

# Ratings of accuracy and insight following dream interpretation by ChatGPT

Mark Blagrove<sup>1</sup>, Laura Roklicer<sup>1</sup>, Jamie Bueno<sup>2</sup>, Oscar Austin<sup>1</sup>, Harriet Foster<sup>1</sup>, and Mea Webb<sup>1</sup>

<sup>1</sup>School of Psychology, Swansea University, UK

<sup>2</sup>School of Psychology, Cardiff University UK

**Summary.** There have been many claims that ChatGPT can provide insightful interpretations of dreams. This study tests whether ChatGPT, if given a participant's report of a recent dream, a report of their recent waking life, and their associations to the dream contents, can achieve outcome ratings by the dreamer comparable to those obtained after discussions of dreams in a lay dream group or with a therapist. The Gains from Dream Interpretation (GDI) questionnaire was used to assess participants' evaluations of the output of ChatGPT regarding their dream. The mean GDI subscale scores were compared to subscale means from previous studies of group and therapist dream interpretations. 30 participants (mean age = 30.72 (SD=13.39)) completed this pre-registered study. Mean GDI insight-exploration subscale score was significantly lower than means for this variable in previous group and therapy studies, and each of the item scores was lower than for those previous studies. However, participants rated the ChatGPT dream interpretations as moderately accurate and as enabling some personal insight. Scores on the exploration-insight, personal insight and interpretation accuracy variables had no significant correlations with word length of the reports inputted to ChatGPT, nor with the individual difference measure Attitude Towards Dreams. As ChatGPT, other generative AI chatbots, and dedicated dream interpretation apps develop, the methods and evaluation questionnaire from the current study are available to assess the value of dream interpretations that the systems produce. This will enable comparisons between different systems and enable longitudinal assessment of individual systems across time and software version releases. .

**Keywords:** Dream, chatbots, ChatGPT, dream interpretation, generative AI

## 1. Introduction

### 1.1. Uses of generative Artificial Intelligence

Generative artificial intelligence, the use of AI to create new content, has had a great impact across numerous sectors, such as industry (Javaid et al., 2023), academia (Rahman & Watanobe, 2023) and healthcare. Modern patient monitoring systems, clinician support tools, therapeutic interventions and treatment outcome predictions are assisted by such generative AI (Colledani et al., 2023; Milne-Ives et al., 2022). This includes interventions for mental health, which has been categorized as a leading cause of global disability (World Health Organization, 2022). The familiar one-to-one form of mental health treatment consultations are quickly becoming inadequate against the recent increase in poor mental health. Utilizing artificial intelligence in mental health care has been suggested as a method to reduce resource constraints, emotional labour and professional burnout of psychotherapists (Wang & Zhang, 2024).

Generative AI has recently been used to enable a new generation of chatbots, which use conversational artificial

intelligence (CAI) and are sophisticated in their understanding of users' questions and intentions. Raile (2024) describes how such modern chatbots can diagnose, give information and advice, provide emotional support with empathy, and assist with skills training. One of these modern chatbots is ChatGPT. Raile asks whether ChatGPT can provide a useful tool for patients and/or psychotherapists and even serve as a substitute for psychotherapists. Such a possibility is supported by findings that AI systems such as ChatGPT possess the ability to process complex human emotions (Cheng et al., 2023; Elyoseph et al., 2023) and can be used to discover new insights from the text of multiple conversations with the same patient (Eshghie & Eshghie, 2023).

For Raile (2024), 'There is no comparable technology as easily accessible as ChatGPT, which is equipped with comprehensive information about mental health problems, real-life answers and can even provide interpretations [of patient material]'. Alanezi (2024) assessed the use of ChatGPT3 for delivering mental health support for at least 15 minutes per day for 14 days to individuals with anxiety, depression or behavioural disorders. More than 50% of the participants stated that ChatGPT was useful for emotional support, self-assessment and monitoring, and psychotherapeutic exercises. Similarly, Siddals et al. (2024) recruited individuals with various mental health or interpersonal issues. In a qualitative analysis responses to ChatGPT showed themes of 'emotional sanctuary', including understanding the user, and 'insightful guidance', especially for relationships and for seeing the other person's perspective in conflict. For many participants there was a positive feeling of 'joy of connection' to the chatbot.

Corresponding address:

Mark Blagrove, School of Psychology, Faculty of Medicine, Health & Life Science, Swansea University, Singleton Park, Swansea, Wales SA2 8PP, UK.

Email: m.t.blagrove@swansea.ac.uk

Submitted for publication: August 2025

Accepted for publication: November 2025

Online first: January, 2026

### 1.2. Generative AI and the interpretation of dreams

There have been many recent accounts of the use of chatbots, including ChatGPT, for dream interpretation (e.g., Beau, 2024; Bulkeley, 2024; Eliot, 2024). In Raile (2024), a patient's dream is described anonymously with his consent, and ChatGPT is asked for an interpretation. Interpretations according to different schools were requested, and Raile concluded that 'the interpretations were not bad, but also not outstanding.... certainly not appropriate for publication in professional journals.' But he concluded that such use of ChatGPT would be a 'good opportunity for people who already have psychotherapy experience to work alone with their material, like dreams or memories. This requires the ability to self-reflect as well as the willingness to deal with discomforting interpretations.' However, although Raile concluded that no comparable technology is as easily accessible as ChatGPT, and that it is equipped with comprehensive information about mental health problems and can give evaluations and real-life answers, he stated there is a limitation in that it fails to ask for more information, such as the biography of the person.

Sheldon (2023) conducted six explorations of the ability of AI to interpret or understand dreams. ChatGPT 3.5 and 4 were used, and were able, in experiment 1, to connect dreams to the dreamers' waking life relationships and situations, and to undertake symbolic analysis. Experiment 2 undertook a Freudian analysis, and experiment 3 a Jungian analysis. A series of dreams from one person was analysed in experiment 4, emphasizing symbolic analysis of the presence of a cat in the dreams. Experiment 5 addressed how characters and interactions in the dream correspond with real-life individuals in the waking life of the dreamer. The interpretation of these correspondences was judged to be 'remarkable', evidencing an understanding by ChatGPT of the continuity hypothesis between waking life and dream content (Domhoff, 2017; Schredl & Hofmann, 2003). Experiment 6 asked ChatGPT to identify changes in dream elements across a series of dreams that mirror changes across time in a wake life relationship. 25 dreams were taken from early in the relationship, and 25 more recently, after the relationship had developed. The changes in dream content identified by ChatGPT were held by Sheldon to show 'a compelling demonstration of the power of generative AI text models.'

In Laureano and Calvo (2024), 22 patients had their dreams analyzed by professional psychoanalysts and ChatGPT, producing a total of 81 analyses, 20 of which were performed by ChatGPT. Responses from ChatGPT were clearly distinguishable from the responses of the human analysts, particularly in the semantic and grammatical categories. Despite these differences, the authors concluded that ChatGPT possesses the ability to produce coherent interpretations of dreams. It is thus plausible that ChatGPT dream interpretations might have some usefulness and some validity. This is especially because ChatGPT can deal with metaphorical language (Puraivan et al., 2024) and because Large Language Models, a form of generative AI, can assess the emotional content of dreams (Bertolini et al., 2024) and assess the match between dream and experimental stimulus content (Mossbridge et al., 2025).

### 1.3. Assessing the interpretation of dreams by ChatGPT

Although the feasibility of this use of ChatGPT for dream interpretation has been shown, to our knowledge the validity of those interpretations has not yet been assessed quantitatively. We therefore designed a study that would provide input to ChatGPT similar to that provided during the Ullman group dream discussion technique (Ullman, 1996). In the Ullman technique the dream-sharer reports a dream, reports recent waking life experiences, events and concerns, whether ostensibly related to the dream or not, and reports free associations to elements of the dream. These are discussed with the group in a series of stages, which culminate in a final 'orchestration' stage in which the information gleaned about the dream and about waking life is compared so as to identify metaphorical mappings between the dream and waking life. In the study proposed here the dreamer would compose the dream, waking life and associations reports after having had a dream, the three reports would then be combined into one document and inputted into ChatGPT. The dreamer would then read and consider the output that ChatGPT gives as its interpretation of the dream. The aim of the study was to compare the dreamers' ratings of the ChatGPT output with ratings found in dream group studies using the Ullman technique, and ratings for discussions of dreams with therapists who use Clara Hill's (1996) cognitive experiential model of dream interpretation.

The data for assessing group interpretations of dreams are taken from Edwards et al. (2013, 2015) and Blagrove et al. (2019). In these studies, dream discussion sessions used the Ullman (1996) dream appreciation technique. Edwards et al. (2013) also reports a metanalysis of the outcomes of studies by Clara Hill and colleagues on dream interpretation using the well-established Hill (1996) therapist-led dream interpretation method. For all these studies outcomes of the group and therapist sessions were assessed using the Gains from Dream Interpretation questionnaire (GDI: Heaton et al., 1998). Edwards et al. (2013) found that exploration-insight gains from Ullman dream groups were comparable to gains found for therapist-run sessions in the many studies by Hill and colleagues. Edwards et al. (2015) found that exploration-insight gains and personal insight gains are higher for discussing dreams than for discussing a recent wake life event, where both discussions use the Ullman technique. Blagrove et al. (2019) found that exploration-insight gains are higher for discussing dreams than for discussing a day-dream. The current study aimed to replicate the method first detailed in Edwards et al. (2013), but using ChatGPT for the dream 'discussion', and to compare exploration-insight and personal insight scores obtained with those reported in the studies cited above. As anecdotal reports and magazine and journal publications are supportive of claims for the abilities of ChatGPT for dream interpretation it was hypothesized that scores for exploration-insight and personal insight from the ChatGPT interpretations will be comparable to those reported in the literature for group and therapy dream discussions. The study also aimed to replicate Edwards et al.'s (2013) finding that positive attitude towards dreams (ATD) is predictive of exploration-insight score after dream discussion, positive ATD and dream interpretation session outcome also having been found to be significantly correlated by Zack and Hill (1998).

## 1.4. Hypotheses

1. Gains from Dream Interpretation questionnaire (GDI; Heaton et al., 1998) exploration-insight subscale score will not differ significantly from mean scores for this variable in studies by Hill (meta-analysis in Edwards et al., 2013), Edwards et al. (2013, 2015), and Blagrove et al. (2019).
2. Personal insight subscale mean score will not differ significantly from the mean score on this subscale in Edwards et al. (2015).
3. GDI Experiential gain subscale mean score may be lower or comparable to means of this subscale in the Hill, Edwards et al. (2023, 2015) and Blagrove et al. (2019) studies; there is thus no prediction here.
4. GDI Action gain subscale mean score may be lower or comparable to means of this subscale in the studies by Hill, Edwards (2013, 2015), and Blagrove et al. (2019); there is thus no prediction here.
5. The Own Attitude Towards Dreams subscale score will be significantly correlated with exploration-insight subscale score as measured by the GDI. No predictions are made for correlations between the other ATD and GDI subscale scores.

## 2. Method

The hypotheses and method were pre-registered on AsPredicted.org on 28th January 2025, 07:32 AM (PT). Link to pre-registration: <https://aspredicted.org/cq74-3d46.pdf>.

### 2.1. Participants

30 participants completed the ChatGPT protocol and answered the GDI questionnaire. The sample size of 30 had been pre-registered. Demographic and ATD questionnaire data were only present for 29 participants. For the 29 participants, males  $n=8$ , females  $n=21$ , mean age = 30.72 ( $SD=13.39$ ). Participants gave informed consent to take part in the study, and the study was approved by the Research Ethics Committee of the School of Psychology, Swansea University. Exclusion criteria were the prospective participant reporting that they recall dreams less than once per week, or reporting having had a trauma in their life.

### 2.2. Materials

The Attitude Towards Dreams questionnaire (ATD; Cernovsky, 1984). This has 16 items which are each scored as true or false. The questionnaire comprises three subscales: (1) own attitude towards dreams, (2) perceptions of attitudes of significant others towards dreams, and (3) perceptions of attitudes of other people in general towards dreams or towards individuals who publicly discuss dreams.

The Gains from Dream Interpretation questionnaire (GDI; Heaton et al., 1998). This has 14 items and uses a 9-point scale for each item, with descriptors 1 = 'strongly disagree' and 9 = 'strongly agree'. The questionnaire comprises three subscales: (1) exploration-insight gains, (2) experiential gains, and (3) action gains. Exploration-insight gains refer to thoroughness of the dream discussion and levels of insight about the relationship of the dream to waking life; experiential gains refer to level of re-experiencing the dream during the discussion; action gains refer to changes that the dreamer intends to make as a result of the dream discussion. Following Edwards et al. (2015) we constituted a fur-

ther personal insight subscale comprising items 5, 6, 7, 10, and 11 of the GDI (see Appendix for the full questionnaire used in the current study).

The following rewordings of the original GDI items demonstrate how items were amended to evaluate the ChatGPT interpretation of the dream:

Original GDI items:

1. I was able to explore my dream thoroughly during the session.
5. I got ideas during the session for how to change some aspect(s) of myself or my life.
11. I will use the things that I learned in this dream interpretation in my life.

Amended items:

1. My dream has been explored thoroughly by the ChatGPT interpretation.
5. I got ideas from the ChatGPT interpretation for how to change some aspect(s) of myself or my life.
11. I will use things that I learned in this ChatGPT dream interpretation in my life.

The original GDI experiential subscale has two items:

- 3) During the session I was able to re-experience the feelings I had in my dream; and
- 9) I felt like I was actually reliving the dream during the session.

These were amended as follows for the current study:

- 3) By typing out my dream and inputting it to ChatGPT and considering the interpretation I was able to re-experience the feelings I had in the dream.
- 9) I felt like I was actually reliving the dream by reporting it, inputting it and by considering the ChatGPT interpretation.

Following the 14 items of the amended GDI, a final question was presented: 'Overall, how accurate do you find the ChatGPT interpretation of your dream?' This was responded to on a 9-point scale, with descriptors 1 = 'extremely inaccurate' and 9 = 'extremely accurate'.

### 2.3. Procedure

Prospective participants accessed the study's first online questionnaire and read instructions there about what taking part in the study would involve. They were asked to report their frequency of recalling dreams and to answer yes or no to a question of whether they had ever had a trauma. If a participant reported recalling dreams less than once per week, or responded yes to the trauma question, they were thanked for their answers but excluded from the study. Those who passed the two screening questions were then asked to give consent to take part in the study. After giving informed consent participants first completed the ATD questionnaire and reported their sex and age. They then slept at home as normal and were asked to make a typed or spoken report of the next dream they have. If the dream report was at least 30 words long they were asked to make written reports of their dream, their recent waking life and experiences, and their associations to the contents of the dream. They combined these three accounts into a single document and submitted the document to a second online questionnaire and to ChatGPT. They then read and considered the ChatGPT interpretation of the dream, uploaded the ChatGPT interpretation to the second online questionnaire,



and then completed online the GDI questionnaire and final interpretation accuracy question. Participants had been instructed to complete the ChatGPT protocol and GDI questionnaire within 48 hours of having their dream: median and mean time duration between dream occurrence and protocol completion for the 30 participants did fall within the 24 – 48 hour period.

## 2.4. Verbatim instructions

The verbatim instructions given to participants regarding the second questionnaire were as follows:

*We are including below the link to the second questionnaire. Before completing the second questionnaire, it is important that you wait until you have a dream that you can recall in as much detail as possible, and which takes at least 30 words to report. Please read the following information carefully about the second questionnaire.*

### Instructions for writing the Word document that you will put into ChatGPT

*1. When you wake up and have been dreaming please try to remember the dream in as much detail as possible. If you can you might voice record your dream report, or write it down soon after you wake. Describe the dream exactly and as fully as you remember it. Your report should contain, whenever possible: a description of the setting of the dream, whether it was familiar to you or not; a description of the people, their age, gender, and relationship to you; and any animals or colours that appeared in the dream. If possible, describe your feelings during the dream and whether it was pleasant or unpleasant. Be sure to tell exactly what happened during the dream to you and the other characters.*

*2. Next, type out your dream report into a Word document. Your dream report needs to be at least 30 words long for it to be used, so if it is shorter than that please wait for further nights/mornings until another dream occurs. In your Word document before the dream paragraph type 'My dream was:'*

*3. Write down in the Word document underneath the dream an account of what you were doing on the evening and day before the dream, including any conversations or incidents or positive or negative experiences you had. State any concerns and emotions you have been having, and anything that has been on your mind. Then write about the days before that, including what have been your negative or positive experiences in those days. In your Word document before the recent waking life paragraph type: 'My recent waking life includes:'*

*4. Next write down anything from your waking life that you associate with parts of the dream. State anything in your waking life that your dream or parts of your dream might be referring to or might be connected to. So if, say, a person or place you know appears in your dream you could say here anything about that person or place, such as what they are like, or when you last saw the person or place. In your Word document before this associations paragraph type: 'My associations to parts of the dream are:'*

*5. You should now have three paragraphs of at least 150 words in total, one paragraph about your dream, one*

*paragraph about your recent waking life, and one paragraph about what associations you can give to parts of the dream, that is, what parts of the dream might refer to or mean in your waking life. The more detail in these three paragraphs the better!*

*6. At the bottom of the three paragraphs type 'What does my dream mean?'*

*7. You will copy and paste this text into the second questionnaire and into ChatGPT.*

*8. You should complete questionnaire 2 within 48 hours of having the dream to ensure that you remember in detail your dream and your recent waking life experiences that occurred before the dream.*

*There will be several steps that you will be asked to follow that will be clearly stated in questionnaire 2, such as following a link to ChatGPT and copy and pasting your paragraphs on your dream and waking life information and associations. ChatGPT will then do an interpretation of the dream which you will read. You will then copy and paste that interpretation back into the questionnaire. You will then be asked questions about the interpretation.*

After reading the above instructions participants slept at home and followed the instructions after they had had a dream the report for which was at least 30 words long. Participants also reported the version of ChatGPT they were using. Responses were: 4o mini, n=6; 4o, n=17; o1-mini, n=0; o1, n=1; version not reported, n=6. Participants were not instructed to use GPT-4-class models as this may not have been available for some.

## 2.5. Statistical analysis

**Outliers/exclusions:** Data are excluded if a dream report is less than 30 words, or if the total length of the three paragraphs (dream report, wake life report, associations to dream items report) is less than 150 words; these exclusion criteria were pre-registered.

### Analyses:

1. Between groups t-tests to compare GDI subscale means from this study with GDI subscale means from Hill (from meta-analysis in Edwards et al., 2013), Edwards et al. (2013, 2015) and Blagrove et al. (2019). The only hypotheses are for the exploration-insight and personal insight subscale mean scores not being significantly different from the mean scores in the previous literature. Where inferential statistics show a clear lack of significant difference between current and previous study means, the Two One-Sided Tests (TOST) procedure is used to test for equivalence of the means.
2. Correlations between outcome variables (GDI subscale and interpretation accuracy scores) and ATD subscales and ChatGPT procedure variables (e.g., word length of dream). The only prediction was for a significant correlation between exploration-insight subscale score and own ATD score.
3. Male / female differences are assessed for the GDI and ATD subscales, interpretation accuracy and ChatGPT procedure variables in case sex needs to be partialled out of the correlations.

**Table 1.** Means and standard deviations for Gains from Dream Interpretation subscales for the current study (n=30) and previous studies, with Welch's (2-tail) t-test comparisons between the current and previous studies. Mean of dream interpretation accuracy rating is presented also.

Variable	Study	Mean	SD	t	df	p
Exploration insight gains	Current Study (n = 30)	6.86	1.46			
	Edwards et al. 2013 (n = 11)	8.17	0.54	4.19	38	<.001
	Edwards et al. 2015 (n = 11)	7.82	0.84	2.61	31	.014
	Blagrove et al. 2019 (n = 30)	7.60	0.88	2.39	47	.022
	All Hill Studies (n = 437)	7.40	1.15	1.98	31	.056
Personal insight gains <sup>a</sup>	Current Study (n = 30)	6.13	2.03			
	Edwards et al. 2015 (n = 11)	6.60	1.43	.83	25	.416
Experiential gains	Current Study (n = 30)	6.24	2.24			
	Edwards et al. 2013 (n = 11)	7.28	1.94	1.46	20	.160
	All Hill Studies (n = 437)	7.03	1.56	1.90	30	.066
Action gains	Current Study (n = 30)	5.48	1.98			
	Edwards et al. 2013 (n = 11)	5.78	1.04	.63	33	.536
	All Hill Studies (n = 437)	6.51	1.34	2.81	30	.008
Accuracy of dream interpretation <sup>b</sup>	Current study (n = 30)	7.33	1.09			

Notes: <sup>a</sup>Two One-Sided Tests equivalence testing. The lower and upper bounds for the procedure were set as for a small effect size, this being 0.2 x pooled SD = 0.38. For lower bound one-sided t-test,  $t(21)=0.158$ ,  $p=.562$ . For upper bound one-sided t-test,  $t(21)= -1.495$ ,  $p=.075$ .

<sup>b</sup>Accuracy of dream interpretation ratings in answer to question 'Overall, how accurate do you find the ChatGPT interpretation of your dream?', ratings from 1 = extremely inaccurate to 9 = extremely accurate. Frequencies for responses: response=4, n=1, 3.33%; response=6, n=4, 13.33%; response=7, n=12, 40.00%; response=8, n=9, 30%; response=9, n=4, 13.33%.

### 3. Results

Table 1 shows means and standard deviations for the Gains from Dream Interpretation subscales for the current study and previous studies. ChatGPT derived exploration-insight is significantly lower than group- or therapy-derived exploration-insight for all previous studies. Personal insight is not significantly different from the group method reported by Edwards et al. (2015). To test this similarity of personal insight scores the Two One-Sided Tests (TOST) procedure was run. On this procedure only one of the null hypoth-

eses was rejected, showing that the two scores were not equivalent as they were not both within the upper and lower bounds (details in Table 1 note a). ChatGPT may thus not be as equally effective as the group Ullman method for personal insight. Finally, the accuracy of ChatGPT dream interpretation mean score suggests that the interpretations on average were rated as moderately accurate.

Table 2 shows exploration-insight subscale and item scores for the current study and for group dream discussions in Blagrove et al. (2019, Table 3). For the subscale

**Table 2.** Exploration-insight subscale and item scores for the current ChatGPT study (n=30) and for Blagrove et al. (2019, n=30) group discussion study, with between-subjects Welch's t-test (2-tail) comparisons.

Category	Group		ChatGPT		t	df	p
	Mean	SD	Mean	SD			
Exploration-insight subscale	7.60	0.88	6.86	1.46	2.38	47	.022
1. My dream has been explored thoroughly by the ChatGPT interpretation	8.35	0.97	7.90	1.03	1.75	58	.084
2. Because of the ChatGPT interpretation I have learned more about what this dream means for me personally	8.00	0.86	7.40	1.77	1.67	41	.102
6. I learned more from the ChatGPT interpretation about how past events influence my present behaviour.	6.57	1.79	6.23	2.32	0.64	54	.526
7. I learned more about issues in my waking life from the ChatGPT interpretation.	6.95	1.74	6.53	2.05	0.86	56	.392
8. I felt like I was very involved in working with the dream by reporting it, inputting it and by considering the ChatGPT interpretation.	7.97	1.28	6.97	2.44	1.99	43	.052
12. I learned things that I would not have thought of on my own.	7.53	1.22	6.47	2.06	2.44	46	.018
13. I was able to make some connections between my dream and issues in my waking life that I had not previously considered.	7.85	0.89	6.50	2.21	3.11	37	.004

Table 3. Descriptive analyses for ChatGPT procedure variables (word length of dreamer reports and word length of ChatGPT output) and Attitude Towards Dreams subscale variables.

Variable	Median	IQR <sup>a</sup>	Min	Max
Word length of dream report	178	219	77	546
Word length of wake life context report	107	104	30	395
Word length of associations report	95	106	31	377
Word length of the 3 combined reports	414	435	157	1063
Word length of ChatGPT output	584	185	350	997
	Mean	SD	Min	Max
Own ATD	5.03	1.76	1	7
Significant others ATD	3.66	1.42	0	5
Society ATD	3.62	0.98	0	4

Note. <sup>a</sup>IQR = Interquartile Range.

mean and for all items ChatGPT scored lower than did group discussions.

Table 3 shows descriptive analyses for the ChatGPT procedure variables (median, interquartile range, minimum, maximum) and Attitude Towards Dreams subscale variables (mean, standard deviation, minimum, maximum).

Table 4 shows the correlations between the procedure/individual difference variables (ChatGPT procedure, ATD, demographic variables) and the outcome variables (exploration-insight, personal insight, accuracy of dream interpretation). The correlations with sex use Pearson's bi-serial correlation. All others are Spearman's correlations because all ATD subscales, age and report length variables were not normally distributed (Shapiro-Wilk tests,  $dfs=29$  for age and ATD subscales,  $dfs=30$  for report lengths, all Shapiro-Wilk test statistics for these variables  $< .910$ , all  $ps < .012$ ). To correct for multiple correlations  $p < .01$  is used as the criterion for significance. The only significant relationships were between age and the exploration-insight and interpretation accuracy variables. There were no male / female differences for the GDI and ATD subscales, nor for interpretation accuracy or for ChatGPT procedure variables, and so sex is not partialled out for the correlations in Table 4.

## 4. Discussion

### 4.1. Testing of hypotheses

Participants rated the accuracy of the ChatGPT dream interpretations predominantly as 7, 8 or 9 on the 9-point scale, equivalent to moderately or very accurate. However, caution is needed for this result because of the Barnum or Forer effect (Forer, 1949), such as can occur when people consider or evaluate horoscopes. This effect is a tendency for people to rate a narrative as being an accurate description of themselves, even though most people will also rate the narrative as an accurate description, thus showing the ability of different people to flexibly map the same narrative onto their lives, and thus spuriously ascribe accuracy to the narrative as an accurate description of the self. The Barnum or Forer effect might have inflated the interpretation accuracy rating and other GDI item ratings in the current study, and might even have affected ratings in the previous in-person comparison studies referred to here.

In the current study, personal insight following the ChatGPT protocol was comparable to that obtained in Edwards et al. (2015) using an Ullman group discussion, although not statistically equivalent on the TOTS procedure. The explo-

Table 4. Correlations between outcome (exploration-insight, personal insight, accuracy of dream interpretation) and procedural/individual difference variables (word length of dream report, word length of wake life report, word length of free associations, total word length of reports, word length of ChatGPT output, ATD subscales, sex and age).

Variable	Exploration-insight	Personal insight	Interpretation accuracy
Word length of dream report	.09	.08	-.02
Word length of wake life report	.25	.25	.27
Word length of free associations	.28	.27	.31
Word length of the 3 reports	.22	.22	.16
Word length of ChatGPT output	.10	.14	-.01
ATD Own	.15	.15	.43
ATD Others	-.15	-.07	.32
ATD Social	-.40	-.26	-.23
Sex	.12	.32	.27
Age	.51*	.44	.53*

Notes. All correlations are Spearman's rho except for point-biserial Pearson correlation with sex. All correlations with report word lengths have  $n=30$ , all other correlations have  $n=29$  (Data for ATD, age and sex were lost for one participant), \*  $p < .01$ , correlations are 2-tailed.

ration-insight subscale score was significantly lower than in previous studies on group- and therapy-based discussions. This subscale combines items regarding thoroughness of exploring and learning about the dream with items that target insights about the relationship of the dreams to waking life: all items had a lower score for the current study compared to previous group- and therapy-based studies.

The current study assessed whether pre-held positive or negative attitudes towards dreams are related to the level of understanding or insight achieved following ChatGPT's interpretation of their dream. Exploration-insight was not related to Own ATD, which may be because there was no active discussion which would mediate the effects of ATD. The absence of a correlation indicates that, unlike for dream group discussions, attitude towards dreams is not related to exploration-insight following ChatGPT dream interpretation. This is a positive finding for ChatGPT, since those who have little interest in dreams would be still able to achieve some insight about their dreams by the use of ChatGPT.

#### 4.2. Comparing use of ChatGPT to dream group sessions and psychotherapy

The environment in which people interact with ChatGPT is very different from that of a discussion with one or more other people. As discussed by Sharma et al. (2023), ChatGPT provides a space where insight into troubling events can be achieved without the need, but having the option to, discuss personal issues. ChatGPT provides a confidential environment to subject one's dreams to somewhat objective and non-judgmental interpretation. The absence of a therapist does, however, prevent a relationship forming between client and therapist, which can be one of the benefits of dream-sharing with a therapist (Pesant & Zadra, 2004). Furthermore, ChatGPT has not historically had the capacity for long-term memory due to technological limitations and the processing power required to provide long-term memory to the 400 million weekly users (Reuters, 2025), but subscription accounts in some countries have now achieved this capability (Weatherbed, 2025).

Ethical and safeguarding challenges may arise when applying ChatGPT to the dreams of vulnerable individuals, such as individuals with post-traumatic stress disorder, or with mental disorders that can lead to nightmares or night terrors. For the current study, people who had experienced a traumatic event were excluded for ethical reasons. Without professional or group supervision, it may be difficult to avoid triggers or implement sufficient safety measures. ChatGPT could display sensitivity in questioning the dreamer so that these risks are ameliorated, but the lack of being able to monitor non-verbal signs of distress, such as crying, does mean that there are greater risks in this area for ChatGPT in contrast to interactions with humans.

The Ullman (1996) technique has the advantage of extended recall and discussion of the dream and of the life circumstances of the dreamer, but requires the group to be assembled, whereas ChatGPT can provide a fast interpretation, if the dream is entered into the program soon after waking. Such immediacy is a very positive feature of this use of ChatGPT. This technological method may result in greater public interest in dreams, as well as providing material and thoughtful suggestions for examining the dream later in a group or in therapy.

#### 4.3. Limitations

Participants were not supervised in undertaking the study and so although some components were obligatory and checked (such as length of reports submitted) participant engagement with the ChatGPT output might not have been thorough for all participants. However, the mean rating of GDI item 1 ('My dream has been explored thoroughly by the ChatGPT interpretation') was 7.90 on the 1 – 9 scale, indicating that participants saw the process as very thorough. As in the Ullman dream group studies reviewed above, participants were not selecting which of their dreams to examine in this study. Participants were asked to work on the first dream they had following completion of the first questionnaire, as long as its report was at least 30 words in length. This means that some dreams addressed in the current study may have had obvious meanings with not much room for AI to make further interpretations, or might have been very mundane, whereas some cryptic or more abstract dreams might have been explored much more with AI and thus produce new insights, and especially for people who have little experience of examining their dreams. Future research could allow participants to select which dream they would like to work on with ChatGPT.

Although the comparisons used here of the current study data with earlier Ullman-group and Hill therapist-method datasets are informative, there is a limitation in this cross-study statistical testing due to potential differences in sample characteristics, procedural details, and contextual factors. This limitation may be reduced in future research by having single omnibus studies where participants are randomly allocated to therapist, Ullman lay-group and ChatGPT conditions.

A final limitation is that the procedure did not amount to a discussion as there was no back and forth with ChatGPT, the design was for a single input and single output. The study was designed in this way so as to reduce random and systematic confounds that may have occurred if participants had been allowed to engage in further conversation with ChatGPT. This does, however, reduce the ecological validity of the study, as real-world AI/Large Language Model use typically involves iterative clarification, follow-up prompts, and refinement. Such multiple exchanges could meaningfully affect interpretation depth, accuracy, and perceived insight. Future studies should incorporate extended conversation with ChatGPT, but controls or measurement of individual differences in such conversations will be required. In favour of the current restricted design, it should be noted that the Ullman procedure is itself highly structured, with the telling of the dream, telling of recent waking life experiences, and associations to the dream, all occurring in stages, which the method of the current study aimed to emulate. The final 'orchestration' stage of the Ullman procedure is, however, a bringing together of the information obtained in the previous stages, and future research could aim to emulate this through an extended conversation with ChatGPT. This suggestion for extended conversation is addressed in the next, Future research section.

#### 4.4. Future research

A protocol with extended conversation should be studied, in which further requests are made of ChatGPT, such as 'Ask me questions so as to make your dream interpretation more accurate.' Such conversations are already pos-



sible with ChatGPT, with requests such as 'now provide a Freudian interpretation of the dream', or 'now provide a Jungian interpretation of the dream.' An extended dialogue with ChatGPT, designed to address experiential-insight and personal insight, may result in greater gains from dream interpretation for these subscales. Also, following the initial ChatGPT dream interpretation, a request to ChatGPT from the dreamer for how knowledge from this dream may be applied to one's life may result in greater action gains. Likewise, greater experiential gains may be achieved by extending the ChatGPT interaction to discuss finer details of the dream, such as for the feelings and sensations experienced during it.

There have been attempts to address the extent to which the benefits of the discussion of dreams are dependent on the dream report itself in contrast to the discussion. For example, Hill et al. (1993) compared outcomes of discussion of one's own dream with discussion of the dream of another person as if it were one's own dream. The own dream condition resulted in more insight than the others' dream condition. However, in therapy and dream discussion groups, and experiments on these, any own dream/others' dream conditions would be confounded by cultural and personal beliefs about dreams (Morewedge & Norton, 2009), such that knowledge that a dream is one's own, or not one's own, may well affect the discussion engagement and outcome. Using ChatGPT for the discussion will diminish this confound if ChatGPT is not told the provenance of the dream. The current protocol will thus enable studies, similar to Hill et al. (1993), that compare gains from the interpretation of one's own dream versus an interpretation of someone else's dream without the confound of knowledge of the provenance of the dream.

Although comparing the outcomes of interpretation of one's own versus others' dreams is practicable and conceptually important, in general the current study design does not allow conclusions about underlying mechanisms for gains from dream interpretation. For example, positive effects may arise from the content of ChatGPT's interpretations, the act of writing and structuring one's dream, or the reflective engagement required by the protocol. Future research should tease apart the relative contributions of these components. This could be done by having conditions of dream report alone, dream with wake life context, dream with wake life context and associations (as in the current study), one's own dream versus others' dream (as stated above), and different levels of reflective engagement, with all these conditions being compared on GDI measures. Such conditions would help to assess the relationship of outcomes to the different components of the protocol. Disentangling these different components could also be useful in explaining the intriguing finding in the current study that the exploration-insight and interpretation accuracy variables had significant correlations with age.

A further extension of the use of AI is the production of visual images of the dream, as explored by Bulkeley (2025). Such images would enable the dreamer to return to the dream at future times, and act as a cue for discussing the dream with other people. Such a benefit of having a visual depiction of a dream is described by Lockheart (2024) for dream paintings she produces while the dream is being discussed.

#### 4.5. Standardized method and questionnaire

As ChatGPT and other AI chatbots develop, the merit and utility of the dream interpretations they produce will doubtless increase, and also be extended to series of dreams. This massively expanding corpus of dream narratives and real-life information and data will provide further training in this regard for generative AI. The assessment methods, protocol and amended GDI questionnaire (see Appendix) described here can be used to assess future improvements in generic ChatGPT and dedicated apps in their capabilities for dream interpretation. It is expected that therapists, research teams and interested members of the public will develop novel and various ways to have extended conversations with AI about the interpretation of dreams. Nevertheless, having a controlled, standardised method and materials for assessing dream interpretations will be of benefit for quantifying differences in capabilities between different AI systems, and improvements in AI capabilities for dream interpretation across time and across software release versions.

#### References

- Alanezi F. (2024). Assessing the Effectiveness of ChatGPT in Delivering Mental Health Support: A Qualitative Study. *Journal of Multidisciplinary Healthcare*, 17, 461–471. <https://doi.org/10.2147/JMDH.S447368>
- Beau, J-B. (2024). A.I. dream analysis -- Can A.I. tools help us understand our dreams. *International Journal of Dream Research*, 17 (Supplement 1), S32. <https://journals.ub.uni-heidelberg.de/index.php/IJoDR/issue/view/7081/1231>
- Bertolini, L., Elce, V., Michalak, A., Widhoelzl, H. S., Bernardi, G., & Weeds, J. (2024). Automatic annotation of dream report's emotional content with large language models. In *CLPsych 2024-9th Workshop on Computational Linguistics and Clinical Psychology, Proceedings of the Workshop* (pp. 92-107). Association for Computational Linguistics (ACL). <https://aclanthology.org/2024.clinpsych-1.7/>
- Blagrove, M., Edwards, C., van Rijn, E., Reid, A., Malinowski, J., Bennett, P., Carr, M., Eichenlaub, J.-B., McGee, S., Evans, K., & Ruby, P. (2019). Insight from the consideration of REM dreams, non-REM dreams, and daydreams. *Psychology of Consciousness: Theory, Research, and Practice*, 6, 138–162. <https://doi.org/10.1037/cns0000167>
- Bulkeley, K. (2024). Key principles for interpreting dreams with AI. *Psychology Today*, March 10, 2024. <https://www.psychologytoday.com/us/blog/dreaming-in-the-digital-age/202403/key-principles-for-interpreting-dreams-with-ai>
- Bulkeley, K. (2025). Using AI Imagery to Interpret Dreams. *Psychology Today*, April 30, 2025. <https://www.psychologytoday.com/us/blog/dreaming-in-the-digital-age/202504/using-ai-imagery-to-interpret-dreams>
- Cernovsky, Z. Z. (1984). Dream recall and attitude towards dreams. *Perceptual and Motor Skills*, 58, 911–914. <https://doi.org/10.2466/pms.1984.58.3.911>
- Cheng, S. W., Chang, C. W., Chang, W. J., Wang, H. W., Liang, C. S., Kishimoto, T., ... & Su, K. P. (2023). The now and future of ChatGPT and GPT in psychiatry. *Psychiatry and Clinical Neurosciences*, 77, 592-596. <https://doi.org/10.1111/pcn.13588>
- Colledani, D., Anselmi, P., & Robusto, E. (2023). Machine learning-decision tree classifiers in psychiatric assessment:



- An application to the diagnosis of major depressive disorder. *Psychiatry Research*, 322, 115127. <https://doi.org/10.1016/j.psychres.2023.115127>
- Domhoff, G. W. (2017). The invasion of the concept snatchers: The origins, distortions, and future of the continuity hypothesis. *Dreaming*, 27, 14-39. <https://doi.org/10.1037/drm0000047>
- Edwards, C.L., Malinowski, J.E., McGee, S.L., Bennett, P.D., Ruby, P.M., & Blagrove, M.T. (2015). Comparing personal insight gains due to consideration of a recent dream and consideration of a recent event using the Ullman and Schredl dream group methods. *Frontiers in Psychology*, 6, 831. <https://doi.org/10.3389/fpsyg.2015.00831>
- Edwards, C.L., Ruby, P.M., Malinowski, J.E., Bennett, P.D., & Blagrove MT (2013). Dreaming and insight. *Frontiers in Psychology*, 4, 979. <https://doi.org/10.3389/fpsyg.2013.00979>
- Eliot, L. (2024). Using generative AI to analyze your sleeping dreams and reveal hidden secrets about yourself. *Forbes*, February 1, 2024. <https://www.forbes.com/sites/lanceeliot/2024/02/01/using-generative-ai-to-analyze-your-sleeping-dreams-and-reveal-hidden-secrets-about-yourself/>
- Elyoseph, Z., Hadar-Shoval, D., Asraf, K., & Lvovsky, M. (2023). ChatGPT outperforms humans in emotional awareness evaluations. *Frontiers in Psychology*, 14, 1199058. <https://doi.org/10.3389/fpsyg.2023.1199058>
- Eshghie M., & Eshghie, M. (2023). ChatGPT as a therapy assistant: a suitability study. *arXiv:2304.09873*. <https://doi.org/10.48550/arXiv.2304.09873>
- Forer, B. R. (1949). The fallacy of personal validation: a classroom demonstration of gullibility. *Journal of Abnormal and Social Psychology*, 44, 118-123. <https://doi.org/10.1037/h0059240>
- Heaton, K. J., Hill, C. E., Petersen, D. A., Rochlen, A. B., & Zack, J. S. (1998). A comparison of therapist-facilitated and self-guided dream interpretation sessions. *Journal of Counselling Psychology*, 45, 115-122. <https://doi.org/10.1037/0022-0167.45.1.115>
- Hill, C. E. (1996). *Working with dreams in psychotherapy*. New York: Guilford Press.
- Hill, C. E., Diemer, R., Hess, S., Hillyer, A., & Seeman, R. (1993). Are the effects of dream interpretation on session quality, insight, and emotions due to the dream itself, to projection, or to the interpretation process? *Dreaming*, 3, 269-280. <https://doi.org/10.1037/h0094385>
- Javaid, M., Haleem, A., & Singh, R. P. (2023). A study on ChatGPT for Industry 4.0: Background, potentials, challenges, and eventualities. *Journal of Economy and Technology*, 1, 127-143. <https://doi.org/10.1016/j.ject.2023.08.001>
- Laureano, M.H., & Calvo, H. (2024). "Computational Study of Dream Interpretations: Psychoanalytic Human vs Artificial Analyses," 2024 IEEE Congress on Evolutionary Computation (CEC), Yokohama, Japan, 2024, pp. 1-9. <https://doi.org/10.1109/CEC60901.2024.10612036>
- Lockheart, J. (2024). Painting and socializing COVID-19 dreams. *International Journal of Surrealism*, 1, 87-108. <https://doi.org/10.1353/ij.s,2024.a922369>
- Milne-Ives, M., Selby, E., Inkster, B., Lam, C., & Meinert, E. (2022). Artificial intelligence and machine learning in mobile apps for mental health: A scoping review. *PLOS Digital Health*, 1, e0000079. <https://doi.org/10.1371/journal.pdig.0000079>
- Morewedge, C. K., & Norton, M. I. (2009). When dreaming is believing: The (motivated) interpretation of dreams. *Journal of Personality and Social Psychology*, 96, 249-264. <https://doi.org/10.1037/a0013264>
- Mossbridge, J. A., Green, D., French, C. C., Pickering, A., & Abraham, D. (2025). Future dreams of electric sheep: Case study of a possibly precognitive lucid dreamer with AI scoring. *International Journal of Dream Research*, 18, 151-168. <https://doi.org/10.11588/ijodr.2025.2.108750>
- Pesant, N., & Zadra, A. (2004). Working with dreams in therapy: what do we know and what should we do? *Clinical Psychology Review*, 24, 489-512. <https://doi.org/10.1016/j.cpr.2004.05.002>
- Puraivan, E., Renau, I. & Riquelme, N. (2024). Metaphor identification and interpretation in Corpora with ChatGPT. *SN Computer Science*, 5, 976. <https://doi.org/10.1007/s42979-024-03331-0>
- Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences*, 13, 5783. <https://doi.org/10.3390/app13095783>
- Raile, P. (2024). The usefulness of ChatGPT for psychotherapists and patients. *Humanities and Social Sciences Communications*, 11, 47. <https://doi.org/10.1057/s41599-023-02567-0>
- Reuters. (2025). OpenAI's weekly active users surpass 400 million. <https://www.reuters.com/technology/artificial-intelligence/openai-weekly-active-users-surpass-400-million-2025-02-20/>
- Schredl, M., & Hofmann, F. (2003). Continuity between waking activities and dream activities. *Consciousness and Cognition*, 12, 298-308. [https://doi.org/10.1016/s1053-8100\(02\)00072-7](https://doi.org/10.1016/s1053-8100(02)00072-7)
- Sharma, A., Lin, I.W., Miner, A.S., Atkins, D.C., & Althoff, T. (2023). Human-AI collaboration enables more empathic conversations in text-based peer-to-peer mental health support. *Nature Machine Intelligence*, 5, 46-57. <https://doi.org/10.1038/s42256-022-00593-2>
- Sheldon, J. (2023). Dreaming with AI. *Poligrafi*, 28, 159-194. <http://doi.org/10.35469/poligrafi.2023.418>
- Siddals, S., Torous, J., & Coxon, A. (2024). "It happened to be the perfect thing": experiences of generative AI chatbots for mental health. *npj Mental Health Research*, 3, 48. <https://doi.org/10.1038/s44184-024-00097-4>
- Ullman, M. (1996). *Appreciating Dreams: A Group Approach*. Thousand Oaks, CA: Sage.
- Wang, D., & Zhang, S. (2024). Large language models in medical and healthcare fields: applications, advances, and challenges. *Artificial Intelligence Review*, 57, 299. <https://doi.org/10.1007/s10462-024-10921-0>
- Weatherbed, J. (2025). ChatGPT will now remember your old conversations. *The Verge*, April 11, 2025. <https://www.theverge.com/news/646968/openai-chatgpt-long-term-memory-upgrade>
- World Health Organization. (2022). *World mental health report: Transforming mental health for all*. World Health Organization.
- Zack, J. S., & Hill, C. E. (1998). Predicting outcome of dream interpretation sessions by dream valence, dream arousal, attitudes toward dreams, and waking life stress. *Dreaming*, 8, 169-185. <https://doi.org/10.1037/h0094487>