

MIHR: A Human-Robot Interaction Model

Jesús Pérez, Jose Aguilar, and Eladio Dapena

Abstract— The interactions between people and social robots have generated positive effects on people of different ages in diverse contexts. A model of the interaction process is important to understand the person who interacts, in order to manage the internal dynamics of interaction in the social robots. There are models that describe the interaction between humans and machines, but they don't integrate the three most important elements to be considered during the interactions by the social robots: the modalities of human communication, the capacity of adaptation, and the expression of emotions. In this paper, a review of the interaction models between people and social robots is made, in order to analyze what has been done about these three important elements of the interaction. Then, it is proposed a Human-Robot Interaction Model (MIHR) based on a Human-Human Interaction Model (MIHH) previously developed, which integrates the main elements to be considered during the interactions by the social robots.

Index Terms— Human-Robot Interaction, Social Interaction, Social Robotics.

I. INTRODUCCIÓN

LA INTERACCIÓN humano-robot es una forma de interacción humano-computador [1], cuyo propósito es proporcionar interacciones intuitivas [2], que permitan la colaboración, comunicación y cooperación entre personas y robots [3]. En general, la interacción humano-robot facilita el intercambio de información entre humanos y varios tipos de máquinas, como los computadores y los dispositivos electrónicos [4].

Actualmente, hay un creciente interés en el desarrollo del ámbito donde los robots interactúan socialmente con los humanos [5-7], porque los robots que realizan tareas útiles para el bienestar de los humanos, interactúan directamente [8]. Algunos investigadores plantean que los robots serán aceptados en la vida cotidiana, si poseen la capacidad de reconocer el comportamiento de las personas y adaptar su comportamiento social para que sea apropiado [9]. En ese sentido, el objetivo es crear sistemas que interactúen con las personas de la manera más natural posible [10], por lo tanto, muchos de los investigadores utilizan modelos y técnicas inspirados en la comunicación entre personas [11]. El modelo presentado en [12] será utilizado como referencia en esta investigación.

La interacción social entre personas y robots ha generado efectos positivos en diferentes personas y contextos, desde incrementar las habilidades sociales en niños con trastorno del espectro autista [13] y motivar el aprendizaje de estudiantes

universitarios [14], hasta incentivar a personas de tercera edad a compartir sus experiencias [15]. De allí, el interés del estudio de la representación de ese proceso de interacción.

En la literatura se encuentran varios modelos que describen la interacción entre humanos y máquinas, desde distintas perspectivas [16-20]. Dicha revisión ha permitido determinar que el comportamiento de los robots se rige por tres elementos principales: la modalidad de la comunicación humana, la adaptación de la interacción y la expresión de emociones. La modalidad de la comunicación humana, según [16], es mediante entradas y salidas multimodales. Por otro lado, [17] describe la comunicación a través de la voz y los gestos, y en [18] se menciona como posibles formas de comunicación la gráfica, hablada, gestual, entre otras. A su vez, la capacidad de adaptación se describe en [16] como la adaptación de los mensajes según el usuario y el contexto de interacción, en [18] se define en términos de los contenidos y presentación de la interfaz al usuario, entre otros aspectos, y en [19] se establece según el modelo del usuario. Finalmente, la expresión de emociones del computador se define en [20] como un atributo de los robots que se reflejan en sus acciones comunicativas, y en [21, 22] se definen en función de los estados posibles de un robot (cargado, sin carga, etc.).

Dado que esos tres elementos principales aparecen dispersos en diferentes modelos, sin que haya uno que los integre, surge la motivación de proponer un Modelo de Interacción Humano-Robot (MIHR) que los incorpore a los tres, para poder ser usado por un robot. Las secciones restantes de este trabajo se organizan de la siguiente manera: la segunda sección describe los principales elementos que rigen los comportamientos de los robots sociales; la sección tres explica el modelo propuesto en este artículo; en la sección cuatro se describe el experimento que verifica las premisas que soportan el modelo interno del robot propuesto y se presentan los resultados; en la quinta sección se discuten los resultados del experimento; en la sexta sección se compara el modelo propuesto con trabajos previos; y finalmente, la última sección contiene las conclusiones de esta investigación.

II. INTERACCIÓN HUMANO-ROBOT

A. Modalidad de Comunicación Humana

Las personas no sólo son capaces de tener relaciones sociales con otras personas, sino que también tienen la habilidad de mantener relaciones sociales con máquinas [11]. Además, las personas responden a los robots sociales de manera similar a como responden a otras personas, especialmente, si los robots se comunican utilizando el mismo lenguaje de las personas [23]. Por lo tanto, las normas sociales de la interacción humana constituyen el núcleo de la interacción humano-robot [24], y los comportamientos de los

J. Pérez, LaSDAI, Universidad de Los Andes, Venezuela, jesuspangulo@ula.ve

J. Aguilar, CEMISID, Universidad de Los Andes, Venezuela, aguilar@ula.ve

E. Dapena, LaSDAI, Universidad de Los Andes, Venezuela, eladio@ula.ve

robots se diseñan según comportamientos humanos [25].

Los algoritmos que gobiernan los robots están orientados a desarrollar interacciones directas que sean seguras para los humanos [4], por lo tanto, no deben ocurrir contactos indeseados o inesperados entre los robots y los humanos [3], es decir, el robot debe garantizar movimientos seguros y socialmente aceptables [26]. En ese sentido, los sensores y los actuadores son diseñados pensando en la seguridad de los humanos. Debido a que la comunicación en lenguaje natural usa diferentes formas para comunicar una misma intención [5], los robots se comunican de diferentes maneras, tales como: voz [8], mirada [2], gestos [10], y contacto físico [9].

B. Adaptación de la Interacción

Las interacciones con personas requieren considerar que una misma persona puede tener un comportamiento inconsistente en el tiempo, incluso, mientras interactúa con el robot [27]. Por lo tanto, para mantener interacciones eficaces y atractivas, los robots deben ser capaces de percibir estas diferencias, y personalizar su comportamiento en consecuencia, en lugar de proporcionar interacciones genéricas que se mantengan constantes en todas las personas o en una misma persona [28]. En ese sentido, los investigadores tratan de encontrar mejores formas de percibir esas diferencias, desde reconocer a la persona [29], hasta estimar sus expectativas y grados de satisfacción [30].

En general, los robots deben adaptar sus comportamientos en aras de establecer interacciones apropiadas. Algunas de las adaptaciones que realizan los robots son: energía de la voz en función de la voz del usuario [31], gestos según lo que percibe por la cámara [32], secuencia de movimientos en función de la postura del usuario [33], tipo de respuesta (positiva, neutral o negativa) en función de las expresiones faciales [34], y momento de ejecutar las acciones en función de los gestos de los usuarios [35].

C. Expresión de Emociones

Las investigaciones proporcionan evidencia de que las personas pueden establecer algún tipo de conexión emocional con los robots [11], por lo tanto, los robots sociales necesitan entender y expresar estados afectivos para mejorar sus interacciones [36]. En general, las emociones pueden ser un factor determinante para el comportamiento y la toma de decisiones del robot [21-22, 37]. Entre las estrategias utilizadas para reconocer emociones están: procesamiento de la voz [38], análisis del lenguaje natural en texto [39] y análisis visual [40]. Las investigaciones en reconocimiento de las emociones han evolucionado del análisis unimodal convencional al análisis multimodal [41-42], mediante la integración de todas las modalidades individuales en una sola representación. Aunque, según [43], simplemente fusionando varias señales no es suficiente para inferir estados emocionales, sino que a través de las interacciones entre ellas un patrón emergerá.

Dado que las emociones juegan un papel importante en la interacción entre personas, éstas esperan que los robots también expresen emociones [44]. La expresión de emociones

por parte de un robot puede tener tres propósitos [23]: primero, ayudar a que el robot comunique su estado emocional; segundo, alentar los comportamientos deseados de las personas; y tercero, ayudar a las personas a conectarse emocionalmente con el robot. Según [44], los canales más utilizados por los robots para la expresión de las emociones son: expresiones faciales, lenguaje natural y gestos. Sin embargo, según [23] se ha estudiado la expresión de emociones tanto unimodal como multimodal mediante otros modos, tales como: colores, sonidos y vibraciones.

III. MODELO PROPUESTO (MIHR)

El proceso de interacción social entre personas y robots debe ser lo más parecido posible al proceso de interacción entre personas, por lo tanto, el punto de partida es el Modelo de Interacción Humano-Humano (MIHH) presentado en [12]. En general, el MIHH está constituido por dos niveles: el nivel interno, que describe el proceso interno de cada persona desde que percibe información hasta que genera respuestas, y el nivel externo, que describe cómo puede ser alterada la información desde que es emitida por una persona hasta que es recibida por otra. En ese mismo orden de ideas, se propone el MIHR como una extensión del MIHH, donde se agrega el nivel interno del robot para describir el proceso que éste realiza desde que recibe información hasta que genera respuestas (ver Fig. 1). En las siguientes secciones se presenta en detalle cada uno de sus componentes, y las interacciones entre ellos.

A. Nivel Interno del Humano

El nivel interno del humano incluye los elementos que hacen que cada persona sea particular durante las interacciones (condiciones fisiológicas, estados afectivos e influencia del inconsciente). Este nivel define tres tipos de comportamientos en las personas: racionales (coherente con los objetivos), irracionales (incoherente con los objetivos) y automáticos (reacciones instintivas). Además, también permite los siguientes estilos de comunicación en cada persona: inhibido (la persona se preocupa por satisfacer a los demás), agresivo (la persona defiende sus derechos de una manera que irrespeta los derechos de los demás), y asertivo (la persona defiende sus derechos y se expresa respetando los derechos de los demás). En [12] se encuentran más detalles.

El nivel interno del humano está constituido por cuatro módulos (ver Fig. 1): físico, afectivo, cognitivo, y conductual. El nivel interno se alimenta de la información que proviene del nivel externo, modificándola según los sentidos de la persona, y activando estados afectivos según la interpretación dada. Luego, se hace una nueva interpretación según los filtros cognitivos de la persona, para determinar la intención de comunicación de la otra persona, activando estados afectivos nuevamente. Después, se genera el objetivo y el estilo de comunicación, a partir de los aportes realizados inconscientemente y conscientemente, y de los estados afectivos. Finalmente, a la respuesta se le asignan los componentes conductuales, según el objetivo y el estilo de comunicación, restringidos por las condiciones fisiológicas de la persona.

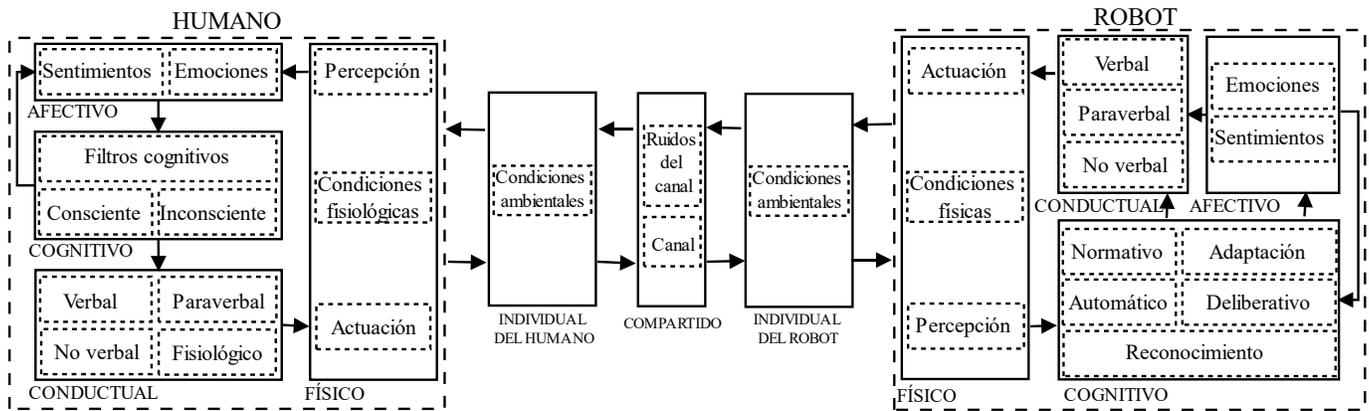


Fig. 1. Modelo de Interacción Humano-Robot (MIHR).

B. Nivel Externo

El proceso de interacción consiste en el envío de información de la persona al robot y viceversa. Esa información es enviada a través de un canal de comunicación, y por lo tanto, puede ser alterada por el canal durante el recorrido. Incluso, la información puede ser alterada durante el camino que recorre desde que es generada por el emisor, hasta que es incorporada al canal de comunicación. Además, puede ser alterada en el proceso de recepción, mientras es desincorporada del canal de comunicación. El nivel externo se comporta de la misma manera para el humano y el robot (ambos cumplen los roles de emisor y receptor).

En el nivel externo se encuentran tres módulos (ver Fig. 1): individual del humano, individual del robot y compartido. El nivel externo se alimenta de las respuestas generadas por las personas y los robots; lo primero que hace es alterar la información de entrada según las condiciones ambientales presentes entre el emisor y el canal (módulo individual del emisor). Luego, incorpora la información al canal de comunicación, y la modifica según los ruidos del canal (módulo compartido). Finalmente, desincorpora la información del canal (módulo individual del receptor), y la modifica según las condiciones ambientales presentes entre el canal y el receptor, lo cual representa la salida de este nivel.

C. Nivel Interno del Robot

La organización de los módulos y componentes del nivel interno del robot está basada en los tres elementos descritos en la sección anterior. Esos elementos se representan mediante tres premisas de comportamiento del robot, que se muestran en la Tabla I, junto con su respectiva justificación. El componente normativo perteneciente al módulo cognitivo, es el principal responsable de que se cumpla la premisa 1, porque le provee al componente deliberativo las normas de interacción. El componente de adaptación del módulo cognitivo es el principal responsable de la premisa 2, siendo el encargado de aplicar las adaptaciones a las respuestas. La premisa 3 es responsabilidad del componente de emociones, porque determina la emoción del robot, y además, envía las indicaciones al módulo conductual para que sean expresadas.

TABLA I
PREMISAS DE COMPORTAMIENTO DEL ROBOT

Premisa	Justificación
1. Los robots interactúan según las normas humanas.	Las normas sociales de la interacción humana constituyen el núcleo de la interacción humano-robot [24].
2. Los robots adaptan las interacciones.	Los robots deben ser capaces de percibir diferencias en las personas, y personalizar su comportamiento en consecuencia [28].
3. Los robots expresan emociones durante las interacciones.	La capacidad de los robots para expresar emociones ayuda a su comunicación con el humano [44].

En el nivel interno del robot, aunque se mantienen los módulos del nivel interno del humano (físico, cognitivo, afectivo y conductual), se modifica la interacción entre ellos (ver Fig. 1). En el nivel interno del humano, el módulo físico alimenta al módulo afectivo; ahora bien, en el caso del robot, en su nivel interno, el módulo físico alimenta al módulo cognitivo, debido a que los robots deben realizar las tareas cognitivas inmediatamente después de obtener las señales a través de los sensores, porque eso es lo que permite darle significado a lo que perciben. Además, algunos de los componentes que constituyen los módulos han sido eliminados (fisiológico y filtros cognitivos), o renombrados (automático por inconsciente y deliberativo por consciente), y otros han sido agregados (reconocimiento, normativo y adaptación).

La entrada de este nivel proviene del nivel externo, y es captada a través de los sensores. Seguidamente, son aplicados algoritmos de reconocimiento, con el propósito de interpretar los datos y descubrir la intención de la comunicación. Se gestionan dos tipos de respuestas: automáticas y deliberadas. Las respuestas automáticas se ejecutan inmediatamente después de descubrir que son requeridas. A diferencia, las respuestas deliberadas están basadas en un objetivo de comunicación, que es determinado según la intención, las normas de interacción y el estado afectivo, para luego adaptarlo según la persona, y generar las formas verbales, paraverbales y no verbales, que serán traducidas en señales que entiendan los actuadores para que sean ejecutadas. En general, los estados afectivos del robot influyen en su comportamiento y en la toma de decisiones. En la Tabla II se presenta la descripción de este nivel.

TABLA II
DESCRIPCIÓN DEL NIVEL INTERNO DEL ROBOT

Entrada	1. Información proveniente del nivel externo.
Proceso	<ol style="list-style-type: none"> 1. Captar las señales multimodales a través de los sensores (módulo físico). 2. Detectar los datos relevantes de las señales captadas mediante un pre-procesamiento (módulo físico). 3. Aplicar los algoritmos de reconocimiento sobre los datos (módulo cognitivo). 4. Si identifica eventos que requieren respuesta automática, aplica los pasos 9, 10 y 11 (módulo cognitivo). 5. Identificar la intención de la comunicación según lo reconocido (módulo cognitivo). 6. Actualizar el estado afectivo según la intención identificada (módulo afectivo). 7. Generar un objetivo de comunicación según la intención, el estado afectivo y las normas de interacción (módulo cognitivo). 8. Adaptar la respuesta según el objetivo generado y la persona (módulo cognitivo). 9. Generar las órdenes de los componentes: verbal, paraverbal y no verbal; de la respuesta (módulo conductual). 10. Traducir las órdenes a señales para los actuadores (módulo físico). 11. Enviar las señales a los actuadores (módulo físico).
Salida	1. Ejecución de los actuadores.

1) Módulo Físico

Este módulo está compuesto por tres componentes: percepción, actuación y condiciones físicas. Si la información proviene del módulo individual del robot del nivel externo, el componente de percepción se encarga de capturar las señales según lo indicado por el componente de condiciones físicas (repetibilidad, sensibilidad, etc.) acerca de los sensores, de manera multimodal, para detectar los datos de interés, y enviarlos al componente de reconocimiento del módulo cognitivo. Por otro lado, si la información que recibe este módulo proviene del módulo conductual, el componente de actuación se encarga de traducir las órdenes en señales comprendidas por los actuadores, y enviarlas para que sean ejecutadas según lo indicado acerca de los actuadores por el componente de condiciones físicas (precisión, velocidad, etc.).

2) Módulo Cognitivo

Está compuesto por cinco componentes: reconocimiento, automático, deliberativo, normativo y adaptación. En el componente de reconocimiento se aplican algoritmos de reconocimiento (de personas, temas de conversación, lugares, estados afectivos, eventos, etc.), usando los datos suministrados por el módulo físico. En el componente automático se identifican los eventos que requieren respuesta inmediata, y se envían las respuestas directamente al módulo conductual, con el propósito de que sean ejecutadas rápidamente. En el componente deliberativo se aplican algoritmos que permiten interpretar lo que se ha reconocido en los diferentes modos o canales, para identificar la intención de comunicación, y generar un objetivo de comunicación según las normas de interacción y el estado afectivo. El componente normativo provee las normas sociales de la interacción. En el componente de adaptación, se adecúa la respuesta según el objetivo y la persona. Adicionalmente, el módulo cognitivo envía los resultados de la intención de comunicación identificada al módulo afectivo.

3) Módulo Afectivo

Este módulo tiene el objetivo de gestionar los estados afectivos del robot en términos de emociones y sentimientos. Ambos aspectos son considerados, ya que en la literatura se ha considerado siempre las emociones como el aspecto a modelar en los robots, pero aunque no son populares, hay modelos formales para representar sentimientos en los robots, como el amor, como también, hay evidencia de que las personas pueden percibir sentimientos en los robots [45]. Así, la entrada de este módulo proviene del módulo cognitivo, el cual le envía la intención de comunicación que ha identificado, para que sean actualizados los estados afectivos del robot en función de esa intención, de la persona y de las normas sociales de interacción. Luego, los estados afectivos son enviados a los módulos conductual y cognitivo.

4) Módulo Conductual

El objetivo de este módulo es distribuir en distintos canales o modos, las respuestas y los estados afectivos que serán expresados por el robot. En ese sentido, está constituido por tres componentes: verbal, paraverbal y no verbal. Cada componente se encarga de generar las órdenes respectivas. El componente verbal relacionado con las palabras; el componente paraverbal asociado al volumen de la voz, tono, timbre, etc.; y el componente no verbal relacionado con las expresiones faciales, movimientos, posturas, etc.

De manera general, según el contexto de aplicación, los componentes arquitectónicos del modelo propuesto requieren formalismos matemáticos para su implementación. Por ejemplo, en [46] se diseñan varios componentes arquitectónicos del modelo propuesto para permitir que un robot social ayude a los niños en el aprendizaje de las tablas de multiplicación. Específicamente, se formalizan las interacciones mediante un Proceso de Decisión de Márkov (PDM) (componente Normativo en nuestro modelo) y la adaptación a través del algoritmo Q-learning (componente de Adaptación en el modelo arquitectónico).

IV. EXPERIMENTO

El objetivo del experimento es validar las tres premisas del comportamiento deseado para un robot social (ver Tabla I), que se reflejan en tres de sus componentes del nivel interno: normativo, adaptativo y emociones. La validación de las premisas se hace a través de la percepción de los estudiantes que interactúan socialmente con un robot, por lo tanto, el experimento consiste en proveer las interacciones, y posteriormente, aplicar encuestas a los estudiantes para validar su percepción de las mismas. Los resultados se consideran satisfactorios si se obtiene en cada ítem de la encuesta un porcentaje mayor o igual a 80% de respuestas positivas. Una respuesta es considerada positiva en el caso de la primera parte de la encuesta cuando se responde “siempre” y “casi siempre”; y en la segunda parte cuando se responde “de acuerdo” y “totalmente de acuerdo”.

En concordancia con [47], en un agente artificial la presencia social para interactuar con personas no se afecta por la presencia física o virtual del mismo, por lo tanto, se diseña una experiencia de interacción entre un estudiante y un agente virtual tele-operado (ver Fig. 2), que es capaz de hablar

utilizando el conocido sintetizador de voz *espeak*, y expresar emociones a través de un rostro genérico [48]. La experiencia de interacción se diseña según las tres premisas de comportamiento del robot, y está constituida por dos partes: primero, el robot desempeña el rol de entrevistador para hacer dos preguntas, y segundo, el robot usa el rol de entrevistado para responder una pregunta. La premisa 1, relacionada con que los robots deben interactuar según normas humanas (componente normativo), aparece en la interacción, porque se hace mediante conversación hablada en lenguaje natural. La premisa 2, asociada a que los robots deben adaptar sus interacciones, se incluye mediante el cambio de roles entre entrevistador y entrevistado, y la utilización reiterada del nombre de la persona durante la interacción (componente de adaptación). La premisa 3, correspondiente a la expresión de emociones, aparece mediante el cambio de emoción del robot durante la interacción (componente de emociones).



Fig. 2. Configuración de la Interacción del Experimento.

En la Fig. 3 se muestra el proceso de interacción seguido en el experimento. En la primera parte, con el rol de entrevistador, el robot comienza la interacción saludando y preguntando el nombre del estudiante. La respuesta del estudiante es representada en la Fig. 3 con la letra A. Luego, el robot hace la segunda pregunta relacionada con el deporte favorito, cuya respuesta del estudiante es representada con la letra B, y es evaluada para responder con alegría o tristeza, tal como se muestra en los pasos 5 y 6. Es importante destacar que durante los pasos 1, 2, 3, y 4, el robot expresa un estado emocional neutral, y según sea el caso, cambia su estado a alegre o triste, en los pasos 5 o 6, respectivamente.

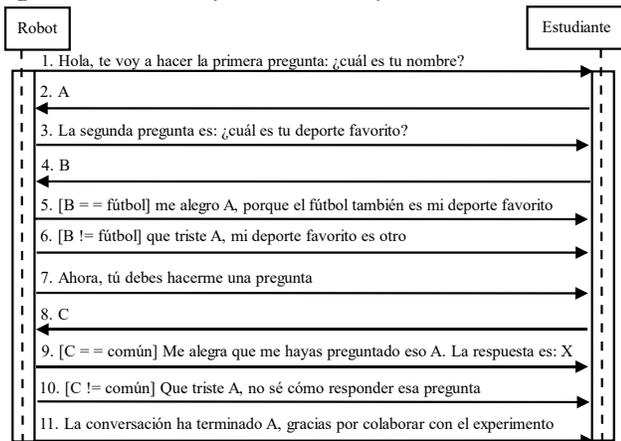


Fig. 3. Diagrama de Interacción del Experimento.

El cambio de rol para la segunda parte sucede en el paso 7, donde el robot le informa al estudiante que ya puede realizar la pregunta, la cual es representada por la letra C, y es evaluada para determinar si la pregunta es común y puede ser respondida rápidamente por el operador, en cuyo caso, el robot

cambia a estado alegre, y responde según lo indicado por el operador. La respuesta es representada con la letra X en el paso 9. En caso contrario, cambia al estado triste (paso 10). Finalmente, en un estado neutral, el robot informa al estudiante que la conversación ha terminado, y le agradece por su participación en el experimento (paso 11).

La encuesta que se aplica está dividida en dos partes. La primera, tiene el objetivo de evaluar la interacción mediante 8 preguntas, que deben ser respondidas según la siguiente escala: siempre, casi siempre, a veces, casi nunca y nunca. La segunda parte, tiene el propósito de conocer la opinión de los estudiantes acerca de las premisas, por lo tanto, se hace una pregunta por cada premisa, que debe ser respondida según la siguiente escala: totalmente de acuerdo, de acuerdo, neutral, en desacuerdo y totalmente en desacuerdo.

El experimento se lleva a cabo en un laboratorio, donde el robot se ubica en el centro de una mesa, y el estudiante se sienta en una silla a una distancia aproximada de 60 cm con respecto al robot (ver Fig. 2). El experimento se aplicó a 30 estudiantes de ingeniería (17 hombres y 13 mujeres), con edades comprendidas entre 19 y 24 años, a quienes no se les mencionó que el robot era tele-operado.

Los resultados de la evaluación de la interacción (primera parte de la encuesta) se presentan en la Tabla III. Las preguntas 1, 2, 3 y 4 están relacionadas con la premisa 1, que está asociada a la interacción de los robots según las normas humanas. Aunque la interacción con el robot proporcionada a los estudiantes ya tiene implícitos aspectos de interacción asociados a las normas humanas, tales como presencia del rostro o la propia conversación hablada, en las preguntas 1 y 2 se desea confirmar la percepción de los estudiantes con respecto al acto de hablar, y en las preguntas 3 y 4 sobre el acto de escuchar y las reglas para una buena conversación.

TABLA III
RESULTADOS DE LA PRIMERA PARTE DE LA ENCUESTA

Pregunta	S	CS	AV	N
1. El robot pronunció correctamente las palabras	43.3%	50%	6.7%	
2. El robot te miró mientras hablaba	90%		10%	
3. El robot prestó atención a lo que dijiste	76.6%	16.7%	6.7%	
4. El robot esperó su turno para hablar	96.7%	3.3%		
5. El robot se dirige a ti con tu nombre	73.3%	10.3%	3.3%	10%
6. El robot realizó bien el rol de entrevistador	80%	10.3%	6.7%	
7. El robot realizó bien el rol de entrevistado	70%	23.3 %	6.7%	
8. El robot expresó claramente sus emociones	73.3%	6.7%	16.7%	3.3%

S (Siempre), CS (Casi Siempre), AV (A Veces), y N (Nunca).

Las preguntas 5, 6 y 7, están asociadas a la premisa 2, la cual especifica que los robots deben adaptar sus interacciones, Por un lado, la pregunta 5 está relacionada con la utilización del nombre de la persona durante la interacción, lo cual es una de las habilidades de los robots que se adaptan a la persona. Por otro lado, las preguntas 6 y 7 están asociadas al cambio de rol que realiza el robot de entrevistador a entrevistado. La pregunta 8 está relacionada a la premisa 3, que indaga sobre las expresiones emocionales que realizó el robot durante las

interacciones, la cual cambió entre neutral, alegre y triste. Es importante destacar que cuando el robot hacía el cambio de expresión facial para el estado alegre, también lo indicaba en el diálogo diciendo: “me alegra”. Análogamente, cuando cambiaba el rostro a triste, decía “que triste”.

En la Tabla IV se presentan los resultados de las opiniones de los estudiantes con respecto a las tres premisas (segunda parte de la encuesta), donde es importante destacar que ningún estudiante estuvo totalmente en desacuerdo en alguna de ellas. La pregunta 9 está asociada a la premisa 1, la pregunta 10 relacionada con la adaptación de las interacciones, y la pregunta 11 a la expresión de emociones de los robots.

TABLA IV
RESULTADOS DE LA SEGUNDA PARTE DE LA ENCUESTA

Pregunta	TA	DA	N	ED
9. Los robots deben interactuar según las normas humanas	56.7%	33.3%	3.3%	6.7%
10. Los robots deben adaptar sus interacciones	56.7%	43.3%		
11. Los robots deben expresar sus emociones	30%	36.7%	26.7%	6.7%

TA (Totalmente de Acuerdo), DA (De Acuerdo), N (Neutral), y ED (En Desacuerdo).

V. DISCUSIÓN DE LOS RESULTADOS

La interacción proporcionada a cada estudiante con el robot, se considera exitosa, porque al menos el 80% de los estudiantes percibió frecuentemente todos los indicadores relacionados con las premisas durante la interacción, lo cual indica que se logró proporcionar una interacción social basada en esas tres premisas. Para la premisa 1, la mayoría de los estudiantes entiende todas las palabras que pronuncia el robot, además, un alto porcentaje percibe que el robot los mira mientras habla. A su vez, la mayoría considera que el robot siempre le prestó atención, y a excepción de un estudiante, todos percibieron que el robot esperó su turno para hablar. También, algunos estudiantes se mostraron ansiosos, esperando respuestas más rápidas, lo cual puede ser atribuido a las características particulares de las personas. Con respecto a la premisa 2, los resultados relacionados a la adaptación del robot indican que: la mayoría considera que el robot siempre se dirige utilizando el nombre del estudiante, y entre los dos roles, el robot se desempeñó mejor en el rol de entrevistado. Con respecto a la premisa 3, la mayoría considera que el robot frecuentemente expresó sus emociones claramente.

Por otra parte, con respecto a las opiniones de los estudiantes sobre las tres premisas, aunque no hay una opinión fuerte donde todos estén totalmente de acuerdo, se obtuvieron resultados satisfactorios. La condición de que al menos el 80% de los estudiantes opinan de manera favorable sobre las premisas, es decir, de acuerdo o totalmente de acuerdo, se cumple para las premisas 1 y 2. Con respecto a la premisa 3, ocurre un fenómeno particular, porque la opinión neutral fue relativamente alta, con un 26.7%, lo cual indica que el tema de las emociones en las máquinas es un aspecto aún sin una posición clara en las personas; sin embargo, dos tercios de los estudiantes opinaron favorablemente con respecto a incluir la expresión de las emociones en los robots. Así, aunque no se obtuvo la aceptación esperada en la premisa 3, los resultados siguen siendo favorables, porque los estudiantes tampoco opinaron de manera desfavorable.

Los resultados apuntan a que se puede mejorar la calidad de las interacciones si se consideran los siguientes factores: utilizar sintetizadores de voz que pronuncien mejor las palabras, usar rostros que cambien la dirección de la mirada en función del contexto, adaptar los tiempos de respuesta según la persona, y expresar las emociones en más modos. Es importante destacar que en las interacciones no necesariamente deben aparecer las tres premisas, ya que eso dependerá del objetivo de la interacción, siendo factible la ausencia total de algunas de las premisas, o las combinaciones posibles entre ellas, por ejemplo, para una interacción corta cuyo objetivo sea recolectar información de las personas mediante sólo una pregunta no sería necesario aplicar la premisa de adaptación. Eso se refleja en nuestro modelo mediante la activación o desactivación de los componentes responsables del cumplimiento de cada premisa (componente normativo, componente adaptación y módulo afectivo, para las premisas 1, 2 y 3, respectivamente).

VI. COMPARACIÓN CON TRABAJOS PREVIOS

La metodología DESMET (acrónimo de *Determining an Evaluation Methodology for Software Methods and Tools*), que ha sido usada para el análisis de características de manera cualitativa en aplicaciones de software, es utilizada para evaluar el MIHR. Dada la posibilidad de comparar con otros modelos, se selecciona el método *Qualitative Screening*, el cual consiste en marcar la existencia de características de interés en cada modelo [49]. Las características de interés que se comparan están divididas en dos grupos. Las características del primer grupo están asociadas a la representación del humano, y son extraídas de la evaluación del MIHH (en [12] se presentan las referencias que justifican cada característica), las cuales son: ruidos en el canal de comunicación, cognición de la persona, condiciones físicas de las personas, influencia de los estados afectivos, e influencia del inconsciente. El segundo grupo de características están relacionadas a la representación del robot, y fueron desarrolladas en este artículo: modalidad de comunicación humana, capacidad de adaptación, y expresión de emociones. En la Tabla V se presenta la comparación.

En general, la única característica que consideran todos los modelos es la influencia de los estados afectivos en las personas. Las características más comunes en los modelos, con aparición en cuatro modelos, son: cognición de la persona, condiciones físicas de la persona, influencia de los estados afectivos de la persona, y capacidad de adaptación. Por otra parte, las características particulares que comparte el MIHR con sólo otro de los demás modelos, son: ruidos en el canal de comunicación, influencia del inconsciente de la persona, y expresión de emociones en la máquina. En ese sentido, el MIHR cumple su propósito de ser un modelo que integra características de otros modelos.

Particularmente, en la característica de modalidad de comunicación, en [16] se hace referencia explícita a entradas y salidas multimodales, y además, es el único modelo que incorpora los conceptos de fusión para interpretar la información recibida de varios canales, y fisión para expresar un mismo mensaje en distintos canales. En [17], la comprensión multimodal se hace explícita sólo sobre la voz y

los gestos, porque son considerados los principales modos de comunicación en las interacciones humanas. Por su parte, en [18] el modelo multimodal está asociado a las posibilidades de comunicación de la máquina (gráfica, hablada, gestual, etc.). Por otro lado, en nuestro modelo también se considera la fusión (en el componente de reconocimiento) y fisión (en el módulo conductual), y además, la comprensión y expresión multimodal incluye todos los canales (verbal, no verbal, paraverbal y fisiológico) de comunicación de las personas.

TABLA V
COMPARACIÓN DE LOS MODELOS

	16	17	18	19	20	MIHR
Ruidos en el canal de comunicación			X			X
Cognición de la persona			X	X	X	X
Condiciones físicas de la persona			X	X	X	X
Influencia de los estados afectivos de la persona	X	X	X	X	X	X
Influencia del inconsciente de la persona	X					X
Modalidad de comunicación humana de la máquina	X	X	X			X
Capacidad de adaptación de la máquina	X		X		X	X
Expresión de emociones de la máquina				X		X

Con respecto a la capacidad de adaptación, en [16] se ajustan los mensajes según el usuario y el contexto de interacción, en [18] en términos de los contenidos e interfaz presentada a los usuarios, y en [19] según el modelo del usuario. En nuestro modelo, la adaptación se hace sobre parámetros que regulan las salidas y las entradas. Por ejemplo, para las salidas se adaptan parámetros como: palabras utilizadas (verbal), velocidad del habla (paraverbal), gestos (no verbal), entre otros; y para las entradas, algunos ejemplos son: adaptación del micrófono e iluminación de la cámara.

A diferencia de los demás modelos, con respecto a la adaptación, nuestro modelo también concibe la adaptación de la arquitectura a nivel de componentes. Por ejemplo, para el contexto donde el robot entrevista a una persona invidente, en el módulo conductual se desactivaría el componente no verbal, es decir, el MIHR permite la selección de los canales de salida apropiados (visual, auditivo, táctil), dependiendo del contexto y las condiciones fisiológicas de la persona. De manera análoga, sucedería con la selección de los canales de percepción, es decir, si el robot se comunica con una persona que no puede hablar, en el componente de percepción del módulo físico se desactivaría el canal auditivo.

Finalmente, en la característica relacionada a la expresión de emociones por parte del robot, en [20] se concibe a las emociones como un atributo de los robots, al mismo nivel que los deseos, conocimientos, creencias, capacidades y limitaciones, y además, las emociones del robot se reflejan en sus acciones comunicativas. En nuestro modelo, los estados afectivos se ubican en un módulo exclusivo, y además, se dividen en emociones (para representar reacciones afectivas espontáneas que surgen ante una situación que aparece de repente) y sentimientos (para estados afectivos estables y duraderos, que son producto de interacciones progresivas que han dejado huella). Los sentimientos no son tan comunes como las emociones en los robots, sin embargo, sentimientos

como el amor han sido implementados en robots y hay evidencia de que las personas son capaces de percibirlos [45]. Adicionalmente, nuestro módulo utiliza los estados afectivos en la expresión del robot y la toma de decisiones.

VII. CONCLUSIONES

La revisión de modelos que describen interacciones entre personas y máquinas, en conjunto con la indagación teórica sobre interacciones entre personas y robots sociales, permitió identificar la ausencia de un modelo que integre los tres elementos principales que representan la tendencia actual del comportamiento de los robots sociales, los cuales son: modalidad de comunicación humana, capacidad de adaptación, y expresión de emociones. MIHR integra esos elementos en aras de proporcionar un modelo más completo.

En función de esos elementos, las tres premisas de comportamiento del robot que considera el modelo propuesto son: los robots interactúan según las normas humanas, los robots adaptan sus interacciones, y los robots expresan emociones durante las interacciones. En el experimento realizado para verificar la aceptación de estas premisas, se obtuvieron respuestas favorables en la mayoría de los participantes. Además, la evaluación de la experiencia de interacción con el robot, permitió verificar que la mayoría de los estudiantes percibieron la aparición de las tres premisas durante la interacción proporcionada en el experimento.

El modelo propuesto es la base para trabajos futuros. Por ejemplo, dependiendo del contexto de aplicación, se deben definir formalismos matemáticos para cada componente del modelo arquitectónico propuesto. Un ejemplo de los anteriores es el trabajo [46], donde se utilizó un PDM para modelar las interacciones de un agente, y el paradigma de aprendizaje Q-learning para modelar el proceso de adaptación del agente.

Otros formalismos se requieren para modelar los estados afectivos, el conocimiento del robot, entre otras cosas. Finalmente, otro trabajo futuro es la especificación de una arquitectura para robots sociales basada en el MIHR, donde se especifiquen aspectos como: representación del conocimiento y protocolos de comunicación.

REFERENCIAS

- [1] T. Ribeiro, A. Pereira, E. Di Tullio and A. Paiva, "The sera ecosystem: Socially expressive robotics architecture for autonomous human-robot interaction", In *Proc. Symposium on Enabling Computing Research in Socially Intelligent Human-Robot Interaction*, pp. 155-163, 2016.
- [2] H. Admoni and B. Scassellati, "Social Eye Gaze in Human-Robot Interaction: A Review", *Journal of HRI*, vol. 6, no. 1, pp. 25-63, 2017.
- [3] P. Lasota, T. Fong, and J. Shah, "A Survey of Methods for Safe Human-Robot Interaction", *Foundations and Trends in Robotics*, vol. 5, no. 4, pp. 261-349, 2017.
- [4] R. Siregar, R. Syahputra, and M. Mustar. "Human-Robot Interaction Based GUI", *Journal of Electrical Technology UMY*, vol. 1, no. 1, pp.10-19, 2017.
- [5] G. Briggs and M. Scheutz, "The Pragmatic Social Robot: Toward Socially-Sensitive Utterance Generation in Human-Robot Interactions", In *Proc. AAAI Fall Symposium Series: Artificial Intelligence for Human-Robot Interaction*, pp. 12-15, 2016.
- [6] A. Ghazali, J. Ham, E. Barakova, and P. Markopoulos, "Effects of robot facial characteristics and gender in persuasive human-robot interaction", *Frontiers in Robotics and AI*, vol. 5, no. 73, pp. 1-16, 2018.
- [7] F. Correia, S. Mascarenhas, R. Prada, F. Melo, and A. Paiva. "Group-based emotions in teams of humans and robots", In *Proc. International Conference on Human-Robot Interaction*, pp. 261-269, 2018.

- [8] K. Luna, E. Palacios, and A. Marin, "A Fuzzy Speed Controller for a Guide Robot Using an HRI Approach", *IEEE Latin America Transactions*, vol. 16, no. 8, pp. 2102-2107, 2018.
- [9] R. Andreasson, B. Alenljung, E. Billing, and R. Lowe, "Affective Touch in Human-Robot Interaction: Conveying Emotion to the Nao Robot", *Journal of Social Robotics*, vol. 10, no. 4, pp. 473-491, 2018.
- [10] Z. Zafar, D. Salazar, S. Al-Darraji, D. Urukalo, K. Berns, and A. Rodić, "Human Robot Interaction Using Dynamic Hand Gestures", In *Proc. International Conference on Robotics*, pp. 649-656, 2017.
- [11] M. Graaf, "An ethical evaluation of human-robot relationships", *Journal of Social Robotics*, vol. 8, no. 4, pp. 589-598, 2016.
- [12] J. Pérez, J. Aguilar, and E. Dapena, "MIHH: Un Modelo de Interacción Humano-Humano", *Revista Venezolana de Computación*, vol. 5, no. 1, pp. 10-19, 2018.
- [13] B. Scassellati et al., "Improving social skills in children with ASD using a long-term, in-home social robot", *Science Robotics*, vol. 3, no. 21, pp. 1-9, 2018.
- [14] J. Pérez and J. Castro, "LRS1: un robot social de bajo costo para la asignatura "Programación 1", *Revista Colombiana de Tecnologías de Avanzada*, vol. 32, no. 2, pp. 68-77, 2018.
- [15] Y. Noguchi, H. Kamide, and F. Tanaka, "Effects on the Self-disclosure of Elderly People by Using a Robot Which Intermediates Remote Communication", In *Proc. 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 612-617, 2018.
- [16] B. Dumas, D. Lalanne, and S. Oviatt, "Multimodal interfaces: A survey of principles, models and frameworks", In *Human Machine Interaction*, Berlin, pp. 3-26, 2009.
- [17] A. Vinciarelli et al., "Open challenges in modelling, analysis and synthesis of human behaviour in human-human and human-machine interactions", *Cognitive Computation*, vol. 7, no. 4, pp. 397-413, 2015.
- [18] J. Hussain et al., "Model-based adaptive user interface based on context and user experience evaluation", *Journal on Multimodal User Interfaces*, vol. 12, no. 1, pp.1-16, 2018.
- [19] G. Martins, L. Santos, and J. Dias, "User-Adaptive Interaction in Social Robots: A Survey Focusing on Non-physical Interaction", *International Journal of Social Robotics*, pp. 1-21, 2018.
- [20] T. Hellström and S. Bensch, "Understandable robots. Paladyn". *Journal of Behavioral Robotics*, vol. 9, no. 1, pp. 110-123, 2018.
- [21] A. Gil, J. Aguilar, E. Puerto, E. Dapena, "Emergence Analysis in a Multi-Robot System," In *Proc. XLIV Latin American Computer Conference (CLEI)*, pp. 662-669, 2018.
- [22] A. Gil, J. Aguilar, E. Dapena, and R. Rivas, "A Control Architecture for Robot Swarms (AMEB)", *Journal Cybernetics and Systems*, vol. 50, no. 3, pp. 300-322, 2019.
- [23] S. Song and S. Yamada, "Expressing emotions through color, sound, and vibration with an appearance-constrained social robot", In *Proc. International Conference on Human-Robot Interaction*, pp. 2-11, 2017.
- [24] I. Brinck, C. Balkenius, and B. Johansson, "Making Place for Social Norms in the Design of Human-Robot Interaction. What Social Robots Can and Should Do", In *Proc. Robophilosophy*, pp. 303-312, 2016.
- [25] R. Pérula et al., "Bioinspired decision-making for a socially interactive robot", *Cognitive Systems Research*, pp. 1-40, 2018.
- [26] D. Herrera, F. Roberti, M. Toibero, and R. Carelli, "Human-Robot Interaction: Legible behavior rules in passing and crossing events", *IEEE Latin America Transactions*, vol. 14, no. 6, pp. 2644-2650, 2016.
- [27] K. Tsiakas, M. Dagioglou, V. Karkaletsis, and F. Makedon, "Adaptive robot assisted therapy using interactive reinforcement learning", In *Proc. International Conference on Social Robotics*, pp. 11-21, 2016.
- [28] S. Strohkorb et al., "Establishing Sustained, Supportive Human-Robot Relationships: Building Blocks and Open Challenges", In *Proc. Spring Symposium on Enabling Computing Research in Socially Intelligent Human-Robot Interaction*, pp. 179-182, 2016.
- [29] A. Araujo, J. Pérez, and W. Rodríguez, "Aplicación de una Red Neuronal Convolutiva para el Reconocimiento de Personas a Través de la Voz", In *Proc. Sexta Conferencia Nacional de Computación, Informática y Sistemas*, pp. 77-81, 2018.
- [30] G. Martins, P. Ferreira, L. Santos, and J. Dias, "A context-aware adaptability model for service robots", In *Proc. Workshop on Autonomous Mobile Service Robots*, pp. 1-7, 2016.
- [31] L. Nichola, W. Erin, and P. Heather, "Effects of voice-adaptation and social dialogue on perceptions of a robotic learning companion", In *Proc. 11th International Conf. on Human-Robot Interaction*, pp. 255-262, 2016.
- [32] Q. Shen, K. Dautenhahn, J. Saunders, and H. Kose, "Can real-time, adaptive human-robot motor coordination improve humans overall perception of a robot?", *IEEE Transactions on Autonomous Mental Development*, vol. 7, no. 1, pp. 52-64, 2015.
- [33] R. Ros, I. Baroni, and Y. Demiris, "Adaptive human-robot interaction in sensorimotor task instruction: from human to robot dance tutors", *Robotics and Autonomous Systems*, vol. 62, no. 6, pp. 707-720, 2014.
- [34] R. Aylett, A. Kappas, G. Castellano, S. Bull, W. Barendregt, A. Paiva, and L. Hall, "I know how that feels—an empathic robot tutor", In *Proc. eChallenges e-2015 Conference*, pp. 1-9, 2015.
- [35] A. Karami et al., "Adaptive artificial companions learning from users feedback", *Adaptive Behavior*, vol. 24, no. 2, pp. 69-86, 2016.
- [36] S. Rukavina, et al., "Affective computing and the impact of gender and age", *PLoS one*, vol. 11, no. 3, pp. 1-20, 2016.
- [37] A. Gil, J. Aguilar, R. Rivas, and E. Dapena, "Behavioral Module in a Control Architecture for Multi-robots", *Revista Ingeniería al Día*, vol. 2, no. 1, pp. 40-57, 2016.
- [38] N. Morán, J. Pérez, and W. Rodríguez, "Reconocimiento de Estados Emocionales de Personas Mediante la Voz Utilizando Algoritmos de Aprendizaje de Máquina", *Revista Venezolana de Computación*, vol. 5, no. 2, pp. 41-52, 2018.
- [39] A. Bandhakavi, N. Wiratunga, S. Massie, and D. Padmanabhan, "Lexicon generation for emotion detection from text", *IEEE Intelligent Systems*, vol. 32, no. 1, pp. 102-108, 2017.
- [40] V. Campos, B. Jou, and X. Giro, "From pixels to sentiment: Fine-tuning cnns for visual sentiment prediction", *Image and Vision Computing*, vol. 65, pp. 15-22, 2017.
- [41] J. Cordero, J. Aguilar, K. Aguilar, D. Chávez, E. Puerto. "Recognition of the Driving Style in Vehicle Drivers". *Sensors*, vol. 20, no. 9, 2020.
- [42] J. Aguilar, K. Aguilar, J. Cordero, D. Chávez, and E. Puerto, "Different Intelligent Approaches for Modeling the Style of Car Driving", In *Proc. International Conference on Informatics in Control, Automation and Robotics*, pp. 284-291, 2017.
- [43] W. Lee and M. Norman, "Affective Computing as Complex Systems Science", *Procedia Computer Science*, vol. 95, pp. 18-23, 2016.
- [44] T. Izui, I. Milleville, S. Sakka, and G. Venture, "Expressing emotions using gait of humanoid robot", In *Proc. International Symposium on Robot and Human Interactive Communication*, pp. 241-245, 2015.
- [45] H. Samani, "The evaluation of affection in human-robot interaction", *Kybernetes*, vol. 45, no. 8, pp. 1257-1272, 2016.
- [46] J. Pérez, J. Aguilar, E. Dapena, and G. Carrillo, "Q-learning based algorithm for learning children's difficulties in multiplication tables", submitted to publication, 2020.
- [47] S. Thellman et al., "Physical vs. Virtual Agent Embodiment and Effects on Social Interaction", In *D. Traum et al. (eds) Intelligent Virtual Agents. Lecture Notes in Computer Science*, Cham: Springer, 2016.
- [48] E. Dapena, J. Pérez, R. Rivas, and A. Guijarro, "Rostro Genérico para Máquinas que Interactúan con Personas", *Revista Científica UNET*, vol. 28, no. 2, pp. 121-130, 2016.
- [49] B. Kitchenham, S. Linkman, and D. Law, "DESMET: A Methodology for Evaluating Software Engineering Methods and Tools", *Computing & Control Engineering Journal*, vol. 8, no. 3, pp. 120-126, 1997.



J. Pérez es Ingeniero Electrónico (2012), Ingeniero de Sistemas (2014), Magister en Educación Superior (2015), Doctor en Ciencias de la Educación (2019) y profesor del Departamento de Computación de la Escuela de Ingeniería de Sistemas de la Universidad de Los Andes. Su línea principal de investigación es la Interacción Humano-Robot.



J. Aguilar es graduado en Ingeniería de Sistemas en 1987 (ULA), MSc. en Ciencias de la Computación en 1991 (Universite Paul Sabatier-Francia), Doctorado en Ciencias de la Computación en 1995 (Universite Rene Descartes-Francia) y Postdoctorado en Ciencias de Computación en 2000 (Universidad de

Houston). Sus líneas son: Inteligencia Artificial, Sistemas Multiagentes, Ambientes Inteligentes, entre otras.



E. Dapena es Ingeniero de Sistemas (1990), Especialista en Automatización Industrial (1997), DEA (2000), Doctor Ingeniero Industrial (2002) y profesor Titular de la Universidad de Los Andes. Sus líneas de investigación son: Robótica móvil y automatización.