WISG

Workshop interdisciplinaire sur la sécurité globale

14 & 15 mars 2024 • Couvent des Jacobins **Rennes**

MENACES ET RISQUES HYBRIDES : LES NOUVEAUX ENJEUX DE LA SÉCURITÉ ET DE LA RÉSILIENCE DE NOS DÉMOCRATIES

Digital phenotyping for psychiatrie disorders from social interaction

Philippe Robert

Pr Émérite de Psychiatrie Université Côte d'Azur

MENACES ET RISQUES HYBRIDES : LES NOUVEAUX ENJEUX DE LA SÉCURITÉ ET DE LA RÉSILIENCE DE NOS DÉMOCRATIES

- Possibilité, probabilité d'un fait, d'un événement considéré comme un mal ou un dommage
- Parole, comportement par lesquels on indique à quelqu'un qu'on a l'intention de lui nuire,Défense, cybersécurité, données personnelles
- Le travail hybride peut engendrer pour les travailleurs du stress, de l'anxiété, de l'isolement ou un burn-out. Le premier facteur de risque est le manque de communication

- Situation dans laquelle quelqu'un, quelque chose n'est exposé à aucun danger
- Capacité d’une personne ou d’un groupe à bien se développer, à continuer à se projeter dans l’avenir, en présence d’événements déstabilisants, de conditions de vie difficiles, de traumatismes parfois sévères
- Système public, forme de gouvernement dans lequel la souveraineté émane du peuple

MePheSTO: Digital Phenotyping for Psychiatric Disorders from Social Interaction

Objective measures of psychiatric symptoms derived from speech and visual features of clinical interaction

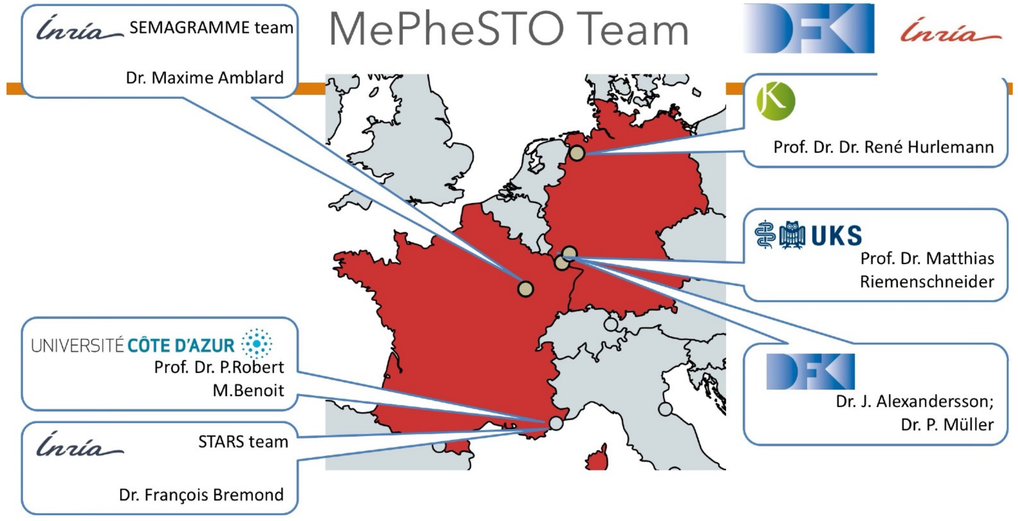
Alexandra König[1,2], Hali Lindsay[3], Philippe Mueller[3], Danilo Postin[4], Eric Ettore[5], Michal Balazia[6], Amandine Lecomte[7], Benedikt Wirth[3]

1 Kinelements
2 Cobtek (Cognition-Behaviour-Technology) Lab, FRIS-University Côte d'azur, Nice, France
3 German Research Center for Artificial Intelligence (DFKI), Saarbrücken, Germany
4 Department of Psychiatry, Carl von Ossietzky University of Oldenburg
5 Department of Psychiatry, Hopital Pasteur, CHU de Nice, France
6 Institut national de recherche en informatique et en automatique, Stars Team, France
7 INRIA, Semagramme Team, Nancy, France

Detection of Depression Severity from Clinical Interactions through Motion Energy Analysis

Alexandra König[1,2], Philippe Mueller[3], Hali Lindsay[3], Danilo Postin[4], Eric Ettore[5], Michal Balazia[6], Benedikt Wirth[3]

1 Kinelements, Saarbrücken, Germany
2 Cobtek (Cognition-Behaviour-Technology) Lab, University Côte d'azur, Nice, France
3 German Research Center for Artificial Intelligence (DFKI), Saarbrücken, Germany
4 Department of Psychiatry, Carl von Ossietzky University of Oldenburg
5 Department of Psychiatry, Hopital Pasteur, CHU de Nice, France
6 Institut national de recherche en informatique et en automatique, Stars Team, France



Automatically Extracted Linguistic Features of MADRS Clinical Interview are associated with Depression Severity - APA 2024 – Abstract

PCN Psychiatry and Clinical Neurosciences

REVIEW ARTICLE

Emotional expression in psychiatric conditions: New technology for clinicians

Karol Grabowski, MD, PhD[1*], Agnieszka Rynkiewicz, MD, MAT, PhD[2,3], Amandine Lassalle, PhD[4], Simon Baron-Cohen, MPhil, PhD[5], Björn Schuller, PhD[6], Nicholas Cummins, PhD[6], Alice Baird, MFA[7], Justyna Podgórska-Bednarz, PhD[8,9], Agata Pieniążek, MA[8,10,11] and Izabela Łucka, MD, PhD[12]

Aim: Emotional expressions are one of the most widely studied topics in neuroscience, from both clinical and non-clinical perspectives. Atypical emotional expressions are seen in various psychiatric conditions, including schizophrenia, depression, and autism spectrum conditions. Understanding the basics of emotional expressions and recognition can be crucial for diagnostic and therapeutic procedures. Emotions can be expressed in the face, gesture, posture, voice, and behavior and affect physiological parameters, such as the heart rate or body temperature. With modern technology, clinicians can use a variety of tools ranging from sophisticated laboratory equipment to smartphones and web cameras. The aim of this paper is to review the currently used tools using modern technology and discuss their usefulness as well as possible future directions in emotional expression research and treatment strategies.

Methods: The authors conducted a literature review in the PubMed, EBSCO, and SCOPUS databases, using the following key words: 'emotions,' 'emotional expression,' 'affective computing,' and 'autism.' The most relevant and up-to-date publications were identified and discussed. Search results were supplemented by the authors' own research in the field of emotional expression.

Results: We present a critical review of the currently available technical diagnostic and therapeutic methods. The most important studies are summarized in a table.

Conclusion: Most of the currently available methods have not been adequately validated in clinical settings. They may be a great help in everyday practice; however, they need further testing. Future directions in this field include more virtual-reality-based and interactive interventions, as well as development and improvement of humanoid robots.

Keywords: *affective computing, autism, emotions, expressed emotion, nonverbal communication.*

http://onlinelibrary.wiley.com/doi/10.1111/pcn.12799/full

Psychiatric Disorders

16% of the population have depression in their lifetime

Schizophrenia affects about 1% of the European population

Severe impairment of quality of life

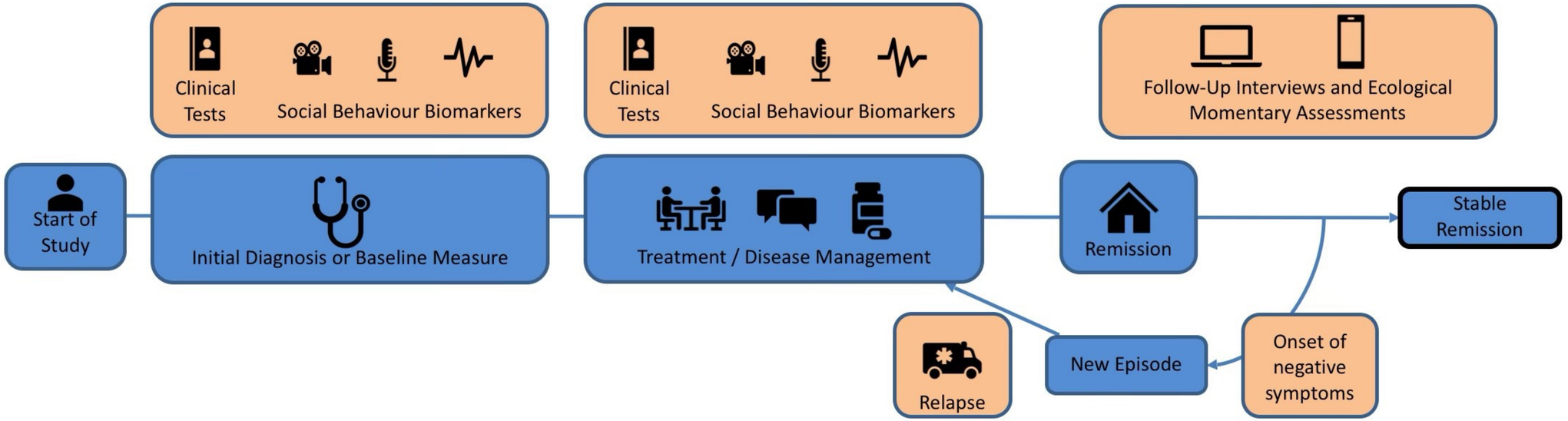
High costs for the health system

Early diagnosis and personalized treatment is crucial!

Major objectives

1. Supporting the **differential diagnosis** of depressive episodes
   - Unipolar depression, bipolar disorder, traumatic experience?
   - Early detection is crucial for choosing the right therapy
2. Objective measurement of **formal thought disorders**
   - Disorganised speech, a major symptom in schizophrenia
   - Discourse cohesion, syntactic complexity, eye movements
3. Quantifying **therapeutic alliance**
   - Therapeutic alliance is an important predictor of therapeutic success
   - Aspects are: Movement synchrony, shared language, affective counter-regulation, shared experience

Methods



Behaviour
in **social interactions** as **markers**
for psychiatric diseases

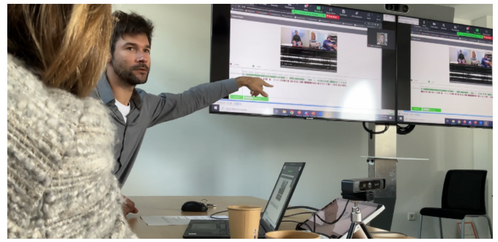
Eye contact

Body language

Speaker:inside change

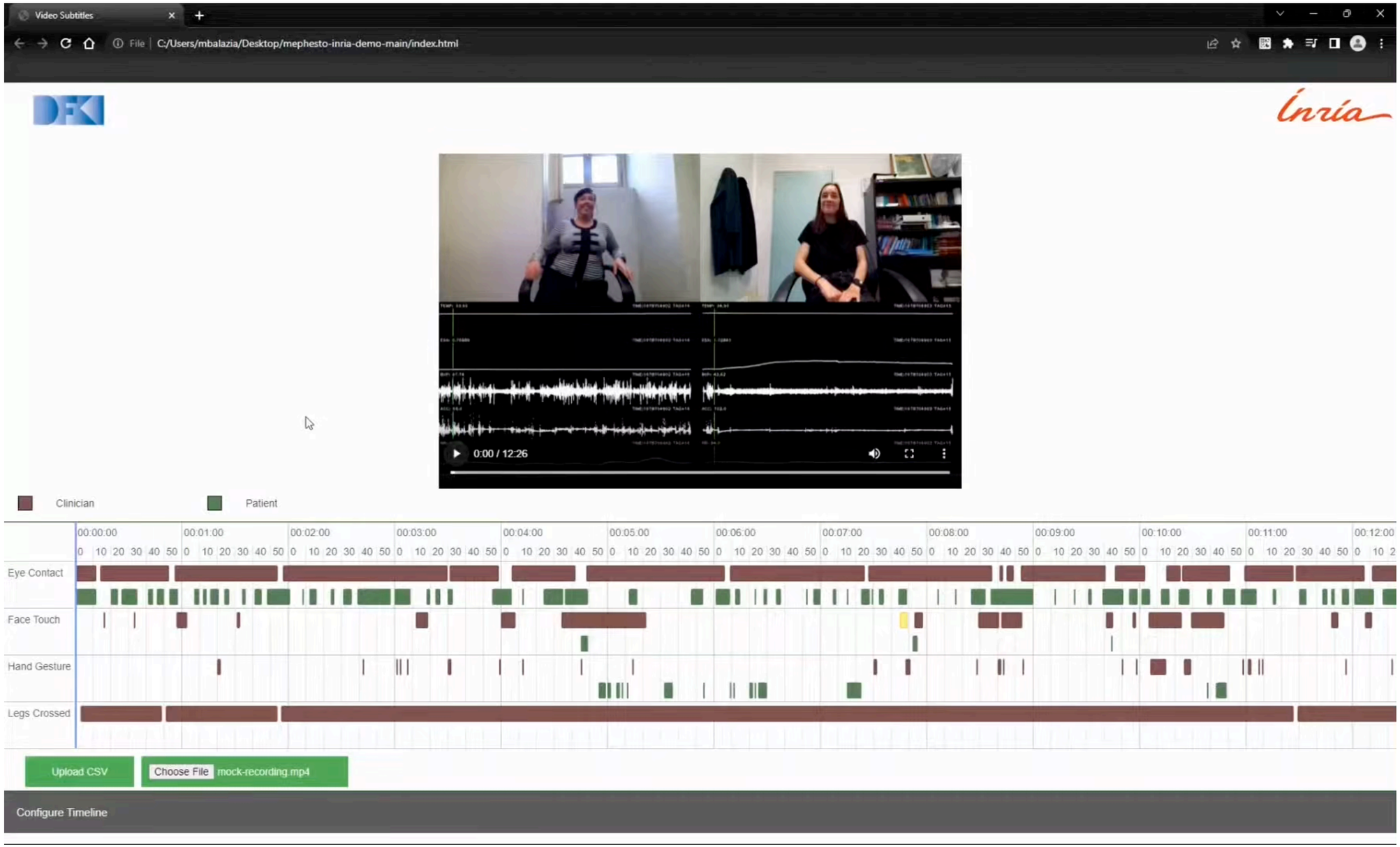
Backchannel

Agreement / Disagreement



Michal Balazia
Inria STARS

http://cobtek.fr/publications/



COBTEK
Cognition Behaviour Technology

COGNITION

MOTIVATION EMOTION

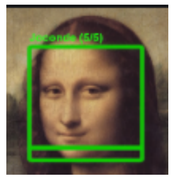
INTERACTION SOCIALE

Application pour l'évaluation

Audio

Vidéo

Un capteur est un dispositif transformant l'état d'une grandeur physique observée en une grandeur utilisable



Audio

Dementia and Geriatric Cognitive Disorders

Dement Geriatr Cogn Disord 2018;45:198–209
DOI: 10.1159/000487852
Accepted: February 21, 2018
Published online: June 8, 2018
© 2018 S. Karger AG, Basel
www.karger.com/dem

Original Research Article

Fully Automatic Speech-Based Analysis of the Semantic Verbal Fluency Task

Alexandra König[a] Nicklas Linz[b] Johannes Tröger[b] Maria Wolters[c] Jan Alexandersson[b] Phillipe Robert[a]

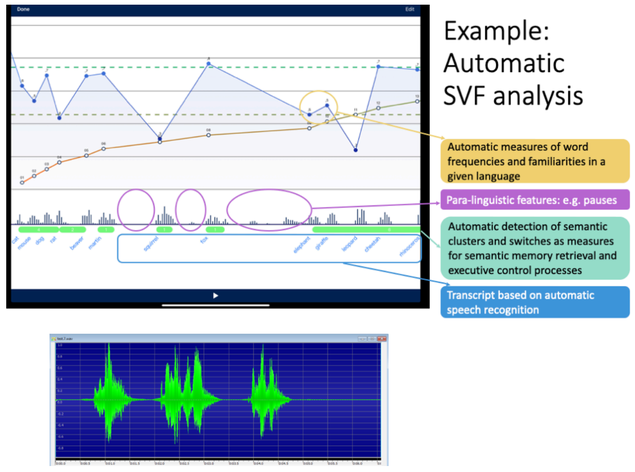
[a]Memory Clinic, Association IA, CoBTek Lab, CHU Université Côte d'Azur, Nice, France; [b]German Research Center for Artificial Intelligence (DFKI), Saarbrücken, Germany; [c]School of Informatics, University of Edinburgh, Edinburgh, UK

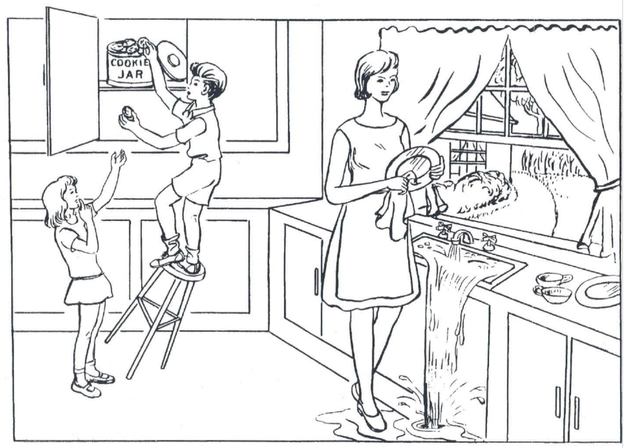
CoBTek Cognition Behaviour Technology

COGNITION

Aide à la standardisation et l'analyse des résultats

Delta





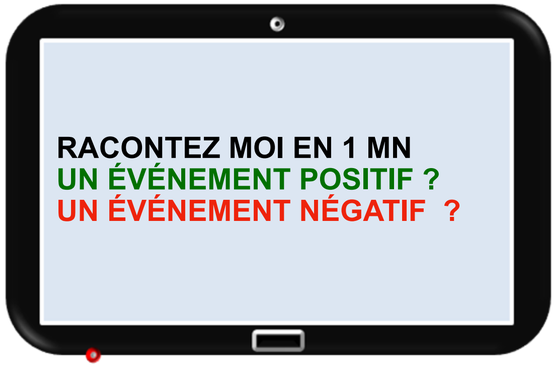
Musée
Limonade
Sauterelle
Passoire
Camion

FLUENCE VERBALE

DESCRIPTION D'IMAGE

RAPPEL

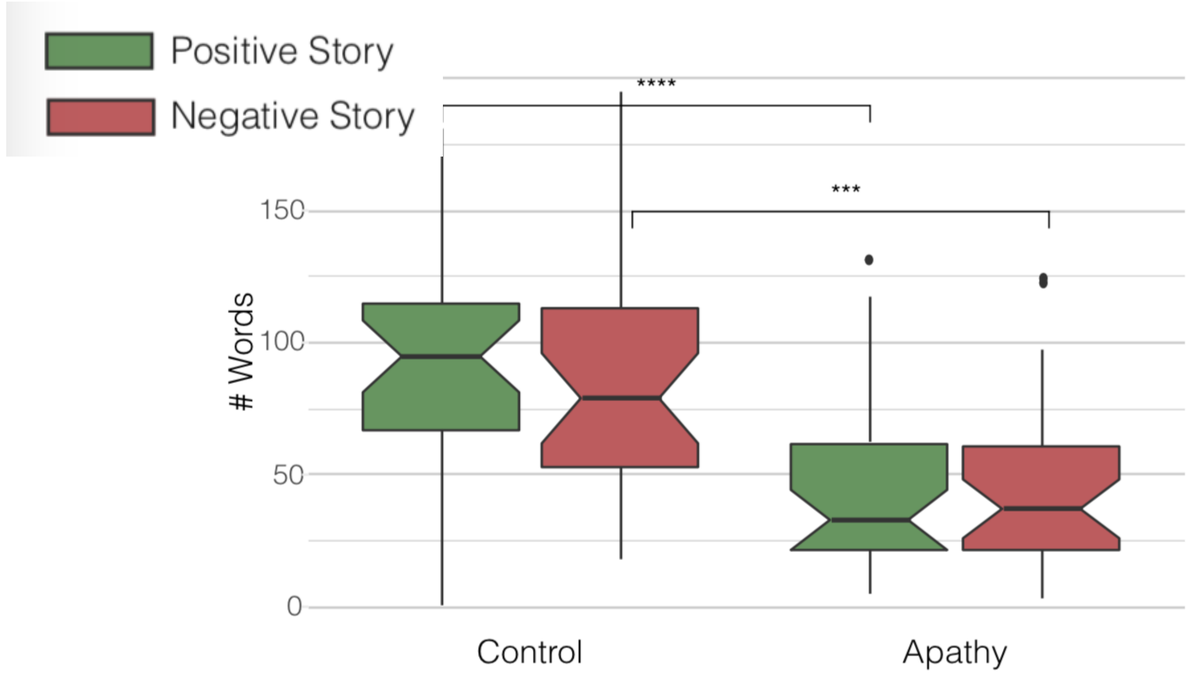
Audio



Detecting Apathy in Older Adults with Cognitive Disorders Using Automatic Speech Analysis

Alexandra König[a,b], Nicklas Linz[c], Radia Zeghari[a], Xenia Klinge[c], Johannes Tröger[c], Jan Alexandersson[c] and Philippe Robert[a]

MOTIVATION
EMOTION



Audio

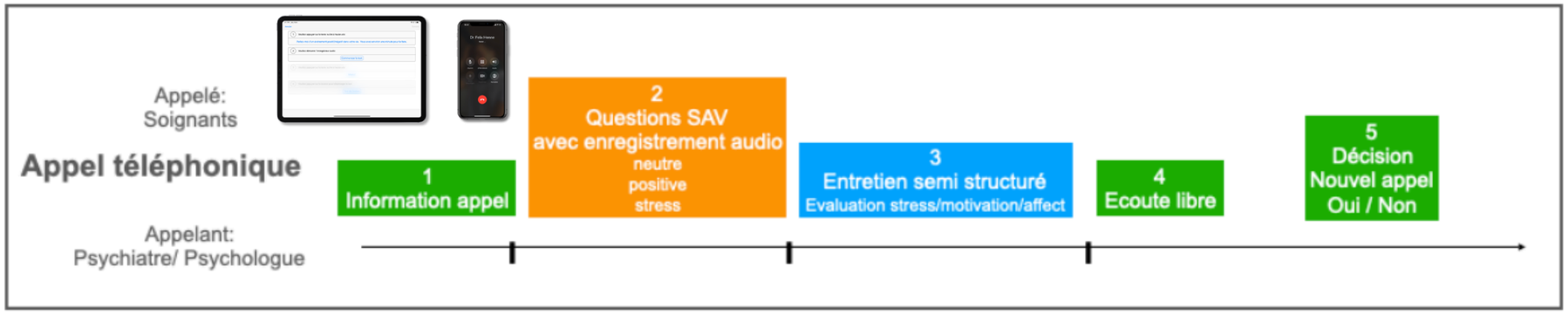
JOURNAL OF MEDICAL INTERNET RESEARCH König et al

Original Paper

Measuring Stress in Health Professionals Over the Phone Using Automatic Speech Analysis During the COVID-19 Pandemic: Observational Pilot Study

Alexandra König[1], PhD; Kevin Riviere[2], MD; Nicklas Linz[3], MSc; Hali Lindsay[4], MSc; Julia Elbaum[2], MD; Roxane Fabre[2], MSc; Alexandre Derreumaux[5], MSc; Philippe Robert[5], MD, PhD

Recherche de mesures de détection du Stress chez les soignants



- Entre le **5 Mai et le 7 Juin 2020**
- 95 appels, **89** appels exploitables du

narqueurs

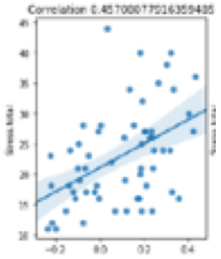
CBTEK
Cognition Behaviour Technology

Marqueurs vocaux prédicteurs du stress

Femme

MFCC

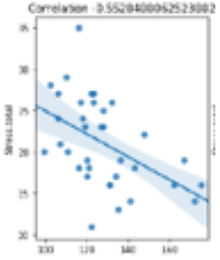
Mel-Fréquence Cepstral Coefficients



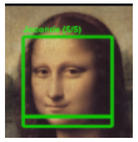
Homme

Fréquence fondamentale

Hauteur du son



Vidéo



COGNITION

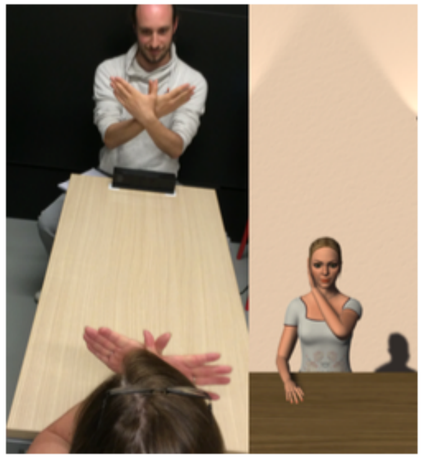


Fig. 7: Examples of challenging cases in Praxis gesture dataset

Fig. 6: The virtual avatar guides patients in a virtual environment.

TABLE I: List of the available gestures in the dataset corresponding information.

Category	Uni/Bimanual	ID	Type	Description
Abstract	Unimanual	A1-1	Static	Left hand on left ear
		A1-2	Static	Left hand on right ear
		A1-3	Static	Right hand on right ear
		A1-4	Static	Right hand on left ear
		A1-5	Static	Index and baby finger on table
	Bimanual	A2-1	Static	Stick together index and baby fingers
		A2-2	Dynamic	Hands on table, twist toward body
		A2-3	Static	Bird
		A2-4	Static	Diamond
		A2-5	Static	ring together
Symbolic	Unimanual	S1-1	Static	Do a military salute
		S1-2	Static	Ask for silence
		S1-3	Static	Show something smells bad
		S1-4	Dynamic	Tell someone is crazy
		S1-5	Dynamic	Blow a kiss
	Bimanual	S2-1	Dynamic	Twiddle your thumbs
		S2-2	Static	Indicate there is unbearable noise
		S2-3	Static	Indicate you want to sleep
		S2-4	Static	Pray
Pantomime	Unimanual	P1-1	Dynamic	Comb hair
		P1-2	Dynamic	Drink a glass of water
		P1-3	Dynamic	Answer the phone
		P1-4	Dynamic	Pick up a needle
		P1-5	Dynamic	Smoke a cigarette
	Bimanual	P2-1	Dynamic	Unscrew a stopper
		P2-2	Dynamic	Play piano
		P2-3	Dynamic	Hammer a nail
		P2-4	Dynamic	Tear up a paper
		P2-5	Dynamic	Strike a match

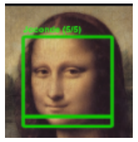
IV. Experiments and Analysis

A. Dataset

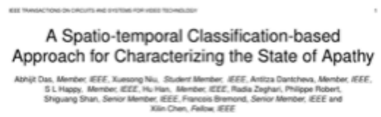
PRAXIS: Towards Automatic Cognitive Assessment Using Gesture Recognition

Farhood Negin, Pau Rodriguez, Michal Koperski, Adlen Kerboua, Jordi Gonzàlez, Jeremy Bourgeois, Emmanuelle Chapoulie, Philippe Robert, and Francois Bremond,

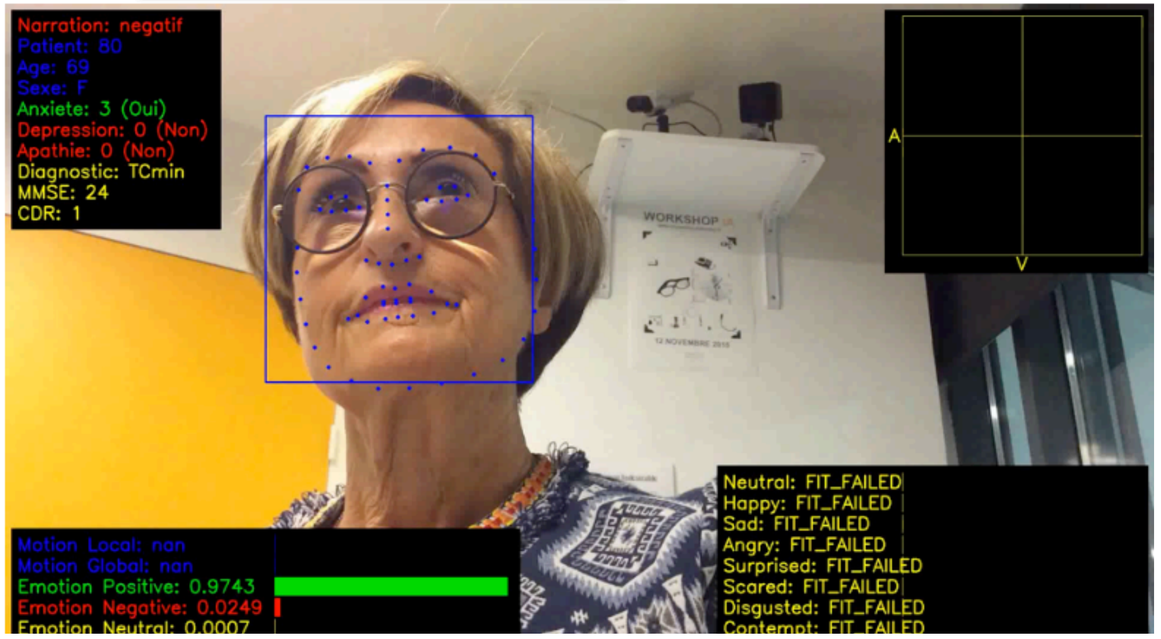
Vidéo



MOTIVATION
EMOTION



A Spatio-temporal Classification-based Approach for Characterizing the State of Apathy



Internet Interventions 34 (2023) 100660

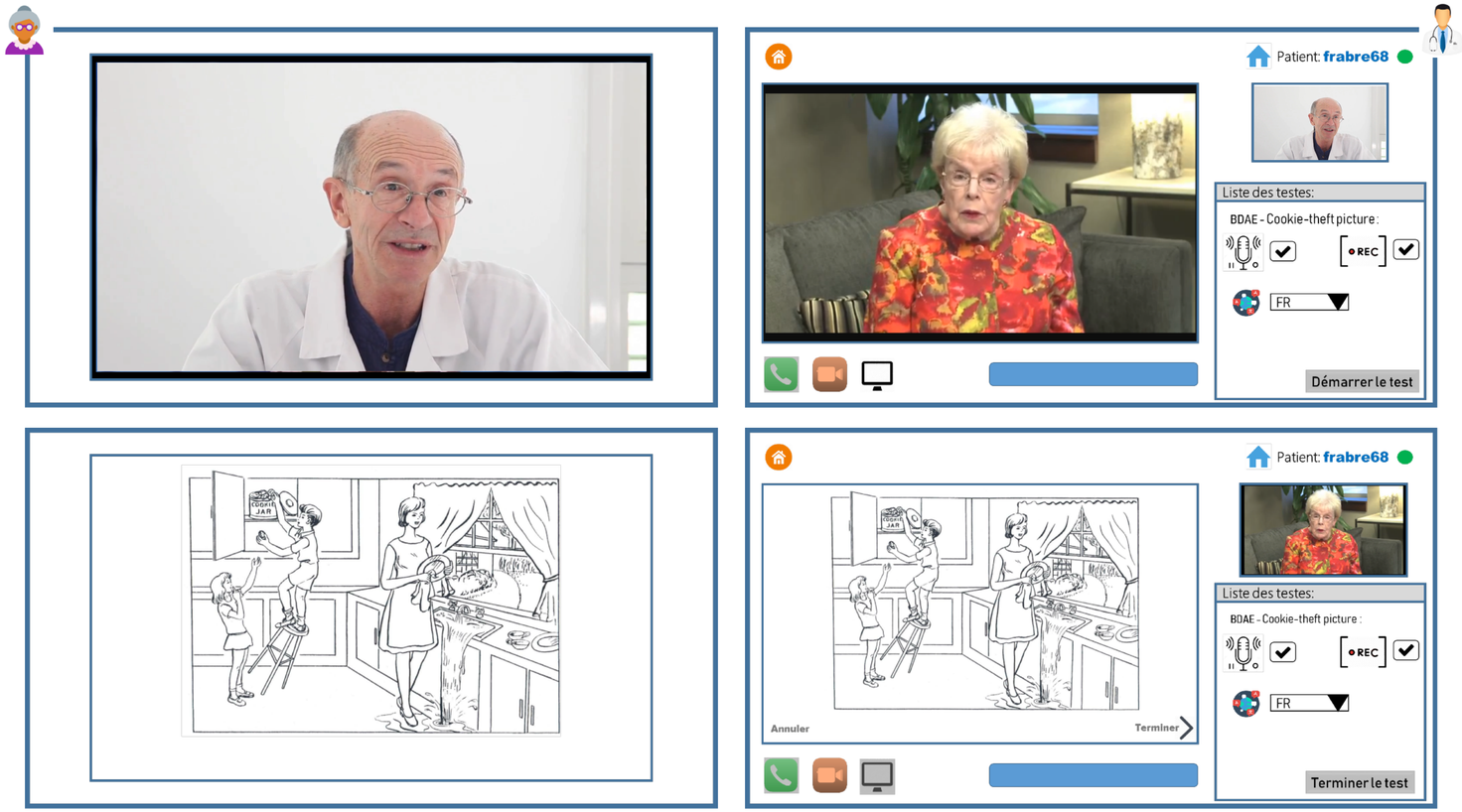
Contents lists available at ScienceDirect

Internet Interventions

journal homepage: www.elsevier.com/locate/invent

Teleconsultations for mental health: Recommendations from a Delphi panel

Valeria Manera [a,b,c,*], Claudia Partos [d], Olivier Beauchet [e,f,g,h], Michel Benoit [a,y], Benjamin Dupetit [j], Julia Elbaum [a,i], Roxane Fabre [k], Morgane Gindt [a,l], Auriane Gros [a,b,c,i], Rachid Guerchouche [a,b,m], Stefan Klöppel [n], Alexandra König [a,b,m], Annick Martin [o], Aurélie Mouton [a,b,i], Marie-Pierre Pancrazi [p], Antonios Politis [q], Gabriel Robert [r,s], Guillaume Sacco [a,b,i], Sabrina Sacconi [t], Kim Sawchuk [u], Fabio Solari [v], Lucille Thiebot [w], Pietro Davide Trimarchi [x], Radia Zeghari [a,l], Philippe Robert [a,b]



Open access Protocol

BMJ Open

Remote cognitive assessment of older adults in rural areas by telemedicine and automatic speech and video analysis: protocol for a cross-over feasibility study

Alexandra König [1,2] Radia Zeghari,[3] Rachid Guerchouche,[1,2] Minh Duc Tran,[1] François Bremond,[1] Nicklas Linz,[3] Hali Lindsay,[4] Kai Langel,[5] Inez Ramakers,[6] Pascale Lemoine,[7] Vincent Bultingaire,[7] Philippe Robert[3]

diagnostics MDPI

Research Article

Feasibility study of an internet-based platform for tele-neuropsychological Assessment of Elderly in Remote areas

Radia Zeghari[1*], Rachid Guerchouche[1,2], Minh Duc Tran[2], François Bremond[1,2], Kai Langel[3], Inez Ramakers[4], Nathalie Amiel[3], Pascale Lemoine[3], Vincent Bultingaire[3], Valeria Manera[1], Philippe Robert[1], Alexandra König[1,2]

Internet Interventions 34 (2023) 100660

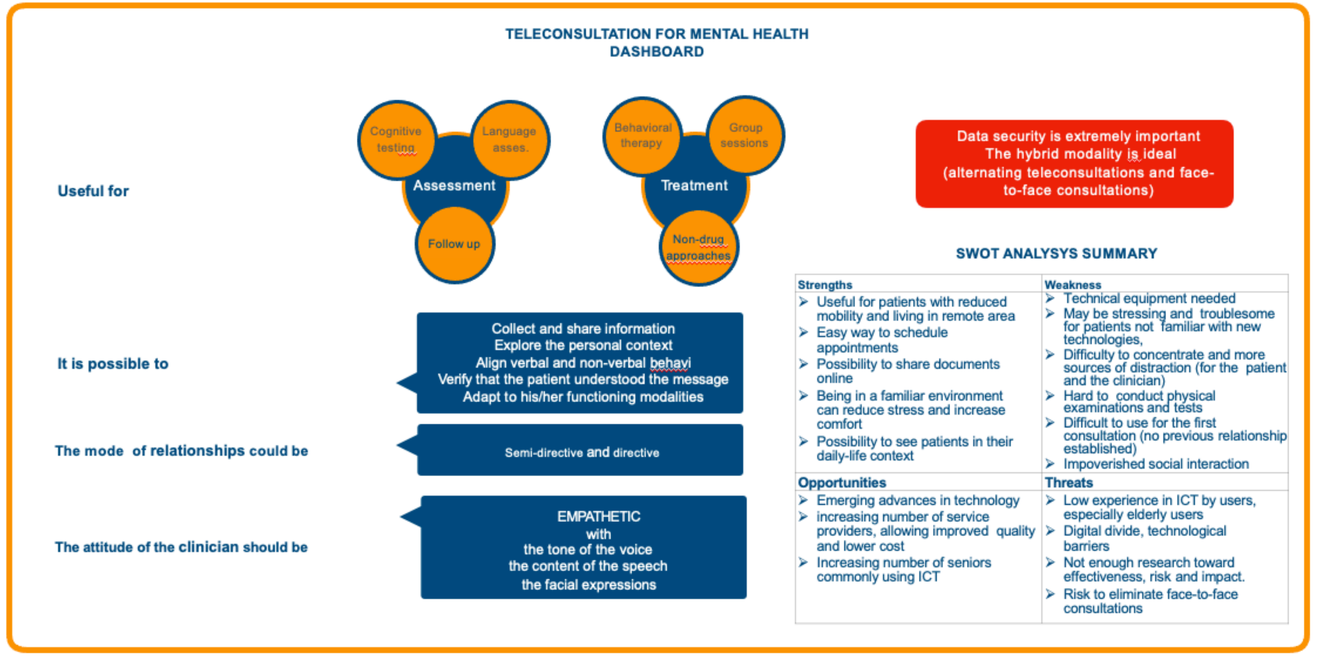
Contents lists available at ScienceDirect

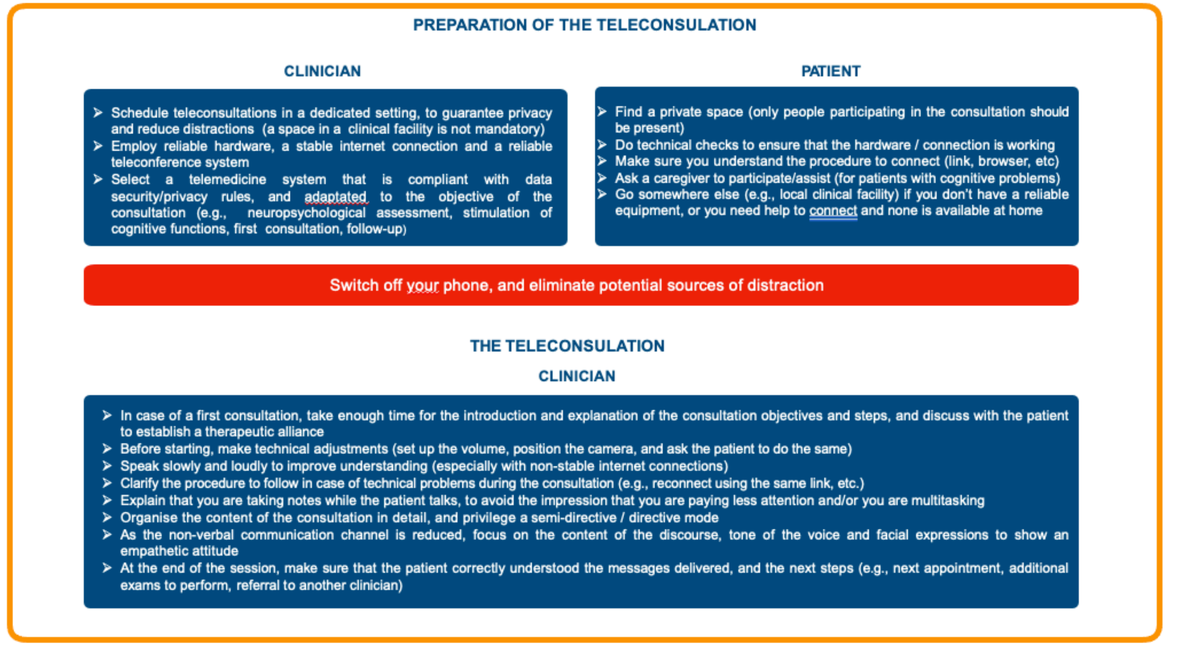
Internet Interventions

journal homepage: www.elsevier.com/locate/invent

Teleconsultations for mental health: Recommendations from a Delphi panel

Valeria Manera [a,b,c,*], Claudia Partos [d], Olivier Beauchet [e,f,g,h], Michel Benoit [a,y], Benjamin Dupetit [j], Julia Elbaum [a,i], Roxane Fabre [k], Morgane Gindt [a,l], Auriane Gros [a,b,c,i], Rachid Guerchouche [a,b,m], Stefan Klöppel [n], Alexandra König [a,b,m], Annick Martin [o], Aurélie Mouton [a,b,i], Marie-Pierre Pancrazi [p], Antonios Politis [q], Gabriel Robert [r,s], Guillaume Sacco [a,b,i], Sabrina Sacconi [t], Kim Sawchuk [u], Fabio Solari [v], Lucille Thiebot [w], Pietro Davide Trimarchi [x], Radia Zeghari [a,l], Philippe Robert [a,b]





L’application TeachMod

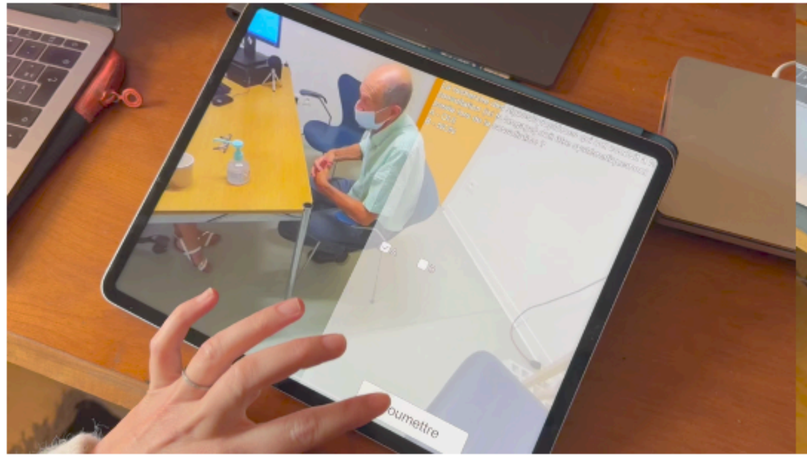
Tournage avec caméra 360° au Centre Mémoire de Ressources et de Recherche à l’Institut Claude Pompidou
1 scénario consultation directement au centre
1 scénario consultation à distance

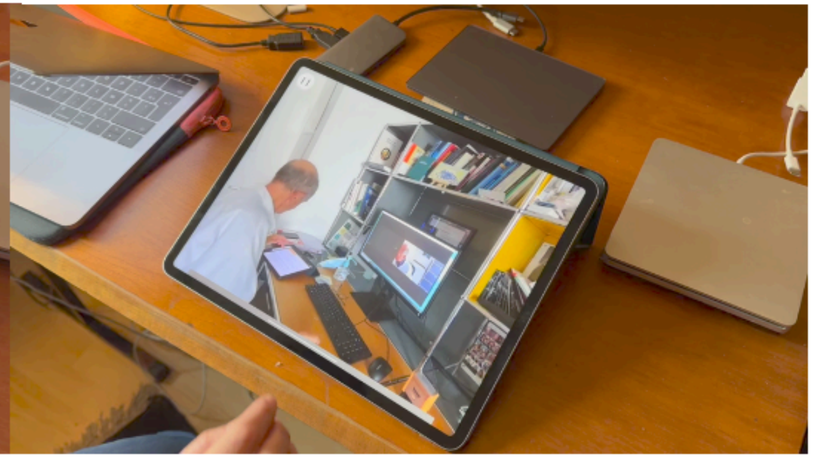
TeachMod

Scénario en présence

Scénario à distance

Durée d vidéo 8mn / scénario complet: 12 mn / Durée cours: 60 mn





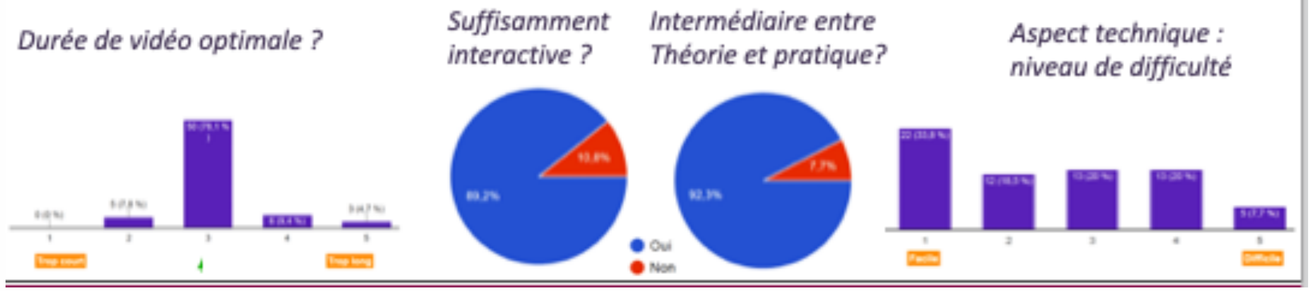
TeachMod

VLab

COBTEK
Cognition Behaviour Technology

TEACHMOD

TeachMod a été utilisé par **65 étudiants du Master 2 d'Orthophonie de l'Université de Côte d'Azur. La présentation de TeachMod et l'ensemble des résultats de l'enquête sont disponibles sur** www.innovation-alzheimer.fr/relation-soigne-soignant/



POUR UTILISER TEACHMOD

- Windows : http://innovation-alzheimer.fr/apps/teachmod-windows-0.3.zip
- Mac, iPhone, iPad : disponible sur l'AppStore https://apps.apple.com/us/app/teachmod/id1589411174
- Android :
  - Sur le Play Store : https://play.google.com/store/apps/details?id=fr.innovationalzheimer.teachmod
  - Télécharger l'APK : http://innovation-alzheimer.fr/apps/teachmod-android-0.3.apk

DISPONIBLE SUR Google Play

Disponible sur App Store

http://cobtek.fr/applications/

Centre de Recherche d'Enseignement pour les Aidants *Sophia Antipolis*

Recherche

Enseignements

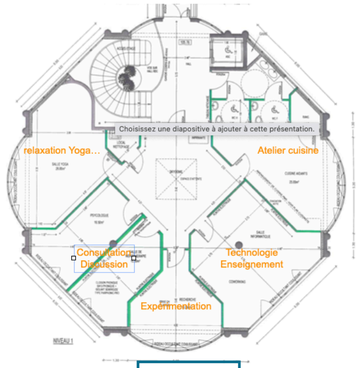
Partage / Événements

Aidants, patients étudiants, professionnels santé, chercheurs

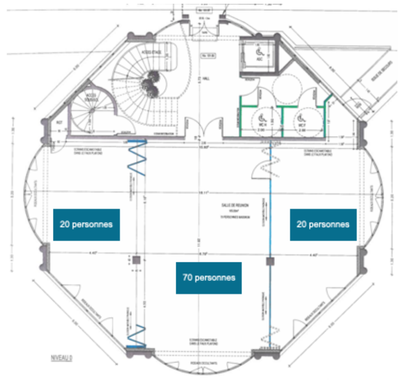
VIRTUEL

SITE WEB

jlncreapolis.fr



Niveau 1



Salles de réunion / meeting rooms / RDC

1°T 2024

1° T 2025

Comment travailler dans le domaine de la santé ?

Digitally augmented, parent-led CBT versus treatment as usual for child anxiety problems in child mental health services in England and Northern Ireland: a pragmatic, non-inferiority, clinical effectiveness and cost-effectiveness randomised controlled trial

Cathy Creswell, Lucy Taylor, Sophie Giles, Sophie Howitt, Lucy Radley, Emily Whitaker, Emma Brooks, Fauzia Knight, Vanessa Raymont, Claire Hill, James van Santen, Nicola Williams, Sam Mort, Victoria Harris, Shuye Yu, Jack Pollard, Mara Violato, Polly Waite*, Ly-Mee Yu**

Summary

Background Anxiety problems are common in children, yet few affected children access evidence-based treatment. Digitally augmented psychological therapies bring potential to increase availability of effective help for children with mental health problems. This study aimed to establish whether therapist-supported, digitally augmented, parent-led cognitive behavioural therapy (CBT) could increase the efficiency of treatment without compromising clinical effectiveness and acceptability.

Methods We conducted a pragmatic, unblinded, two-arm, multisite, randomised controlled non-inferiority trial to evaluate the clinical effectiveness and cost-effectiveness of therapist-supported, parent-led CBT using the Online Support and Intervention (OSI) for child anxiety platform compared with treatment as usual for child (aged 5–12 years) anxiety problems in 34 Child and Adolescent Mental Health Services in England and Northern Ireland. We examined acceptability of OSI plus therapist support via qualitative interviews. Participants were randomly assigned (1:1) to OSI plus therapist support or treatment as usual, minimised by child age, gender, service type, and baseline child anxiety interference. Outcomes were assessed at week 14 and week 26 after randomisation. The primary clinical outcome was parent-reported interference caused by child anxiety at week 26 assessment, using the Child Anxiety Impact Scale–parent report (CAIS-P). The primary measure of health economic effect was quality-adjusted life-years (QALYs). Outcome analyses were conducted blind in the intention-to-treat (ITT) population with a standardised non-inferiority margin of 0·33 for clinical analyses. The trial was registered with ISRCTN, 12890382.

Findings Between Dec 5, 2020, and Aug 3, 2022, 706 families (706 children and their parents or carers) were referred to the study information. 444 families were enrolled. Parents reported 255 (58%) child participants' gender to be female, 184 (41%) male, three (<1%) other, and one (<1%) preferred not to report their child's gender. 400 (90%) children were White and the mean age was 9·20 years (SD 1·79). 85% of families for whom clinicians provided information in the treatment as usual group received CBT. OSI plus therapist support was non-inferior for parent-reported anxiety interference on the CAIS-P (SMD 0·01, 95% CI −0·15 to 0·17; p<0·0001) and all secondary outcomes. The mean difference in QALYs across trial arms approximated to zero, and OSI plus therapist support was associated with lower costs than treatment as usual. OSI plus therapist support was likely to be cost effective under certain scenarios, but uncertainty was high. OSI plus therapist support acceptability was good. No serious adverse events were reported.

Interpretation Digitally augmented intervention brought promising savings without compromising outcomes and as such presents a valuable tool for increasing access to psychological therapies and meeting the demand for treatment of child anxiety problems.

Lancet Psychiatry 2024; 11: 193–209

Published Online February 6, 2024 https://doi.org/10.1016/S2215-0366(23)00429-7

See **Comment** page 161

*Joint senior authors

Departments of Experimental Psychology and Psychiatry (Prof C Creswell PhD, L Taylor MSc, L Radley MSc, E Whitaker MSc, P Waite PhD), **Department of Psychiatry, Warneford Hospital** (E Brooks MSc, V Raymont MB ChB), **Nuffield Department of Primary Care Health Science** (J van Santen MSc, S Mort PGCert, V Harris PhD, N Williams MSc, L-M Yu DPhil), **and Nuffield Department of Population Health** (S Yu PhD, J Pollard MSc, M Violato PhD), **University of Oxford, Oxford, UK; Sussex Partnership NHS Foundation Trust, Worthing, UK** (S Giles MSc); **Oxford Health NHS Foundation Trust, Abingdon, UK** (S Howitt MSc); **Centre for Psychological Sciences, University of Westminster, London, UK** (F Knight PhD); **School of Psychology & Clinical Language Sciences, University of Reading, Reading, UK** (C Hill PhD)

Quelles sont les compétences techniques d'un développeur IA ?

L'exercice du métier développeur data IA nécessite des compétences techniques solides en intelligence artificielle, en programmation et en logique.

Ce professionnel doit :

- Savoir visualiser, analyser et **exploiter les données** ;
- Être à l'aise dans la conception et dans l'optimisation de base de données ;
- Maîtriser les langages de programmation orientés objet ;
- Maîtriser le **fonctionnement des algorithmes** de machine learning ;
- Avoir des connaissances solides dans l'orchestration de big data - Savoir développer des applications à l'aide des technologies de l'intelligence artificielle ;
- Maîtriser des langues de programmation populaires, notamment R, Python, Java et C++ ;
- Développer des compétences dans la machine learning et dans le deep learning ;
- Avoir une bonne connaissance des **réseaux de neurones** ;
- Être à l'aise avec les mathématiques, les probabilités et les statistiques ;
- Connaitre les normes et les droits sur l'intelligence artificielle
- Développer des modèles d'apprentissage.

Quelles sont les qualités requises pour réussir dans ce domaine ?

Réussir sa carrière en intelligence artificielle nécessite :

- Un sens aigu de l'organisation pour être à l'aise dans la programmation ;
- Une bonne aptitude dans l'analyse des données ;
- De la rigueur, du pragmatisme et de la patience ;
- Des aptitudes pour travailler en équipe ;
- Un bon sens du relationnel ;
- De la curiosité pour être au courant des **dernières avancées technologiques** de l'intelligence artificielle ;
- De la créativité pour créer des solutions innovantes ;
- Des aptitudes pour travailler sous pression.

Medical News & Perspectives | AI IN MEDICINE

AI Developers Should Understand the Risks of Deploying Their Clinical Tools, MIT Expert Says

Samantha Anderer; Yulin Hswen, ScD, MPH

This conversation is part of a series of interviews in which JAMA *Editor in Chief Kirsten Bibbins-Domingo, PhD, MD, MAS, and expert guests explore issues surrounding the rapidly evolving intersection of artificial intelligence (AI) and medicine.*

AI applications for health care should be designed to function well in different settings and across different populations, says Marzyeh Ghassemi, PhD (**Video**), whose work at the Massachusetts Institute of Technology (MIT) focuses on creating "healthy" machine learning (ML) models that are "robust, private, and fair." The way AI-generated clinical advice is presented to physicians is also important for reducing harms, according to Ghassemi, who is an assistant professor at MIT's Department of Electrical Engineering and Computer Science and Institute for Medical Engineering and Science. And, she says, developers should be aware that they have a responsibility to clinicians and patients who could one day be affected by their tools.

Multimedia

Medical News website

JAMA Editor in Chief Kirsten Bibbins-Domingo, PhD, MD, MAS, recently spoke with Ghassemi about "ethical machine learning," the computer scientist's decision to opt out of AI in her own health care, and more.

The following interview has been edited for clarity and length.

DR BIBBINS-DOMINGO: You have a research lab, Healthy ML. It specializes in examining biases in artificial intelligence, and you're specifically interested in its applications in clinical practice. I'd love to hear how you got into the very specific area.

DR GHASSEMI: At the end of my PhD, we found out that [machine learning] models tend not to work as well in all groups. And that really informs what we do here in my lab today, focusing on how we make sure that models that are developed work robustly. And if you think about robustness, that could mean that it works well in a new environment or across different kinds of people.

DR BIBBINS-DOMINGO: How do you think about the range of reasons why a model might not perform well in one setting vs another or in one group of people vs another?

DR GHASSEMI: I try to think about it within the pipeline that all models are developed in. And this is not just in health care. This is for any machine learning model that might be developed and deployed in any human-facing setting. You choose a problem, collect some data, define a label, develop an algorithm, and then deploy it. In each part of that pipeline, there are reasons that your model might not perform as well. For problem selection, what we choose to fund and what we choose to work on is often biased. We tend to look at problems that are easy to address where there are more data readily available that can be correlated with different metrics of social status, or privilege, or just where funding tends to be allocated to.

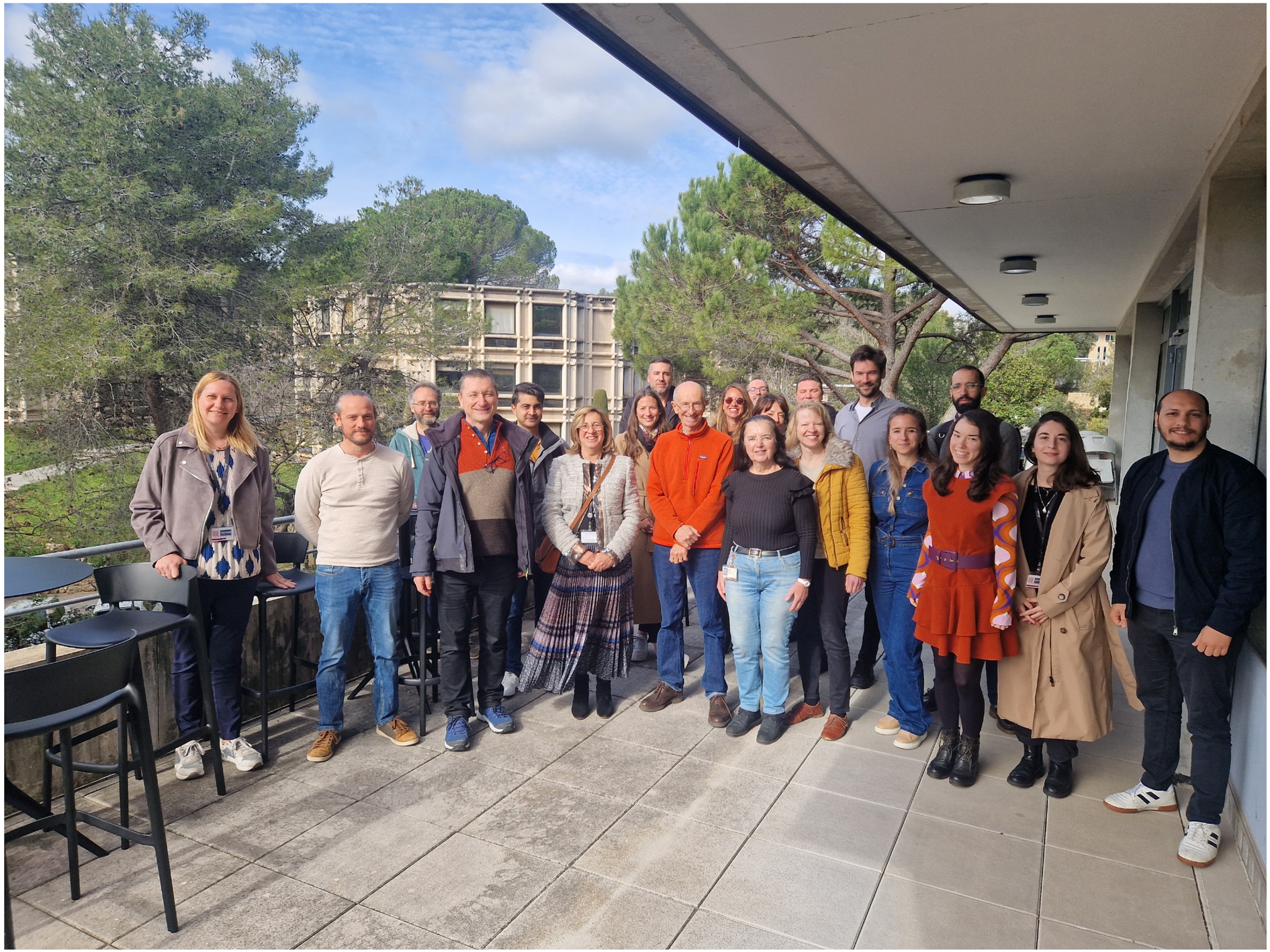
For example, diseases that are disproportionately affecting people who are biologically female at birth tend to be understudied. And if we're collecting data from these human sources, it's probably going to have some bias in it just because of the way that humans interact with one another. Just by collecting data from a human process, you're going to have some potential performance issues. We probably want machine learning models to replicate the very best health care practices that we see now, but if we take a random sample of data from thousands of hospitals and say, "Perform the way that an average doctor is performing on an average day," we might get some behaviors that we don't want to extend.

When we define a label, that's another way that bias can be injected into the learning process. It's a true-false label. We never contextualize it with the choice that's being made or the human rule that's being applied. When you collect labels in this descriptive way but then train a machine learning model, all of those machine learning models become much harsher. They have a much higher false-positive rate.

DR BIBBINS-DOMINGO: You use the term *ethical machine learning*. I'd love you to define what that term means for you and help us to understand it in the context of medical practice.

DR GHASSEMI: I think for me as a technical person, *ethical machine learning* means recognizing your responsibility to end users that might potentially be impacted by the models that you're developing, the technology that you're releasing. And I think there

CBTEK
Cognition Behaviour Technology



Clinique

Technologie