

Routledge
Taylor & Francis Group

Journal of Social Work Practice

Psychotherapeutic Approaches in Health, Welfare and the Community

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/cjsw20

Social workers' evaluation of ChatGPT for solving ethical dilemmas within the limits of confidentiality

Michal Segal

To cite this article: Michal Segal (20 Mar 2025): Social workers' evaluation of ChatGPT for solving ethical dilemmas within the limits of confidentiality, Journal of Social Work Practice, DOI: 10.1080/02650533.2025.2480092

To link to this article: https://doi.org/10.1080/02650533.2025.2480092

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 20 Mar 2025.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=cjsw20

JOURNAL OF SOCIAL WORK PRACTICE
https://doi.org/10.1080/02650533.2025.2480092

OPEN ACCESS Check for updates

Social workers' evaluation of ChatGPT for solving ethical dilemmas within the limits of confidentiality

Michal Segal

Department of Social Work, Tel-Hai College, Qiryat Shemona, Israel; Research Center for Innovation in Social Work, Tel-Hai College; Qiryat Shemona, Israel

ABSTRACT

This study examined how social workers evaluate the use of ChatGPT in solving ethical dilemmas related to confidentiality in social work. The study employed qualitative document analysis based on a phenomenological approach, analysing assignments of 70 master's degree students in social work. Participants compared their own solutions to ethical dilemmas with those proposed by ChatGPT. The analysis revealed three main issues: disparities in analysing complex professional dilemmas, differences in reference to ethical codes and laws, and disparities in emotional aspects and empathy. Study participants perceived their responses as more comprehensive, legally based, and empathetic compared to ChatGPT's solutions, which were perceived as superficial and lacking emotional understanding. While ChatGPT offers potential as a supportive tool, it cannot replace social workers' ethical judgement and empathy. Recommendations include developing AI training workshops and examining AI tools for ethical training while maintaining the profession's human element.

ARTICLE HISTORY
Received 21 September 2024
Accepted 24 February 2025

KEYWORDS
ChatGPT; social work ethics; confidentiality; artificial intelligence

Introduction

Complex ethical dilemmas are common in social work, whether they stem from issues with clients, colleagues in the profession, the professional framework itself, or the employers. Social workers find themselves operating in challenging organisational environments, communicating with vulnerable populations and handling sensitive information, while carefully navigating between personal and professional values, legal frameworks, diverse requirements and conflicting interests (Banks, 2021, Banks et al., 2020; Congress, 2000; Linzer et al., 2003; Reamer, 2018).

The professional literature offers a variety of models for dealing with ethical dilemmas in the field of social work. These models share similar principles regarding the ethical decision-making process – identifying the ethical values relevant to social work, understanding the client's values, mapping the organisations and groups that may be affected by any decision, in-depth examination of the code of ethics and the relevant legislation, reviewing all action options while weighing the advantages and disadvantages of each one, consultation with professional colleagues, formulation of a final decision, and

CONTACT Michal Segal michalseg@telhai.ac.il

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

detailed documentation in the client's file of the decision-making process and the reasons that led to it (Congress, 2000; Reamer, 2018).

A significant part of the decision-making process in ethical dilemmas is the integration of the social worker's personal values but also the need to identify possible conflicts between those personal values (cultural, political, ethnic values, etc.) and the values of clients or colleagues. Therefore, social workers must be aware of their personal values and how they may influence a decision-making process, while maintaining transparency and constant reflection (Banks, 2021; Congress, 2000; Reamer, 2018).

Boundaries of confidentiality in social work

The challenge of maintaining confidentiality is one of the most common confronting social workers (Abrams & Dettlaff, 2020, Banks, 2021; Linzer et al., 2003). A study on ethical dilemmas in different areas of social work intervention in Spain reported that confidentiality was the most common dilemma faced regardless of whether the social workers dealt with health services and hospitals, children, families, minority residents or prisoners (Viscarret et al., 2020). Additionally, *[Author's]* qualitative study that examined the coping of Israeli social workers with ethical and legal dilemmas during COVID-19 revealed a central theme involving the potential for damage to clients' privacy and confidentiality due to the unique circumstances of the pandemic (Banks et al., 2020; Segal & Gur, 2023).

Legal and ethical aspect of confidentiality

Confidentiality is an essential part of the social worker-client relationship and the basis for trust between them. The client assumes that the social worker is committed to confidentiality at every stage of the service, and therefore feels safety in cooperation, in revealing hidden and intimate secrets and not hiding information. This principle is further emphasised in the Global Social Work Statement of Ethical Principles, which declares: 'Social workers respect and work in accordance with people's rights to confidentiality and privacy unless there is risk of harm to the self or to others or other statutory restrictions' (Global Social Work Statement of Ethical Principles, 2018, Section 6.1, Respect for Confidentiality and Privacy).

The Israeli Social Work Act, 1996–1574 (SWA) defined the duty of confidentiality for the first time in legislation regarding those engaged in social work having information about their clients, and also delineated the cases in which this duty may be violated. Israel's SWA mentions two features of the duty of professional confidentiality:

(1) *Maintaining the confidentiality of information provided by the client*, detailed in section 8 of the SWA. 'A social worker is obligated to keep private the information about a person given to him as part of his profession work . . .' Only in limited cases provided by law may the social worker violate the duty of confidentiality – upon belief the client will harm himself or to prevent injury to another person [Section 8 (a)(4)] *or* when obligated under law to disclose information [Section 8 (a)(5)] (i.e. the obligation to report abuse of a minor or a helpless person in accordance with section 368D of the Penal Code Act, 5737-1977). The legislature

considered it appropriate to allow the court to sanction disclosure of information if there are special circumstances that justify it [Section 8 (a)(8)]; that is, in Israel the maintenance of confidentiality is not absolute and can be violated to protect other social interests (Reamer, 2018).

(2) *The disclosure of information to the client*. The confines of that information which the client is entitled to receive from the social worker are defined in section 7 of the SWA. Section 7 (a) stipulates that 'a person is entitled to receive from a social worker information regarding the social worker's handling of his affairs, except for information concerning another person and personal notes of the social worker'. Similar to section 8 of the SWA, section 7 (b) stipulates that the social worker may refuse to provide information to the client 'if in his opinion, the provision of the information may cause serious harm to the applicant or another person'.

Integration of Artificial Intelligence (AI) in decision-making processes in social work: possibilities and challenges

In November 2022, the Open-AI company launched ChatGPT - (Chat Generative Pre-trained Transformer) which has since become an integral part of our lives (Faggella, 2020; Hodgson et al., 2021; Hu, 2023). ChatGPT is described by Jürgen et al., (2023) as an advanced conversational model 'created to generate text that is indistinguishable from text written by humans. It is able to conduct conversations with users in a natural and sometimes intuitive way' (Jürgen et al., 2023, p. 342).

Recent studies that examined the potential of ChatGPT in the field of social work have shown interesting and thought-provoking results. Victor et al. (2023) conducted a comprehensive study that tested ChatGPT's ability to answer questions from the American Board of Social Work (ASWB) licensure tests. The results of the study were surprising, as the model was able to correctly answer a high percentage of the questions. These findings raise significant questions about the validity of the existing tests and offer a fascinating possibility to utilise this technology as an alternative to traditional licencing tests.

Netto (2023) then investigated the use of ChatGPT in evaluations in social work programmes, revealing interesting insights into the capabilities and limitations of AI in this context. While the AI tool was able to provide general and correct answers, it had significant difficulty in combining specific details of test cases and drawing indirect conclusions. These findings highlight the importance of using complex and detailed test cases as a means of evaluation, which still represent a significant challenge for AI systems.

Another study examining the use of ChatGPT was conducted by Segal (2024), who investigated the ability of ChatGPT to assist social workers in solving complex ethical dilemmas. Although she concluded that ChatGPT had significant potential as a support tool for understanding the complexity of ethical dilemmas, her study emphasised that despite its capabilities, ChatGPT could not replace the expertise and judgement of a human ethics committee.

Despite ChatGPT's great potential presented by these studies, the use of AI in social work also raises significant challenges that require much attention. There are

understandable concerns about a possible violation of customer privacy, algorithmic biases that may affect decision-making, and the making of inequality between different population groups (Goldkind, 2021; Segal, 2024). There is also a danger that excessive focus on technical approaches may harm the core values of the social work profession such as promoting social justice and creating significant social change (James et al., 2023; Morley et al., 2019), as well as a deep concern that the widespread use of AI technologies will harm the human and emotional involvement, which is the core of the social work profession. Forming a human connection, understanding emotional nuances, and demonstrating empathy are all critical components of effective social work, and careful consideration must be given to how to integrate AI technologies without compromising them (James et al., 2023).

The possibility of using AI in social work must continue to be explored, and the issue of the degree of confidentiality in social work is a ubiquitous dilemma; therefore, this study sought to examine how social workers perceive the solutions offered by ChatGPT as compared with their own solutions, in dealing with ethical dilemmas within the limits of maintaining confidentiality. The aim of the research was to contribute a better understanding of both the potential and the challenges of integrating AI technologies in decision-making processes in social work, and to provide essential insights may help develop integrative approaches combining the benefits of AI with human expertise and judgement.

Methods

The current study employed a qualitative research method based on the phenomenological approach of document analysis (Denzin & Lincoln, 2017; Patton, 2002). This methodology allows researchers to present the perspectives of the research participants in relation to the phenomenon being studied and examine how they view the phenomenon, how they experience it, and what meaning they attribute to it (Braun & Clarke, 2022; Patton, 2002).

Research tools and procedures

The research participants were master's degree students in social work currently enrolled in the 'Advanced Law and Ethics' course mandatory in the second year. The course is designed to present a sequence of examples in social work involving dilemmas of values and legal aspects that highlight the values of the social work profession and allow the student to become familiar with the legal system that regulates the duties and powers of the social worker in issues of professional confidentiality, disclosure of information and immunity. The course is delivered online and students are required to submit assignments on the course topics. For the purpose of the current study, we examined task #5 which was given to students as part of the course requirements. The task read as follows:

(1) You must choose one dilemma from the three presented to you. After choosing the dilemma, write how you would approach the dilemma as a social worker responsible for handling the case. Carefully consider the various considerations and options open to you and explain why you chose the approach that you did. Be

sure to refer in your answer to the legal aspects relevant to the chosen dilemma as expressed in section 8 and/or 7 of the Social Workers Law.
(2) Access ChatGPT at the following link: https://openai.com/chatgpt[1]
(3) Input the instruction provided by the prompt engineer for your chosen dilemma as a prompt into ChatGPT.
(4) Copy and paste the output by ChatGPT.
(5) Explain whether you think this is an appropriate solution from an ethical-professional point of view and summarise what you learned from the comparison between your solution and ChatGPT's suggestion concerning dilemmas in the field of confidentiality.

All 70 students who took part in the course consented to participate in the research. The students were divided into 20 groups, with each submitting one assignment for analysis. Each group chose one dilemma from the three dilemmas offered as detailed in Table 1. The lecturer wrote the format (prompt engineer) that the students entered into ChatGPT in a similar manner for each dilemma. For example, in dilemma #1 the format was:

I own a private clinic and treat Jacob, a 41-year-old insurance agent, once per week, who is in distress due to problems in his marriage. Now I ask you, as an external consultant who is a social worker by training, to provide us with another perspective on the events. What would you recommend I do in this situation? What steps would you take? Explain and justify your position while referring to the ethical and therapeutic

Table 1. The dilemmas presented to the students as part of the assignment.

Dilemma 1 - Threat of violence from the client
You own a private clinic and treat Jacob, a 41-year-old insurance agent, who is in distress due to problems in his marriage, once per week. According to Jacob, for about a year now his wife has been 'hostile and distant towards him. She hardly talks to him anymore and he doesn't know what happened between them'. At one of the meetings, Jacob became unusually nervous and tense. He told you that two days earlier he had discovered that his wife was having an affair with a man she knew through her work. Jacob spoke at length about the feelings of being betrayed and angry. Towards the end of the meeting, Jacob said: 'I may regret it in retrospect, but I think I could kill this man'. In the short time left in the meeting, you asked Jacob specific questions about his desire to hurt his wife's lover. By the end of the meeting, you weren't sure if Jacob really intended to attack the man who was apparently involved with his wife.
Dilemma 2 - Hiding information in couples therapy
You are a social worker at a family care centre. You handle couples therapy for Eran and Yael, a couple who began therapy when Yael told Eran that she was considering asking him for a divorce. Eran and Yael agreed that Eran has a serious gambling problem. The couple is in financial distress as a result of Eran's gambling, and Yael recently said: 'I just can't take it anymore. Either Eran gets help and stops gambling, or we break up. I'm tired of it'. Eran started therapy in a support group for compulsive gamblers and spoke in the couple meetings with you of how much the meetings have helped him and how glad he is that he stopped gambling. One afternoon, Eran called you and said he needed to talk to you privately the next day; he then informed you that he had lied about participating in 'compulsive gamblers' meetings, his gambling problem is worse than ever, and he was feeling desperate. Eran asked you for help and begged you not to tell his wife about his lies.
Dilemma 3 - Hiding a diagnosis from a client at the request of the family
You are social workers in the surgical department of Haifa General Hospital, where you work with cancer patients and their families. One of the patients is a 68-year-old man, Shmuel, who was hospitalised for exploratory surgery on his abdomen. Shmuel, weak and tired, is described by his attending physician as 'very emotional'. One afternoon, Shmuel's daughter, Irena, approached you and said that her father's doctor informed her that lab results from the exploratory surgery showed that her father has terminal cancer. Irena said that she and the family were in shock, noting: 'We never imagined that he was so sick. We just can't tell him that his life will end soon. He won't be able to handle it'. Irena said she wanted us to help her explain this to the attending physician and the hospital staff involved in her father's care.

dilemmas involved in the decision. Your answer will serve as a reflective tool for examining and improving our professional practice.

The dilemmas presented to the students were taken from Reamer (2018), a book in which ethical dilemmas that may confront social workers are presented.

The three ethical dilemmas demonstrated the limits of confidentiality in social work, referring to sections 7 and 8 of the Social Workers Law. The first dilemma involved the threat of violence on the part of the client, raising the question of whether the social worker's duty of confidentiality should be violated to protect a third party from potential danger. The second dilemma concerned the concealment of information in couples therapy, testing the balance between maintaining the confidentiality of the client and the social worker's professional duty towards the other partner in therapy who is engaged in sharing information for the benefit of the therapy. The third dilemma concerns hiding a medical diagnosis from a client at the request of the family, raising ethical questions regarding the client/patient's right to information about his/her medical condition versus the family's wishes and claims of consideration for the client/patient's welfare. Each of the dilemmas requires the social worker to exercise in-depth professional judgement on balancing the duty of confidentiality, protecting the client and the public, and respecting the client/patient's autonomy, all within the framework of the law and professional ethics.

The assignments were written in Hebrew, as that was the language of instruction. For the purposes of this study, all student responses were professionally translated into English by a certified translator.

Data analysis

The students' assignments were reviewed methodologically and systematically to identify 'units of meaning' (Shkedi, 2003). In stage 2, students' assignments were reorganised according to meaning units to ascertain the commonality of all the meaning units present in each text (Braun & Clarke, 2022; Denzin & Lincoln, 2002). During stage 3, the main themes were coded and structured, initially employing an 'open coding' system, followed by axial coding. Seven primary themes emerged during the open-coding stage.[2] In the axial coding stage, certain themes were amalgamated while others were discarded based on their relative significance.

Prerequisites were subsequently established for each category. In the fourth stage, the final themes were delineated by utilising the prerequisites of each theme and categorising each meaning unit accordingly. The various themes were then integrated, creating the final three main themes.

The reliability and validity of the research

Reliability and validity of qualitative research are examined with different tools than those used to examine quantitative research (Patton, 2002). Following is a description of how the four components of the trustworthiness principle are applied in order to examine the reliability of the qualitative research performed in the data analysis from the judgements in this study (Braun & Clarke, 2022; Denzin & Lincoln, 2002).

The first element is the reliability of the research expressed in the researcher's prolonged and in-depth engagement in the research topic, his/her involvement in the research and intensive observation of the findings, as well as in the demonstration of proficiency in the research findings upon presentation. To maintain credibility of the research, the researcher must be reflexive, aware of his/her own perceptions, attitudes and personal values regarding the subject under investigation.

The second component concerns the ability to transfer and generalise the research. In doing so, the researcher made sure not to take the text out of its original context. The third element is dependency; to reduce the dependence of the study as much as possible, the researcher described the method of data analysis with full transparency. The fourth element is the appropriateness of the study. In the current study, the phase of formulating the themes included a careful and separate examination of each work and a constant comparison between its findings and those that emerged from the other works.

Pertaining the current study, the researcher holds a PhD in law and has been teaching the 'Law and Ethics' course in the Department of Social Work of an academic college for over 12 years. She is an expert in analysing ethical dilemmas in social work through the legal prism. To increase the credibility of the research and provide an additional perspective to the findings, two social workers participated in data analysis: one with 20 years of experience in the field of welfare services, and the other with 10 years of experience in the field of social work

During the analysis of the research findings, the researcher reflected constantly on her professional experiences by adding reflective notes to the research findings. The credibility of the research was strengthened by a thick description of the findings including representative quotes, allowing readers to assess the results clearly and accurately (Denzin & Lincoln, 2002; Patton, 2002).

Ethics

Upon completion of the course and publication of the grades, the researcher contacted the students by email and attached an informed consent form and an explanation of the present study. The students were asked to confirm by return email their agreement to the use for research purposes of the assignment they had submitted as part of the course. All students ($N = 70$) agreed to participate in the study and confirmed their consent via email. All students were guaranteed complete confidentiality and anonymity. It was made clear to the students that they were free to refuse participation in the study at any time without any negative consequences, and that their agreement or refusal would not affect their grade in the course or the continuation of their studies. The original dilemmas presented by Reamer were further adapted by changing the names to Hebrew names. The information was then fed into ChatGPT for analysis and learning assistance. During the analysis and writing process, all identifying details of the study participants were removed, and pseudonyms were used throughout the study. The research proposal was presented to and approved by the Ethics Committee of our college.

Results

From the analysis of the comparisons made by the participants in the study between their own answers and those offered by ChatGPT, three main themes emerged. These themes reflected the way in which the study participants read, examined and evaluated the ChatGPT responses in relation to the ethical dilemmas that had been presented.

Analysis of the 20 group assignments revealed that the majority of groups (80%) emphasised ChatGPT's superficial handling of complex dilemmas (Theme 1). Most groups (70%) noted that their responses were more firmly grounded in ethical codes and relevant laws (Theme 2), and a significant number of groups (60%) highlighted ChatGPT's lack of empathy and emotional understanding (Theme 2). These themes reflected the way in which the study participants read, examined, and evaluated the ChatGPT responses to the ethical dilemmas that had been presented.

Theme 1: disparity in the ability to analyze and deal with the complexity of professional dilemmas

This theme involves the way in which the study participants identified differences between their own and ChatGPT's ability to deal with the complexity of the cases. The participants perceived their own answers as better in dealing with complex issues, while the ChatGPT responses were perceived as more superficial and not comprehensive enough. Thus, for example, in assignment 6 (dilemma #2), the participants remarked about the ChatGPT response:

> The artificial intelligence did talk about the different layers of the problem, but in a superficial way that does not demonstrate the complexity of the case, does not let us see how the situation unfolds, or understand the conscious and unconscious elements of the situation and exactly the order of priorities in couples therapy and addiction treatment, the 'sting' and the consequences of everything.

And similarly, in assignment 19 (dilemma #1) the research participants emphasised:

> From a comparison of the ChatGPT answer to the solution we proposed, we realize that the way to solving the dilemma may be complicated and there are no shortcuts. In taking the situation seriously, we as social workers must do our best to make all the inquiries and investigations with the client and with all professional parties involved so that we may reach a decision that will benefit all parties to the dilemma.

In assignment 4 (dilemma #3) the participants described the superficiality of ChatGPT's answer:

> It seems that the artificial intelligence site analyzed the case in a one-dimensional way and only looked at the motives, needs and concerns of Irena and the family without taking into account Shmuel's side as a client, as the first person who should have himself received the information.

Theme 2: disparities in reference to the code of ethics and relevant laws

Following the disparities identified in ChatGPT's responses that bring into doubt its ability to deal with the complexity of the dilemmas, this theme focuses on the way in

which the study participants identified discrepancies between their own answers based on the code of ethics and specific laws known to them, and ChatGPT's answers which were perceived to be more general. The participants therefore perceived their answers as having been based more on the ethical code and the relevant laws, while the ChatGPT responses were perceived as being general and less anchored in specific provisions of the law and professional ethics.

Thus, for example, in assignment 1 (dilemma #2) the study participants stated:

> In our opinion, the solution proposed by Chat GPT is inappropriate from an ethical and professional stance; the solution is not based on laws involving patient privacy and does not refer to the code of ethics of social workers.

At the same time, they describe their own approach to solving the dilemma as, in fact, based upon the laws relevant to that dilemma:

> Comparing our solution to that proposed by ChatGPT, we learn that ours is based on laws that deal with the subject of patient privacy and the issue of confidentiality limitations. The rules by which social workers act and the code of ethics provide a professional and appropriate answer for dealing with dilemmas between these two issues and with other ethical dilemmas that confront us during our work.

The study participants also emphasised their reliance on the code of ethics of the Association of Social Workers in Israel (2018) in providing a solution to the dilemma and their view that this is ignored in the ChatGPT response. For example, in assignment 16 (dilemma #2):

> The website does not refer to our duty towards Yael to conduct ourselves in a manner worthy of trust and honesty, as stated in the code of ethics for social workers.

Similarly, in assignment 18 (dilemma #2) the participants emphasised the importance of relying on the code of ethics and the law: 'According to the Fundamental Law on Human Dignity and Freedom, Article 7, which states that "every person has the right to confidentiality and the privacy of his life"'. The right to privacy also appears in the social workers' code of ethics and the social workers' law of 1996. And ChatGPT ignores these aspects: 'It seems that the artificial intelligence site does not take into account all the ethical and legal issues in the framework of considerations for such a complex decision'.

Theme 3: disparities in references to emotional aspects and empathy in the therapeutic process

The previous themes examined the disparities between the solutions in addressing the complexity of each dilemma and in references to ethical and legal codes; this theme focuses on another aspect of social work – emotional empathy. Theme #3 highlights the differences between the participants' approach emphasising the importance of empathy, emotional sensitivity and understanding the emotional complexity of the clients, and the ChatGPT approach which is perceived as lacking in these aspects. The participants view empathy and emotional understanding as critical tools in the therapeutic process and maintain that these are lacking in the ChatGPT answers.

For example, in assignment 8 (dilemma #2), the participants emphasised the lack of empathy in the ChatGPT response:

> In ChatGPT's answer there was no empathic regard for the client, which is the basis of the relationship between therapist and client, and which may harm the needs of the client and even the therapeutic process [. . .] The comparison shows us that the artificial intelligence site manages to examine different options in a comprehensive way, but at the same time the personal intuition, ethics and conscience of a therapist are all absent. Also, as therapists one of the useful tools is our ability to empathize with clients, this is our ability. . .

Similarly, in assignment 6 (dilemma #2), the participants indicated the inability of ChatGPT to understand the emotional complexity: 'Artificial intelligence still cannot see the complete therapeutic picture, i.e. the unconscious elements. . .and the empathic way of looking at things'.

In assignment 9 (dilemma #1) the participants point out the lack of ChatGPT's reference to emotional states: 'It [ChatGPT] does not reduce the risk of potential violence and does not refer to the emotional state of the client, the fragile state'.

And finally in assignment 7 (dilemma #2) the participants refer to the empathy they felt towards Keren in contrast to the ChatGPT answer:

> When we discussed the dilemma, we felt gut feelings of empathy for Keren and at the same time a desire not to harm her, that is, the emotional world greatly influenced our decisions and our methods of action as social workers. On the other hand, ChatGPT is an artificial intelligence site and is not motivated by emotional drives.

In this central theme, the importance of showing sensitivity and empathy is expressed, especially when the dilemmas concerning the limits of confidentiality are complex. An essential element in the social worker's effort is the trust built between him/her and the client, a trust based on the promise of confidentiality. There are, however, situations in which the breach of confidentiality is required by law or rules of professional ethics, as in cases of danger to the client or another person. This balance between maintaining the client's trust and fulfilling legal and ethical obligations confronts the social worker with a complex dilemma. Showing empathy and sensitivity in such a situation is a critical skill in social work that emphasises the importance of human consideration in ethical decision-making processes.

Discussion

Artificial intelligence (AI) technologies are developing at a fast pace and are now an integral part of our lives (Faggella, 2020; Hodgson et al., 2021; Hu, 2023). However, the adoption of artificial intelligence tools such as ChatGPT in social work decision-making raises challenges, as reflected in the current research findings where social workers weighed the differences between their own solutions to a dilemma and those proposed by ChatGPT.

In the *first theme*, 80% of the participants expressed what they felt was the superficiality of ChatGPT and its inability to deal with the complexity of professional dilemmas in a comprehensive and in-depth manner. The findings of this study are also consistent with those of Netto's (2023) study, which demonstrated that while ChatGPT is capable of providing general information, it can have difficulty integrating specific details and understanding more complex sociocultural contexts.

In the second theme, 70% of the participants maintained that references to the code of ethics and relevant laws differed in their own solutions versus those of ChatGPT, as the latter tended not to emphasise the legal and ethical basis in its answers. These findings are consistent with the study of Segal (2024) and reflect the difficulty of ChatGPT to be precise in the legal corpus required when solving ethical dilemmas within the limits of confidentiality in social work (Segal, 2024).

Our findings highlight that despite developments in technology, the professional and legal knowledge of social workers remains essential and vital for ethical decision-making. Victor et al. (2023) emphasised the need to develop comprehensive ethical frameworks for the use of AI in the field of care and welfare. As of now, ChatGPT is not suitable to be used as a direct tool in making ethical decisions, mainly due to issues relating to the violation of customer privacy and information security (Segal, 2024). These challenges underscore the importance for a careful and informed approach in integrating AI technologies in the practice of social work while maintaining the ethical and professional standards of social work. Offering targeted training workshops or guidance sessions on the limitations of the use of artificial intelligence (like ChatGPT) and the ethical use of artificial intelligence in social work may provide a practical approach. Such initiatives should be developed collaboratively, involving social workers from the field, government representatives from the Ministry of Welfare in each country, and experts in the field of AI. These efforts should focus on addressing the challenges highlighted in this study such as privacy and confidentiality, and providing practical tools for integrating AI technologies in a manner aligned with the ethical and professional standards of social work.

Despite its current limitations, ChatGPT has significant potential as an instructional tool. The development of a chatbot that simulates ethical dilemmas may provide an effective platform for practice and learning for social workers while maintaining clients' privacy and confidentiality, as was done in the present study. It is possible to use generic dilemmas taken from published professional literature (Reamer, 2018) or, as suggested by Segal (2024), to use questions published by the ethics committee of a professional union of social workers. This approach allows social workers to experience ethical decision-making processes in a safe space as they deal with authentic and relevant dilemmas without risking the privacy of real clients. Using these dilemmas in combination with AI technologies provides a unique opportunity for professional training while confronting the ethical and practical challenges identified in the research.

In the third theme, 60% of the participants emphasised the importance of showing empathy and sensitivity in the day-to-day practice of social work, aspects they claimed were lacking in the ChatGPT answers (James et al., 2023). Thus, a fundamental question arises from the research findings regarding the role of artificial intelligence in the emotional context of social work. Is it possible and appropriate to use AI as a tool to neutralise emotional biases and stereotypes, thus providing a more objective point of view?

Despite the study participants' observation that ChatGPT lacks the emotional component, this may actually be a supportive factor in the process of making ethical decisions, representing a more objective point of view and serving as a reflective 'mirror' that helps social workers avoid the risk of emotional involvement that may harm the therapeutic process (LaBrenz et al., 2022; Rogerson et al., 2021). However, the claim that emotional neutrality brings good objectivity may be naïve. While emotional engagement

might indeed skew a social worker's response (LaBrenz et al., 2022; Rogerson et al., 2021), algorithmic decision-making based on no emotional engagement at all can also be skewed in relation to race, ethnicity, gender, sexual orientation, gender expression, and other vulnerable or protected categories (Netto, 2023; Ray, 2023; Reamer, 2023)

Conclusions

Research findings indicate the need for a thorough examination of the possibility for a balanced and ethical integration of artificial intelligence technologies such as ChatGPT into the individual work of social workers. At the same time, we clearly recognise that artificial intelligence is not capable of replacing the professional and ethical judgement of social workers (Netto, 2023; Segal, 2024). In this context, the words of a research participant (Assignment 6) emerge as an insight that reflects the complexity of integrating AI into social work and the fear of losing the sensitive human element that is critical to social work:

> The artificial intelligence solution shows that there is still room for hope regarding the therapeutic profession we have chosen. Artificial intelligence is already weakening almost every possible field, conquering it and as a result collapsing the professionalism, their creative and logical sense, in the technological age. Therefore, we are glad that AI cannot yet see the complete therapeutic picture, that is, the unconscious elements, the integration of the narrative together with the ethics, and the empathic look at things.

Limitations of the study

First, this study utilised a qualitative method and its findings cannot be generalised to a wider population. Second, the study was based on the use of the basic and free version of ChatGPT, which may affect the quality and depth of the answers received. It is possible that a more advanced version of ChatGPT another AI platform such as Claude, would have yielded different results. These limitations emphasise the need for follow-up studies that will examine a wider variety of AI tools and their impact on ethical decision-making processes in social work. Third, while steps were taken to ensure authentic student analysis, we acknowledge that students could technically have used AI tools to analyse ChatGPT responses. Future research may benefit from implementing more stringent controls such as requiring students to complete analyses in controlled environments without internet access or incorporating real-time discussions and presentations of their analytical processes.

Notes

1. At the time of the research, free access to the most advanced model of ChatGPT (GPT-4) was not yet available to the general public. To maintain uniformity in the results and ensure that all participants use the same version of the tool, the students participating in the study were asked to register and use the free version of ChatGPT-3.5.
2. 1. Artificial Intelligence will not replace therapists.
   2. ChatGPT does not rely on the code of ethics and organisational procedures.
   3. ChatGPT is misleading and does not rely on laws.
   4. ChatGPT lacks consideration of situational complexity.

5. ChatGPT lacks emotional intuition and empathy.
6. Do not rely on ChatGPT and consult with colleagues.
7. Importance of client involvement and transparency.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributor

Michal Segal, an attorney specialising in Human Rights, lectures at the Social Work Department of Tel Hai College, Israel. Among her specialities are Law and Ethic, Legal Aspects of Social Work and Policy Practice in Social Work.

ORCID

Michal Segal http://orcid.org/0000-0001-8234-1439

Data availability statement

Data and materials supporting the results or analyses presented in their paper available from the corresponding author upon reasonable request, provided that they do not compromise the identity of the study participants

References

Abrams, L. S., & Dettlaff, A. J. (2020). Voices from the frontlines: Social workers confront the COVID-19 pandemic. *Social Work*, *65*(3), 302–305. https://doi.org/10.1093/sw/swaa030

The Association of Social Workers in Israel. (2018). Retrieved March 9, 2025,from https://www.socialwork.org.il/prdFiles/%D7%A7%D7%95%D7%93%20%D7%94%D7%90%D7%AA%D7%99%D7%A7%D7%94%20%D7%91%D7%90%D7%A0%D7%92%D7%9C%D7%99%D7%AA.pdf

Banks, S. (2021). *Ethics and values in social Work* (5th ed.). Bloomsbury.

Banks, S., Cai, T., De Jonge, E., Shears, J., Shum, M., Sobočan, A. M., Strom, K., Truell, R., Úriz, M. J., & Weinberg, M. (2020). Practising ethically during COVID-19: Social work challenges and responses. *International Social Work*, *63*(5), 569–583. https://doi.org/10.1177/0020872820949614

Braun, V., & Clarke, V. (2022). Thematic analysis: Apractical Guide. *QMiP Bulletin*, *1*(33), 46–50. https://doi.org/10.53841/bpsqmip.2022.1.33.46

Congress, E. P. (2000). What social workers should know about ethics: Understanding and resolving ethical dilemmas. *Advances in Social Work*, *1*(1), 1–25. https://doi.org/10.18060/124

Denzin, N. K., & Lincoln, Y. S. (2002). *The qualitative inquiry Reader*. SAGE.

Denzin, N. K., & Lincoln, Y. S. (2017). *The SAGE handbook of qualitative research* (5th ed.). SAGE Publications, Inc.

Faggella, D. (2020). *What is machine learning?* Emerj Artificial Intelligence Research. https://emerj.com/ai-glossary-terms/what-is-machine-learning

Goldkind, L. (2021). Social work and artificial intelligence: Into the matrix. *Social Work*, *66*(4), 372–374.

Hodgson, D., Goldingay, S., Boddy, J., Nipperess, S., & Watts, L. (2021). Problematising artificial intelligence in Social Work Education: Challenges, issues and possibilities. *British Journal of Social Work*, *52*(4), 1878–1895. https://doi.org/10.1093/bjsw/bcab168

Hu, K. (2023, February 2). ChatGPT sets record for fastest-growing user base - analyst note. *Reuters*. https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01

James, P., Lal, J., Liao, A., Magee, L., & Soldatic, K. (2023). Algorithmic decision-making in social work practice and pedagogy: Confronting the competency/critique dilemma. *Social Work Education*, 1–18. https://doi.org/10.1080/02615479.2023.2195425

Jürgen, R., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, *6*(1), 342–363. https://doi.org/10.37074/jalt.2023.6.1.9

LaBrenz, C. A., Reyes-Quilodran, C., Padilla-Medina, D., Contreras, M. A., & Piñones, L. C. (2022). Deconstructing bias: The decision-making process among child protective service workers in Chile. *International Social Work*, *66*(3), 697–714. https://doi.org/10.1177/00208728211068927

Linzer, N., Conboy, A., & Ain, E. (2003). Ethical dilemmas of Israeli social workers. *International Social Work*, *46*(1), 7–21. https://doi.org/10.1177/0020872803046001590

Morley, C., Ablett, P., & Stenhouse, K. (2019). Technicist education: Paving the way for the rise of the social work robots? *Critical and Radical Social Work*, *7*(2), 139–154. https://doi.org/10.1332/204986019x15491042559718

Netto, N. R. (2023). Use of case studies in social work assessments - ChatGPT's kryptonite? *Social Work Education*, 1–12. https://doi.org/10.1080/02615479.2023.2266461

Patton, M. Q. (2002). *Qualitative research & evaluation methods*. Sage.

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, *3*, 121–154.. https://doi.org/10.1016/j.iotcps.2023.04.003

Reamer, F. G. (2018). *Social work values and ethics (Foundations of Social Work Knowledge)* (5th ed.). Columbia University Press.

Reamer, F. G. (2023). Artificial intelligence in social work: Emerging ethical issues. *The International Journal of Social Work Values and Ethics*, *20*(2), 52–71. https://doi.org/10.55521/10-020-205

Rogerson, C. V., Prescott, D. E., & Howard, H. G. (2021). Teaching social work students, the influence of explicit and implicit bias: Promoting ethical reflection in practice. *Social Work Education*, *41*(5), 1035–1046. https://doi.org/10.1080/02615479.2021.1910652

Segal, M. (2024, Jul 20). Confronting and managing ethical dilemmas in social work using ChatGPT. *European Journal of Social Work*. https://doi.org/10.1080/13691457.2024.2377786

Segal, M., & Gur, A. (2023). Ethical and legal dilemmas experienced by Israeli social workers during the COVID-19 pandemic. *Journal of Social Work*, *23*(3), 411–427.

Shkedi, A. (2003). *Words attempting to touch: Qualitative research, theory and practice* [Hebrew]. Ramot.

Victor, B. G., Sokol, R. L., Goldkind, L., & Perron, B. E. (2023). Recommendations for social work researchers and journal editors on the use of generative AI and large language models. *Journal of the Society for Social Work and Research*, *14*(3), 563–577. https://doi.org/10.1086/726021

Viscarret, J. J., Idareta, F., Ballestero, A., & Úriz, M. J. (2020). Ethical dilemmas and areas of social work intervention in Spain. *Journal of Social Service Research*, *46*(1), 55–70. https://doi.org/10.1080/01488376.2018.1524813