This is a preprint of:

McAleavey, A. A. (2021). On the importance of context-sensitive interventions for college student mental health: Commentary on “Prevention and treatment of mental health and psychosocial problems in college students: An umbrella review”. *Clinical Psychology: Science and Practice*, 28(3), p. 245-249. doi: <https://doi.org/10.1037/cps0000034>

©American Psychological Association, 2021. This paper is not the copy of record and may not exactly replicate the authoritative document published in the APA journal. The final article is available, upon publication, at: <https://doi.org/10.1037/cps0000034>

**On the importance of context-sensitive interventions for college student mental health:
Commentary on “Prevention and treatment of mental health and psychosocial problems in college
students: An umbrella review”**

Andrew A. McAleavey

Center for Health Research, Helse Førde Hospital Trust

Department of Psychiatry, Weill Cornell Medical College

Author Note

Andrew A. McAleavey <https://orcid.org/0000-0001-5986-2033>

No known conflicts of interest to disclose.

Correspondence concerning this article should be addressed to Andrew A. McAleavey, 45
Tefrevegen, 6812 Førde, Norway. Email: andrew.athan.mcaleavey@helse-forde.no

In this issue, Cuijpers et al. (2021) provide an overview of the research on interventions to prevent and treat mental health problems among students in higher education. This valuable service to the field provides a moment to ask what we have accomplished so far, and what remains. While no study can identify all evidence relevant to any given research question, this review goes farther than most towards summarizing the available data on treatment and prevention of mental distress in college students. It will, in all likelihood, serve as an important introduction to the topic for trainees, a map of the field for clinicians, and mandatory reading for policy makers. The authors have truly done yeoman's work to bring coherence to a large volume of literature. Given the breadth of the research summarized here, however, it is daunting to draw conclusions about the state of research on college student mental health. Any attempts to summarize, including those below, will inevitably oversimplify.

This is, perhaps, a good state of affairs for the field as a whole. This review depicts an area of active research with diverse goals and outcomes, which is exactly what is needed to improve and refine interventions in the future. Reviews like this are probably our best opportunities to assess the accumulated evidence and consider what is needed next. As such, it is just as important to identify the limits our confidence as it is to articulate our confident findings.

In this commentary, I pose three questions regarding the present umbrella review:

1. What conclusions, if any, can we make with relative certainty?
2. What are the remaining areas of uncertainty?
3. What priorities would best support mental health in college students going forward?

The answers to these questions below are only intended as suggestions, while the real value of this review will be its breadth and ability to inform diverse perspectives on these and similar questions. All empirically-minded interventionists who wish to improve the wellbeing of students in higher education would be well advised to critically appraise the findings of this review with these questions in mind. In particular, acknowledging the limits of the summarized research can only help guide future efforts towards college student mental health.

1. What conclusions, if any, can we make with relative certainty?

While there are many reasons to be cautious in interpretation, there are some conclusions that seem well-justified from this review. First, there are clearly interventions for prevention and treatment of common mental health conditions that are, on average, more effective than control conditions for students in higher education. Or, in a more falsificationist phrasing: We can reasonably confidently rule out the idea that the tested interventions produce no effect when compared to the controls used in these studies. My interpretation would be that this is a reasonable conclusion despite the risk of bias discussed below, given the consistency of findings and the fact that an intervention being more effective than control is relatively low bar. Nevertheless, this is encouraging. It may not be surprising – there is evidence to support similar interventions in many settings and populations besides higher education – but novelty is far from the most important contribution for a review-of-reviews like this. As a result of this umbrella review, we can at minimum conclude that it seems possible to provide efficacious interventions to students in higher education. This gives the field a standard to aim for when delivering interventions. We should expect that students who receive interventions have a good chance to experience some benefit. When our interventions are not successful, we can use this review to direct our attempts to improve future outcomes, rather than infer that beneficial interventions are yet to be created.

However, the meta-analyses summarized here contain studies with very high levels of heterogeneity. While there were some patterns in this – substance use studies had somewhat lower levels than depression and anxiety studies, though both had a wide range and virtually all 95% CIs for I^2 as summarized in Tables 3, 4, and 5 extend beyond 0.5 – moderate to highly heterogeneous effects were present in every meta-analytic category. Reasonable people may disagree on how to interpret the observed heterogeneity, and my own interpretation varies somewhat from that of the review authors. The I^2 statistic used here is an estimate of the percent of total variance across studies due to between study variance, and when it is large, we can conclude that the effects being summarized may well represent different true effects. It does not tell us precisely how different another effect would be in another sample, so we cannot be sure that the differences are large in absolute value and therefore we cannot know exactly how meaningful in everyday life such heterogeneity is. Nonetheless, it provides important information about between-study

effects: many effects summarized here have substantial between-study differences, relative to the sampling error. Heterogeneity may have many possible causes, including publication bias, different treatment effects and/or sample definitions across studies, inherent complexity of the problems and treatments being summarized, or numerous other factors. We can be reasonably confident, based on this, that the efficacy of an intervention in one trial will have limited predictive value in estimating the effects of similar interventions (as grouped by the original meta-analysts) in other studies or settings, in spite of the general efficacy of these interventions already discussed. Heterogeneity should limit our willingness to generalize findings across studies, and therefore we should not assume that the aggregated effect sizes will recur in future studies, settings, or individuals. While we may still be able to make some conclusions about average effects, better conclusions may be possible with as-yet-unknown context-specific information.

And finally, we can confidently state that the quality of meta-analyses and randomized trials in this field is lower than desired. Obviously, we would desire extremely high-quality research, but this review reveals that the field as a whole suffers from a lack of rigor in the conduct and reporting of research. Fewer than half of the studies that had quality ratings satisfied half of the quality metrics that were assessed. It appears that low-quality trials have become the norm in this field. If we wish to make informed choices for individual students and for larger groups, we require greater attention to the quality of the trials. We do not need more poorly conducted, low-powered trials without sufficient controls; we need pre-registered, well-documented, large-N trials with open publication to lower the risks of bias. Meta-analysts should also note that the quality of meta-analyses was not rated much higher than the quality of the summarized primary studies. Some issues related to meta-analysis quality ratings might be considered “simple” incomplete reporting of results (e.g., only two of 31 meta-analyses reported the funding sources of their studies), but others are quite serious omissions (e.g., fewer than half of the meta-analyses adequately assessed risk of bias). These limitations in the conduct and reporting of research in the field of college student mental health are frustrating, and deserve decisive actions on the part of

researchers and funders. Unfortunately, they make further inferences regarding the major findings of this umbrella review difficult.

2. What are the remaining areas of uncertainty?

Since science largely progresses by identifying uncertainty and subsequently narrowing down its possible causes, any uncertainty that remains after an expansive review like this are important targets for future research. I will highlight two major points of uncertainty that should be considered substantial contributions of the current study, because they remain to be clarified.

1. What explains the heterogeneity of intervention effects across studies?
2. How effective are the actually-available interventions for college student populations?

What explains the heterogeneity of intervention effects across studies? On the limited basis of the I^2 statistics reported here, we can see considerable differences in observed effects across studies. We must ask: Why are some interventions more efficacious than others, even when they seem to be similar in form, apply to similar individuals, and use similar strategies?

The list of potential causes of heterogeneity is infinite and most are not mutually exclusive. Broadly, heterogeneous effects might be attributable to differences between patient samples (some patients might have better responses than others because of some trait(s) or state(s), and those trait(s) and state(s) are non-randomly variable across studies), within and/or between interventions (some interventions may be more efficacious than other interventions and/or one intervention may be more efficacious under some circumstances than others), within and/or between study designs (e.g., different measurement instruments, variability across sites, control groups, timing of assessments). In addition, heterogeneity might occur as a result of oversimplification in study or meta-analytic design (asking research questions that ignore essential differences, as summarizing a main effect when clinically meaningful interactions dominate it), publication bias, or the interactions of several of these factors. Because this review pertains to college students, we must also consider that many of these potential causes overlap with the major clustering variable determining inclusion in this review at all: differences between institutions which may relate to all of the other factors.

Clearly, this covers a huge range of open questions, which is precisely why the observed heterogeneity may be one of the most important findings from this umbrella review. Aggregating main between-group effects across comparative efficacy trials may be asking an overly simple question of a complicated reality. Put another way, any trial with a significant finding for one intervention compared to control appears to convey limited information about the efficacy of seemingly similar interventions in seemingly similar samples of students. This is a humbling but meaningful insight.

It would be ideal to have synthesized moderator analyses, perhaps from meta-regression, individual patient data meta-analyses, or other methods, that might help explain some of this heterogeneity. Some of the meta-analyses included here already contain such analyses, but not at the extreme high-level view of this umbrella review. To ensure this is possible in future reviews of reviews, we must improve the quality of research and publication in this population. Primary studies on college students would be most likely to contribute to such meta-findings if they report as much detail on the samples, methods, interventions, data, and measurements as possible. It is increasingly apparent that transparency in research practices is necessary for meaningful scientific progress. Researchers who are interested in maximizing the potential for positive impact on college students should make their studies more amenable to meta-analytic synthesis by making more aspects of their study available to public view, from protocol to data analysis. The easier it becomes to aggregate data across diverse studies with measured (rather than unknown) variability, the swifter progress of our interventions will become.

How effective are the actually available interventions for college student populations? It is vital that we recognize that the randomized trials summarized here likely do not represent, and may not even resemble, the actual experiences of most students seeking services in higher education. Especially in the context of non-universal interventions, it is likely that students participating in clinical research trials differ from their unstudied peers, and by definition anyone participating in a randomized trial is experiencing something out of the ordinary. In this context, while it is encouraging that this review demonstrates efficacy for many interventions in controlled trials, this umbrella review does not inform us

as to the effectiveness of the types of interventions a given student would be likely to encounter “in the wild.”

In virtually every interventional science, there is a difference between the effects of interventions in controlled research and application of interventions in widespread use. At a minimum, this will entail a delay between development of efficacious interventions and their dissemination, such that actual practices adopt state of the art interventions after they are shown to be efficacious and safe. In the worst case, the effects of an intervention may be considerably lesser when applied in actual practice than in research, due to a wide range of causes, some of which are attributable to poor research practices (e.g., reliance on small and unrepresentative samples) and others to issues in practice and service delivery (e.g., limited resources for training in novel interventions). There are also serious problems with access to mental healthcare in real world settings which have more to do with systemic barriers like poverty, racism, payment structures, and politics than clinical services per se, but nonetheless must be overcome for an intervention to have meaningful and equitable effects. These obstacles mean that even after efficacious treatments are disseminated, their availability and effects may differ from those observed in trials as underlying disparities perpetuate. For all these reasons and more, we do not, on the basis of this umbrella review, know how effective the actually-available interventions for college students are. It is worth observing that, as Kazdin (2013) noted, since so many individuals in need of mental healthcare cannot or do not access any services, perhaps the term “usual care” (particularly in the United States) should actually refer to no treatment. It is very important that we understand what services are actually available to college students, and how effective those services actually are, prior to drawing conclusions from efficacy research, because the capacity for a novel (or newly evidenced-based) intervention to help individuals depends on both demonstrating efficacy and the identified limitations in existing interventions.

By definition, the most direct research on actual interventions comes from nonrandomized observational studies, which have several limitations. Most clearly, observational studies are limited in their ability to support causal inferences, and unmeasured causes and confounders are difficult to rule out.

These prevent strong conclusions, but not descriptive findings. In higher education, we can learn a great deal from examining some of the most commonly-available services for college students, namely university and college counseling centers (UCCs). At many institutions of higher learning, UCCs are specifically mandated to improve the wellbeing of the student body through both prevention and treatment interventions, often with relatively low costs at point of service due to drawing funds directly from institutional sources. Particularly by examining large samples of diverse institutional providers, we can build confidence in the representativeness of findings for individuals in these and similar settings. It is fortunate that many UCCs have, for a long time, participated in structured data collection and sharing programs, which enable some interrogation of the needs of the individuals in their care as well as the strengths and weaknesses of the care available.

The Center for Collegiate Mental Health (CCMH) is one international group of UCCs that routinely collects and aggregates structured data directly during service provision, rather than in survey form. Research from CCMH has documented increasing demand for services at UCCs over a five-year period and increasing self-reported levels of distress and prior mental healthcare among those seeking services (Xiao et al., 2017), suggesting that there are more students seeking these routine treatments and those who do so are increasingly distressed. Moreover, CCMH research has found substantial variability among UCCs in service models and capacities, such that some UCCs appear to offer a treatment model similar to those common in research (i.e., roughly weekly individual care with substantial treatment lengths without substantial cost at point of service) while other UCCs tend to provide “diluted” care on schedules that are decidedly not evidence-based (fewer and less frequent sessions with frequent referrals to outside care and additional costs for service) due to limited resources (Center for Collegiate Mental Health, 2021). These differences in treatment models across centers mean that the kind of care that most college students can expect to receive in practice is likely highly variable, with many students likely going underserved, particularly in specific institutions, and others having relatively easy access to substantial low-cost care.

We can also examine treatment outcome in UCCs, since many UCCs have been longstanding users of routine outcome monitoring (ROM). In ROM, service providers collect structured data, usually in the

form of standardized quantitative self-report outcome questionnaires, at regular intervals during treatment in order to track patient progress towards goals and overall service productivity. Based on ROM data, it is possible to provide some sense of the effectiveness of these widely-available treatment interventions for college students. Formal methods for evaluating routine outcomes have been developed that allow for statistical comparisons between routine care and meta-analytically-derived benchmarks (see Minami et al., 2009). The clear limitations of observational studies still apply, but because UCCs provide some of the most widely available treatments in this setting, it would be useful to know if students initiating care tend to show improvement or not.

The research so far on effectiveness of routine psychological interventions provided by UCCs seems to suggest that treatment is associated with substantial pre-post changes in symptoms, often with effect sizes that are statistically indistinguishable from the pre-post changes observed in well-controlled trials. Minami et al. (2009) compared outcomes from one counseling center to the outcomes of meta-analytically synthesized research on the treatment of Major Depression. They found that their sample of 2,672 college students self-reported standardized changes in general distress ($d = .98$) that were equivalent to treatment conditions and superior to waitlist conditions in published trials. McAleavey et al. (2019), in a sample of CCMH data from 108 UCCs, report similar findings: large pre-post effect sizes that are statistically and clinically significantly larger than waitlist benchmarks and clinically equivalent to meta-analytic benchmarks from clinical trials for students with elevated initial scores. This finding held across measures of depression ($d = 1.90$, $N = 4,896$), generalized anxiety ($d = 1.63$, $N = 4,152$), social anxiety ($d = 1.52$, $N = 5,754$), eating pathology ($d = .93$, $N = 2,701$), hostility/anger ($d = 1.31$, $N = 2,974$), with slightly less promising results for alcohol use ($d = 1.00$, $N = 2,493$), in which the effect was superior to no treatment but not clinically equivalent to the treatment benchmarks. These studies had the benefit of routinely-completed outcome measures and include individuals who completed only one session at a UCC, so provide a relatively complete picture of the outcomes for individuals who initiate care at UCCs, and are similar to an intent-to-treat sample. While these are only two papers among others, the available

observational studies in college students appear to show substantial pre-post changes associated with treatment across UCCs.

There are many reasons to treat findings from routine care settings with caution, but as descriptive information they are quite useful. We can see evidence of unmet needs of college students, disproportionately distributed at certain institutions with limited access. The observed effect sizes from treatments are substantial but also variable, apparently depending on factors like setting, sample characteristics, and measurement tools. This, like the between-study heterogeneity observed in the present umbrella review, may be an essential component of the results. That would mean that we should not only seek to make claims about treatment's overall effectiveness, but rather better understand the factors predicting greater or lesser effectiveness in practice, which may vary from one setting to another. Properly addressing the limitations in existing care, therefore, requires context-sensitive interventions.

3. What priorities would best support mental health in college students going forward?

It seems that we can conclude, based on this umbrella review, that efficacious interventions for the treatment and prevention of depression, anxiety, stress, substance use, and some other common problems in college students exist, though we do not fully understand why these interventions are more or less effective in some studies. We can see, from other areas of research, that access to mental healthcare is limited in many higher education settings, and that the treatment interventions that are available at UCCs appear at least somewhat effective but with considerable variability. Given this, what is needed next?

My suggestion would be that we need to address the serious issues of access to mental healthcare, and improve the existing services, but ensure that we do both in context-sensitive ways. In general, we should prioritize research and interventions that provide the largest chance for positive impact on the population as a whole, rather than just those who complete a given type of treatment course. This is probably more achievable, in most settings, through deployment of scalable interventions and addressing barriers to access than through improving quality of care among people who are already likely to receive interventions. However, there will not be one solution at every institution of higher education. The barriers to access to care will be different at a commuter school in a large city than they will at a

residential college in a rural setting, as will the quality and quantity of standard mental health services. There are also substantial differences across national borders, demographic and cultural categories, and type of education being pursued which cannot be ignored. We must be cautious in generalizing across settings and individuals.

To ensure interventions are appropriately context sensitive, it is important to have close collaboration among the major stakeholders in those contexts. In some settings, there may be a need for universal preventative interventions which are likely relevant only to a small number of students in another context (e.g., there are some universities where alcohol use is relatively common and others where it is much less so, those institutions should have different strategies to reduce the burdens of substance use). Researchers, clinical service providers, university administration, and students should all be involved in decision-making processes. Together, these stakeholders can identify the specific needs of their students, choose among the possible interventions to add to the existing resources in place, and develop effective implementation plans. One model for such collaborations is CCMH, which operates as a practice-research network. In this model, clinicians from member UCCs collaborate with researchers to identify assessment and intervention targets requiring study. Each member UCC can choose interventions and assessments that are most relevant to their needs, but the aggregated standardized data and resources from other institutions facilitates a number of important functions. For instance, UCC administrators can identify ways their clinical service delivery model deviates from other centers, and how their clinical population differs from average (Center for Collegiate Mental Health, 2021). Such collaboration, if combined with increased availability of scalable resources, could be an efficient and context-sensitive way to improve mental healthcare for college students.

Conclusion

In their umbrella review, Cuijpers et al. (2021) have provided an excellent opportunity to assess the output of a substantial portion of research on college student mental health. Their findings are complex and limited, but overall hopeful: efficacious interventions have been developed in this population for a wide range of issues. The findings also appear to contain substantial heterogeneity, especially for non-

substance use targets, suggesting that these positive effects are meaningfully context- and study-specific. In addition, we should not forget that many of the challenges facing most potential users of mental health interventions – extremely limited access to care and variable effectiveness of actually-available interventions – is not well-captured in this review. To improve the well-being of the greatest number of students in the future, we likely need scalable interventions to be deployed in targeted, context-specific ways. This may be best accomplished through collaboration among researchers, clinical providers, student representatives, and university leaders, who can together identify those students not being adequately served by existing interventions and develop contextually-optimal additive interventions to complement existing strengths in each context. We should not expect one solution to resolve all problems in all contexts, especially given large differences in effects across studies and large observed differences in existing services across institutions. However, this review provides hope that when interventions are properly targeted in these settings, they can be efficacious.

References

- Center for Collegiate Mental Health. (2021). *2020 Annual Report* [Publication No. STA 21-045].
<https://ccmh.psu.edu/assets/docs/2020%20CCMH%20Annual%20Report.pdf>
- Cuijpers, P., Miguel, C., Ciharova, M., Aalten, P., Batelaan, N., Salemink, E., Spinhoven, P., Struijs, S., de Wit, L., Gentili, C., Ebert, D., Harrer, M., Bruffaerts, R., Kessler, R. C., & Karyotaki, E. (2021). Prevention and treatment of mental health and psychosocial problems in college students: An umbrella review of meta-analyses. *Clinical Psychology: Science and Practice*.
- Kazdin, A. E. (2013). Evidence-Based Treatment and Usual Care: Cautions and Qualifications. *JAMA Psychiatry*, 70(7), 666–667. <https://doi.org/10.1001/jamapsychiatry.2013.2112>
- McAleavey, A. A., Youn, S. J., Xiao, H., Castonguay, L. G., Hayes, J. A., & Locke, B. D. (2019). Effectiveness of routine psychotherapy: Method matters. *Psychotherapy Research*, 29(2), 139–156. <https://doi.org/10.1080/10503307.2017.1395921>
- Minami, T., Davies, D. R., Tierney, S. C., Bettmann, J. E., McAward, S. M., Averill, L. A., Huebner, L. A., Weitzman, L. M., Benbrook, A. R., Serlin, R. C., & Wampold, B. E. (2009). Preliminary evidence on the effectiveness of psychological treatments delivered at a university counseling center. *Journal of Counseling Psychology*, 56(2), 309–320. <https://doi.org/10.1037/a0015398>
- Xiao, H., Carney, D. M., Youn, S. J., Janis, R. A., Castonguay, L. G., Hayes, J. A., & Locke, B. D. (2017). Are we in crisis? National mental health and treatment trends in college counseling centers. *Psychological Services*, 14(4), 407.