users: Effectiveness of a universal prevention program. <i>Prevention Science</i> , 8(1), 35–49.
5	Marsiglia, F. F., Kulis, S. S., Booth, J. M., Nuño-Gutierrez, B. L., & Robbins, D. E. (2015). Long-Term Effects of the keepin' it REAL Model Program in Mexico: Substance Use Trajectories of Guadalajara Middle School Students. <i>Journal of Primary Prevention</i> , 36(2), 93–104.

Critical Appraisal of Selected Studies

Weight of Evidence

Gough's (2007) Weight of Evidence (WoE) approach was used to critically consider the five remaining studies. This approach considers the methodological quality of the studies (WoE A), the methodological relevance of the studies to the review question (WoE B) and the topic relevance of the study to the review question (WoE C).

In order to consider WoE A, the methodological quality, for each study in regard to how well each study has been designed and conducted; Kratochwill's (2003) American Psychological Association Task Force coding protocol was used. This coding protocol is designed for group-based design studies such as those used in this review paper. Adaptations to this coding protocol were made in order to meet the needs of this review. These adaptations are detailed in Appendix C and the completed protocols are provided in Appendix D. Further information regarding WoE A can be found in Appendix E.

The relevance of the methodology (WoE B) and the topic (WoE C) in regards to the review question were then considered and rated accordingly (See Appendices F and G); before an average was taken for each study's ratings across all three areas to provide WoE D (See Appendix H for further detail), the overall WoE (see Table 5).

Table 5

Overview of Weight of Evidence

Study Reference	Author	WoE A: methodological quality	WoE B: methodological relevance	WoE C: topic relevance	WoE D: overall weight of evidence
1	Elek et al. (2010)	Low (1.00)	Low (1.00)	Medium (2.00)	Low (1.33)
2	Hecht et al. (2003)	Low (1.75)	Medium (2.00)	Medium (2.00)	Medium (1.92)
3	Kulis et al. (2016)	Low (1.25)	Medium (2.00)	Low (1.00)	Low (1.42)
4	Kulis et al. (2007)	Low (1.33)	Medium (2.00)	Medium (2.00)	Medium (1.78)
5	Marsiglia et al. (2015)	Medium (2.00)	Medium (2.00)	Low (1.00)	Medium (1.67)

Research Design

All included studies used a group-based, between participants approach to consider the effectiveness of the “kiR” intervention. Quantitative pre- and post- data was collected via self-report methods across all studies. None of the included studies utilised other sources, limiting the information collected to solely the young person’s self-report and only one questionnaire was used rather than a variety of methods. This is reflected in the WoE A, B and C ratings, in which all studies scored lower due to not collecting multi-method or multi-source data. This is a limitation on the reliability of this information, particularly with the topic of drug and substance use, as participants may have been cautious about their responses and demonstrated social desirability response bias (Randall & Fernandes, 1991).

All studies used block randomisation based on prior selection of schools, followed by random assignment to conditions. “Active” control/comparison groups were utilised through the schools’ already existing curriculum programmes or alternative intervention implementation. Hecht et al. (2003); Kulis, Nieri, Yabiku, Stromwall and Marsiglia (2007); Kulis, Ayers and Harthun (2016) and Marsiglia, Kulis, Booth, Nuño-Gutierrez and Robbins (2015) looked at group equivalency after collecting post-intervention data and carried out statistical processes to accommodate any differences and missing data. However, Elek, Wagstaff and Hecht (2010) did not do this, which may result in non-equivalent comparisons and inaccurate assumptions. The use of group equivalency statistical analysis is accounted for in WoE A.

Intervention

All studies included in this review used a version of the “kiR” substance use intervention programme that was equivalent to the original programme detailed by Miller-Day and Hecht (kiR - keepin’ it REAL, n.d.; Gosin, Marsiglia & Hecht, 2003). These included “kiR-Plus” and “kiR-AE” as adaptations for the specific target population. Additionally, there were some culturally adapted versions which utilised the original programme but adapted to the language and cultural norms of the sample, these included “kiR-Mexican American version”, “kiR-Black/White version”, and “kiR-Multicultural version”. Within Kulis et al. (2016) the “kiR” original programme was compared to the “Living in 2 Worlds” cultural variation of the “kiR” programme for Native Americans. The terminology ‘Native American’ has been used in this review rather than ‘American Indian’ as used by Kulis et al. (2016) due to the former being deemed more culturally appropriate in the current United Kingdom (UK) context. For the purpose of this review, the “kiR” data was treated as the intervention group data and considered against the “Living in 2 Worlds” programme as “kiR” is the focus of the review question.

All but one study (Kulis et al., 2007) reported using a trained class teacher to administer the intervention. However, it is worth noting that Kulis et al. (2016) used a specialist Native American teacher working at the school as the facilitator during ‘enhancement’ classes during the normal school day. The training provided was one day for all teachers and provided written materials as a curriculum programme. Elek et al. (2010) and Kulis et al. (2007) do not report any form of monitoring for fidelity. This is reflected in the WoE A ratings for these two studies with the other three studies scoring ‘medium’ for these aspects as, without fidelity checks, there is no

guarantee that the curriculum is being delivered as intended and to a high level. The other studies report good fidelity across intervention sessions.

Participants

The total number of participants across all of these studies was 9921, ranging considerably in size across the studies from 107 (Kulis et al., 2016) to 6035 (Hecht et al., 2003). All except one study (Kulis et al., 2016) had good statistical power in their studies, this is considered in WoE B ratings. The age of these participants primarily ranged from 11 to 18 years as the focus of the review was on secondary aged students. However, Elek et al. (2010) included younger students in their study and despite these students not being considered in this review, no breakdown of ages was provided to differentiate the number of secondary aged students from the primary aged students. No gender imbalance was found, with females making up 48.9% of the participants.

All students were attending a mainstream setting, as per the inclusion criteria, and were participating in a group, school-based version of the “kiR” programme. The sample was, to a large extent, biased in terms of demographic characteristics with a very small minority identifying as white heritage (9.5% across four of the studies). Kulis et al. (2016) only included Native American participants, reducing the white heritage to 1.9%, if considered as a whole. Excluding Kulis et al. (2016), the remaining four studies consisted of a majority Mexican/Mexican American heritage, at 77.25% of the participants. This impacts on the generalisability of this research to other ethnic majority provisions. Additionally, the majority of participants were entitled to free or reduced price school meals (85% across the four studies that

reported this information). These demographics were primarily a result of the location used for this research, with all but one study being conducted in the South Western area of the USA, which borders Mexico; and Kulis et al. (2016) conducted their research in a Mexican school. Further demographic information is detailed in Appendix B.

Measures

All studies utilised a different self-report questionnaire or survey based approach to measure use of substances. These were often designed for use in the study and varied vastly. Elek et al. (2010); Hecht et al. (2003) and Kulis et al. (2016) reported reliability measures such as Cronbach's alpha with the majority being at an 'acceptable' to 'good' reliability level (Barker, Pistrang & Elliott, 2016). However (Marsiglia et al., 2015) do not report the reliability of their measure. Due to the event history method of measurement chosen by Kulis et al. (2007) reliability measures were not conducted.

None of the studies were able to rate highly on the WoE A for 'measurement' due to a lack of use of more than one method of measurement. All studies created a measurement questionnaire that met the needs of the study without triangulating this data.

The studies varied in the data collected from simple rating of 'how much use over the previous 30 days' (Kulis et al., 2007), to an in depth exploration of use, frequency and beliefs about substance use (Elek et al., 2010; Hecht et al., 2003). As not all measures were related to the review question, only measures that related to

frequency and amount of use were considered. Further exploration could be conducted of the use of the “kiR” strategies and intrinsic aspects of substance use from other data within these studies.

Findings

The data used in the studies included, allowed effect sizes using Cohen's d , to be calculated using Chi square data, t-statistic data and mean difference data.

Elek et al. (2010) provided Cohen's d effect sizes; due to insufficient reporting of group means and standard deviations these were used within this report, however the accuracy of these is unknown. Hecht et al. (2003) reported the mean difference between the intervention and control group along with the standard error, these were used to calculate the t-statistic, which was then used to calculate Cohen's d . To calculate Cohen's d for the Kulis et al. (2016) paper the t-test value and sample sizes were used. For Kulis et al. (2007) the Chi square value and sample size were used to calculate Cohen's d and for Marsiglia et al. (2015) the unstandardised regression coefficient beta, b , was used along with the standard error and sample sizes to calculate Cohen's d . Table 6 shows a summary of the effect sizes. An effect size of less than 0.2 shows no or very little effect that is practically not significant; an effect size of 0.2 but less than 0.5 shows a 'small' effect size, more than 0.5 and less than 0.8 shows a 'medium' effect size and 0.8 or higher is a 'large' effect size (Dancey & Reidy, 2011).

Reduced use of alcohol, cigarettes and marijuana

Findings showed a mix of positive and negative effect sizes. Hecht et al. (2003) found no practically significant effects, at any of the time points, for the "kiR" intervention within each version provided for any of the substances; alcohol, cigarettes and marijuana. An overall effect was also not found when looking at all substances as a whole. Additionally, Kulis et al. (2007) found a small positive effect

size in reduction of alcohol use for those in the “kiR” intervention; however no effect for reduction in cigarette or marijuana use. For cigarette and marijuana amount, Marsiglia et al. (2015) found no effect, showing no reduction in use for those in the “kiR” intervention. However a small positive effect was found for alcohol amount, demonstrating a small increase in alcohol use in “kiR” participants. These studies were all weighted at the ‘medium’ level on WoE D and should be considered first as they are more reliable in their design and findings than those that rated as ‘low’. Kulis et al. (2016) found a small positive effect size for alcohol, cigarettes and marijuana use showing an increased use for those in the “kiR” intervention. Elek et al. (2010) looked at substance use prevalence over the last month, over a number of time points, within the two “kiR” programmes studied and found a small negative effect for both groups. This showed a small reduction for those in the “kiR” conditions; this did not consider substances individually. However, these studies weighted ‘low’ on WoE D and therefore should be interpreted with caution due to less reliable methodology and reduced relevance to the review question.

Discontinued use of alcohol, cigarettes and marijuana

Kulis et al. (2007) looked at the discontinuation of each substance individually and any substance as a whole, however effect sizes were not practically significant for any of the four aspects.

Table 6

Summary of Effect Sizes

Outcome	Study ref.	Authors	WoE D	Sample size	Time 1 effect size	Effect size descriptor	Time 2 effect size	Effect size descriptor	Post-data effect size	Post-data effect size descriptor
Past month's substance use prevalence (Plus version)	1	Elek et al. (2010)	Low	1984					-0.21	Small
Past month's substance use prevalence (AE version)	1	Elek et al. (2010)	Low	1984					-0.23	Small
Recent substance use Alcohol Cigarettes Marijuana	2	Hecht et al. (2003)	Medium	6035	-0.05 -0.10 -0.03 0.01	No effect No effect No effect No effect	-0.08 -0.07 -0.06 -0.05	No effect No effect No effect No effect	-0.11 -0.11 -0.04 -0.11	No effect No effect No effect No effect
Alcohol amount	3	Kulis et al. (2016)	Low	107					0.38	Small
Cigarette amount reduced	3	Kulis et al. (2016)	Low	107					0.41	Small
Marijuana amount reduced	3	Kulis et al. (2016)	Low	107					0.43	Small

Alcohol reduced use	4	Kulis et al. (2007)	Medium	1028	0.22	Small
Cigarette reduced use	4	Kulis et al. (2007)	Medium	544	0.05	No effect
Marijuana reduced use	4	Kulis et al. (2007)	Medium	614	0.04	No effect
Alcohol discontinued use	4	Kulis et al. (2007)	Medium	1028	0.18	No effect
Cigarette discontinued use	4	Kulis et al. (2007)	Medium	544	0.10	No effect
Marijuana discontinued use	4	Kulis et al. (2007)	Medium	614	0.02	No effect
Any discontinued use (all substances)	4	Kulis et al. (2007)	Medium	1364	0.13	No effect
Alcohol amount	5	Marsiglia et al. (2015)	Medium	431	0.44	Small
Cigarette amount	5	Marsiglia et al. (2015)	Medium	431	0.00	No effect
Marijuana amount	5	Marsiglia et al. (2015)	Medium	431	-0.02	No effect

Conclusion and Recommendations

The effect sizes for alcohol reduction were inconsistent across studies and therefore an effect of the “kiR” programme cannot be assumed from this data. The significant effect sizes found were small for both increased and reduced use and therefore it is likely that the “kiR” programme is unlikely to reduce alcohol use amongst secondary school students. Additionally, the more reliable studies as weighted on WoE D showed no effect for cigarette or marijuana reduction and it is unlikely that there is a benefit of “kiR” in reducing use of these substances.

When considering the effect of “kiR” on reduction in any of the substances as a whole, there was a small effect found for Elek et al. (2010) and no effect for Hecht et al. (2003). Elek et al. (2010) was weighted ‘low’ on WoE D whilst Hecht et al. (2003) was rated ‘medium’. It is unlikely that the “kiR” programme has any effect on overall substance use reduction, however any effect it does have is likely to be small.

No effect was found in the one study (Kulis et al., 2007) that considered discontinuation; and based on the small effect in reduction of substances across the studies reviewed, it is unlikely that discontinuation would occur as a result of the “kiR” intervention.

It can therefore be seen, that in contrast to publicised recommendations for school-based interventions that recommend “kiR” as an effective intervention (Early Intervention Foundation, 2015); this review lacks evidence of the programme as an effective programme for reducing substance use amongst secondary school aged pupils. However, these findings should be considered and generalised with care due

to the distinct lack of research into the use of the “kiR” programme in the UK and other English speaking countries. The studies included were heavily biased towards Mexican and Mexican American heritage (77.25% of participants) and the “kiR” programme is designed to be culturally sensitive (Hecht et al., 2003). However, this further raises the question of the effectiveness of “kiR” as the studies were conducted primarily on the target population for the intervention and yielded disappointing evidence for effectiveness.

Further research is required for considering the effectiveness of the “kiR” programme in the UK. Life Skills Education are providing a UK version of the “kiR” programme that has been adapted over the course of three years to be culturally appropriate for the UK education system and society (Life Skills Education, 2017). However, despite providing this intervention, no research currently exists for this intervention on the UK secondary student population.

Limitations in the design of these studies may have an impact on the reports of effectiveness and it may be that given a multi-method and multi-source design, data may have provided further information into substance use amongst youths. The self-report method of data collection lacks reliability (Randall & Fernandes, 1991) and may have benefitted from additional data from parents, teachers or peers where appropriate. It is also likely that participants’ estimates of their own substance use was inaccurate due to the time-delay in reporting. Many of the questionnaire data relied on students to estimate their use of substances over the past 30 days; this can lead to bias due to a lack of intrinsic awareness of one’s own use (Stone, Shiffman, Atienza & Nebeling, 2007). Further research may look at use of online or paper

diaries to record use of substances, making the data collection more reliable and providing more accurate data about the effectiveness of the “kiR” intervention. This would then inform future applicability of “kiR” within school-based settings.

A final consideration in critically considering the studies in this review, is that the majority of the studies within this research area have been conducted by the same research team, whom are all members of the “kiR” development team. This is important when considering conflict of interest and personal investment in positive outcomes. However, from this critical analysis of the effectiveness of “kiR” on reduction of substance use, this does not appear to have been an influential factor. This is a possible area of future research for educational psychologists working in secondary schools in the UK who are not affiliated with the “kiR” research team. This would be vital research for progressing forward with the use of “kiR” within secondary schools in the UK.

References

- Barker, C., Pistrang, N., & Elliott, R. (2016). *Research Methods in Clinical Psychology : An Introduction for Students and Practitioners (Third ed.)*. England, West Sussex: John Wiley & Sons Limited.
- Campbell Collaboration (2017). Effect Size Calculator. Retrieved from <https://www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-ESTypes.php>
- Chang, C. K., Hayes, R.D., Broadbent, M., Fernandes, A. C., Lee, W., Hotopf, M., & Stewart, R. (2010). All-Cause Mortality Among People with Serious Mental Illness (SMI), Substance Use Disorders, and Depressive Disorders in Southeast London: a Cohort Study. *BMC Psychiatry, 10*, 77.
- Dancey, C., & Reidy, J. (2011). *Statistics Without Maths for Psychology (Fourth ed.)*. England, Essex: Pearson Education Limited.
- Early Intervention Foundation (2015). What Works in Enhancing Social and Emotional Skills Development During Childhood and Adolescence? Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/411492/What_works_in_enhancing_social_and_emotional_skills_development_during_childhood_and_adolescence.pdf
- Elek, E., Wagstaff, D. A., & Hecht, M. L. (2010). Effects of the 5th and 7th Grade Enhanced Versions of the Keepin' it REAL Substance Use Prevention Curriculum. *Journal of Drug Education, 40*(1), 61–79.

- Gosin, M., Marsiglia, F. F., & Hecht, M. L. (2003). Keepin' it R.E.A.L.: A Drug Resistance Curriculum Tailored to the Strengths and Need of Pre-Adolescents of the Southwest. *Journal of Drug Education*, 33(2), 119-142.
- Gough, D. (2007). Weight of Evidence: A Framework for the Appraisal of the Quality and Relevance of Evidence. *Applied and Practice-Based Research. Special Edition of Research Papers in Education*, 22(2), 213-228.
- Gulliford, A., & Miller, A. (2015). Managing Classroom Behaviour: Perspectives from Psychology. In A. Gulliford, & S. Birch (Second ed.), *Educational Psychology: Topics in Applied Psychology* (pp.223-256). England, East Sussex: Routledge.
- Hecht, M. L., Colby, M., & Miller-Day, M. (2010). The Dissemination of Keepin' it REAL Through D.A.R.E. America: A Lesson in Disseminating Health Messages. *Health Communication*, 25(6-7), 585-586.
- Hecht, M. L., Marsiglia, F. F., Elek, E., Wagstaff, D. A., Kulis, S., Dustman, P., & Miller-Day, M. (2003). Culturally Grounded Substance Use Prevention: An Evaluation of the 'Keepin' it R.E.A.L.' Curriculum. *Prevention Science*, 4(4), 233-248.
- HM Government (2017a). Preventing Suicide in England: Third Progress Report of the Cross-Government Outcomes Strategy to Save Lives. Retrieved from <https://www.gov.uk/government/publications/suicide-prevention-third-annual-report>
- HM Government (2017b). The Government's Response to the Five Year Forward View for Mental Health. Retrieved from <https://www.gov.uk/government/publications/five-year-forward-view-for-mental-health-government-response>

kiR- Keepin' it REAL at Penn State University: An Effective, Multicultural Middle School Drug Prevention Program. (n.d.). Retrieved from <http://www.kir.psu.edu/index.shtml>

Kratochwill, T. R. (2003). Task Force On Evidence-Based Interventions in School Psychology, 608.

Kulis, S., Marsiglia, F. F., Elek, E., Dustman, P., Wagstaff, D. A., & Hecht, M. L. (2005). Mexican/Mexican American Adolescents and Keepin' it REAL: An Evidence-Based Substance Use Prevention Program. *Children & Schools*, 27(3), 133-145.

Kulis, S., Nieri, T., Yabiku, S., Stromwall, L. K., & Marsiglia, F. F. (2007). Promoting Reduced and Discontinued Substance Use Among Adolescent Substance Users: Effectiveness of a Universal Prevention Program. *Prevention Science*, 8(1), 35-49.

Kulis, S., Yabiku, S. T., Marsiglia, F. F., Nieri, T., & Crossman, A. (2007). Differences By Gender, Ethnicity, and Acculturation in the Efficacy of the "keepin' it REAL" Model Prevention Programme. *Journal of Drug Education*, 37, 123-144.

Kulis, S. S., Ayers, S. L., & Harthun, M. L. (2016). Substance Use Prevention for Urban American Indian Youth: An Efficacy Trial of the Culturally Adapted Living in 2 Worlds Program. *The Journal of Primary Prevention*, 1-22.

Life Skills Education (2017). Life Skills Education. Retrieved from <https://lifeskills-education.co.uk/keepin-it-real/>

Marsiglia, F. F., Ayers, S. L., Baldwin-White, A., & Booth, J. (2016). Changing Latino

Adolescents' Substance Use Norms and Behaviors: The Effects of Synchronized Youth and Parent Drug Use Prevention Interventions. *Prevention Science*, 17(1), 1-12.

Marsiglia, F. F., Booth, J. M., Ayers, S. L., Nuño-Gutierrez, B. L., Kulis, S., & Hoffman, S. (2014). Short-Term Effects on Substance Use of the Keepin' it REAL Pilot Prevention Program: Linguistically Adapted for Youth in Jalisco, Mexico. *Prevention Science*, 15(5), 694-704.

Marsiglia, F. F., Kulis, S. S., Booth, J. M., Nuño-Gutierrez, B. L., & Robbins, D. E. (2015). Long-Term Effects of the Keepin' it REAL Model Program in Mexico: Substance Use Trajectories of Guadalajara Middle School Students. *Journal of Primary Prevention*, 36(2), 93-104.

Marsiglia, F. F., Kulis, S., Rodriguez, G. M., Becerra, D., & Castillo, J. (2009). Culturally Specific Youth Substance Abuse Resistance Skills: Applicability Across the U. S. Mexico Border. *Research on Social Work Practice*, 19(2), 152-164.

Marsiglia, F. F., Kulis, S., Yabiko, S. T., Nieri, T. A., & Coleman, E. (2011). When to Intervene: Elementary School, Middle School, or Both? Effects of Keepin' it REAL on Substance Use Trajectories of Mexican Heritage Youth. *Prevention Science*, 12(1), 48-62.

Miller, P. H. (2002). *Theories of Developmental Psychology (Fourth ed.)*. New York, N. Y.: Worth Publishers.

Randall, D. M., & Fernandes, M. F. (1991). The Social Desirability Response Bias in Ethics Research. *Journal of Business Ethics*, 10(11), 805-817.

Stone, A., Shiffman, S., Atienza, A. A., & Nebeling, L. (2007). *The Science of Real-Time Data Capture: Self-Reports in Health Research*. New York: Oxford University Press.

Sussman, S., Skara, S., & Ames, S. L. (2008). Substance Abuse Among Adolescents. *Substance Use & Misuse*, 43, 1802-1828.

Volkow, N., Baler, R., Compton, W., & Weiss, S. (2014). Adverse Health Effects of Marijuana Use. *New England Journal of Medicine*, 370, 2219-2227.

Williams, L. R., Ayers, S., Baldwin, A., & Marsiglia, F. F. (2016). Delaying Youth Substance-Use Initiation: A Cluster Randomized Controlled Trial of Complementary Youth and Parenting Interventions. *Journal of the Society for Social Work & Research*, 7(1), 177-200.

Appendices

Appendix A

Studies Excluded During the Full Text Screen and Rationale

Excluded Studies	Rationale for Exclusion
Hecht, M. L., Colby, M., & Miller-Day, M. (2010). The Dissemination of Keepin' it REAL Through D.A.R.E. America: A Lesson in Disseminating Health Messages. <i>Health Communication, 25</i> (6-7), 585-586.	Exclusion criteria: 3a
Kulis, S., Marsiglia, F. F., Elek, E., Dustman, P., Wagstaff, D. A., & Hecht, M. L. (2005). Mexican/Mexican American Adolescents and Keepin' it REAL: An Evidence-Based Substance Use Prevention Program. <i>Children & Schools, 27</i> (3), 133-145.	Exclusion criteria: duplicated data
Kulis, S., Yabiku, S. T., Marsiglia, F. F., Nieri, T., & Crossman, A. (2007). Differences By Gender, Ethnicity, and Acculturation in the Efficacy of the "keepin' it REAL" Model Prevention Programme. <i>Journal of Drug Education, 37</i> , 123-144.	Exclusion criteria: 1b
Marsiglia, F. F., Ayers, S. L., Baldwin-White, A., & Booth, J. (2016). Changing Latino Adolescents' Substance Use Norms and Behaviors: The Effects of Synchronized Youth and Parent Drug Use Prevention Interventions. <i>Prevention Science, 17</i> (1), 1-12.	Exclusion criteria: duplicated data
Marsiglia, F. F., Booth, J. M., Ayers, S. L., Nuño-Gutierrez, B. L., Kulis, S., & Hoffman, S. (2014). Short-Term Effects on Substance Use of the Keepin' it REAL Pilot Prevention Program: Linguistically Adapted for Youth in Jalisco, Mexico. <i>Prevention Science, 15</i> (5), 694-704.	Exclusion criteria: duplicated data
Marsiglia, F. F., Kulis, S., Rodriguez, G. M., Becerra, D., & Castillo, J. (2009). Culturally Specific Youth Substance Abuse Resistance Skills: Applicability Across the U. S. Mexico Border. <i>Research on Social Work Practice,</i>	Exclusion criteria: 3b

19(2), 152-164.

Marsiglia, F. F., Kulis, S., Yabiko, S. T., Nieri, T. A., & Coleman, E. (2011). When to Intervene: Elementary School, Middle School, or Both? Effects of Keepin' it REAL on Substance Use Trajectories of Mexican Heritage Youth. *Prevention Science, 12*(1), 48-62. Exclusion criteria: duplicated data

Williams, L. R., Ayers, S., Baldwin, A., & Marsiglia, F. F. (2016). Delaying Youth Substance-Use Initiation: A Cluster Randomized Controlled Trial of Complementary Youth and Parenting Interventions. *Journal of the Society for Social Work & Research, 7*(1), 177-200. Exclusion criteria: 3a

Appendix B

Summary of the Studies Included

Study	Study Design	Study Location	Sample Characteristics					Intervention Details	Measures	Outcomes and Key Findings
			Size	Age Range	Gender	SES ¹	Ethnicity			
Elek, Wagstaff & Hecht (2010) Study ID: 1	Randomised block design (between participants) 7 conditions: 1) 5 th grade kiR-Plus 2) 5 th grade kiR-AE 3) 7 th grade kiR-Plus 4) 7 th grade kiR-AE 5) 5 th and 7 th grade kiR-Plus 6) 5 th and 7 th grade kiR-AE 7) Control group – existing school based intervention	Arizona, USA Urban area	1984 29 public elementary schools	Ave: 10.4 years	50% (f) 50% (m)	Reports details on access to free school meals: 75% free school lunch 17% reduced price school lunch	75% Mexican/Mexican American 9% African American 5 % Non-Hispanic White 3% Hispanic (other) 3% Native American	<u>Programme Duration:</u> 12x45minute lessons and 5 boosters in 6 th and 8 th grades <u>Intervention Group Size:</u> Not reported <u>Control Group Size:</u> Not reported <u>Structure:</u> 6 waves: W1 = 5 th grade (baseline); W6 = 5 th follow up in 8 th grade. Implemented by trained class teacher.	Self-report questionnaire: 45 minutes long; 104 (W1) – 177 (W6) items; Completed at each time point. Socio-demographic characteristics, lifetime substance use prevalence, past month's prevalence, intention to use substances, refusal efficacy, hypothetical alcohol resistance, number of substance use resistance strategies, personal anti-drug norms, descriptive substance use norms, positive	Lifetime prevalence: Control baseline to W6 did not increase significantly to that reported by intervention groups, except 5 th grade kiR-AE Last month's use: Control group was significantly less than 5 th grade kiR-Plus and 5 th grade kiR-AE Refusal: Control group significantly more confident in ability to refuse than 5 th grade kiR-Plus, 5 th grade kiR-AE and 7 th grade kiR-AE Strategies: Control report significantly greater increase than 5 th grade kiR-Plus but not greater than other intervention groups. Peer use rejection: 7 th

									substance use expectancies.	grade kiR-AE significantly decreased in rejection of peer use than control group; 5 th and 7 th grade kiR-Plus significantly increased their rejection of peer use than control group.
Hecht, Marsiglia, Elek, Wagstaff, Kulis, Dustman & Miller-Day (2003)	Randomised block design (between participants)	USA	6035	Ave: 12.53 years Range: 11-18 years (in Grade 7)	47% (f) 53% (m)	Reports details on access to free school meals: 74% free school lunch 8% reduced price school lunch	3318 Mexican/American 527 African American 1049 Non-Hispanic White 1141 Latino (other) or multi-ethnic Latino origin	<u>Programme</u> <u>Duration:</u> 10 lessons; bilingual public service announcements and billboard campaigns; booster activities between Waves 3 and 4 <u>Intervention</u> <u>Group Size:</u> 25 schools <u>Control Group</u> <u>Size:</u> 10 schools <u>Structure:</u> 4 waves: W1 = 7 th grade baseline; W2 = 2 month questionnaire; W3 = 8 month questionnaire (following campaigns); W4 = 14 month post questionnaire.	Self-report questionnaire: 45 minutes long completed at each time point. Behavioural variables: 30 day use and frequency of substance use and use of strategies Psychosocial variables: Self-efficacy; intent to accept; perception of positive consequence of substance use and Focus theory of norms	Self-report questionnaire for use of substances; use of strategies; and norms Findings: Better behavioural and social outcomes for intervention than control; increase in substance use over time for both groups but significantly less for intervention group; more strategies used for intervention group; no significant differences between intervention groups; little substantial support for the cultural matching hypothesis (i.e. ethnicity matched to intervention version).
Study ID: 2	4 conditions: 1) Mexican American version 2) Black/White version 3) Multicultural version 4) Control group – existing school based intervention	Urban area	35 public middle schools							

								Implemented by trained class teacher.		
Kulis, Ayers & Harthun (2016)	Randomised block design (between participants)	Phoenix, Arizona, USA	107	Ave: 12.4 years	49.5% (f) 50.5% (m)	50% living with both parents 40% living with one parent 10% living with no parent (usually with grandpa rents)	All Native American 79.4% mother Native American heritage 62.1% father Native American (NA) heritage	<u>Programme</u> <u>Duration:</u> 12 lessons <u>Intervention</u> <u>Group Size:</u> 85 <u>Comparison</u> <u>Group Size:</u> 22 <u>Structure:</u> pre- test, intervention, post-test. Implemented by trained teacher.	Self-report questionnaire: 50 minutes long, pre- and post- intervention	Repeated self-report questionnaire t-test within participant change baseline adjusted general linear models for group comparison Findings: General increase in substance use and stronger pro-drug attitudes and exposure. Increased use of the strategies across both groups, with the exception of a decrease in the use of the 'leave' strategy for the L2W group. L2W: increase in marijuana use; decrease in agreement of substance use being harmless; kiR: reported increase in substance use outcomes (alcohol frequency, cigarette frequency, cigarette amount, marijuana frequency and
Study ID: 3	2 conditions: 1) L2W version of kiR (adapted for Native American students) 2) kiR (comparison group)	Urban area	3 middle schools	80 % 12-13 year olds		All but 6% reported belongin g to one of 17 Native America n tribes Reports details on access to free school meals: 67% free school				

						lunch 17.5% reduced price school lunch				marijuana amount) Group differences: More positive change for cigarette frequency, NA spirituality and NA cultural traditions for L2W over kiR students expanded use of the leave strategy more than L2W group.
Kulis, Nieri, Yabiku, Stromwall & Marsiglia (2007) Study ID: 4	Randomised block design (between participants) 2 conditions: 1) kiR intervention 2) Control group – existing school based intervention	USA, large south- western city Urban area	1364 35 middle schools	Ave: 13 years Range: 11-16 years	43% (f) 57% (m)	82% free or reduced price school lunch	77% Latino heritage: 95% of which identified as either Mexican, Mexican American or Chicano 47% Spanish language dominant 13% identified as White or Anglo heritage	<u>Programme</u> <u>Duration:</u> 10 lessons <u>Intervention</u> <u>Group Size:</u> 1050 <u>Control Group</u> <u>Size:</u> 314 <u>Structure:</u> pre- test at wave 1, repeat questionnaires at wave 2 (2 months after pre-data), wave 3 (8 months after pre-data) and wave 4 (14 months post intervention).	Self-reported reduced or discontinued use in the previous 30 days (event history analysis). Collected data at pre- intervention, wave 2, wave 3 and wave 4.	Alcohol reduction and discontinuation rates were significantly higher in the intervention group than the control group. 40% of prior alcohol users reduced their use, and 32% discontinued use in the course of the study. Cigarette smoking reduction and discontinuation rates were non-significantly higher in the intervention group than the control group. 35% of cigarette smokers reduced their use and 29% discontinued use. 26% discontinued use of any and all substances, this was

										significantly higher in the intervention group.
Marsiglia, Kulis, Booth, Nuño-Gutierrez & Robbins (2014)	Randomised block design (between participants) 2 conditions: 1) kiR intervention adapted for Mexican students 2) Control group – existing school based intervention	Guadalajara, Jalisco, Mexico	431 2 secundarias (equivalent to USA middle school)	Ave: 13.01 years old Range: 12-15 years	55% (f) 45% (m)	Not detailed	Mexican heritage	<u>Programme Duration:</u> 10 lessons <u>Intervention Group Size:</u> 225 <u>Control Group Size:</u> 206 <u>Structure:</u> pre-test at wave 1, post-test immediately after intervention, follow up post-test 8 months post intervention. Implemented by trained class teacher.	Self-reported survey of substance use at three time points	Self-reported use across all groups increased from pre- to post- and follow up for alcohol and marijuana, and declined for cigarettes. Females in the intervention group showed a slower initiation of alcohol use compared to the control group. Males in the intervention group showed little change in marijuana use and frequency, whilst the male control group participants showed a significant increase.

¹ Socioeconomic status

Appendix C

Rationale for Eliminations Made to Coding Protocol

Eliminations	Rationale
Section I. General Characteristics For studies using qualitative research methods B7. Coding B8. Interactive process	Studies were looking at quantitative data as specified in the inclusion criteria
Section II. Key Features for Coding Studies and Rating Level of Evidence/Support C. Primary/Secondary Outcomes Are Statistically Significant	Review process will look at outcomes separately through considering effect sizes and WoE D
Section II. Key Features for Coding Studies and Rating Level of Evidence/Support D. Educational/Clinical Significance	This also considers outcomes which will be looked at separately
Section II. Key Features for Coding Studies and Rating Level of Evidence/Support E. Identifiable Components	This looks at the primary/secondary outcomes; which have not been looked at (Section II. C.)
Section II. Key Features for Coding Studies and Rating Level of Evidence/Support G. Replication	Not relevant for this review
Section II. Key Features for Coding Studies and Rating Level of Evidence/Support H. Site of Implementation	No variation in site across the studies as the inclusion criteria specified school as the site
Section III. Other Descriptive or Supplemental Criteria to Consider	This information is summarised in

A2. Participant Characteristics Specified for Treatment and Control Group 'Appendix B: Summary of the Studies Included'

Section III. Other Descriptive or Supplemental Criteria to Consider
A4. Receptivity/acceptance by target participant population (treatment group) Not relevant for the review question

Section III. Other Descriptive or Supplemental Criteria to Consider
D. Dosage Response Not relevant for the review question

-

Appendix D

Coding Protocol

[Adapted from Task Force on Evidence-Based Interventions in School Psychology, American Psychology Association, Kratochwill, T.R. (2003)]

Coding Protocol

Name of Coder:

Date: 07.02.2017

Full Study Reference in proper format: Elek, E., Wagstaff, D. A., Hecht, M. L. (2010). Effects of the 5th and 7th grade enhanced versions of the keepin' it REAL substance use prevention curriculum. *Journal of Drug Education*, 40(1), 61-79.

Intervention Name (description of study): "keepin' it REAL-Plus and keepin' it REAL-Acculturation Enhanced

Study ID Number: 1

- Type of Publication:
 Book/Monograph
 Journal Article
 Book Chapter
 Other (specify):

1. General Characteristics

A. General Design Characteristics

A1. Random assignment designs (if random assignment design, select one of the following)

- A1.1 Completely randomized design
 A1.2 Randomized block design (between participants, e.g., matched classrooms)
 A1.3 Randomized block design (within participants)
 A1.4 Randomized hierarchical design (nested treatments)

A2. Nonrandomized designs (if non-random assignment design, select one of the following)

- A2.1 Nonrandomized design
 A2.2 Nonrandomized block design (between participants)
 A2.3 Nonrandomized block design (within participants)
 A2.4 Nonrandomized hierarchical design
 A2.5 Optional coding for Quasi-experimental designs

A3. Overall confidence of judgment on how participants were assigned (select one of the following)

- A3.1 Very low (little basis)
 A3.2 Low (guess)
 A3.3 Moderate (weak inference)
 A3.4 High (strong inference)
 A3.5 Very high (explicitly stated)

- A3.6 N/A
 A3.7 Unknown/unable to code

B. Statistical Treatment

- B1. Appropriate unit of analysis Yes No
 B2. Familywise error rate controlled Yes No Unknown/NA
 B3. Sufficiently large N Yes No
 Statistical Test: Difference between two mean (independent)
 _ level: 0.05
 ES: 0.20
 N required: 788

B4. Total size of sample (start of the study): 1984
 N

B5. Intervention group sample size: Unknown
 N

B6. Control group sample size: Unknown
 N

~~**For studies using qualitative research methods, code B7 and B8**~~

~~B7. Coding~~

- ~~B7.1 Coding scheme linked to study's theoretical-empirical basis (select one) Yes No~~
~~B7.2 Procedures for ensuring consistency of coding are used (select one) Yes No~~
 ~~No~~

~~Describe procedures: _____~~

- ~~B7.3 Progression from abstract concepts to empirical exemplars is clearly articulated (select one) Yes No~~

~~B8. Interactive process followed (select one) Yes No~~

~~Describe process: _____~~

C. Type of Program

- C1. Universal prevention program
 C2. Selective prevention program
 C3. Targeted prevention program
 C4. Intervention/Treatment
 C5. Unknown

D. Stage of Program

- D1. Model/demonstration programs
 D2. Early stage programs
 D3. Established/institutionalized programs
 D4. Unknown

E. Concurrent or Historical Intervention Exposure

- E1. Current exposure
 E2. Prior exposure

E3. Unknown

2. Key Features for Coding Studies and Rating Level of Evidence/Support

(Rating Scale: 3= Strong Evidence, 2=Promising Evidence, 1=Weak Evidence, 0=No Evidence)

A. Measurement (Estimating the quality of the measures used to establish effects)

A1. Use of the outcome measures produce reliable scores for the majority of the primary outcomes. The table for Primary/Secondary Outcomes Statistically Significant allows for listing separate outcomes and will facilitate decision making regarding measurement (select one of the following)

A1.1 Yes

A1.2 No

A1.3 Unknown/unable to code

A2. Multi-method (at least two assessment methods used)

A2.1 Yes

A2.2 No

A2.3 N/A

A2.4 Unknown/unable to code

A3. Multi-source (at least two sources used self-reports, teachers etc.)

A3.1 Yes

A3.2 No

A3.3 N/A

A3.4 Unknown/unable to code

A4. Validity of measures reported (well-known or standardized or norm-referenced are considered good, consider any cultural considerations)

A4.1 Yes validated with specific target group

A4.2 In part, validated for general population only

A4.3 No

A4.4 Unknown/unable to code

Rating for Measurement (select 0, 1, 2 or 3):

3

2

1

0

B. Comparison Group

B1. Type of Comparison Group (Select one of the following)

B1.1 Typical intervention (typical intervention for that setting, without additions that make up the intervention being evaluated)

B1.2 Attention placebo

B1.3 Intervention element placebo

B1.4 Alternative intervention

B1.5 Pharmacotherapy

B1.6 No intervention

B1.7 Wait list/delayed intervention

B1.8 Minimal contact

B1.9 Unable to identify type of comparison

B2. Overall confidence of judgment on type of comparison group

- B2.1 Very low (little basis)
- B2.2 Low (guess)
- B2.3 Moderate (weak inference)
- B2.4 High (strong inference)
- B2.5 Very high (explicitly stated)
- B2.6 Unable to identify comparison group

B3. Counterbalancing of change agent (participants who receive intervention from a single therapist/teacher, etc. were counter-balanced across intervention)

- B3.1 By change agent
- B3.2 Statistical (analyse includes a test for intervention)
- B3.3 Other: class teacher implementation
- B3.4 Not reported/None

B4. Group equivalence established (select one of the following)

- B4.1 Random assignment
- B4.2 Posthoc matched set
- B4.3 Statistical matching
- B4.4 Post hoc test for group equivalence

B5. Equivalent mortality

- B5.1 Low attrition (less than 20 % for post)
- B5.2 Low attrition (less than 30% for follow-up)
- B5.3 Intent to intervene analysis carried out? None
Findings_____

Rating for Comparison Group (select 0, 1, 2 or 3): 3 2 1 0

~~C. Primary/Secondary Outcomes Are Statistically Significant~~

~~C1. Evidence of appropriate statistical analysis for primary outcomes~~

- ~~C1.1 Appropriate unit of analysis (rate from previous code)~~
- ~~C1.2 Familywise/experimentwise error rate controlled when applicable (rate from previous code)~~
- ~~C1.3 Sufficiently large N (rate from previous code)~~

~~C2. Percentage of primary outcomes that are significant~~

- ~~_____ C2.1 Significant primary outcomes for at least 75% of the total primary outcome measures for each key construct~~
- ~~_____ C2.2 Significant primary outcomes for between 50% and 74% of the total primary outcome measures for each key construct~~
- ~~_____ C2.3 Significant primary outcomes for between 25% and 49% of the total primary outcome measures for any key construct~~

~~Rating for Primary Outcomes Statistically Significant (select 0, 1, 2 or 3): 3 2 1 0~~

~~C3. Evidence of appropriate statistical analysis for primary outcomes~~

- ~~C3.1 Appropriate unit of analysis (rate from previous code)~~
- ~~C3.2 Familywise/experimentwise error rate controlled when applicable (rate from previous code)~~
- ~~C3.3 Sufficiently large N (rate from previous code)~~

~~C4. Percentage of secondary outcomes that are~~

C4.1 Significant secondary outcomes for at least 75% of the total secondary outcome measures for each key construct

C4.2 Significant secondary outcomes for between 50% and 74% of the total secondary outcome measures for each key construct

C4.3 Significant secondary outcomes for between 25% and 49% of the total secondary outcome measures for each key construct

Rating for Secondary Outcomes Statistically Significant (select 0, 1, 2 or 3): 3 2
1 0

C5. Overall Summary of Questions Investigated

C5.1 Main effect analyses conducted (select one) Yes No

C5.2 Moderator effect analyses conducted (select one) Yes No

Specify results: _____

C5.3 Mediator analyses conducted (select one) Yes No

Specify results: _____

C. Primary/Secondary Outcomes Statistically Significant (only list $p \leq .05$)

(list primary outcomes first in alphabetical order, followed by secondary outcomes in alphabetical order)

Outcomes	Primary vs. Secondary	Who Changed	What Changed	Source	Treatment Information	Outcome Measure Used	Reliability	ES	(1- →)
Outcome #1:	<input type="checkbox"/> Primary <input type="checkbox"/> Secondary <input type="checkbox"/> Unknown	<input type="checkbox"/> Child <input type="checkbox"/> Teacher <input type="checkbox"/> Parent/sign. adult <input type="checkbox"/> Ecology <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Behavior <input type="checkbox"/> Attitude <input type="checkbox"/> Knowledge <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Self Report <input type="checkbox"/> Parent Report <input type="checkbox"/> Teacher Report <input type="checkbox"/> Observation <input type="checkbox"/> Test <input type="checkbox"/> Other <input type="checkbox"/> Unknown					
Outcome #2	<input type="checkbox"/> Primary <input type="checkbox"/> Secondary <input type="checkbox"/> Unknown	<input type="checkbox"/> Child <input type="checkbox"/> Teacher <input type="checkbox"/> Parent/sign. adult <input type="checkbox"/> Ecology <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Behavior <input type="checkbox"/> Attitude <input type="checkbox"/> Knowledge <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Self Report <input type="checkbox"/> Parent Report <input type="checkbox"/> Teacher Report <input type="checkbox"/> Observation <input type="checkbox"/> Test <input type="checkbox"/> Other <input type="checkbox"/> Unknown					
Outcome #3:	<input type="checkbox"/> Primary <input type="checkbox"/> Secondary <input type="checkbox"/> Unknown	<input type="checkbox"/> Child <input type="checkbox"/> Teacher <input type="checkbox"/> Parent/sign. adult <input type="checkbox"/> Ecology <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Behavior <input type="checkbox"/> Attitude <input type="checkbox"/> Knowledge <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Self Report <input type="checkbox"/> Parent Report <input type="checkbox"/> Teacher Report <input type="checkbox"/> Observation <input type="checkbox"/> Test <input type="checkbox"/> Other <input type="checkbox"/> Unknown					
Outcome #4:	<input type="checkbox"/> Primary <input type="checkbox"/> Secondary <input type="checkbox"/> Unknown	<input type="checkbox"/> Child <input type="checkbox"/> Teacher <input type="checkbox"/> Parent/sign. adult <input type="checkbox"/> Ecology <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Behavior <input type="checkbox"/> Attitude <input type="checkbox"/> Knowledge <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Self Report <input type="checkbox"/> Parent Report <input type="checkbox"/> Teacher Report <input type="checkbox"/> Observation <input type="checkbox"/> Test <input type="checkbox"/> Other <input type="checkbox"/> Unknown					
				<input type="checkbox"/> Self Report					

Outcome #5:	<input type="checkbox"/> Primary <input type="checkbox"/> Secondary <input type="checkbox"/> Unknown	<input type="checkbox"/> Child <input type="checkbox"/> Teacher <input type="checkbox"/> Parent/sign. adult <input type="checkbox"/> Ecology <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Behavior <input type="checkbox"/> Attitude <input type="checkbox"/> Knowledge <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Parent Report <input type="checkbox"/> Teacher Report <input type="checkbox"/> Observation <input type="checkbox"/> Test <input type="checkbox"/> Other <input type="checkbox"/> Unknown					
-------------	--	---	--	--	--	--	--	--	--

Null Findings/Negative Outcomes Associated with the Intervention (listed alphabetically by outcome)

Outcomes	Primary vs. Secondary	Who Was Targeted for Change	What Was Targeted for Change	Source	Note null/negative outcomes	Outcome Measure Used	Reliability	ES
Outcome #1:	<input type="checkbox"/> Primary <input type="checkbox"/> Secondary <input type="checkbox"/> Unknown	<input type="checkbox"/> Child <input type="checkbox"/> Teacher <input type="checkbox"/> Parent/sign. adult <input type="checkbox"/> Ecology <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Behavior <input type="checkbox"/> Attitude <input type="checkbox"/> Knowledge <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Self Report <input type="checkbox"/> Parent Report <input type="checkbox"/> Teacher Report <input type="checkbox"/> Observation <input type="checkbox"/> Test <input type="checkbox"/> Other <input type="checkbox"/> Unknown				
Outcome #2	<input type="checkbox"/> Primary <input type="checkbox"/> Secondary <input type="checkbox"/> Unknown	<input type="checkbox"/> Child <input type="checkbox"/> Teacher <input type="checkbox"/> Parent/sign. adult <input type="checkbox"/> Ecology <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Behavior <input type="checkbox"/> Attitude <input type="checkbox"/> Knowledge <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Self Report <input type="checkbox"/> Parent Report <input type="checkbox"/> Teacher Report <input type="checkbox"/> Observation <input type="checkbox"/> Test <input type="checkbox"/> Other <input type="checkbox"/> Unknown				
Outcome #3:	<input type="checkbox"/> Primary <input type="checkbox"/> Secondary <input type="checkbox"/> Unknown	<input type="checkbox"/> Child <input type="checkbox"/> Teacher <input type="checkbox"/> Parent/sign. adult <input type="checkbox"/> Ecology <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Behavior <input type="checkbox"/> Attitude <input type="checkbox"/> Knowledge <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Self Report <input type="checkbox"/> Parent Report <input type="checkbox"/> Teacher Report <input type="checkbox"/> Observation <input type="checkbox"/> Test <input type="checkbox"/> Other <input type="checkbox"/> Unknown				
Outcome #4:	<input type="checkbox"/> Primary <input type="checkbox"/> Secondary <input type="checkbox"/> Unknown	<input type="checkbox"/> Child <input type="checkbox"/> Teacher <input type="checkbox"/> Parent/sign. adult <input type="checkbox"/> Ecology <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Behavior <input type="checkbox"/> Attitude <input type="checkbox"/> Knowledge <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Self Report <input type="checkbox"/> Parent Report <input type="checkbox"/> Teacher Report <input type="checkbox"/> Observation <input type="checkbox"/> Test <input type="checkbox"/> Other <input type="checkbox"/> Unknown				
Outcome #5:	<input type="checkbox"/> Primary <input type="checkbox"/> Secondary	<input type="checkbox"/> Child <input type="checkbox"/> Teacher	<input type="checkbox"/> Behavior <input type="checkbox"/> Attitude	<input type="checkbox"/> Self Report <input type="checkbox"/> Parent Report <input type="checkbox"/> Teacher				

<input type="checkbox"/> Unknown	<input type="checkbox"/> Parent/sign. adult <input type="checkbox"/> Ecology <input type="checkbox"/> Other <input type="checkbox"/> Unknown	<input type="checkbox"/> Knowledge <input type="checkbox"/> Other <input type="checkbox"/> Unknown	Report <input type="checkbox"/> Observation <input type="checkbox"/> Test <input type="checkbox"/> Other <input type="checkbox"/> Unknown					
----------------------------------	---	--	---	--	--	--	--	--

Type of Data Effect Size is Based On	Confidence Rating in ES Computation
(check all that apply) <input type="checkbox"/> Means and SDs <input type="checkbox"/> t - value or F - value <input type="checkbox"/> Chi-square (df = 1) <input type="checkbox"/> Frequencies or proportions (dichotomous) <input type="checkbox"/> Frequencies or proportions (polytomous) <input type="checkbox"/> Other (specify): <input type="checkbox"/> Unknown	(select one of the following) <input type="checkbox"/> Highly estimated (e.g., only have N p value) <input type="checkbox"/> Moderate estimation (e.g., have complex but complete statistics) <input type="checkbox"/> Some estimation (e.g., unconventional statistics that require conversion) <input type="checkbox"/> Slight estimation (e.g., use significance testing statistics rather than descriptives) <input type="checkbox"/> No estimation (e.g., all descriptive data is present)

D. Educational/Clinical Significance

Outcome Variables:	Pretest	Posttest	Follow Up
D1. Categorical Diagnosis Data	Diagnostic information regarding inclusion into the study presented: Yes No Unknown	Positive change in diagnostic criteria from pre to posttest: Yes No Unknown	Positive change in diagnostic criteria from posttest to follow up: Yes No Unknown
D2. Outcome Assessed via continuous Variables		Positive change in percentage of participants showing clinical improvement from pre to posttest: Yes No Unknown	Positive change in percentage of participants showing clinical improvement from posttest to follow up: Yes No Unknown
D3. Subjective Evaluation: The importance of behavior change is evaluated by individuals in direct contact with the participant.	Importance of behavior change is evaluated: Yes No Unknown	Importance of behavior change from pre to posttest is evaluated positively by individuals in direct contact with the participant: Yes No Unknown	Importance of behavior change from posttest to follow up is evaluated positively by individuals in direct contact with the participant: Yes No Unknown
D4. Social Comparison: Behavior of participant at pre, post, and follow up is	Participant's behavior is compared to normative data	Participant's behavior has improved from pre to posttest when compared to normative data:	Participant's behavior has improved from posttest to follow up when compared to normative data:

compared to normative data (e.g., a typical peer).	Yes	No	Unknown	Yes	No	Unknown	Yes	No	Unknown
	<input type="checkbox"/>								

Rating for Educational/Clinical Significance (select 0, 1, 2 or 3): 3 2 1 0

E. Identifiable Components

E1. Evidence for primary outcomes (rate from previous code): one 3 2 1 0

E2. Design allows for analysis of identifiable components (select one) Yes No

E3. Total number of components: _____
N

E4. Number of components linked to primary outcomes: _____
N

Additional criteria to code descriptively:

E5. Clear documentation of essential components (select one) Yes No

E6. Procedures for adapting the intervention are described in detail (select one) Yes No

E7. Contextual features of the intervention are documented (select one) Yes No

Rating for Identifiable Components (select 0, 1, 2 or 3): 3 2 1 0

F. Implementation Fidelity

F1. Evidence of Acceptable Adherence

F1.1 Ongoing supervision/consultation

F1.2 Coding intervention sessions/lessons or procedures

F1.3 Audio/video tape implementation

F1.3.1 Entire intervention

F1.3.2 Part of intervention

None detailed

F2. Manualization

F2.1 Written material involving a detailed account of the exact procedures and the sequence in which they are to be used

F2.2 Formal training session that includes a detailed account of the exact procedures and the sequence in which they are to be used

F2.3 Written material involving an overview of broad principles and a description of the intervention phases

F2.4 Formal or informal training session involving an overview of broad principles and a description of the intervention phases

Rating for Implementation Fidelity (select 0, 1, 2 or 3): 3 2 1 0

G. Replication

G1. Same Intervention

G2. Same Target Problem

G3. Independent Evaluation

Rating for Replication (select 0, 1, 2 or 3): 3 2 1 0

H. Site of Implementation

H1. School (if school is the site, select one of the following options)

H1.1 Public

H1.2 Private

H1.3 Charter

H1.4 University Affiliated

H1.5 Alternative

H1.6 Not specified/unknown

H2. Non School (if it is a non school site, select one of the following options)

- H2.1 Home
- H2.2 University Clinic
- H2.3 Summer Program
- H2.4 Outpatient Hospital
- H2.5 Partial inpatient/Intervention Program
- H2.6 Inpatient Hospital
- H2.7 Private Practice
- H2.8 Mental Health Centre
- H2.9 Residential Treatment Facility
- H2.10 Other (specify): _____
- H2.11 Unknown/insufficient information provided

Rating for Site of Implementation (select 0, 1, 2 or 3): 3 2 1 0

I. Follow Up Assessment

- Timing of follow up assessment: specify : 8th grade (wave 6) follow up/post questionnaires
- Number of participants included in the follow up assessment: specify _____
- Consistency of assessment method used: specify _____

Rating for Follow Up Assessment (select 0, 1, 2 or 3): 3 2 1 0

3. Other Descriptive or Supplemental Criteria to Consider

A. External Validity Indicators

- A1. Sampling procedures described in detail Yes No
 Specify rationale for selection: _____
 Specify rationale for sample size: _Opportunistic sample size due to school attendees and participation_
- A1.1 Inclusion/exclusion criteria specified Yes No
- A1.2 Inclusion/exclusion criteria similar to school practice Yes No
- A1.3 Specified criteria related to concern Yes No
- A2. Participant Characteristics Specified for Treatment and Control Group

Participants from Treatment Group	Grade /age	Gender	Ethnicity or Multi-ethnic	Ethnic Identity	Race(s)	Acculturation	Primary Language	SES	Family Structure	Locale	Disability	Functional Descriptors
<input type="checkbox"/> Child/Student <input type="checkbox"/> Parent/caregiver <input type="checkbox"/> Teacher <input type="checkbox"/> School <input type="checkbox"/> Other												
<input type="checkbox"/> Child/Student <input type="checkbox"/> Parent/caregiver <input type="checkbox"/> Teacher <input type="checkbox"/> School <input type="checkbox"/> Other												
<input type="checkbox"/> Child/Student <input type="checkbox"/> Parent/caregiver <input type="checkbox"/> Teacher <input type="checkbox"/> School <input type="checkbox"/> Other												
<input type="checkbox"/> Child/Student <input type="checkbox"/> Parent/caregiver <input type="checkbox"/> Teacher <input type="checkbox"/> School <input type="checkbox"/> Other												
<input type="checkbox"/> Child/Student <input type="checkbox"/> Parent/caregiver <input type="checkbox"/> Teacher <input type="checkbox"/> School <input type="checkbox"/> Other												
Participants from Control Group	Grade /age	Gender	Ethnicity or Multi-ethnic	Ethnic Identity	Race(s)	Acculturation	Primary Language	SES	Family Structure	Locale	Disability	Functional Descriptors

<input type="checkbox"/> Child/Student <input type="checkbox"/> Parent/caregiver <input type="checkbox"/> Teacher <input type="checkbox"/> School <input type="checkbox"/> Other												
<input type="checkbox"/> Child/Student <input type="checkbox"/> Parent/caregiver <input type="checkbox"/> Teacher <input type="checkbox"/> School <input type="checkbox"/> Other												
<input type="checkbox"/> Child/Student <input type="checkbox"/> Parent/caregiver <input type="checkbox"/> Teacher <input type="checkbox"/> School <input type="checkbox"/> Other												
<input type="checkbox"/> Child/Student <input type="checkbox"/> Parent/caregiver <input type="checkbox"/> Teacher <input type="checkbox"/> School <input type="checkbox"/> Other												

A3. Details are provided regarding demographic variables that:

A3.1 Have differential relevance for intended outcomes Yes No

Specify: _____

A3.2 Have relevance to inclusion criteria Yes No

Specify: _____

A4. Receptivity/acceptance by target participant population (treatment group)

Participants from Treatment Group	Results (What person reported to have gained from participation in program)	General Rating
<input type="checkbox"/> Child/Student <input type="checkbox"/> Parent/caregiver <input type="checkbox"/> Teacher <input type="checkbox"/> School <input type="checkbox"/> Other		<input type="checkbox"/> Participants reported benefiting overall from the intervention <input type="checkbox"/> Participants reported not benefiting overall from the intervention
<input type="checkbox"/> Child/Student <input type="checkbox"/> Parent/caregiver <input type="checkbox"/> Teacher <input type="checkbox"/> School <input type="checkbox"/> Other		<input type="checkbox"/> Participants reported benefiting overall from the intervention <input type="checkbox"/> Participants reported not benefiting overall from the intervention
<input type="checkbox"/> Child/Student <input type="checkbox"/> Parent/caregiver <input type="checkbox"/> Teacher <input type="checkbox"/> School <input type="checkbox"/> Other		<input type="checkbox"/> Participants reported benefiting overall from the intervention <input type="checkbox"/> Participants reported not benefiting overall from the intervention

A5. Generalization of Effects:

A5.1 Generalization over time

A5.1.1 Evidence is provided regarding the sustainability of outcomes after intervention is terminated Yes No

Specify: _____

A5.1.2 Procedures for maintaining outcomes are specified Yes No

No

Specify: _____

A5.2 Generalization across settings

A5.1.1 Evidence is provided regarding the extent to which outcomes are manifested in contexts that are different from the intervention context Yes No

Specify: _____

A5.2.2 Documentation of efforts to ensure application of intervention to other settings

Yes No

Specify: ___School-based curriculum-based intervention conducted as intended in the program___

A5.2.3 Impact on implementers or context is sustained Yes No

Specify: ___teacher is trained to implement the intervention___

A5.3 Generalization across persons

A5.1.1 Evidence is provided regarding the degree to which outcomes are manifested with participants who are different than the original group of participants for which the intervention was evaluated

Yes No

Specify: _____

B. Length of Intervention

- B1. Unknown/insufficient information provided
- B2. Information provided (if information is provided, specify one of the following)
 - B2.1 weeks 10 N
 - B2.2 months _____ N
 - B2.3 years _____ N
 - B2.4 other _____ N

C. Intensity/dosage of intervention

- C1. Unknown/insufficient information provided
- C2. Information provided (if information is provided, specify both of the following)
 - C2.1 length of intervention session 45 mins N
 - C2.2 frequency of intervention session 1/week N

~~D. Dosage Response~~

- ~~D1. Unknown/insufficient information provided~~
- ~~D2. Information provided (if information is provided, answer D2.1)~~
 - ~~B2.1 Describe positive outcomes associated with higher dosage:~~

E. Program Implementer

- E1. Research Staff
- E2. School Speciality Staff
- E3. Teachers
- E4. Educational Assistants
- E5. Parents
- E6. College Students
- E7. Peers
- E8. Other
- E9. Unknown/insufficient information provided

F. Characteristics of the Intervener

- F1. Highly similar to target participants on key variables (e.g., race, gender, SES)
- F2. Somewhat similar to target participants on key variables
- F3. Different from target participants on key variables

G. Intervention Style or Orientation

- G1. Behavioural
- G2. Cognitive-behavioural
- G3. Experiential
- G4. Humanistic/interpersonal
- G5. Psychodynamic/insight oriented
- G6. Other (specify): _____
- G7. Unknown/insufficient information provided

H. Cost Analysis Data

- H1. Unknown/insufficient information provided
- H2. Information provided (if information is provided, answer H2.1)
 - H2.1 Estimated Cost of Implementation: _____

I. Training and Support Resources

- I1. Simple orientation given to change agents

- I2. Training workshops conducted
 # of Workshops provided ____ Not detailed ____
 Average length of training ____ Not detailed ____
 Who conducted training (select all that apply)
 - I2.1 Project Director
 - I2.2 Graduate/project assistant
 - I2.3 Other (please specify):
 - I2.4 Unknown
- I3. Ongoing technical support
- I4. Program materials obtained
- I5. Special Facilities
- I6. Other

J. Feasibility

- J1. Level of difficulty in training intervention agents (select one of the following)
 - J1.1 High
 - J1.2 Moderate
 - J1.3 Low
 - J1.4 Unknown
- J2. Cost to train intervention agents (specify if known): _____
- J3. Rating of cost to train intervention agents (select one of the following)
 - J3.1 High
 - J3.2 Moderate
 - J3.3 Low
 - J3.4 Unknown

Summary of Evidence

Indicator	Overall evidence rating 0-3	Description of evidence Strong Promising Weak No/limited evidence Or Descriptive ratings
General Characteristics		
General Design Characteristics		Promising
Statistical Treatment		Promising
Type of Program		Promising
Stage of Program		Promising
Concurrent/ Historical Intervention Exposure		No/Limited Evidence
Key features		
Measurement	1	Weak
Comparison Group	1	Weak
Implementation Fidelity	1	Weak
Follow Up Assessment	1	Weak
Descriptive or Supplemental Criteria		
External Validity Indicators		Promising
Length of Intervention		10 weeks
Intensity/Dosage of Intervention		45 mins 1/week
Program Implementer		Teacher
Characteristics of the Intervener		Unknown
Intervention Style of Orientation		Interpersonal, Experiential
Cost Analysis Data		Unknown
Training and Support Resources		Course Materials
Feasibility		Unknown

Appendix E

Weight of Evidence A

Kratochwill's (2003) American Psychological Association Task Force on Evidence-Based Interventions in School Psychology outlines a coding protocol that can be used to critically consider studies with a group-based design. Following initial adaptations being made to ensure it was suited to this paper's review question; the protocol was used to produce WoE A criteria. Table 1, Table 2, Table 3 and Table 4 show the criteria required to obtain differing ratings based on the requirements detailed in Kratochwill (2003). These ratings were then averaged to create an 'Overall WoE A', which was converted to a 'high', 'medium' or 'low' WoE A rating (See Table 5).

Table 1

Measurement

Weighting	Criteria
Strong (3)	<ul style="list-style-type: none"> - Reliability statistic used and reported with a reliability coefficient of 0.85 or higher for most outcomes - Multiple methods used for data collection and across multiple sources - Validity of data is reported - If multiple primary outcome measures have been used; criteria must be met for all of them
Promising (2)	<ul style="list-style-type: none"> - Reliability statistic used and reported with a reliability coefficient of 0.70 or higher for most outcomes - Multiple methods used for data collection or across multiple sources - If multiple primary outcome measures have been used; criteria must be met for >75% of them
Weak (1)	<ul style="list-style-type: none"> - Reliability statistic used and reported with a reliability coefficient of 0.50 or higher for most outcomes - If multiple primary outcome measures have been used; criteria must be met for >50% of them
No/Limited Evidence (0)	<ul style="list-style-type: none"> - Reliability statistic used and reported with a reliability coefficient of less than 0.50

Table 2

Comparison Group

Weighting Criteria

Weighting	Criteria
Strong (3)	<ul style="list-style-type: none"> - Minimum of one type of “active” comparison group is used and described - Initial group equivalency is demonstrated through statistical analysis - Evidence of counterbalancing of change agents - Criteria for equivalent mortality are met and low attrition is reported
Promising (2)	<ul style="list-style-type: none"> - Minimum of a “no intervention group” comparison - At least two of the following must be demonstrated: <ul style="list-style-type: none"> - Change agents are counterbalanced - Group equivalency - Equivalent mortality with low attrition (or an intent-to-intervene analysis with no significant group differences; if equivalent mortality is not demonstrated)
Weak (1)	<ul style="list-style-type: none"> - A comparison group - At least one of the following must be demonstrated: <ul style="list-style-type: none"> - Change agents are counterbalanced - Group equivalency - Equivalent mortality with low attrition (or an intent-to-intervene analysis with no significant group differences; if equivalent mortality is not demonstrated)
No/Limited Evidence (0)	<ul style="list-style-type: none"> - No efforts made to ensure group equivalence

Table 3

Implementation Fidelity

Weighting	Criteria
Strong (3)	<ul style="list-style-type: none"> - Strong evidence of acceptable adherence measured through at least two of the following: <ul style="list-style-type: none"> - Ongoing supervision/consultation - Coding sessions - Audio/Video tapes - AND use of an intervention manual that is either written or provided through a formal training session (must include a detailed account of the exact procedures and sequence of use)
Promising (2)	<ul style="list-style-type: none"> - Evidence of acceptable adherence measured through at least two of the following: <ul style="list-style-type: none"> - Ongoing supervision/consultation - Coding sessions - Audio/Video tapes - AND information provided through either a written material or a formal or informal training session (must include an overview of principles and a description of intervention phases)
Weak (1)	<ul style="list-style-type: none"> - Evidence of acceptable adherence measured through at least one of the following: <ul style="list-style-type: none"> - Ongoing supervision/consultation - Coding sessions - Audio/Video tapes - OR use of a manual
No/Limited Evidence (0)	<ul style="list-style-type: none"> - No attempts to ensure implementation fidelity or evidence of unacceptable adherence

Table 4

Follow Up Assessment

Weighting	Criteria
Strong (3)	<ul style="list-style-type: none"> - Strong evidence of acceptable adherence measured through at least two of the following: <ul style="list-style-type: none"> - Ongoing supervision/consultation - Coding sessions - Audio/Video tapes - AND use of an intervention manual that is either written or provided through a formal training session (must include a detailed account of the exact procedures and sequence of use)
Promising (2)	<ul style="list-style-type: none"> - Evidence of acceptable adherence measured through at least two of the following: <ul style="list-style-type: none"> - Ongoing supervision/consultation - Coding sessions

- Audio/Video tapes
 - AND information provided through either a written material or a formal or informal training session (must include an overview of principles and a description of intervention phases)
- Weak (1)
- Evidence of acceptable adherence measured through at least one of the following:
 - Ongoing supervision/consultation
 - Coding sessions
 - Audio/Video tapes
 - OR use of a manual
- No/Limited Evidence (0)
- No attempts to ensure implementation fidelity or evidence of unacceptable adherence

Table 5

Overall WoE A

I D	Author	Indicators				Overall I WoE A
		Measure- ment	Compariso n group	Implement -ation Fidelity	Follow up assessme nt	
1	Elek et al. (2010)	1	1	1	1	1
2	Hecht et al. (2003)	1	2	2	2	1.75
3	Kulis et al. (2016)	1	2	2	0	1.25
4	Kulis et al. (2007)	0	1	1	2	1.33
5	Marsiglia et al. (2015)	1	3	2	2	2

Appendix F

Weight of Evidence B

Weight of Evidence B considers the methodological relevance of each study to the review question. This is therefore looking to critically consider whether the design undertaken in the study is appropriate for looking at effectiveness of the “keepin’ it REAL” intervention as a universal approach for reducing substance use. This is based on work by Guyatt et al. (1995, 2008) into evidence hierarchies in which threat to internal validity is considered. When there is a higher threat to internal validity, lower weighting is given, whilst a higher weighting is given to those studies with low threat to internal validity. Table 1 shows the criteria that must be met for each weighting.

Table 1

Weighting	Criteria
Strong (3)	<ul style="list-style-type: none"> - Must have an “active” comparison group with randomisation in group allocation - Outcomes collected pre- and post- intervention for both/all groups - Multiple methods and multiple sources of outcome data are obtained for each outcome - Statistical power should be met
Promising (2)	<ul style="list-style-type: none"> - Must have a comparison group with randomisation in group allocation - Outcomes collected pre- and post- intervention for both/all groups - Multiple methods and/or multiple sources of outcome data are obtained for each outcome - Statistical power should be met
Weak (1)	<ul style="list-style-type: none"> - No evidence of a comparison group - Outcomes collected pre- and post- intervention for intervention group - One method and/or source of outcome data is obtained for each outcome

Table 2

Study	Author	Overall WoE B
1	Elek et al. (2010)	1
2	Hecht et al. (2003)	2
3	Kulis et al. (2016)	2
4	Kulis et al. (2007)	2
5	Marsiglia et al. (2015)	2

Appendix G

Weight of Evidence C

Weight of Evidence C considers the topic relevance of each study to the review question. This is therefore looking to critically consider whether the study and the findings matched to the review question. The criteria is detailed in Table 1 and the following details the rationale for these criteria:

- The level of training, support and intervention monitoring is described in the study and follows the curriculum specified by Miller-Day and Hecht. This would ensure that the curriculum was being delivered as designed to ensure fidelity.
- Data should be triangulated through collecting information from multiple sources and through multiple methods to enhance the validity of the findings (Barker, Pistrang & Elliott, 2002).
- The sample population should reflect the population specified in the review question. Therefore the sample population should be a universal population within school. This would ensure a larger generalisability across secondary schools.

Table 1 shows the criteria that must be met for each weighting.

Table 1

Weighting	Criteria
Strong (3)	<p>Study meets all of the below criteria:</p> <ul style="list-style-type: none"> - Must contain a detailed description of the “keepin’ it REAL” curriculum delivered to the participants which matches the curriculum originally described by Miller-Day and Hecht - Must contain information about how fidelity was ensured through monitoring - Data is triangulated - The sample population is the whole class/school - The intervention should be implemented in a Native English origin country
Promising (2)	<p>Study meets three of the below criteria:</p> <ul style="list-style-type: none"> - Must contain a detailed description of the “keepin’ it REAL” curriculum delivered to the participants which matches the curriculum originally described by Miller-Day and Hecht - Must contain information about how fidelity was ensured through monitoring - Data is triangulated

- The sample population is the whole class/school
- The intervention should be implemented in a Native English origin country

Weak
(1)

Study meets at least one of the below criteria:

- Must contain a detailed description of the “keepin’ it REAL” curriculum delivered to the participants which matches the curriculum originally described by Miller-Day and Hecht
- Must contain information about how fidelity was ensured through monitoring
- Data is triangulated
- The sample population is the whole class/school
- The intervention should be implemented in a Native English origin country

Table 2

Study	Author	Overall WoE C
1	Elek et al. (2010)	2
2	Hecht et al. (2003)	2
3	Kulis et al. (2016)	1
4	Kulis et al. (2007)	2
5	Marsiglia et al. (2015)	1

- Appendix H
- Banding Criteria for Weight of Evidence D

Weight of Evidence D is the overall average weighting for each study. This therefore takes into consideration all of the above criteria and provides an overall score that reflects the extent to which the study answers the review question.

Table 1 shows the weightings criteria and their corresponding categories of 'high', 'medium' or 'low'.

Table 1

<u>Weighting criteria</u>	<u>Description</u>
≤1.4	Low
1.5-2.4	Medium
≥ 2.5	High