Anthropomorphic Character Animations Versus Digital Chalk Talks in a Resident Diabetes Pharmacotherapy Curriculum: a Randomized Controlled Trial

Bryan Brown, MD MHS[1] ⓘ, Katherine A. Gielissen, MD MHS[2], Sarita Soares, MD[2], Catherine A. Gao, MD[3], Jeremy Moeller, MD MSc[4], and Donna Windish, MD MPH[2]

[1]Office of Medical Education, University of Hawai'i John A. Burns School of Medicine, Honolulu, HI, USA; [2]Department of Internal Medicine, Yale University School of Medicine, New Haven, CT, USA; [3]Division of Pulmonary & Critical Care, Northwestern Feinberg School of Medicine, Chicago, IL, USA; [4]Department of Neurology, Yale University School of Medicine, New Haven, CT, USA.

BACKGROUND: Animation in medical education has boomed over the past two decades, and demand for distance learning technologies will likely continue in the context of the COVID-19 pandemic. However, experimental data guiding best practices for animation in medical education are scarce.

OBJECTIVE: To compare the efficacy of two animated video styles in a diabetes pharmacotherapy curriculum for internal medicine residents.

DESIGN: Learners were randomized to receive one of two versions of the same multimodal didactic curriculum. They received identical lectures, group activities, and quizzes, but were randomized to either digital chalk talk (DCT) videos or *Sugar-Coated Science* (SCS). SCS is an animated series using anthropomorphic characters, stories, and mnemonics to communicate knowledge.

PARTICIPANTS: Ninety-two internal medicine residents at a single academic medical center received the curriculum within ambulatory medicine didactics.

MAIN MEASURES: Knowledge was measured at multiple time points, as was residents' self-reported comfort using each medication class covered. Surveys assessed video acceptability and telepresence. Key themes were identified from open-ended feedback.

KEY RESULTS: Baseline knowledge was low, consistent with prior needs assessments. On immediate posttest, mean scores were higher with SCS than DCT (74.8% versus 68.4%), but the difference was not statistically significant, $p = 0.10$. Subgroup analyses revealed increased knowledge in the SCS group for specific medication classes. Delayed posttest showed significant knowledge gains averaging 17.6% across all participants ($p < 0.05$); these gains were similar between animation types. SCS achieved significantly higher telepresence, entertainment, and acceptability scores than DCT. Qualitative data suggested that residents prioritize well-designed, multimodal curricula over specific animation characteristics.

CONCLUSION: SCS and DCTs both led to learning within a multimodal curriculum, but SCS significantly enhanced learner experience. Animation techniques exemplified by both SCS and DCTs have roles in the medical educator toolkit. Selection between them should incorporate context, learner factors, and production resources.

KEY WORDS: animation; multimedia; graduate medical education; diabetes.

© The Author(s), under exclusive licence to Society of General Internal Medicine 2022

INTRODUCTION

Medical education animation (MEA) is popular among medical students and trainees, as evidenced by the prevalence of publicly and commercially available study products.[1–4] MEA use is likely to increase with the need for distance learning strategies in the setting of the COVID-19 pandemic and the learning preferences of today's medical learners.[5,6] Despite this high demand, best practices for MEA creation, curation, and implementation are lacking in the literature.[7] Animation is broadly defined as "simulated motion pictures depicting movement of drawn (or simulated) objects."[8] MEA research is therefore complicated by the diversity of technologies and aesthetics, including digital chalk talks (DCTs),[1,4,9] two-dimensional (2D) animation,[10] and three-dimensional (3D) computer-generated models.[11]

Without additional experimental medical animation research, medical educators are left to extrapolate from evidence in non-medical disciplines[12] and foundational cognitive theories.[8,13–15] For example, Paivio's dual coding theory[16] suggested in the late 1980s that combining visual and auditory information may expand the capacity of working memory to handle greater cognitive load. The cognitive theory of multimedia learning was subsequently described by Mayer and Moreno, which added additional guidance based on experimental evidence.[8] Recommendations included juxtaposing

related concepts in space and time; avoiding extraneous words, sounds, or visuals; avoiding redundancy between auditory narration and on-screen text; and speaking naturally and conversationally rather than formally.

Numerous commercially successful educational resources for medical trainees employ strategies to enhance learning through use of narrative, characters, and emotional design. These resources take diverse forms including illustrated books,[17] interactive web-based interfaces,[2] and videos.[3] Medical science characters have also spread beyond studying and into entertainment products such as a card game[18] and graphic novel turned animated television series.[19] Characters in these products sometimes have names which phonetically resemble biomedical entities like drugs, molecules, or pathogens, thereby drawing on a mnemonic strategy.[15] Even when their names do not function as mnemonics, characters may still serve as advance organizers, or schema that precede and facilitate information processing.[20,21] Finally, "emotional designs," including application of anthropomorphic features such as faces or limbs, have the potential to engender positive learner emotions and subsequent positive learning gains.[13]

The media industry and educators have also leveraged the benefits of telepresence. Telepresence is a construct describing the degree to which the user experiences "arrival" in the artificial world presented by the media, and degree of "departure" from the real world where the user actually exists.[14] Telepresence can occur with media ranging from written text to virtual reality, and the degree of telepresence achieved from educational multimedia may enhance memory and persuasion.[14]

Meanwhile, management of type 2 diabetes mellitus (T2DM) requires knowledge of numerous medication classes, posing a challenge for some medical trainees.[22,23] Workforce comfort and knowledge with newer T2DM medication classes is particularly vital, given better side effect profiles as well as evidence that the sodium-glucose transporter-2 inhibitors (SGLT2i)[24–26] and glucagon-like peptide 1 receptor agonists (GLP1ra)[27,28] provide important cardiovascular and renal benefits. Recently, a flipped-classroom curriculum for medical residents focused on outpatient T2DM management.[9] The curriculum included asynchronous DCT viewing followed by 45-min synchronous lecture, and demonstrated positive impact on knowledge quizzes at short- and long-term follow-up, as well as improved confidence using medication classes other than insulins, sulfonylureas, and metformin. Other published educational interventions focusing on various aspects of T2DM have included interprofessional electives,[23] interactive seminars,[29] and designated theme days.[30] Despite these advances, best strategies for teaching trainees how to incorporate newer medication classes like GLP1ra, SGLT2i, and dipeptidyl peptidase 4 inhibitors (DPP4i) remain poorly defined.

We sought to define a T2DM MEA approach that is most acceptable and effective for internal medicine residents. We hypothesized that an animated video series rich in metaphorical characters, stories, and comical dialogue would be superior to the DCT style in its impact on knowledge and attitudes surrounding T2DM. This randomized controlled trial makes a uniquely granular comparison, interchanging two distinct forms of MEA in the same curriculum.

METHODS

Sugar-Coated Science Development

The authors' novel approach to building MEA capacity among clinician-educators by combining Kern's Six Steps of Curricular Development with animation industry techniques has been previously published.[20] The product of this interdisciplinary exploration was an animated video series entitled *Sugar-Coated Science* (SCS) (see Appendix 1 for video sample). SCS features anthropomorphic cartoon characters to represent drugs. T2DM content was transformed into stories and scripts of each episode (Table 1), to allow learning to occur during the viewing experience. The videos' overarching objective was that learners choose appropriate diabetes agents based on host factors such as comorbidities, side effect profiles/preferences, and potential added benefits. This goal was subdivided into four learning objectives (Table 1). A total of four SCS episodes were created, each focusing on a different T2DM medication class. Given curricular time limitations, as well as a high level of baseline knowledge and comfort with metformin in the baseline assessment (below) and in a prior pilot study,[20] *Episode 1* was excluded from the curriculum.

Digital Chalk Talk Development

The DCTs were designed to resemble popular existing DCTs.[1,4] DCTs were scripted based on the same learning objectives and content outlines as SCS. A digital slide deck was created for each of the three medication classes in the trial. Each was organized by "mechanism of Action," "examples of medications in the class," "a1c benefit," "added benefits," "adverse effects / contraindications," and "combining agents". Each concluded with an "in summary" section that recapitulated the prior information. A team member (BB) narrated the videos, while transitioning through and drawing on the slides in real time using a drawing tablet and QuickTime Player (see Appendix 1 for video sample).

Knowledge Assessment Development

Twenty case-based multiple-choice questions (MCQs) (Appendix 2) were developed to target four specific learning objectives as shown in Table 1. These questions were written and crosschecked by five authors (BB, CG, SS, KG, DW) and feedback was provided by a T2DM expert (SI). Response process validity was obtained by administering the MCQs to three internal medicine physician volunteers. For additional item validation, the DPP4i and GLP1ra episodes were piloted as a didactic conference to a separate group of internal medicine residents in a primary care program at the same

Table 1 A Description of the Three *Sugar-Coated Science* Episodes Administered in the RCT

	Episode 2: Keeping Fresh with the Flozins		Episode 3: Terror from the Deep		Episode 4: The Pen-Ultimate Stand	
Medication class	SGLT2i		DPP4i		GLP1ra	
Metaphorical premise	Seagull brothers GliFlo, Empa, and Dapper promote sweet tea business, while attempting to downplay their side effects.		Western-style story of roadrunner named City Gliptin defends the Pancreatic Plains and its incretin lizards from "deep-pp4" sand worms.		Gila monsters Lyre Aglutide and Axe Enatide are summoned to boost the effects of incretins on the Gastric Valley and Isles of Langerhans.	
Animated module duration (min)	15.4		9.35		11.1	
DCT module duration (min)	13.9		9.5		12.2	
Relative learning objective representation by medication class on pretest and delayed posttest						
Learning objective	Pretest	Delayed posttest	Pretest	Delayed posttest	Pretest	Delayed posttest
Choose agents based on GFR	0	1	1	1	1	0
Choose agents based on contraindications & adverse effects	2	1	0	1	1	1
Choose agents based on added benefits, patient goals, and magnitude of A1c reduction	1	1	1	1	1	1
Explain mechanism of action in simple terms	0	1	1	0	1	1
Total number of questions per medication	3	4	3	3	4	3

As shown, efforts were made to make corresponding DCTs similar in viewing duration, and to create even representation of medication classes and learning objectives across pretest and delayed posttest
DCT, digital chalk talk group; SGLT2i, sodium-glucose transporter-2 inhibitor; DPP4i, dipeptidyl peptidase-4 inhibitor; GLP1ra, glucagon-like peptide-1 receptor agonist; GFR, glomerular filtration rate

institution, and a subset of the MCQs were discussed in real time with this learner group to gather additional item feedback. Final questions were divided into the pretest and delayed posttest 5 months later, with objectives and medication classes represented as evenly as possible between the two.

In addition to the pretest and delayed posttest MCQs, module viewing was always followed by an immediate posttest that assessed knowledge delivery from the preceding video, as well as baseline knowledge of a non-pharmacology topic about to be presented in class. The role of these immediate posttests will be further elucidated in the curricular format description below. For immediate posttests, three team members (BB, KG, SS) drafted case-based vignettes and question sets, then exchanged with one another to suggest revisions (Appendix 3).

Additional Survey Instruments

Additional data was collected at the pretest and delayed posttest time points. First, residents self-assessed their confidence using T2DM medications on a digital slider from 0 to 100%. Specifically, they were asked to rate comfort with the medications in the upcoming videos, but also with metformin to confirm the instructors' suspicion that metformin review could be omitted from the session.

Second, participants rated their video experience on a novel Likert scale questionnaire assessing four parameters of video acceptability: clarity, attention, usefulness, and entertainment (1 = strongly disagree, 5 = strongly agree; see Appendix 4). These subconstructs were based on a list of "fractal communication elements" that are important to medical multimedia in the Internet age,[31,32] and the authors considered them congruous in this context with the "acceptability" construct from the field of

implementation science.[33] These four items were combined into a composite acceptability score used to compare SCS to DCT viewing experiences. Self-assessment of viewer telepresence was assessed using three items derived from a preexisting questionnaire (Appendix 4).[14] Given the limited time for synchronous teaching, these psychometric instruments could not be administered at the immediate posttest time point.

Participants

The study was administered to all residents of a single internal medicine residency program who attended their mandatory ambulatory didactic curriculum on one of four dates. These four dates were randomly and evenly assigned to either the SCS or DCT group by random number generator. All other curricular components were identical for the two groups. Participants were informed of their freedom to withhold responses from any MCQs and survey items. They were asked to provide anonymous identifiers to facilitate paired analyses.

Session Format

The entire experience, including video viewing, occurred within a 3-h ambulatory didactic session. The session began with the pretest, administered via Qualtrics™. Then, a 50-min multimodal format was repeated three times. After 3 h, the residents had received videos and teaching on three T2DM medication classes, and lectures on three non-pharmacotherapy topics. Each cohort remained in the SCS or DCT group throughout their experience.

Delayed Posttest Administration

Delayed posttests were scheduled to be administered 24 weeks following the curriculum, when each resident group returned

for additional ambulatory didactics. However, the coronavirus pandemic in Spring of 2020 resulted in cancellation of all in-person didactics at that time point. One of two DCT cohorts received designated time for posttest completion as part of a video-conference didactic early in the pandemic. However, the remaining three quarters of learners received posttests by email. These delayed posttests included the remaining ten pharmacotherapy MCQs, two additional non-pharmacotherapy MCQs, repeat self-reported confidence with medication classes, and additional survey items as above.

Statistical Analysis

Percent of correct responses on knowledge MCQs between pretest and delayed posttest and between SCS and DCT groups were compared by chi-square analyses. Video acceptability survey items were assessed for relatedness by Cronbach's reliability coefficient, and then analyzed individually and in aggregate by Mann-Whitney U tests. The same was done for the three-item telepresence questionnaire. Analyses were performed in Microsoft™ Excel and Graphpad™ Prism.

RESULTS

Demographics

Baseline characteristics of participants are summarized in Table 2. Note that sample sizes vary at different stages given learners' ability to opt in or out of the experience at these different points.

Video Engagement Metadata

The three DCT videos had view count (mean ± SD) of 46.3 ± 0.58 participants across the two DCT cohorts. The three SCS videos had view counts of 45.7 ± 3.2 across the two SCS cohorts. The mean percentages of each video's full duration viewed by trainees ranged from 91.0 to 97.7%, and these percentages did not differ significantly between DCT and SCS for any medication class, all $p > 0.05$.

Pretest Findings

Participants' baseline self-reported confidence with each medication class is shown in Table 2; there were no significant intergroup differences, $p > 0.05$. Comfort with metformin was significantly higher than comfort with the other three classes in both groups by one-way ANOVA, both $p < 0.01$. Baseline performance on the knowledge MCQ did not differ between DCT and SCS groups, respectively ($p = 0.40$; Table 2).

Immediate Posttest Findings

For the pharmacotherapy-focused questions across the immediate posttest activities, mean scores were 74.8% for SCS compared to 68.4% for DCTs, $p = 0.10$, with average response rates of 96% and 93% respectively. When analyzed by individual medication class, the SCS group scored significantly higher than the DCT group on DPP4i items at 87.4% vs 71.8%, $p = 0.01$. The other two medication classes did not show significant differences (88.0% vs 95.5%, $p = 0.072$, for SGLT2i and 48.3% vs 36.1%, $p = 0.11$, for GLP1ra). Given

Table 2 Summary of Baseline Demographics, Self-confidence, and Knowledge at Time of Pretest

	Total	SCS	DCT	*p*-value
Number of pretest respondents	92	47	45	
Post graduate year (PGY)				
PGY-1	32	14	18	
PGY-2	34	18	16	
PGY-3	26	15	11	
Post-residency career goals				
Endocrinology	0			
General internal medicine—inpatient	7			
Primary care	0			
Other subspecialty	70			
Undecided/other	14			
Amount of prior teaching on T2DM medications in residency				*p = 0.99*
Never	6	3	3	
<1 h	35	18	17	
1–5 h	51	26	25	
>5 h	0	0	0	
Prior exposure to *SCS*				*p = 0.96*
Episode 1: Metformin Rises	3	2	1	
Episode 2: Keeping Fresh with the Flozins	6	1	5	
Both	9	7	2	
No prior exposure	74	37	37	
Baseline self-reported confidence with medication class, from 0 to 100 (average ± SD)				
Metformin		82.9 ± 15.9	79.6 ± 18.3	*p = 0.37*
SGLT2i		49.4 ± 25.4	46.3 ± 23.6	*p = 0.56*
DPP4i		40.6 ± 22.4	40.5 ± 24.6	*p = 0.98*
GLP1ra		42.0 ± 22.7	43.3 ± 24.8	*p = 0.80*
Baseline diabetes medication knowledge scores (% ± SD)		40.1.0±21.2	36.2±19.5	*p = .40*

SCS, Sugar-Coated Science group; DCT, digital chalk talk group; SD, standard deviation

unexpectedly high scores on SGLT2i questions for both groups (raising concern for inadequate question difficulty regarding this medication class), analysis was also performed excluding SGLT2i questions. This revealed significantly higher scores for SCS over DCT when looking at DPP4i and GLP1ra knowledge, 68.5% compared to 58.2%, $p = 0.04$. There was no significant group difference in knowledge on non-pharmacotherapy topics not discussed in the videos, $p = 0.80$.

Open-ended Feedback

Learners provided feedback at two time points: in general evaluations by the course director immediately following the session, and at the delayed posttest. Free text responses from both sources were compiled and reviewed by two authors, BB and KG. Consensus was reached on a set of consistent themes, most focusing on the curriculum macrostructure rather than specific animation elements. These themes, including the need for consolidative resources such as handouts, the value of a multimodal approach, the importance of optimized activity time allocation and organization, and appropriateness of knowledge assessments, are summarized in Table 3. Multiple learners in both groups stated that they "liked" or "loved" the videos. Learners in either group rarely offered any spontaneous feedback on specific audio or visual elements or animation strategies.

Delayed Posttest Findings

In the setting of the COVID-19 pandemic, response rates at the time of delayed posttest were impacted by the recruitment of residents to alternative clinical activities. Sample sizes (percent pretest sample size) at delayed posttest were 21 (46%) DCT group and 11 (23%) SCS group participants.

Respondent comfort with key medication class use showed significant increases from pretest for SGLT2i, DPP4i, and GLP1ra (all $p < 0.05$). Increases were similar for both groups, but significance was not achieved for the SCS group in the setting of lower response rate.
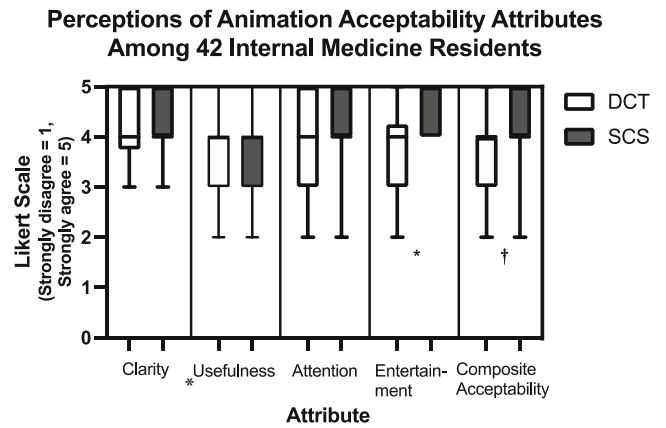

Figure 1 Learners' ratings (*n* = 22 and 10 participants for DCT and SCS, respectively) of video's acceptability in terms of clarity, usefulness, attention impact, and entertainment. *$p = 0.015$ entertainment ($p = 0.015$) and the † composite ($p = 0.034$).

Respondent ratings of video's clarity, usefulness, entertainment, and ability to maintain attention showed a Cronbach's reliability coefficient of 0.87 within learners. The composite acceptability score was significantly higher for SCS compared to DCT, median 4.0 [interquartile range (IQR) 4.0 to 5.0] compared to 4.0 (IQR 3.0 to 4.0), $p = 0.03$ (Fig. 1). Among the individual analyses of these four items, only entertainment independently reached significance, 5.0 (IQR 4.0 to 5.0) compared to 4.0 (IQR 3.0–4.3), $p = 0.02$.

Responses to three items, each with a 5-point scale, addressing the construct of telepresence showed a Cronbach's reliability coefficient of 0.87, and showed a significantly higher rating for SCS compared to DCT, 3.5 (IQR 2.8 to 4.0) vs 3.0 (IQR 2.0–3.0), $p = 0.02$. Each individual items had higher IQR for the SCS group, but none achieved statistical significance.

Ten-item pharmacotherapy knowledge score among all respondents improved from pretest to delayed posttest, 38.0% compared to 55.6%, $p < .05$. Average knowledge scores improved from 36.2 to 56.6% for DCT group, and 39.0 to 53.7% for SCS group, both $p < 0.05$ (Fig. 2). On two control, non-pharmacotherapy questions based on lectures

---

Table 3 Themes and Corresponding Quotes Identified in Open-ended Feedback

| Theme	Examples
Desire for Consolidative Resources	[Requests for] "practice directed algorithm," "concise/short table with highlights", "a handout or optional bibliography on the major studies", and "a sheet on takeaways from the videos".
Diversified modalities within the curriculum	"[I] liked the mix of learning from videos/questions/brief lecture-like talks and working with peers." "Multimodal teaching approach"
Organization and time allocation	"Did not find watching the videos during the session to be that helpful" [in that they sacrificed time to] "interact with the lecturer and ask questions". "I thought the videos made more sense as a flipped classroom activity, although I can imagine some didn't watch them prior to the session last year."
Knowledge assessment appropriateness	[While] "this was the best diabetes education I've ever had" and "I certainly know much more than I would have otherwise," the posttest was likely to "underestimate the true value of this learning experience given that they test a fairly high level of understanding." "More straight forward questions" [to] "identify the big learning points."
Video-specific reactions	[Learners in both groups] "liked" and "loved the videos!!". Outlier respondent: [The SCS videos felt] "a little childish to me". Of note, no comments specifically focused on discrete elements of art or animation style nor audio assets such as sound effects, voice acting, or music.

Knowledge Scores at Pretest and Delayed Posttest
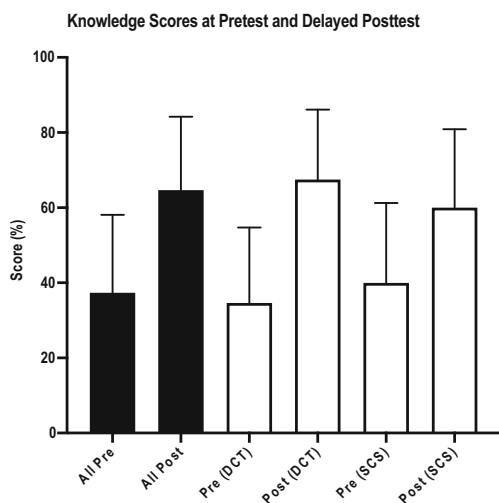
Figure 2 Among respondents (*n* = 21 and 11 participants for DCT and SCS, respectively) who completed the pretest and delayed posttest, knowledge scores rose statistically significantly for the overall class, and for those within each study arm. No between-group differences were detected.

rather than videos, the two groups scored similarly, averaging 0.81 vs 0.86 out of 2.0 points, *p* = 0.9. Given the unexpected loss of response rate, direct comparison of knowledge score improvements between the DCT and SCS groups was not performed.

Within the sample size of the delayed posttest, 24 learners (16 DCT, 8 SCS) had provided an anonymous identifier that could be linked to a pretest and had completed all ten knowledge questions at both time points. Overall, these 24 learners improved significantly from a mean score of 38.7 to 54.3%, *p* < 0.01. On paired analysis, the DCT group's scores significantly improved by an average of 18%, *p* < .01, while the SCS group improved by an average of 19%, but significance was not achieved, *p* = 0.1.

DISCUSSION

Sugar-Coated Science elicited significantly higher composite acceptability scores from residents compared to the DCT style, with entertainment as the largest driver of this difference. *Sugar-Coated Science* also achieved higher telepresence, consistent with the authors' hypothesis that videos more closely resembling popular entertainment media would make learners feel more immersed in the video's artificial world. This finding is encouraging given data supporting correlation of telepresence with memory and persuasion.[14] These data suggest that while both video formats were positively received by learners, MEA resembling *Sugar-Coated Science* may be particularly appealing and engaging to residents.

The *Sugar-Coated Science* animated series uses anthropomorphic characters intended to serve as mnemonics, advance organizers, and emotional designs to enhance resident knowledge and attitudes around T2DM management. This work integrated Kern's Six Steps of Curriculum Development[34] with workflows adapted from the animation craft[20] to create an effective and feasible curriculum. Despite the skills and technology required, the techniques in *Sugar-Coated Science* remain generalizable, as they were implemented by clinician-educators without formal training in animation.

Knowledge gains were assessed at multiple time points, and the overall multimodal curriculum significantly benefitted long-term knowledge. However, a difference in long-term knowledge based on animation style viewed was not detected. Residents' low baseline knowledge and self-confidence with prescribing these key medications confirms the imperative to enhance T2DM pharmacotherapy education among internal medicine residents, especially given that the knowledge items were designed to emulate realistic clinic scenarios. At the immediate posttest, there was some evidence of higher knowledge for *Sugar-Coated Science*, particularly for DPP4i and GLP1ra content. The inability to detect a benefit on immediate SGLT2i knowledge likely related to inadequate SGLT2i question difficulty. Furthermore, the fact that the DCT group almost significantly outperformed SCS for that class may provide insight into the differing roles: DCTs delivered simple knowledge such as the risk of fungal infections with SGLT2i, while SCS aided learners with challenging questions requiring higher application of knowledge. The degree to which question complexity modifies animation style impact is an area for additional study. Outside of direct impacts on knowledge transfer, animation style may have indirectly increased correct responses through heightened engagement in activities immediately following viewing.

This study had multiple limitations, the largest of which was the decline in response rate for the delayed posttest. The authors feel this directly related to the COVID-19 pandemic, the cancellation of synchronous delayed posttest activities, and the extraneous cognitive load at the time of delayed posttest. There may have also been a reduction in outpatient diabetes clinical experiences that may have impacted how knowledge was consolidated. Despite showing significant improvement in correct responses from pretest to delayed posttest in each video group and overall, residents still only answered approximately half of questions correctly at the latter time point. Content misalignment between learning activities and questions is possible but less likely given the extensive efforts to design valid questions based on specific video learning objectives.

Further investigation of different MEA styles in graduate medical education is warranted. First, this single-center study would benefit from replication in additional residency programs to support generalizability. Importantly, each DCT required approximately three person-hours to record, compared to over fifty person-hours of animation per *Sugar-Coated Science* episode. However, the graphic and animation files

of SCS—and not of DCTs—are saved and editable, meaning that content can be easily revised should content needs evolve. Additional research is needed to elucidate contexts in which animation resembling *Sugar-Coated Science* should be implemented, either independently or somehow hybridized with simpler techniques like DCTs. We hypothesize that smaller exposures to character animations could likely achieve similar benefits towards the learners' experience without exhausting feasible production resources. Additional experimentation with such a hybrid format is needed. Additionally, studies controlling for individual animation elements, such as the presence compared to absence of characters, plots, humor, metaphors, or signaling, may be beneficial in identifying best practices, but may require large sample sizes for sufficient power.

In addition to contributing to knowledge about animation specifically, our findings support the existing literature on the benefits of deliberate curriculum development,[34] and particularly on the use of blended learning.[35] The combination of independent video interaction, live polling, synchronous lectures, and group work created a successful didactic experience. The praise as well as constructive, open-ended feedback suggests that residents place more value—at least consciously—on deliberate curricular formatting than on specific multimedia stylistic elements.

In summary, the authors recommend the judicious yet open-minded application and study of techniques including characters, stories, metaphors, and humor in MEA for resident education. Additional research is needed to further our understanding of these techniques' effect among different learner groups, subject areas, and session formats.

Corresponding Author: *Bryan Brown, MD MHS; Office of Medical Education, University of Hawai'i John A. Burns School of Medicine, Honolulu, HI, USA (e-mail: bbrown3@hawaii.edu).*

Data Availability The datasets during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations:

Ethical Approval: *All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee (Deemed Exempt by Yale University Institutional Review Board, ID2000021015) and with* the 1964 Helsinki declaration and its later amendments or comparable ethical standards."

REFERENCES

1. Sattar HA *Fundamentals of Pathology: Medical Course and Step 1 Review.* (Pathoma LLC, 2018).
2. Goel H *et al.* The Picmonic Learning System: enhancing memory retention of medical sciences, using an audiovisual mnemonic Web-based learning platform. AMEP 125 (2014) https://doi.org/10.2147/AMEP.S61875.
3. Wong AL *Sketchymedicine.* (2020).
4. Khan Academy. https://www.khanacademy.org/ (2020).
5. DiLullo C., McGee, P. & Kriebel, R. M. Demystifying the Millennial student: a reassessment in measures of character and engagement in professional education. Anat Sci Ed **4**, 214–226 (2011).
6. Eckleberry-Hunt J, Lick D & Hunt R Is medical education ready for Generation Z? J Graduate Med Educ **10**, 378–381 (2018).
7. Ruiz JG, Cook DA & Levinson AJ Computer animations in medical education: a critical literature review. [Review] [80 refs]. Med Educ **43**, 838–846 (2009).
8. Mayer RE & Moreno R Animation as an aid to multimedia learning. Educ Psychol Rev 13 (2001).
9. Blair RA, Caton JB & Hamnvik OR A flipped classroom in graduate medical education. Clin Teach tct.13091 (2019) https://doi.org/10.1111/tct.13091.
10. Strømme TAA & Mork SM Students' conceptual sense-making of animations and static visualizations of protein synthesis: a sociocultural hypothesis explaining why animations may be beneficial for student learning. Res Sci Educ (2020) https://doi.org/10.1007/s11165-020-09920-2.
11. Daly CJ, Bulloch JM, Ma M & Aidulis D A comparison of animated versus static images in an instructional multimedia presentation. Adv Physiol Educ **40**, 201–205 (2016).
12. Berney S & Bétrancourt M Does animation enhance learning? A meta-analysis Comput Educ **101**, 150–167 (2016).
13. Wong RM & Adesope OO Meta-analysis of emotional designs in multimedia learning: a replication and extension study. Educ Psychol Rev (2020) https://doi.org/10.1007/s10648-020-09545-x.
14. Kim T & Biocca F Telepresence via television: two dimensions of telepresence may have different connections to memory and persuasion.[1]. J Comput-Mediated Commun **3**, 0–0 (2006).
15. Scientific mnemonomies: methods for maximizing more than memory. 22 (2020).
16. Paivio A. *Mental representations: a dual coding approach.* (Oxford University Press ; Clarendon Press, 1986).
17. Gladwin M. *Clinicl Microbiology Made Ridiculously Simple.* (MedMaster, 2011).
18. Healing Blade. http://www.healingblade.com/ (2010).
19. Suzuki K & Ogura H *Cells at Work!* (Aniplex of America, 2018).
20. Brown B *et al.* Becoming clinician-animators: a toolkit and pilot study for novel animated content development in a medical education curriculum. Med.Sci.Educ. **30**, 977–988 (2020).
21. Ausubel DP The use of advance organizers in the learning and retention of meaningful verbal material. J Educ Psychol **51**, 267–272 (1960).
22. Lansang MC & Harrell H Knowledge on inpatient diabetes among fourth-year medical students. Diabetes Care **30**, 1088–1091 (2007).
23. Fazel M, Cooley J, Kurdi S & Fazel M A co-curricular diabetes-specific elective with interprofessional students and faculty. Curr Pharm TeachLearn **11**, 172–177 (2019).
24. Zinman B *et al.* Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. N Engl J Med **373**, 2117–2128 (2015).
25. Neal B *et al.* Canagliflozin and cardiovascular and renal events in type 2 diabetes. N Engl J Med **377**, 644–657 (2017).
26. Wiviott SD *et al.* Dapagliflozin and cardiovascular outcomes in type 2 diabetes. N Engl J Med **380**, 347–357 (2019).
27. Marso SP *et al.* Semaglutide and cardiovascular outcomes in patients with type 2 diabetes. N Engl J Med **375**, 1834–1844 (2016).
28. Marso SP *et al.* Liraglutide and cardiovascular outcomes in type 2 diabetes. N Engl J Med **375**, 311–322 (2016).
29. Bodnar TW, Iyengar JJ, Patil PV & Gianchandani RY Can a single interactive seminar durably improve knowledge and confidence of hospital diabetes management? Clin Diabetes Endocrinol **2**, 20 (2016).
30. MacEwen AW, Carty DM, McConnachie A, McKay GA & Boyle JG A "Diabetes Acute Care Day" for medical students increases their

knowledge and confidence of diabetes care: a pilot study. BMC Med Educ **16**, 88 (2016).

31. Meryn S Multimedia communication: quo vadis ? Medical Teacher **20**, 87–90 (1998).

32. Hurtubise L, Martin B, Gilliland A & Mahan J To play or not to play: leveraging video in medical education. J Grad Med Educ **5**, 13–18 (2013).

33. Proctor E *et al.* Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health **38**, 65–76 (2011).

34. Thomas PA, Kern DE, Hughes MT & Chen B *Curriculum Development for Medical Education, A Six-Step Approach.* (Johns Hopkins University Press, 2016).

35. Betihavas V, Bridgman H, Kornhaber R & Cross M The evidence for 'flipping out': a systematic review of the flipped classroom in nursing education. Nurse Educ Today **38**, 15–21 (2016).