



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIA E INGENIERÍA DE LA COMPUTACIÓN

**“AJUSTE DINÁMICO DE DIFICULTAD DE UNA APLICACIÓN INTERACTIVA PARA
NEURO-REHABILITACIÓN MEDIANTE APRENDIZAJE POR REFORZAMIENTO”**

TESIS

**QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIA E INGENIERIA DE LA COMPUTACIÓN**

PRESENTA:

YOÁS SAIMON RAMÍREZ GRAULLERA

TUTOR:

DR. MIGUEL ÁNGEL PADILLA CASTAÑEDA
ICAT, UNAM

CIUDAD UNIVERSITARIA, CD. MX., FEBRERO DE 2022



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos

En primer lugar quiero agradecer a mi madre, por su incondicional apoyo y cariño a lo largo de toda mi vida. Todo lo que he logrado se lo debo a ella.

A mi tía Vero, por todos sus consejos y por siempre animarme a superarme.

A la M en C. Ana María Escalante Gonzalbo, responsable del Laboratorio de Investigación y Desarrollo de Aplicaciones Interactivas para la Neuro-Rehabilitación (LANR) en el Instituto de Fisiología Celular y también mi superior, por apoyarme desde un inicio para ingresar al programa de maestría y darme todas las facilidades para poder realizar mis estudios de posgrado al mismo tiempo que mis responsabilidades académicas con el LANR.

A la especialista en neuro-rehabilitación María Isabel Heredia, quien me asesoró en la definición de los ejercicios que debía contener la aplicación interactiva que se presenta en este trabajo.

A mi tutor de maestría el Dr. Miguel Ángel Padilla Castañeda, por sus enseñanzas durante la maestría y por su apoyo en la elaboración del presente proyecto de tesis. Extiendo este agradecimiento también a todos los sinodales, cuyos comentarios y correcciones enriquecieron este trabajo.

Finalmente, quiero agradecer a la Universidad Nacional Autónoma de México, institución que me ha formado desde el bachillerato. En sus aulas he adquirido muchos conocimientos y habilidades, pero también valores y valiosas amistades, además de que aún hoy en día me permite seguir desarrollándome académica, cultural y profesionalmente.

Resumen

En este trabajo se presenta la aplicación interactiva Penal Madness y un algoritmo que ajusta dinámicamente su dificultad, que se basa en métodos de aprendizaje por reforzamiento como Q-learning y Dyna-Q. Penal Madness es un videojuego serio enfocado en la rehabilitación de miembros superiores de pacientes que han sufrido una Enfermedad Vascul ar Cerebral (EVC), su objetivo es incrementar el rango de movilidad articular y fuerza muscular del hombro en los planos frontal y sagital. En este juego el paciente asume el rol de un portero y debe tratar de detener la mayor cantidad de penaltis posibles; para detectar los movimientos del paciente se emplea el sensor Kinect.

El algoritmo detecta el lado parético del paciente y asigna 10 ejercicios de forma aleatoria con la siguiente proporción: 6 deben realizarse con el brazo parético, 2 con el brazo no parético y 2 serán ejercicios bimanuales. El algoritmo determinará automáticamente la dificultad del ejercicio y el tipo de movimiento que se debe realizar (abducción o flexión del hombro). Durante la partida se asignarán ejercicios con la intención de que el paciente falle, para ello el algoritmo buscará que estos ejercicios se encuentren dentro del límite de las capacidades del paciente, de esta forma se intenta evitar la frustración del jugador y se busca motivarlo para que se esfuerce un poco más y así, eventualmente, los logre realizar. Ya sea que el paciente incremente su desempeño o lo disminuya a lo largo de su sesión, el algoritmo adaptará la dificultad del juego para lograr que al final de cada partida se obtengan 6 aciertos y 4 fallos, lo que, de acuerdo a algunos estudios, indica un grado óptimo de disfrute y de rendimiento.

Para probar la eficacia del algoritmo se realizaron simulaciones en Python a través de la plataforma Google Colab, donde se modelaron numerosos pacientes con distintas capacidades, representadas a través de probabilidades. En la mayoría de los escenarios simulados se observó que en más del 70% de los casos se lograban partidas con 6 aciertos y 4 fallos, tanto con pacientes que “mejoraban” a lo largo del tiempo, como en aquellos que “empeoraban”. Los mejores resultados se lograron cuando se simulaba pacientes que no mostraban ningún cambio en su comportamiento, logrando la meta en más del 90% de las simulaciones. También se realizaron tres experimentos con un sujeto sano, tratando de simular tres comportamientos diferentes que podría presentar un paciente, en cada experimento se le indicó cuál sería su rango máximo de movimiento en cada partida, observando también resultados favorables.

La investigación acerca de algoritmos que adapten la dificultad de videojuegos orientados a la neuro-rehabilitación es un área poco explorada, y es aún más reducida la cantidad de estudios acerca de la aplicación de técnicas de aprendizaje por reforzamiento para este fin. La aplicación de este tipo de algoritmos en los juegos serios es particularmente relevante en la tele rehabilitación, una de las modalidades en las que se buscará emplear este desarrollo en el futuro, pues de esta manera el juego podrá adaptarse a la evolución de cada paciente, de forma personalizada y con la oportunidad que se requiere.

Índice

1. Introducción	1
1.1 Objetivo	1
1.2 Objetivos específicos	1
1.3 Hipótesis.....	1
1.4 Metodología	1
1.5 Relevancia y contribución.....	2
1.6 Estructura del trabajo	3
2. Enfermedad Vascular Cerebral	5
2.1 Clasificación.....	6
2.1.1 Isquemia Cerebral.....	7
2.1.2 EVC Hemorrágico.....	12
2.2 Campañas de identificación de síntomas	14
2.3 Fases de evolución de un EVC.....	14
2.4 Complicaciones post EVC.....	15
2.4.1 Déficits motores	15
2.4.2 Déficits de comunicación	16
2.4.3 Déficits cognitivos.....	17
2.4.4 Déficits emocionales.....	18
2.4.5 Otros déficits	18
3. Rehabilitación posterior a un EVC.....	20
3.1 Valoración del déficit funcional	20
3.1.1 Área motora.....	21
3.1.2 Área perceptiva, sensitiva y visual	23
3.1.3 Comunicación y lenguaje	24
3.1.4 Dominio cognitivo.....	24
3.1.5 Área emocional.....	24
3.1.6 Discapacidad instrumental	24
3.2 Plasticidad neuronal después de un EVC	24
3.3 Pronóstico de la recuperación motora	26
3.4 Principios generales de rehabilitación	29

3.5	Objetivos en las fases de rehabilitación.....	31
3.5.1	Periodo agudo	31
3.5.2	Periodo subagudo.....	32
3.5.3	Periodo crónico o de estado	32
3.6	Tratamientos de rehabilitación motora	33
3.6.1	Terapia física.....	33
3.6.2	Compensación y recuperación.....	35
3.6.3	Estrategias complementarias para la rehabilitación motora	36
4.	Rehabilitación virtual	44
4.1	Clasificación	46
4.1.1	Rehabilitación musculo-esquelética	46
4.1.2	Rehabilitación cognitiva.....	46
4.1.3	Neuro-rehabilitación.....	47
4.2	Interfaces sensoriales y motoras en la rehabilitación virtual.....	47
4.2.1	Sistemas de seguimiento de movimiento humano.....	49
4.3	Beneficios de la rehabilitación virtual.....	54
4.4	Retos y limitaciones de la rehabilitación virtual	55
4.5	Factores en el diseño de juegos serios para rehabilitación	57
4.6	Retroalimentación positiva o afectiva	60
4.7	Transferencia de habilidades al mundo real.....	63
4.8	Flow y ajuste dinámico de dificultad	64
4.9	Plataformas con ajuste dinámico de dificultad.....	65
4.9.1	Sistema de rehabilitación PERCRO-BRANDO	66
4.9.2	Rehabilitation Gaming System (RGS)	69
4.9.3	Videojuego de rehabilitación con ajuste dinámico de dificultad con MPRL.....	72
5.	Aprendizaje por Reforzamiento.....	76
5.1	Elementos del aprendizaje por reforzamiento	77
5.2	El problema del bandido multibrazo	80
5.3	Procesos de Decisión de Markov	83
5.4	Retornos esperados en los episodios	84
5.5	Políticas y funciones de valor.....	86
5.6	La ecuación de Bellman y funciones óptimas	88

5.7	Aprendizaje por diferencias temporales.....	90
5.8	Q-learning.....	91
5.9	Planeación con métodos tabulares.....	93
6.	Desarrollo.....	97
6.1	Aplicación interactiva Penal Madness	97
6.1.1	Mecánica de juego.....	97
6.1.2	Interfaz de clínica.....	102
6.1.3	Interfaz del terapeuta	106
6.1.4	Interfaz del paciente	108
6.2	Penal Madness con aprendizaje por reforzamiento	109
6.2.1	Definición de la función de valor estado-acción $Q(S,A)$	110
6.2.2	Definición de la función de recompensa.....	113
6.2.3	Planeación	116
6.3	Simulaciones.....	118
6.4	Transgresiones.....	119
7.	Resultados.....	120
7.1	Experimentos exploratorios	120
7.1.1	Experimento 1: Comportamiento del algoritmo sin planeación	120
7.1.2	Experimento 2: Comportamiento del algoritmo con planeación	122
7.1.3	Experimento 3: Planeación en ambientes estocásticos	124
7.1.4	Experimento 4: Comportamiento del algoritmo ante diversos pacientes	126
7.2	Experimentos simulando casos específicos de pacientes	129
7.2.1	Simulación de un paciente EVC que mejora	129
7.2.2	Simulación de un paciente EVC que no muestra cambios	132
7.2.3	Simulación de un paciente EVC que decae	136
7.3	Experimentos en un ambiente real.....	139
7.3.1	Comportamiento del algoritmo ante un participante que mejora	139
7.3.2	Comportamiento del algoritmo ante un participante que decae	141
7.3.3	Comportamiento del algoritmo ante un participante que no muestra cambios	143
8.	Conclusiones y trabajo a futuro	145
	Bibliografía.....	148

Glosario de acrónimos

ACA	Arteria Cerebral Anterior
ACI	Arteria Carótida Interna
ACAI	Arteria Cerebral Anterior Inferior
ACM	Arteria Cerebral Media
ACP	Arteria Cerebral Posterior
ACPI	Arteria Cerebral Posterior Inferior
ACS	Arteria Cerebral Superior
ADD	Ajuste Dinámico de Dificultad
AHA-SOC	<i>American Heart Association-Stroke Outcome Classification</i>
AIT	Ataque Isquémico Transitorio
AVD	Actividades de la Vida Diaria
CAHAI	<i>Chedoke Arm and Hand Activity Inventory</i>
EEF	Estimulación Eléctrica Funcional
EENM	Estimulación Eléctrica Neuromuscular
EENT	Estimulación Eléctrica Neuromuscular Transcutánea
EMT	Estimulación Magnética Transcraneal
EVC	Enfermedad Vascul ar Cerebral o Evento Vascul ar Cerebral
FE	Flexión y Extensión
IC	Infarto Cerebral
MDA	Modelo <i>Mechanics-Dynamics-Aesthetics</i>
MDP	Proceso de Decisión de Markov
ML	Aprendizaje Automático o Aprendizaje de Máquina (<i>Machine Learning</i>)
MPRL	Aprendizaje por Reforzamiento de Múltiples Periodos
PS	Pronación y Supinación
PTM	Módulo de Entrenamiento Personalizado
RGS	Rehabilitation Gaming System
RL	Aprendizaje por Reforzamiento (<i>Reinforcement Learning</i>)
RV	Realidad Virtual
TD	Aprendizaje por Diferencias Temporales (<i>Temporal Difference learning</i>)
TMIR	Terapia de Movimiento Inducida por Restricción
TOAST	<i>Trial of Org 10172 in Acute Stroke Registry</i>

Capítulo 1

Introducción

1.1 Objetivo

Implementar y evaluar la funcionalidad de ajuste dinámico de dificultad (ADD) en una aplicación interactiva enfocada en la rehabilitación de miembros superiores para pacientes con Enfermedad Vasculat Cerebral (EVC), mediante técnicas de aprendizaje por reforzamiento.

1.2 Objetivos específicos

Implementar un algoritmo para el juego serio Penal Madness, que ajuste automáticamente y en tiempo real su dificultad, asignando ejercicios que se encuentren dentro del límite de las capacidades del paciente y enfocándose en aquellos movimientos que le resulten más complicados. El algoritmo tratará de que al final de una partida de 10 ejercicios el paciente haya realizado 6 de ellos correctamente, de esta forma se espera mantener motivado al paciente al presentar una experiencia ni muy fácil ni muy difícil, evitando la frustración y el aburrimiento.

Realizar simulaciones que permitan evaluar el algoritmo a través de diferentes modelos que traten de representar el comportamiento y evolución que tendría un paciente EVC al utilizar los juegos.

1.3 Hipótesis

Un algoritmo basado en técnicas de aprendizaje por reforzamiento, implementado en una aplicación interactiva para rehabilitación, será capaz de asignar ejercicios de manera óptima en tiempo real y de mantener un grado de reto que favorezca el compromiso del paciente con la aplicación, manteniéndolo motivado y buscando superar los límites de sus capacidades.

1.4 Metodología

Primero realicé una investigación acerca del EVC y las principales afecciones motrices que puede ocasionar en una persona, así como los métodos tradicionales de rehabilitación y los principios que subyacen su efectividad. Posteriormente recopilé información sobre diversos estudios en donde se emplearán sensores de posición y movimiento, como el Kinect, en conjunto con aplicaciones interactivas con el objetivo de apoyar en la rehabilitación de miembros superiores de pacientes EVC, sobre todo estudios recientes que implementarán algún método que ajustara automáticamente la dificultad del programa.

Me documenté en el diseño del desarrollo de videojuegos serios, los elementos que debe considerar un juego de esta naturaleza y la forma en la que se pueden aplicar los principios de control motor y aprendizaje motor en un videojuego. A continuación, desarrollé la aplicación interactiva Penal Madness con el motor de videojuegos Unity recibiendo asesoría de una especialista en rehabilitación para el diseño de los ejercicios dentro del juego, la aplicación está asociada al sensor Kinect para la detección de movimientos y también incluye funciones con miras a poder ser usada de forma remota a través de la tele-rehabilitación.

Después comencé el diseño de un algoritmo de ajuste dinámico e dificultad para Penal Madness empleando técnicas de aprendizaje por reforzamiento, se eligió este acercamiento por sobre otras técnicas de aprendizaje de máquina gracias a la forma natural en la que el algoritmo aprende y se ajusta a las necesidades y capacidades del paciente, similar a como lo haría un humano. Otros métodos como el aprendizaje supervisado o los modelos psicométricos requieren de mucha información *a priori*, lo cual es complicado debido a la cantidad requerida de pacientes.

Para la evaluación del algoritmo se desarrollaron múltiples simulaciones de pacientes en Google Colab. A través de una estructura de datos se definía la probabilidad que tenía el paciente simulado de realizar alguna de las acciones del juego. Esta probabilidad cambiaba después de un determinado número de partidas, según lo que se quisiera simular, el paciente podía mejorar, empeorar o no mostrar ningún cambio. Se realizaron diversos experimentos para observar el comportamiento del algoritmo, en cada experimento se registraba el marcador final que se obtenía en cada partida, así como los ejercicios que el algoritmo asignaba. Finalmente, se realizaron tres experimentos con un sujeto sano, en cada experimento también se simuló un comportamiento diferente (mejora, empeoramiento y estabilidad).

1.5 Relevancia y contribución

La implementación aplicaciones interactivas para neuro-rehabilitación, desarrolladas específicamente para el entrenamiento de las funciones motoras de pacientes que han sufrido un EVC, que ofrezcan distintos tipo de retroalimentación en un ambiente virtual enriquecido, asociadas a sensores de posición y movimiento, pueden representar un cambio de paradigma en lo que se refiere a la rehabilitación de pacientes con hemiparesia a consecuencia de un EVC, brindándoles nuevas oportunidades de recuperación, más allá de lo que podrían obtener con los esquemas tradicionales de rehabilitación.

El desarrollo que se propone permitirá que los pacientes que han sufrido un EVC comiencen a realizar sus ejercicios desde casa lo más pronto posible y con una alta frecuencia, aprovechando la ventana de neuroplasticidad aumentada y recuperación espontánea. También les permitirá ser más independientes, pues el paciente ya no dependería de la disponibilidad de citas y no tendría que trasladarse tan frecuentemente hasta un hospital.

Con este desarrollo se pretende lograr una mejor recuperación motriz del hombro mediante ejercicios en los planos frontal y sagital, siguiendo el esquema tradicional que prioriza la recuperación de miembros proximales antes que miembros distales. Al ajustar la dificultad del

juego a las capacidades del paciente, éste se mantendrá motivado al realizar sus ejercicios a través de un videojuego e idealmente entrará en un estado donde el disfrute de la actividad realizada y el rendimiento físico sea óptimo.

Cada paciente es diferente y son muchos los factores que determinarán su grado de mejora y evolución a través del tiempo, por esta razón, la dificultad del juego y las tareas que el mismo programa asigne deben ser personalizadas, similar a como lo haría un terapeuta. Un enfoque psicométrico requiere realizar pruebas con muchos pacientes para obtener una muestra aceptable que represente el comportamiento de los pacientes EVC y así obtener un modelo funcional; estos estudios pueden requerir mucho tiempo en reunir el tamaño de cohorte necesario y, aun así, siempre existirán casos en los que los pacientes que no se ajusten a los datos. Por esta razón resulta natural emplear un enfoque de aprendizaje por reforzamiento (RL por sus siglas en inglés), en donde el algoritmo aprende directamente de las interacciones con el ambiente, en este caso de las interacciones con el paciente, y se va ajustando de acuerdo al objetivo que se busca, sin necesidad de contar con una gran cantidad de datos a priori. El algoritmo si necesitará una gran cantidad de datos para incrementar su efectividad, pero serán generados por el mismo paciente conforme realice sus ejercicios con el videojuego, de esta forma el algoritmo aprenderá a tomar mejores decisiones.

Es importante resaltar que este tipo de desarrollos, que forman parte de un campo conocido como rehabilitación virtual, no pretenden sustituir las terapias convencionales ni el trabajo de los terapeutas, sino servir como un complemento que les permita a los pacientes realizar sus ejercicios lo más temprano posible, con la exigencia adecuada pero motivante con el fin de incrementar su compromiso con el tratamiento de rehabilitación.

Finalmente, el uso de técnicas de RL para ajustar la dificultad de los juegos ha sido muy poco explorado en videojuegos, en su lugar, los estudios relacionados a este tipo de técnicas se han enfocado en lograr que un agente aprenda a jugar un juego. Si se limita el problema a aplicaciones de RL en videojuegos para neuro-rehabilitación los estudios relacionados a ello son prácticamente inexistentes, por lo que los resultados de su implementación pueden ser un aporte interesante en el campo de la inteligencia artificial aplicada a la rehabilitación virtual.

1.6 Estructura del trabajo

En el segundo capítulo se describe la Enfermedad Vascul ar Cerebral, su clasificación y las principales afecciones que pueden presentarse a causa de este padecimiento.

En el tercer capítulo se presentan los principales aspectos que se abordan en la rehabilitación de un paciente que ha sufrido un EVC, comenzando con una correcta valoración de los déficits, los principios que debe seguir un plan terapéutico, las diferentes etapas de evolución de un paciente EVC, la neuroplasticidad como mecanismo endógeno de recuperación y, finalmente, se presentan varios métodos enfocados en la recuperación motora posterior a un EVC.

En el cuarto capítulo se describe un tipo de terapia conocida como rehabilitación virtual, su relación con la Realidad Virtual (RV) y los videojuegos, sus distintas aplicaciones, las principales interfaces sensoriales y motoras que se emplean en este campo, sus potenciales beneficios y retos cuando se usa como tratamiento complementario en la rehabilitación de pacientes EVC, así como una breve revisión de tres desarrollos de rehabilitación virtual, que forman parte del estado del arte, los cuales logran implementar efectivamente un mecanismo que ajusta dinámicamente la dificultad de los programas interactivos de acuerdo con el desempeño del paciente.

En el quinto capítulo, Aprendizaje por Reforzamiento, se exponen los principales componentes de esta rama del aprendizaje de máquina, como el agente, el ambiente, las acciones, la política del agente, y la función e recompensa, así como el marco de trabajo de Procesos de Decisión de Markov, bajo el cual se suelen modelar los problemas que se abordan con este tipo de métodos. Finalmente, se describen brevemente dos de los algoritmos más característicos del aprendizaje por reforzamiento, Q-learning y Dyna-Q, este último implementa las ventajas de Q-learning junto con un proceso conocido como planificación, que le permite al algoritmo converger mucho más rápido.

En el sexto capítulo se presenta el desarrollo del trabajo, se describe la mecánica del juego Penal Madness y todas sus funcionalidades, así como la forma en la que se diseñó, implementó y evaluó un algoritmo que ajustara dinámicamente la dificultad de la aplicación interactiva empleando técnicas de aprendizaje por reforzamiento.

En el séptimo capítulo se presentan y analizan los distintos resultados experimentales de las simulaciones realizadas, así como los resultados obtenidos con un sujeto sano que realiza tres comportamientos diferentes al utilizar el juego: mejora, estabilidad y empeoramiento.

En el último capítulo se exponen las conclusiones del trabajo desarrollado y el trabajo a futuro que se puede realizar.

Capítulo 2

Enfermedad Vascular Cerebral

La Enfermedad Vascular Cerebral (EVC) o enfermedad cerebrovascular se refiere a un trastorno en el cual un área del encéfalo se ve afectada de forma transitoria o permanente, debido a la afectación de uno o más vasos sanguíneos por un proceso patológico. De acuerdo con la Organización Mundial de la Salud (OMS) “Un EVC es un accidente con aparición rápida de signos clínicos de alteración focal o global de la función cerebral, con síntomas que duran 24 horas o más o que conducen a la muerte, sin una causa aparente aparte del origen vascular” (Warlow, 1998).

Ictus (golpe en latín) es un término que engloba el infarto cerebral (IC) y la hemorragia cerebral, ambos procesos son de carácter brusco y súbito. Sus manifestaciones son variables, dependiendo de la ubicación y la duración del EVC. Algunos otros nombres por los que se conoce a esta enfermedad son accidente cerebrovascular, evento vascular cerebral, ataque cerebrovascular y apoplejía.

De acuerdo con cifras de la Organización Mundial de la Enfermedad Vascular Cerebral, el EVC constituye la segunda causa de muerte (9.7%) y la primera de discapacidad adquirida en adultos a nivel mundial, siendo uno de los principales problemas de salud pública en todos los países industrializados, pues su alta prevalencia e incidencia suponen notables costos humanos y económicos. Anualmente ocurren 13.7 millones de casos EVC, con más de 80 millones de personas sobrevivientes (World Stroke Organization, 2020).

En México, la Secretaría de Salud señala que la incidencia de esta enfermedad ha aumentado desde el año 2000, especialmente en personas menores de 65 años (Arauz & Ruiz-Franco, 2012). De acuerdo con datos del ISSSTE, en el 2018 el EVC fue la sexta causa de mortalidad hospitalaria (5.65%) y la doceava causa de morbilidad hospitalaria (1.8%), mientras que en el 2019 el EVC fue la quinta causa de mortalidad hospitalaria (6%) y se mantuvo como la doceava de morbilidad (2%) (ISSSTE, 2020).

Sin embargo, en México no se conoce la magnitud real del problema, ya que no se cuentan con registros nacionales confiables, aunado al hecho de que esta patología se enmascara con otros padecimientos que pueden ser condicionantes de la misma. Por lo anterior, es probable que existan muchos casos de EVC no registrados en nuestro país (Chiquete et al., 2011).

La edad promedio de un mexicano que sufre un ataque cerebral es de 73 años, esto se atribuye a que durante el envejecimiento ocurren cambios fisiológicos en las estructuras vasculares, lo que predispone a eventos oclusivos. Además, la incidencia de un EVC se duplica cada 10 años después de los 35 años de edad. El 20% de las personas que sufren un accidente cerebrovascular mueren al

mes siguiente, mientras que el 75% de los supervivientes tendrán secuelas definitivas a lo largo de su vida. (De la Garza-Longoria et al., 2018)

Al igual que en México, la incidencia del EVC ha crecido significativamente en diversos países (Benjamin et al., 2017) lo que significa más sobrevivientes con algún grado de discapacidad física o cognitiva. Esta transición epidemiológica está relacionada al incremento en la esperanza de vida de la población, así como al incremento de factores de riesgo asociados al EVC, como la hipertensión, diabetes, obesidad y dislipidemia, entre otros. Si las tendencias actuales continúan, para 2050 habrá 30 millones de casos de EVC cada año y 200 millones de sobrevivientes de este padecimiento en todo el mundo (World Stroke Organization, 2020).

Se ha reportado una relación entre la infección por SARS-Cov2 y la ocurrencia de infartos cerebrales (Hernández-Fernández et al., 2020). Los reportes médicos mundiales señalan un aumento de la incidencia de EVC en sobrevivientes de Covid-19, que podría estar relacionado con una mayor adherencia de la sangre y propensión a la formación de trombos o coágulos (Beyrouiti et al., 2020). Existe el temor de que en los próximos años se presente un repunte de infartos cerebrales y padecimientos asociados.

2.1 Clasificación

A lo largo del tiempo se han propuesto diversas clasificaciones sobre el EVC, pero actualmente las más relevantes son el *Trial of Org 10172 in Acute Stroke Registry* (TOAST) (Chung et al., 2014) y la clasificación desarrollada por el *Oxfordshire Community Stroke Project* (OCSP) (Yang et al., 2016), también conocida como clasificación de Oxford o clasificación de Bamford.

De acuerdo con la clasificación TOAST, el EVC puede agruparse de acuerdo con su origen en isquémico y hemorrágico, estos dos grandes grupos se dividen a su vez en subgrupos que toman en cuenta otros criterios como la localización del EVC, el mecanismo que lo produce, entre otros.

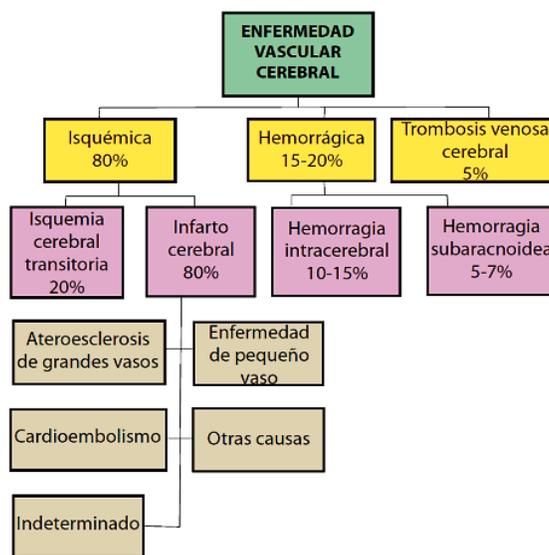


Figura 2.1 Frecuencia de los subtipos de EVC de acuerdo con la clasificación TOAST. (Arauz & Ruiz-Franco, 2012)

Como se muestra en la figura 2.1, en México el 80% de los casos reportados de EVC son de tipo isquémico, de los cuales el 20% son Isquemias Cerebrales Transitorias, también conocidas como Ataques Isquémicos Transitorios (AIT) y 80% son Infartos Cerebrales (IC). El EVC hemorrágico representa el 15-20% del total de casos. Dentro de este grupo el 10-15% se asocia a una hemorragia intracerebral y el 5-7% a una hemorragia subaracnoidea. Adicionalmente, existe un tercer grupo donde el EVC es provocado por una trombosis venosa cerebral, pero su incidencia a nivel mundial es muy baja (Arauz & Ruiz-Franco, 2012).

2.1.1 Isquemia Cerebral

Este tipo de EVC es provocado por una obstrucción total o parcial de un vaso del cerebro, lo que provoca una ausencia de riego sanguíneo. Dependiendo del tiempo de duración de los síntomas de la isquemia, si existe recuperación espontánea o no, y de estudios de radiología, se divide a la Isquemia Cerebral en AIT y en IC.

2.1.1.1 Isquemia Cerebral Transitoria

Su principal característica es que los síntomas no tendrán una duración mayor a 60 minutos. Dependiendo del área cerebral afectada los síntomas pueden ser motores o sensitivos puros, como la pérdida de movilidad o de sensibilidad unilateral, alteraciones del campo visual, disartria¹ y/o afasia². Posterior al evento, el paciente se recupera de manera espontánea, sin que exista un daño neuronal permanente. Suele ocurrir en las arterias carótidas y en las arterias vertebro basilares. Su aparición debe tomarse con cautela, ya que el 20% de pacientes con AIT que no sean identificados ni tratados pueden culminar en un IC en un periodo de 3 meses, siendo las dos semanas posteriores al evento el periodo de mayor riesgo (Arauz & Ruiz-Franco, 2012).

2.1.1.2 Infarto Cerebral

Son el tipo de EVC de mayor prevalencia y al igual que en el AIT, existe una disfunción cerebral focal asociada al área donde inicia la isquemia, pero en este caso el flujo sanguíneo cerebral no se restablece de forma espontánea. La zona donde ocurre la oclusión se conoce como foco o núcleo isquémico, mientras que a la zona aledaña se le denomina zona de penumbra isquémica, ambos conceptos se ilustran en la figura 2.2. Estas denominaciones se encuentran asociadas a la cantidad de flujo sanguíneo que recibe la zona del cerebro afectada. El flujo sanguíneo cerebral de una persona normal es de 50ml a 55ml por cada 100g de tejido encefálico por minuto. La zona de penumbra es un tejido que recibe entre 10 y 40ml por 100g por minuto, en este caso las neuronas presentarán alteraciones metabólicas e iónicas, pero aún conservarán su integridad estructural. Debido a que las neuronas de esta región aún se pueden recuperar, la zona de penumbra debe ser el objetivo a rescatar para disminuir el daño que pueda sufrir el paciente. Por otra parte, el foco

¹ Trastorno que afecta el control de los músculos de la boca, provocando alteraciones del habla.

² Trastorno del lenguaje que afecta la capacidad de comunicación de una persona.

isquémico es un tejido que recibe menos de 10ml por 100g por minuto, se trata de una zona infartada donde ha ocurrido muerte neuronal y ya no es rescatable.

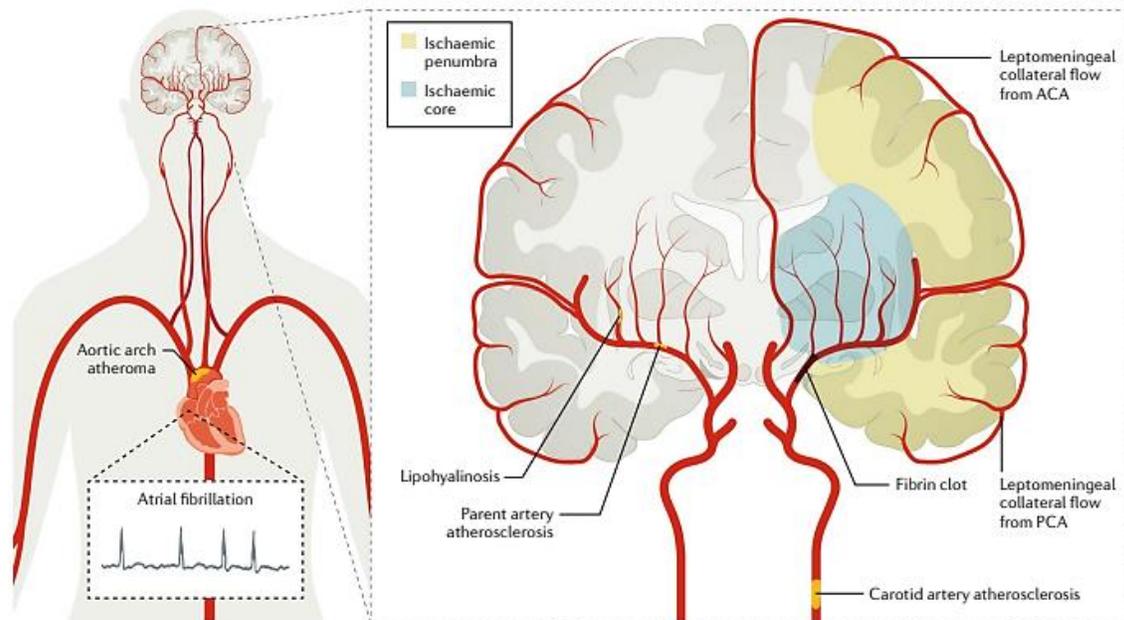


Figura 2.2. Se muestran algunas de las causas de un IC, como la aterosclerosis, oclusiones de pequeños vasos, así como émbolos originados en placas de ateroma en el corazón (cardioembolismo). También se ilustra el núcleo isquémico de la lesión, donde ocurre muerte neuronal, y la zona aledaña conocida como zona de penumbra. (Campbell et al., 2019).

La clasificación TOAST define 5 grupos de infartos cerebrales:

- a) Aterosclerosis de grandes vasos: Es el más común y suele ser el resultado de una oclusión trombótica o tromboembólica de los vasos. En esta última se forma una placa vulnerable en alguna arteria del cerebro que tiene riesgo de fragmentarse y embolizar³ de una arteria a otra arteria. La aterosclerosis extracraneal afecta principalmente la bifurcación carotídea, la porción proximal de la carótida interna y el origen de las arterias vertebrales. La aterosclerosis intracraneal suele ocurrir en la arteria cerebral media.
- b) Cardioembolismo: Oclusión de una arteria por un émbolo formado en el corazón. Algunas causas pueden ser un infarto al miocardio, lesiones valvulares, estenosis, endocarditis, fibrilación atrial, etc.
- c) Enfermedad de pequeño vaso cerebral: También conocido como infarto lacunar, son oclusiones pequeñas (zona de infarto menor a 15mm de diámetro) y profundas. Sucede en el 25% de los IC. Ocurre principalmente en arterias lenticuloestriadas y talamoperforantes, aquellas que irrigan las zonas más profundas del cerebro.

³ Proceso en el que un coágulo de sangre o un pedazo de placa se desprendió y viajó desde un lugar en el cuerpo a otro hasta bloquear el flujo sanguíneo de un vaso.

- d) Otras causas: Vasculitis, enfermedades autoinmunes como Lupus, obstrucción por hemoglobinopatías, hiperviscosidad de la sangre, hipercoagulabilidad, tumores, etc.
- e) Indeterminado: A pesar de una evaluación completa no se puede determinar la causa o se tiene una evaluación incompleta.

2.1.1.3 Isquemia cerebral global

Se emplea este término cuando la disminución del flujo sanguíneo cerebral se produce en todo el encéfalo de manera simultánea debido a una hipotensión arterial marcada. Este tipo de isquemia afecta a los hemisferios cerebrales de forma difusa, asociado o no a una lesión del tronco del encéfalo y/o cerebelo. Sus causas más frecuentes son el paro cardíaco y el bloqueo de ambas arterias carótidas, aunque cualquier cuadro de shock prolongado lo puede producir (e.g. ahogamiento, falta de oxigenación por intoxicación de CO₂, etc.). Las consecuencias clínicas abarcan desde un déficit cognitivo leve (memoria, atención) hasta el estado vegetativo persistente por necrosis neocortical, e incluso, si la duración es elevada, muerte cerebral por afectación troncoencefálica (Arauz & Ruiz-Franco, 2012).

2.1.1.4 Fisiopatología de la Isquemia Cerebral

Una vez que ocurre la oclusión arterial el flujo cerebral comienza a reducirse, provocando que las neuronas y la glía dejen de recibir la cantidad de oxígeno y glucosa que necesitan para trabajar. En el foco isquémico las neuronas dejan de sintetizar adecuadamente ATP y las bombas de sodio-potasio fallan, esto propicia que las neuronas comiencen a tener un exceso de sodio y se despolaricen bruscamente, liberando diversos neurotransmisores excitadores, entre ellos el glutamato. La retención de agua y sodio en los astrocitos provoca que se hinchen y ocurra un edema citotóxico. El exceso de glutamato extracelular se conoce como excitotoxicidad. El glutamato favorece la activación de los receptores NMDA, canales comúnmente bloqueados por magnesio y por donde ingresará calcio a las neuronas. Con los canales NMDA libres las neuronas presentan un exceso de calcio y sodio. El exceso de calcio provoca la producción de especies reactivas de oxígeno y otros mecanismos dependientes de calcio que provocan apoptosis. La excitotoxicidad viaja como una onda excitando a las neuronas de la zona de penumbra, esta activación eleva los requerimientos metabólicos de las neuronas, las cuales, al no tener suficiente oxígeno para suplir estos requerimientos, paulatinamente mueren y el núcleo isquémico comienza a crecer. Por otra parte, las especies reactivas de oxígeno provocan disfunción celular y se pierden conexiones intracelulares, lo que aumenta la permeabilidad de la barrera hematoencefálica, provocando un aumento en el volumen de agua y plasma en el espacio extracelular, estado conocido como edema vasogénico, con peligro de hemorragia al restablecerse el flujo sanguíneo (Campbell et al., 2019).

2.1.1.5 Síntomas principales iniciales de la Isquemia Cerebral

En la tabla 2.1 se resume la sintomatología que puede presentar una persona que sufre un EVC isquémico, dependiendo si la lesión se localiza en una arteria principal o en un pequeño vaso. El hemisferio del cerebro donde ocurre la lesión se denomina ipsilateral de la lesión o ipsilesional, mientras que el hemisferio no afectado se conoce como contralateral a la lesión o contralesional. Es común que los síntomas sensorimotores se presenten en los miembros contralaterales a la lesión debido a la decusación de múltiples fibras nerviosas que ocurren entre el bulbo raquídeo y la médula espinal, generando la vía piramidal, mejor conocida como tracto corticoespinal lateral, lo que permite que gran parte del control de un lado de nuestro cuerpo sea controlado por el hemisferio opuesto del cerebro.

Tabla 2.1 Síntomas principales que se presentan de acuerdo con la localización del EVC y el área cerebral comprometida (Arias Cuadrado, 2009; Campbell et al., 2019; Dobkin, 2011).

Arteria Cerebral	Área de lesión	Síntomas principales
Anterior Izquierda (ACA)	Áreas frontales y parasagitales del lóbulo frontal izq. (Corteza Motora y somatosensorial)	Debilidad y entumecimiento de miembros inferiores derechos. Pérdida de sensibilidad de miembros inferiores derechos. Afasia motora transcortical. Apraxia ideomotora (dificultad para coordinar).
Anterior Derecha (ACA)	Áreas frontales y parasagitales del lóbulo frontal der. (Corteza Motora y somatosensorial)	Debilidad y entumecimiento de miembros inferiores izquierdos. Pérdida de sensibilidad de miembros inferiores izquierdos. Apraxia ideomotora. Apatía. Alteraciones del razonamiento.
Media Izquierda (ACM)	Lóbulo frontal izq. (Corteza motora) Lóbulo parietal izq. (Corteza somato sensorial) Lóbulo temporal izq. (área de Wernicke) Lóbulo frontal izq. (área de Broca)	Afasia Hemianopsia derecha. Debilidad y entumecimiento principalmente de cara y brazo derecho y en un grado menor de la pierna derecha. Desviación de la mirada al lado izquierdo.
Media Derecha (ACM)	Lóbulo frontal der. (Corteza motora) Lóbulo parietal der. (Corteza somato sensorial) Estructuras subcorticales.	Debilidad y entumecimiento principalmente de cara y brazo izquierdo y en un grado menor de la pierna izquierda. Negligencia hemiespacial izquierda. Desviación de la mirada al lado derecho. Astereognosia (dificultad para reconocer objetos a través del tacto).
Posterior Izquierda (ACP)	Mesencéfalo superior. Corteza y subcorteza occipital izq. Tálamo. Lóbulo temporal medial izq.	Pérdida de la visión completa o parcial del lado derecho. Si el mesencéfalo se ve afectado parálisis y pérdida de sensibilidad del lado derecho (síndrome Weber). Alteraciones de la memoria.
Posterior Derecha (ACP)	Mesencéfalo superior. Corteza y subcorteza occipital der. Tálamo. Lóbulo temporal medial derecho.	Pérdida de la visión completa o parcial del lado izquierdo. Si el mesencéfalo se ve afectado parálisis y pérdida de sensibilidad del lado izquierdo (síndrome Weber) Alteraciones de la memoria.

Superior (ACS)	Tracto corticoespinal lateral. Lemniscos mediales.	Hemiparesis de miembros superiores e inferiores contralaterales. Pérdida de propiocepción. Pérdida de equilibrio.
Posterior Inferior (ACPI)	Médula lateral-núcleo vestibular. Tracto lateral espinotalámico. Núcleo trigémino espinal. Pedúnculo cerebeloso inferior.	Vértigo. Disminución de sensibilidad ante el dolor o temperatura del lado contralateral del cuerpo.
Anterior Inferior (ACAI)	Pons lateral. Núcleo vestibular. Núcleo facial. Núcleo coclear. Pedúnculos cerebelosos medial e inferior.	Vértigo. Sordera ipsilesional. Debilidad facial ipsilesional.
Basilar	Pons. Mesencéfalo inferior. Médula. Cerebelo. Tracto cortico espinal.	Parálisis total de brazos y piernas (cuadriplejía). Náuseas, vértigo, vómito. Coma. Deterioro del equilibrio y coordinación (ataxia). Pérdida de movimientos faciales voluntarios (disartria).
Arterias lenticuloestriadas de ACM o arterias perforantes de la arteria Basilar	Cápsula interna. Pons.	Únicamente hemiparesis motora del lado contralateral.
Arterias lenticuloestriadas de ACM o arterias tálamo perforantes de la ACP	Tálamo posterior.	Únicamente pérdida de sensibilidad del lado contralateral.
Arterias lenticuloestriadas de ACM	Cápsula interna. Materia blanca. Tálamo.	Únicamente pérdida motora y sensitiva del lado contralateral.
Arterias perforantes de la arteria Basilar	Cápsula interna. Materia blanca.	Debilidad en la mano y disartria.
Arterias lenticuloestriadas de ACM o arterias perforantes de la arteria Basilar	Cápsula interna. Base del pons.	Pérdida de control motor voluntario principalmente de la pierna, pero también puede ser de brazo (ataxia).

2.1.1.6 Factores de riesgo de la Isquemia Cerebral

Los factores de riesgo asociados a un EVC isquémico se pueden agrupar en no modificables y modificables.

No modificables:

- Edad. Existen diferencias entre países desarrollados (prevalencia en mayores de 49 años) y en vías de desarrollo (prevalencia en mayores de 39 años). Aunque un EVC isquémico puede producirse en todos los grupos de edades estudios muestran que el riesgo de tener un Infarto Cerebral se duplica por cada década entre los 55 y los 85 años (Dobkin, 2011).
- Género. En el estudio Global Burden of Disease en 2013 se observó que en hombres se presentaron 133 casos por cada 100,000 personas, mientras que en mujeres ocurrieron 99 casos por cada 100,000 personas (Arias Cuadrado, 2009).
- Genética. Estudios han encontrado que hasta 42% de pacientes EVC Isquémico tienen un familiar que también ha presentado un EVC. Muchos factores de riesgo modificables son

resultado de factores genéticos, como la presión arterial alta o la diabetes (Arauz & Ruiz-Franco, 2012).

Modificables:

- Hipertensión arterial. Considerado el factor de riesgo de mayor impacto.
- Baja actividad física regular (vida sedentaria).
- Una alta proporción de apolipoproteína B (ApoB) respecto a ApoA1.
- Dieta.
- Estrés psicosocial y depresión.
- Tabaquismo.
- Causas cardiacas (e.g. fibrilación atrial, previos infartos al miocardio).
- Alto consumo de alcohol.
- Diabetes melitus.
- Otros: apnea del sueño, enfermedades crónicas del riñón, enfermedades periodontales, anticonceptivos orales, exposición a aire contaminado.

2.1.2 EVC Hemorrágico

Este tipo de EVC se caracteriza por la ruptura de un vaso sanguíneo del cerebro (figura 2.3). La sangre contiene sustancias lipofílicas que son potencialmente neurotóxicas y al entrar en contacto con las neuronas llega a provocar su muerte.

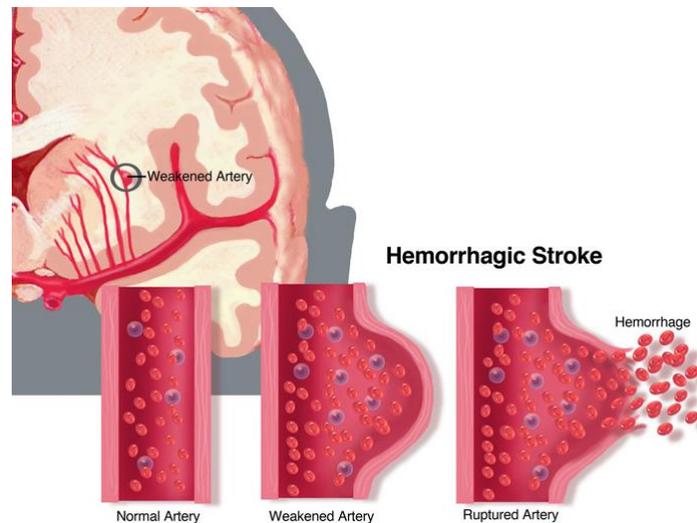


Figura 2.3 Ilustración de un EVC hemorrágico causado por la ruptura de un vaso sanguíneo debilitado. (Basic-Science, 2020).

Gran parte del daño ocasionado por este EVC también es causado por la presión que ejerce la sangre al comprimir a las pequeñas arterias vecinas al aumentar el volumen intracraneal y limitar el flujo sanguíneo de esa zona. Dependiendo de la localización de la lesión se puede dividir en los siguientes grupos:

2.1.2.1 Hemorragia intracerebral

También conocido como hemorragia intraparenquimatosa. Ocurre por la ruptura espontánea de un vaso sanguíneo profundo del cerebro, formando un hematoma. Su incidencia es de 10 a 20 casos por cada 100,000 habitantes al año y se duplica cada 10 años después de los 35 años. Este tipo de EVC presenta una alta mortalidad, ya que sólo el 38% de los casos sobrevive al pasar 1 año, mientras que el 30% logra ser independiente a los 3 meses (Arauz & Ruiz-Franco, 2012). Se denomina hemorragia primaria cuando, ocurrido el EVC, sólo se presenta sangre en los ventrículos. Las causas de la hemorragia primaria suelen ser una hipertensión arterial sistémica o una angiopatía amiloide, donde las proteínas β -amiloides se acumulan en las paredes de las arterias cerebrales, debilitándolas y provocando su posterior ruptura. Se le conoce como hemorragia secundaria cuando la irrupción de sangre procede del parénquima cerebral, por ejemplo, cuando se transforma un EVC isquémico en hemorrágico.

Una de las causas de la hemorragia intracerebral es la ruptura de aneurismas conocidos como aneurismas de Charcot y Bouchard, los cuales se presentan en capilares cerebrales no mayores de 300 μ m ubicados en los vasos lenticuloestriados y talamicoperforantes que irrigan el putamen y el núcleo caudado, su ruptura es comúnmente causada por la hipertensión.

La hemorragia intracerebral ocurre en un 50% de los casos en ganglios basales, 35% en ganglios lobares y en un 6% en tronco cerebral o cerebelo (Arauz & Ruiz-Franco, 2012).

Los síntomas que puede presentar una persona que sufre este tipo de EVC son variados y se caracterizan por su aparición súbita y rápidamente progresiva, entre ellos se encuentran un déficit sensitivo y/o motor del lado contralateral de la lesión, hemiplejía, ataxia o dismetría, una fuerte cefalea, náuseas, vómito, así como crisis convulsivas en el 15% de los casos (De la Garza-Longoria et al., 2018).

2.1.2.2 Hemorragia subaracnoidea

Este tipo de hemorragia se da principalmente por la invasión o la presencia de sangre en el espacio subaracnoideo debido principalmente a la ruptura de un aneurisma intracraneal. Representa entre el 4 al 7% de todos los EVCs (Arauz & Ruiz-Franco, 2012). Tiene un alto índice de mortalidad, el 45% de los pacientes fallece en los primeros 30 días y el 50% de los supervivientes evolucionan con secuelas irreversibles. Su incidencia es de 10.5 casos por 100,000 personas al año y afecta principalmente a la población menor de 65 años (Chiquete et al., 2011). En este tipo de EVC se le denomina hemorragia primaria cuando el sangrado tiene lugar directamente en el espacio subaracnoideo y secundaria cuando el sangrado se produjo inicialmente en otro lugar, como el parénquima cerebral. Su causa principal es un traumatismo. El 80% de las causas de origen no traumático es la ruptura de un aneurisma, principalmente en las bifurcaciones que se presentan en el polígono de Willis. Entre el 30-35% de la ruptura de los aneurismas se presentan en la Arteria Cerebral Media (ACM), entre 30-35% en la bifurcación entre la Arteria Cerebral Posterior (ACP) y la Arteria Carótida Interna (ACI), entre 30-35% se presentan en la Arteria Cerebral Anterior, 5% se ubican en el tope basilar del polígono de Willis (Dobkin, 2011). Como causas mucho menos frecuentes se encuentran malformaciones vasculares y neoplastias.

Los síntomas iniciales más característicos de un EVC hemorrágico son:

- Presencia súbita de una cefalea intensa.
- Náuseas, vómito, fotofobia y alteración de la conciencia.
- Pérdida de control de esfínteres o abulia.
- Hemiparesia, afasia o negligencia visoespacial.

2.2 Campañas de identificación de síntomas

En cualquier tipo de EVC es de vital importancia reconocer los signos de alerta que puede presentar una persona para identificar si está sufriendo un EVC. Por esta razón se creó en México la campaña CAMALEON, cuyas siglas ayudan a identificar los síntomas de una persona que está sufriendo un EVC. CA hace referencia a si la persona presenta cara colgada, MA si alguna de las manos la siente pesada, LE hace referencia al lenguaje, si la persona tiene problemas para articular palabras, finalmente si se presenta alguno de estos síntomas ON se refiere a actuar inmediatamente y llamar al 911.

En Estados Unidos existe una campaña similar conocida como FAST, cuyas siglas también hacen alusión a los síntomas de un EVC, siendo F referencia a *face droops*, A significa *arm weakness*, S si se observa *speech difficulty* y finalmente, T asociado a *time is critical*.

En el caso de un EVC isquémico, entre más rápido se logre restablecer y controlar el flujo sanguíneo cerebral, así como reducir la presión intracraneal, disminuirá el avance de la zona de muerte neuronal y aumentarán las probabilidades de recuperación del paciente. Los medios para deshacer el coágulo sanguíneo pueden ser mecánicos (trombectomía) o a través de fármacos (trombólisis). Por otra parte, cuando se trata de un EVC hemorrágico lo más importante es detener la hemorragia lo antes posible, para lograr esto comúnmente se recurre a una neurocirugía.

2.3 Fases de evolución de un EVC

Las etapas más ampliamente aceptadas en la comunidad para clasificar a un paciente EVC de acuerdo con el tiempo transcurrido desde que presenta síntomas (Arias Cuadrado, 2009; Figueroa et al., 2014) son:

- Fase hiperaguda. Cuando han pasado máximo 6 horas desde que se presentaron los primeros síntomas.
- Fase aguda. Cuando han transcurrido entre 6 y 48 horas después del evento.
- Fase subaguda. Cuando el paciente presenta una estabilidad clínica hasta menos de 6 meses de evolución.
- Fase crónica. Son aquellos pacientes de más de 6 meses de evolución.

Cabe resaltar que estas fases son una aproximación, cada caso es diferente y la duración de cada fase estará definida por la exploración física y la observación de cambios, más que por un criterio temporal.

2.4 Complicaciones post EVC

Una persona que ha sufrido un EVC puede presentar una gran variedad de déficits y discapacidades como consecuencia de múltiples factores, como el tipo de EVC, lugar y tamaño de la lesión, dominancia (izquierda, derecha o ambidiestro), variantes anatómicas vasculares, etcétera. La *American Heart Association-Stroke Outcome Classification* (AHA-SOC) sistematiza los déficits neurológicos en 6 dominios o áreas: motora, sensitiva, comunicación, visual, cognitiva y emocional.

2.4.1 Déficits motores

El déficit motor suele ser unilateral y es la principal causa de discapacidad física. Algunos de los déficits pueden ser:

- Apraxia. Es una alteración de las habilidades motoras voluntarias, por ejemplo, peinarse o abrocharse los botones de una camisa. Los pacientes pueden realizar algunas acciones automáticamente, pero no de manera voluntaria.
- Hemiplejia. Parálisis total de un lado del cuerpo (brazo, pierna y tronco). Generalmente la hemiplejia se presenta en la fase aguda y pasa a volverse hemiparesia a partir de la fase subaguda.
- Hemiparesia. Debilidad o dificultad para realizar movimientos voluntarios. Existe un mayor grado de movimiento respecto a la hemiplejia.
- Afectaciones en la función motora del tronco, así como pérdida o disminución del control postural.
- Ataxia. Ocurre cuando hay daño en el cerebelo o en algunas de sus vías aferentes o eferentes. Provoca hipotonía, y alteraciones en la coordinación de movimientos complejos, así como dificultad para realizar movimientos secuenciales y simultáneos.
- Disartria hipocinética. Ocurre cuando hay lesión que afecta a los ganglios de la base o a sus vías aferentes o eferentes. Se caracteriza por una reducción en la velocidad de movimientos debido a la rigidez muscular.
- Disartria hipercinética. También consecuencia de daño a los ganglios basales. Ocurren movimientos anormales involuntarios e impredecibles, aparecen en reposo, durante posturas mantenidas o mientras se realizan movimientos voluntarios.

La mayoría de estas complicaciones suelen observarse cuando existe daño en la vía piramidal o también llamada tracto cortico espinal, a este déficit se le conoce como Síndrome de la neurona motora superior (Dobkin, 2011). Si el daño ocurre en la cápsula interna del cerebro es probable que se presente una parálisis en ambos brazos y piernas (tetraparesia).

2.4.1.1 Espasticidad

La espasticidad es una alteración donde existe una pérdida del balance entre la contracción y relajación de los músculos antagonistas y agonistas, que lleva a un estado de rigidez y espasmos musculares involuntarios resultantes de estímulos mínimos internos o externos. Se caracteriza por una tensión muscular, rigidez y exageración en los reflejos tendinosos profundos que interfieren con la actividad muscular, la marcha, el movimiento o el lenguaje. También existe pérdida de reflejos cutáneos, por ejemplo el reflejo de Babinski (Dobkin, 2004).

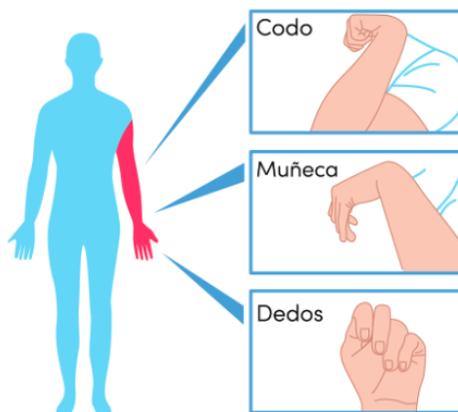


Figura 2.4 Se ilustran las alteraciones que se pueden presentar en miembro superior con espasticidad. (LANR, 2020).

Es común que la espasticidad en miembros superiores se presente como flexión de muñeca, mano cerrada atrapando el dedo pulgar y codo flexionado (figura 2.4). En el caso de los miembros inferiores suele presentarse como una flexión de rodilla y tobillo.

La espasticidad puede afectar cualquier músculo del cuerpo, haciendo que el movimiento sea difícil y torpe.

Su prevención debe iniciar precozmente con movilizaciones lentas en sentido contrario al reflejo de estiramiento, en progresión próximo distal y con posturas de inhibición antiespástica o férulas. En el ámbito clínico y de investigación, una de las herramientas más utilizadas para su evaluación es la Escala de Ashworth Modificada (Bohannon & Smith, 1987)

2.4.2 Déficit de comunicación

De acuerdo con la literatura (Arias Cuadrado, 2009; Dobkin, 2011; Figueroa et al., 2014), los problemas de comunicación asociados con frecuencia al EVC son:

- Disfasia. Afectación en la comprensión del habla o al uso de las palabras correctas.
- Afasia. Incapacidad para expresarse o de comprender el lenguaje, ya sea escrito o hablado.
- Disartria de músculos de la boca y lengua. Es un problema de articulación y puede provocar un arrastre del habla.

El trastorno más frecuente es la afasia. Más específicamente, se han identificado diversos tipos de afasia que se pueden dividir en:

- Afasia de Broca. Dificulta la capacidad para expresarse.
- Afasia de Wernicke. Dificultad para comprender.
- Afasia por disartria. Dificultad para articular palabras debido a una afección en los músculos de la boca y lengua, provocando problemas de articulación y arrastres del habla.
- Alexia. Incapacidad para leer.
- Agrafia. Incapacidad para escribir.
- Acalculia. Dificultad para realizar cálculos.

2.4.3 Déficits cognitivos

Muchos pacientes presentan cierta pérdida cognitiva en la fase aguda, sin embargo, en la fase subaguda, cuando el paciente ha sido estabilizado y regresa a casa, cobran importancia funciones más complejas y se evidencian problemas al intentar realizar las actividades que antes efectuaba cotidianamente.

Las principales afecciones cognitivas durante la fase aguda (Figuroa et al., 2014) son:

- Orientación. Dificultad para expresar conocimiento sobre sí mismo y dónde está temporal y espacialmente.
- Atención e Inhibición. Dificultad para seleccionar el estímulo que le sea útil para la tarea o proceso mental que está llevando a cabo. Particularmente se pueden observar déficits en la atención focalizada, atención mantenida y capacidad de alerta. Es una de las afecciones con mayor prioridad en ser atendida debido a que su alteración afecta al resto de las funciones cognitivas y motrices.
- Lenguaje (e.g. afasias). En la fase aguda el 21-33% de los pacientes presentan alteraciones del lenguaje acompañadas de afectación sensoriomotora derecha o hemianopsia.

Principales afecciones cognitivas en la fase subaguda (Arias Cuadrado, 2009; Dobkin, 2004; Figuroa et al., 2014):

- Memoria de trabajo, memoria declarativa, memoria retrógrada, memoria autobiográfica, memoria semántica, memoria anterógrada y memoria no declarativa o implícita.
- Atención selectiva y atención alternante.
- Agnosia auditiva, sensitiva o visual y trastornos visoperceptivos (No son problemas visuales, si no de reconocimiento).
- Funciones ejecutivas como la planificación, velocidad de procesamiento, toma de decisiones y fluidez verbal.
- Apraxia ideomotora e ideativa. No se trata de una incapacidad física, si no de la dificultad para planificar el acto motor.
- Habilidades de cálculo y constructivas, así como disminución en la habilidad de aprender y retener información.

2.4.3.1 Demencia vascular

Es la demencia secundaria a lesiones cerebrales causadas por enfermedad cerebrovascular. Su frecuencia es aproximadamente del 10 al 20% del total de las demencias (Arias Cuadrado, 2009). Sus características clínicas son:

- Deterioro cognitivo progresivo en los tres meses siguientes a un ictus con posterior evolución fluctuante o escalonada. Esto ocurre en el 30% de los sobrevivientes.
- Historia de alteración de la marcha y caídas frecuentes.
- Signos de focalidad en la exploración neurológica.
- Predominio de déficit cognitivo subcortical.
- Presencia de factores de riesgo vascular.

Se debe realizar siempre el diagnóstico diferencial con otras demencias, fundamentalmente con la enfermedad de Alzheimer y con aquellas que presenten características mixtas entre ésta y la demencia vascular.

2.4.4 Déficit emocionales

- Depresión: Es un problema clínico complejo. Depende de muchos factores como edad, deterioro cognitivo, grado de discapacidad, apoyo familiar, secuelas, etc.
- Labilidad emocional: Presencia de reacciones de llanto, en menos frecuencia risa, de intensidad desmedida, al margen de control voluntario y provocado por estímulos menores.
- Reacción catastrófica: Reacciones súbitas de llanto, enfado y temor ante el hecho de fracasar al afrontar una tarea.
- Apatía: Se presenta principalmente cuando por lesión en el área prefrontal dorsolateral. Ausencia de iniciativa y actividad, además de ausencia de emociones.
- Anosognosia: Falta de conciencia sobre la discapacidad y sus consecuencias.
- Ansiedad.
- Síndrome de desinhibición: Se produce tras una lesión en el área orbito frontal, que está relacionada con la respuesta emocional y de inhibir la conducta. Puede manifestarse como irritabilidad, agresividad, inflexibilidad, euforia y conducta inapropiada socialmente.
- Egocentrismo: Dificultad para empatizar con otras personas

2.4.5 Otros déficits

- Alteraciones sensitivas contralaterales.
- Heminegligencia. Es la dificultad que presenta el individuo para orientarse, actuar o responder a estímulos o acciones que ocurren en el lado contralateral a una lesión hemisférica, que no es debida a trastornos elementales sensoriales o motores (López Arguelles et al., 2012).
- Deficiencias visuales contralaterales como pérdida de visión en la mitad del campo visual (hemianopsia), del movimiento del ojo, anomalías del campo visual y dificultades de percepción visual.
- Incontinencia urinaria, ocurre en el 60% de los pacientes en la primera semana.

- Dolor en hombro parético, presente entre el 5-50% de los pacientes. El dolor incrementa la hipertonicidad y provoca espasmos flexores y extensores, así como la flexión de codo y muñeca.
- Dolor lumbar, cervical, en la cadera, rodillas o talón, principalmente debido a sobreuso, sobre estiramiento, o a malas posturas al descansar, así como a un mal manejo del paciente.
- Es común que si la lesión ocurrió en el núcleo ventroposterolateral del tálamo se presente dolor central espontáneo (síndrome del dolor central).
- Sentimiento de fatiga o cansancio. Se desconoce la causa, pero una explicación probable es la relación con el desacondicionamiento físico. En ocasiones puede deberse a un desorden psicológico.
- Disfagia. Dificultad para tragar líquidos y sólidos. Aproximadamente la mitad de las personas que sufren un EVC presentan disfagia.
- Úlceras en la piel.
- Insomnio o apnea del sueño.

Capítulo 3

Rehabilitación posterior a un EVC

El proceso de rehabilitación en pacientes que han sufrido un EVC debe iniciar de forma precoz en cuanto se encuentren estables o establecidos para optimizar la recuperación. Se considera a un paciente estable cuando permanece sin cambios más de 24 horas en caso de un EVC en territorio carotideo y más de 72 horas para los vertebro basilares. El EVC establecido o permanente se considera a las 3 semanas. No está indicado iniciar la rehabilitación en un EVC donde existe una progresión de síntomas o cambios en el déficit neurológico, sobre todo durante las primeras 24 horas (Arias Cuadrado, 2009).

Para llevar a cabo un adecuado programa de rehabilitación del paciente con ictus cerebral se deben considerar 3 aspectos fundamentales:

1. Valorar las lesiones y el déficit funcional antes de la intervención y durante su evolución.
2. Hacer una estimación del pronóstico más probable de recuperación.
3. Con base en lo anterior, establecer un plan terapéutico individualizado para cada paciente.

Los objetivos principales del plan terapéutico de rehabilitación deben ser los siguientes:

1. Prevenir y tratar complicaciones.
2. Mantener o recuperar las funciones orgánicas.
3. Recuperar las capacidades funcionales perdidas.
4. En caso de que no sea posible la recuperación, se debe trabajar en una adaptación a las funciones residuales.

El proceso de rehabilitación debe abordarse a través de un equipo multidisciplinario y coordinado que aborde el conjunto de áreas afectadas. Idealmente este equipo estaría integrado por médicos especialistas, fisioterapeutas, terapeutas ocupacionales, logopedas, neuropsicólogos, personal de enfermería y asistentes sociales. Los familiares y cuidadores serán una parte esencial del equipo de rehabilitación, pues ellos serán imprescindibles para la toma de decisiones, el apoyo emocional, la aceptación social y el mantenimiento de las ganancias obtenidas con el tratamiento (Figuerola et al., 2014).

3.1 Valoración del déficit funcional

La valoración del paciente debe ser amplia, dada la gran variedad de déficits y discapacidades que un EVC puede provocar a consecuencia de múltiples factores como el tipo de EVC, lugar y tamaño de la lesión, dominancia, variantes anatómicas vasculares, etcétera. En la fase aguda es aconsejable utilizar una escala específica para ictus, que permita una valoración global de déficits

neurrológicos, junto con una escala de valoración de actividades de la vida diaria (AVD) básicas; de esta forma se podrán detectar las áreas que requieren mayor atención y se determinará la discapacidad resultante. Las escalas de valoración global recomendadas por la AHA-SOC son la *National Institute of Health Stroke Scale* (NIH-SS) y la *Canadian Neurological Scale*, que evalúan el nivel de consciencia, paresia facial, de extremidades superiores y de extremidades inferiores, ataxia, sensibilidad, lenguaje, disartria y heminegligencia. Por otra parte, el Índice de Barthel es una escala de uso habitual para valorar la discapacidad en AVD, pues ha demostrado ser válida, fiable y de aplicación sencilla. Se trata de un cuestionario donde se le pide al evaluado que evalúe su capacidad para realizar tareas como comer, lavarse, vestirse, caminar, etc. con una puntuación de 0 (dependiente), 5 (necesita ayuda para algunas cosas), o 10 (totalmente independiente). Además, es una escala sensible a cambios, aunque tiene un efecto techo y no logra detectar cambios en los niveles más altos de función (Stone et al., 1994).

En la fase subaguda se realiza la valoración de déficits por áreas de función, que aporten información más específica a través de escalas de medición estandarizadas. Estos instrumentos traducen la valoración clínica en resultados objetivos y cuantificables. Las principales características que deben de cumplir estas escalas son: contar con propiedades psicométricas de alto nivel, es decir, que tengan una alta fiabilidad inter e intra observador, que demuestren validez y reproducibilidad, que sean prácticas, sencillas y muy específicas con el método de administración y puntuación; además, también deben ser aceptadas por la comunidad clínica y deben estar adaptadas para el medio en el que se desenvuelve el paciente.

En el caso específico de la rehabilitación, las escalas de medición permiten:

- Identificar problemas y establecer objetivos.
- Evaluar la eficacia de las intervenciones.
- Dar un valor pronóstico.
- Realizar un seguimiento para observar si los cambios obtenidos se mantuvieron después de la rehabilitación.

A continuación, se mencionan las áreas sistematizadas por la AHA-SOC y algunas escalas comunes que se emplean para su evaluación.

3.1.1 Área motora

Como se mencionó en el capítulo anterior, el déficit motor suele ser unilateral y contralateral al área de la lesión. La hemiparesia y la hemiplejia son los déficits más prevalente tras un EVC y son la principal causa de discapacidad física en adultos (Krakauer & Cortés, 2018). La recuperación de la actividad motora suele seguir un orden de los miembros más proximales al tronco a los más distales (Verheyden & Ashburn, 2013).

Existe una plétora de escalas enfocadas en la función motriz, algunas de las más importantes son las siguientes:

Motricity Index (Collin & Wade, 1990). Evalúa la actividad motora voluntaria o fuerza muscular isométrica máxima del brazo y la pierna hemiparéticos. Esta prueba se evalúa en posición sentada y examina la prensión de la pinza, la flexión del codo y la abducción del hombro, así como la flexión dorsal del tobillo, la extensión de la rodilla y la flexión de la cadera. Estos 6 movimientos se gradúan en una escala ordinal de 0 a 33.

Trunk Control Test (Collin & Wade, 1990). Evalúa el control del tronco pidiendo al paciente que realice cuatro maniobras: desde decúbito supino, rodar hacia el lado débil, la segunda tarea consiste en rodar hacia el lado menos afectado, la tercera tarea consiste en pasar a sedestación desde tumbado y la última tarea es mantener el equilibrio estando sentado al borde de la cama. Cada maniobra se califica en una escala ordinal de 3 puntos, que va de 0 a 25 puntos.

Berg Balance Test (Berg et al., 1995). Evalúa de modo funcional el equilibrio estático y dinámico. La escala consta de 14 actividades puntuadas en una escala ordinal de 5 puntos (0 a 4). Las tareas son pasar de sentado a levantado, mantenerse levantado sin ayuda, mantenerse sentado sin ayuda, pasar de levantado a sentado, transferencias de peso, mantenerse levantado con los ojos cerrados, mantenerse levantado con los pies juntos, inclinarse hacia adelante, levantar un objeto del suelo, girarse para mirar atrás, dar una vuelta completa sobre su propio eje, colocar de forma alternada un pie sobre un taburete, mantenerse levantado con un pie adelante y mantenerse levantado sobre una pierna. La puntuación total varía entre 0 y 56 puntos.

Fugl-Meyer (Fugl-Meyer et al., 1975). Es una de las más utilizadas en el ámbito clínico por su alta validez y confiabilidad. Evalúa y mide la recuperación en pacientes EVC en 5 dominios: movimiento voluntario, movimiento pasivo, coordinación, sensibilidad y dolor. Se divide en miembro superior y en miembro inferior. En general, cada tarea se califica con un 0 si no la puede realizar, 1 si la realiza parcialmente y 2 si la realiza completamente. Tanto la sección enfocada en miembros superiores, como la de extremidades inferiores, evalúan la actividad refleja, el movimiento voluntario dentro de sinergias⁴ sin ayuda de la gravedad, movimiento voluntario mezclando sinergias sin compensación, movimiento voluntario con poca o ninguna sinergia y la actividad refleja normal.

Wolf Motor Function Test (WMFT) (Wolf et al., 2005). Cuantifica la habilidad motriz del miembro superior parético a través de 17 tareas funcionales cronometradas. Se puntúa en una escala ordinal que abarca desde la imposibilidad de ejecutar la tarea (puntuación=0) hasta la normalidad (puntuación=5). Todas las actividades se deben realizar tan rápido como sea posible y si alguna tarea no se puede realizar se penaliza con un tiempo máximo de 120s.

Chedoke Arm and Hand Activity Inventory (CAHAI) (Barreca et al., 2004). Esta escala evalúa la habilidad funcional del miembro superior hemipléjico para realizar AVD. Cada tarea se evalúa con una escala de 7 puntos, donde el mínimo 1 indica que el paciente requiere una asistencia total y el máximo 7 se traduce en una independencia completa en completar la tarea. En la versión CAHAI-7, las tareas a realizar son abrir un frasco de café, llamar al 911, dibujar una línea con una regla,

⁴ Activación de un grupo de músculos para contribuir a la ejecución de un movimiento particular.

servir un vaso de agua, exprimir un trapo, abotonarse una camisa y secarse la espalda con una toalla. También existe una versión más extensa con 13 tareas (CAHAI-13).

Ashworth Modified Scale (Bohannon & Smith, 1987). Es una graduación empleada para valorar la espasticidad muscular, consiste en realizar un movimiento de flexión o extensión pasiva del miembro espástico. La escala comprende 6 marcas. Se asigna la marca 0 cuando no hay cambios en la respuesta del músculo en los movimientos de flexión o extensión. Se asigna 1 cuando existe un ligero aumento en la respuesta del músculo al movimiento. Se asigna 1+ cuando ocurre un ligero aumento en la resistencia del músculo en menos de la mitad del arco de movimiento. Una marca de 2 se asocia a un notable incremento en la resistencia del músculo durante la mayor parte del arco de movimiento. Una marca de 3 significa que el paciente presenta una marcada resistencia del músculo, causando que el movimiento pasivo sea difícil. Finalmente, se asigna 4 cuando las partes afectadas están rígidas en flexión o extensión cuando se mueven pasivamente.

Nine Hole Peg Test (9HPT) (Kellor et al., 1971). Esta prueba mide la destreza de los dedos. Contrario al resto de las pruebas aquí no hay una escala, en cambio, esta prueba es cronometrada y el tiempo que tarde el paciente en completarla indicará su grado de destreza. La prueba consiste en tomar 9 clavijas de un contenedor, una por una, y colocarlas en los hoyos de un tablero, posteriormente el evaluado debe quitarlas y devolverlas a su contenedor original una por una. El paciente debe realizar la tarea lo más rápido posible y utilizar únicamente la mano que se está evaluando.

Functional Ambulation Category (Holden et al., 1986). Evalúa el grado de dependencia ambulatoria. Existen 6 categorías y el paciente es asignado a una de ellas acorde con sus capacidades, donde la categoría 0 significa que es incapaz de caminar, 1 significa que el paciente necesita ayuda para llevar peso y equilibrio, 2 se refiere a que necesita ayuda de una persona para equilibrio y coordinación, 3 es para quien necesita solo supervisión verbal y ayuda sin contacto físico, 4 es la categoría donde el paciente necesita ayuda sólo en escaleras, pendientes y superficies irregulares, finalmente, la categoría 6 es para los pacientes que pueden caminar en cualquier superficie de forma independiente.

3.1.2 Área perceptiva, sensitiva y visual

Para la evaluación de este tipo de déficits no existe ningún test específico. Los reflejos, la disimetría, somestesia⁵ y negligencias se exploran de manera habitual. Este tipo de déficits se detectan rápidamente con la escala NIH-SS en la etapa aguda, también existen algunos test psicométricos que evalúan, aunque no de manera aislada, la heminegligencia, tales como los test de cancelación o de tachado como el Test de las Campanas (Gauthier et al., 1989).

⁵ Capacidad de percibir sensaciones con todo el cuerpo, como la temperatura, la presión y el dolor.

3.1.3 Comunicación y lenguaje

Cuando se aprecian signos de disfasia puede analizarse el lenguaje en sus vertientes de comprensión, denominación, repetición, fluencia, lectura y escritura. El test de Boston es el más recomendado para la valoración exhaustiva de la afasia, aunque es una prueba de aplicación lenta y compleja. Otros test que permiten identificar y valorar afasias son el Token Test, la Batería Western y el Test de despistaje de afasias de Friedrich (Figuerola et al., 2014).

3.1.4 Dominio cognitivo

Un EVC puede dañar la memoria, la atención, la orientación visoespacial, las habilidades de cálculo y constructivas, así como la habilidad de aprender y retener información. Dos de las pruebas más sencillas y rápidas de aplicar son el *Mini Mental State Examination* (MMSE) y el *Montreal Cognitive Assessment* (MoCA), aunque los resultados de ambas pueden verse sesgados por el nivel de escolaridad del paciente y si se encuentra en situación de afasia. Algunos de los test más exhaustivos son WEIS y la Batería neuropsicológica BANFE. Uno de los aspectos que más se tiene que vigilar en pacientes que han sufrido un ictus es el deterioro de las funciones cognitivas, ya que esto puede indicar el desarrollo de una demencia vascular (Devesa Gutiérrez et al., 2014).

3.1.5 Área emocional

A pesar de que pueden presentarse diversos cambios emocionales tales como irritabilidad y apatía, el problema clínico más común y complejo que se presenta es la depresión. En ese respecto la Escala de Depresión Geriátrica (GDS) es breve y fácil de aplicar, pero tiene una alta tasa de falsos negativos en casos de depresión menor. Otra prueba comúnmente usada para detectar y evaluar la gravedad de la depresión es el Inventario de Depresión de Beck (BDI-II). La depresión debe ser confirmada o descartada mediante una entrevista clínica (Arias Cuadrado, 2009).

3.1.6 Discapacidad instrumental

Las AVD instrumentales se refieren a la posibilidad de usar transporte, usar un teléfono, realizar compras, asearse, vestirse, comer, preparar comida o hacer las tareas de la casa. Algunas escalas que evalúan la capacidad de realizar estas tareas son el Índice de Barthel, la Medida de Independencia Funcional (FIM) y el *Frenchay Activity Index* (Verheyden & Ashburn, 2013).

3.2 Plasticidad neuronal después de un EVC

El cerebro no es un órgano rígido, incluso después de haber alcanzado la edad adulta siempre está constantemente cambiando ante los distintos estímulos que recibe del ambiente, generando nuevas conexiones neuronales y perdiendo otras, lo que nos permite aprender y adaptarnos ante nuevas situaciones. A esta propiedad tan fundamental, aunque no exclusiva, del cerebro se le conoce como plasticidad neuronal o neuroplasticidad, que puede definirse de forma general como la capacidad adaptativa del sistema nervioso central para modificar su propia organización estructural y funcional, lo que le permite tener una respuesta adaptativa (o maladaptativa) a la

demanda funcional (Krakauer & Cortés, 2018). Esta capacidad le permite al cerebro minimizar los efectos de las lesiones a través de la modificación de su propia organización estructural y funcional, sin ello cualquier daño al sistema nervioso central sería irreversible.

Las bases de la adquisición, retención y recuperación de la información en un cerebro sano no son diferentes a las de un cerebro que ha sufrido un EVC. La potenciación a largo plazo es uno de los mecanismos moleculares en el que las sinapsis y un grupo de neuronas codifican nueva información para representar una habilidad motriz (memoria procedural) o retener información acerca de hechos y eventos (memoria declarativa). La potenciación a largo plazo se desarrolla gracias a la repetición de estímulos asociados hacia neuronas en la corteza motora mientras participan en una actividad, por ejemplo, cuando se practica una nueva secuencia de movimientos de los dedos para tocar el piano. Estas neuronas, junto con otras que están asociadas con ellas en regiones sensorimotoras, representarán los movimientos de la habilidad adquirida, mientras que esas mismas neuronas también participan en la ejecución de otros movimientos. Se formarán conexiones sinápticas más fuertes entre grupos de neuronas que representen movimientos más complejos y que reciban el mismo estímulo constantemente, de ahí la famosa frase de Donald Hebb "*Neurons that fire together, wire together*". Por otra parte, las conexiones sinápticas entre neuronas pueden debilitarse e incluso desaparecer con el tiempo si los grupos de neuronas dejan de recibir estímulos, proceso conocido como depresión a largo plazo (Hoyer & Celnik, 2011).

Una vez que ha ocurrido el EVC y se ha logrado controlar el proceso de muerte neuronal, el tejido dañado queda inflamado (edema cerebral) y puede suceder un fenómeno conocido como diasquisis, que ocurre cuando se ve afectado el funcionamiento de áreas distales a la lesión, inclusive de hemisferios opuestos. Los déficits causados por la inflamación disminuirán conforme ésta ceda. Otro factor que provocará diasquisis es el decremento en la actividad metabólica inmediatamente después del evento (Hylin et al., 2017).

Mientras que el EVC desata eventos devastadores para el cerebro, también induce una cascada de eventos que les permite a las neuronas que se encuentran tanto cerca como lejos del foco de la lesión sobrevivir, repararse y formar nuevas conexiones. Algunas investigaciones (Kerr et al., 2011; Kolb & Gibb, 2011) incluso han propuesto que varios de los eventos que ocurren después de la lesión son similares a los que se observan durante el desarrollo normal de crecimiento, sugiriendo que el cerebro tiene una capacidad intrínseca de reaccionar ante el cambio. Los cambios en la zona de penumbra que incluyen una reorganización cortical, la neurogénesis (sólo ocurre en áreas muy específicas), los brotes axonales, la plasticidad dendrítica y la angiogénesis, han sido relacionados con una recuperación endógena de diferentes déficits conductuales después de la resolución de la diasquisis (Hylin et al., 2017).

Dentro del área perilesional (área aledaña al núcleo de la lesión), también ocurre un incremento en la expresión de genes promotores de crecimiento (factores neurotróficos) poco después del evento, esto crea condiciones que permiten un rebrote axonal en la región, lo cual se correlaciona con un incremento de actividad neuronal tanto en la región de la lesión como en otras zonas corticales asociadas a ella, como es el caso del área contralesional. Ambos son mecanismos que

subyacen la reorganización cortical en el área perilesional. Es decir, gracias al proceso de neuroplasticidad, las neuronas sobrevivientes “modifican” sus funciones originales. El área perilesional no es la única que se verá alterada en respuesta a la lesión, se ha encontrado que el hemisferio contralesional también experimenta un incremento en el crecimiento dendrítico, así como un incremento en el brote de axones corticoestriados. Este incremento en el número de sinapsis también propiciará una reorganización cortical, aunque se ha sugerido que este fenómeno obedece a comportamientos compensatorios (Sleimen-Malkoun et al., 2011) .

Debido a la resolución de la diasquisis y a una neuroplasticidad aumentada que involucra mecanismos como el reclutamiento de caminos funcionales homólogos, la desinhibición de conexiones neuronales redundantes y la formación de nuevas redes neuronales que sustituyan las funciones de las áreas afectadas, se observará que un paciente que ha sufrido un EVC tendrá un proceso de recuperación endógena o también llamado recuperación espontánea, en el que recuperará varias de las funciones afectadas independientemente del tratamiento de rehabilitación (Hoyer & Celnik, 2011). Sin embargo, uno de los objetivos de la rehabilitación es encaminar este proceso de recuperación espontánea para maximizarlo y enfocarlo en la recuperación de los déficits; una incorrecta rehabilitación propiciará que los cambios anatómicos que ocurran en el cerebro se consoliden en mayor medida hacia la compensación más que a la recuperación. Se ha observado que este periodo de neuroplasticidad aumentada puede llegar a durar hasta 6 meses después de ocurrido el EVC, aunque generalmente este fenómeno será más notable durante los primeros 3 meses, periodo durante el cual se estima que el paciente logre una recuperación motora en una proporción fija respecto a su máximo potencial de recuperación (Krakauer & Cortés, 2018).

Se debe resaltar que el problema principal, y lo que en verdad se está trabajando en la rehabilitación, es de índole neuronal, no muscular, es por eso que se emplea el término neuro-rehabilitación. Además, como ya se ha mencionado, los mecanismos que se tratarán de explotar para lograr la recuperación de los déficits son esencialmente los mismos que funcionan para el aprendizaje. Para favorecer una reorganización cortical y potenciar la creación de nuevas sinapsis se deben repetir tareas específicas de forma intensa, tal y como cuando queremos aprender algo nuevo, de ahí que también se mencione que la terapia es un proceso de reaprendizaje. La primera vez que aprendemos alguna habilidad motriz, como manipular algún instrumento, se activan varias regiones del cerebro de forma simultánea, pues aún no se ha definido la “ruta neuronal” más eficiente para desempeñar la actividad, es sólo a través de numerosas repeticiones que se fortalecen determinadas conexiones neuronales y eventualmente, gracias a la neuroplasticidad, cuando realicemos esa misma actividad sólo se activarán regiones muy específicas del cerebro. Un proceso muy similar es el que ocurre durante el proceso de rehabilitación.

3.3 Pronóstico de la recuperación motora

La valoración precoz de un pronóstico de función a medio y largo plazo resulta esencial para diseñar un objetivo de rehabilitación. Los supervivientes a un EVC hemorrágico suelen presentar un pronóstico funcional más favorable que los supervivientes a uno isquémico, probablemente

porque el tejido dañado es menor, aunque este precepto no siempre se cumple. En cualquier tipo de EVC, el peor pronóstico se correlaciona con las localizaciones talámicas o putaminales que destruyen la cápsula interna, comprometiendo numerosas funciones cerebrales (Figuroa et al., 2014).

La evolución típica del EVC sigue una curva ascendente con una pendiente que se vuelve progresivamente menor (figura 3.1). En un estudio comunitario de Copenhague se muestra que el 95% de la recuperación se habrá logrado hacia el tercer mes, durante el primer mes y medio se logra la recuperación más rápida (85%); entre el cuarto y sexto mes la pendiente de recuperación es leve y a partir del sexto mes apenas se detecta una mejoría palpable, por lo que a partir de ese momento se suele dar por estabilizado el cuadro (Arias Cuadrado, 2009). Sin embargo, este comportamiento no es absoluto, cada paciente es diferente y también se debe tomar en consideración que los momentos de máxima recuperación suelen ser proporcionales a la intensidad de los déficits, de tal forma que un EVC con déficits leves encontraría su cenit de recuperación pronto, mientras que cuadros más graves hasta los 5 ó 6 meses. La mejoría durante los primeros días se atribuye a la recuperación del tejido ubicado en la zona de penumbra de la lesión debido a la restauración del flujo sanguíneo, así como a la resolución de la diasquisis y a la neuroplasticidad, conceptos que se mencionarán más adelante.

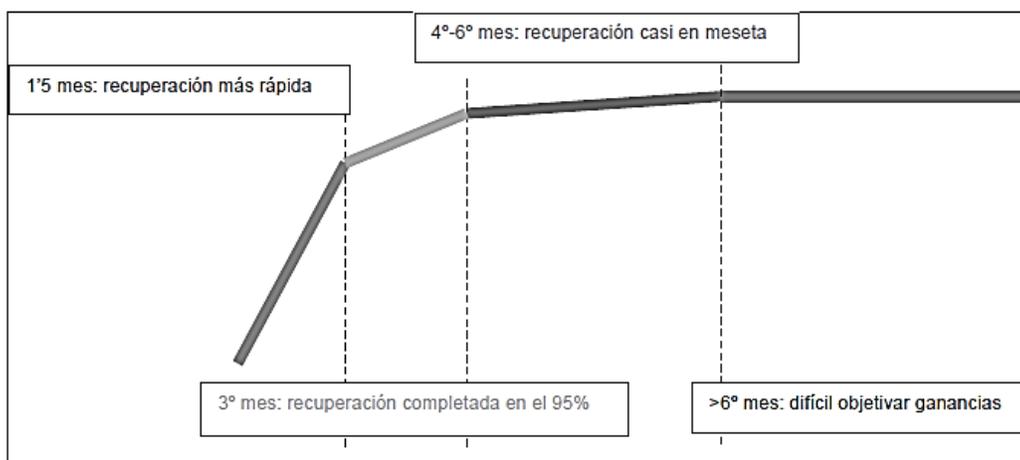


Figura 3.1. Curva de recuperación esperada en un paciente EVC. (Arias Cuadrado, 2009).

Algunos factores que pueden dar indicios sobre la evolución que tendrá el paciente son:

1. Retraso en la mejoría: La ausencia de mejoría en los primeros 3 meses sugiere un mal pronóstico. Además, cuanto más tardía sea la mejoría, mayor riesgo habrá de desarrollar patrones no deseados en la recuperación.
2. Intensidad de las deficiencias: A mayor intensidad de los síntomas se espera una evolución menos favorable.
3. Reserva cognitiva: La "capacidad de reserva cerebral" hace referencia al número de sinapsis entre neuronas y la "reserva cognitiva" es un término que se refiere a utilizar estrategias de procesamiento pre existentes para maximizar el desempeño de tareas

cognitivas ante una neuropatología, como el EVC, así como para reclutar redes neuronales alternativas o redes poco usadas para compensar la disrupción neurológica (compensación neural). Actividades como aprender un nuevo idioma, ejercitarse, leer y realizar manualidades incrementan la reserva cognitiva. Aunque aún falta más investigación al respecto, después de una lesión cerebral la reserva cognitiva parece influir en la adaptación y mejora de nuevas estrategias de procesamiento cognitivo, mejora la sinaptogénesis, y aumenta la regulación y expresión de factores neurotrópicos, responsables de facilitar la neuroplasticidad (Rosenich et al., 2020). Es por ello que factores como el grado de estudios del paciente, el dominio de varios idiomas, así como sus actividades físicas, sociales y de ocio previas al ictus, podrían incidir en su recuperación.

Otros factores a tomar en cuenta son la edad, pues se espera que los mecanismos endógenos de recuperación como la neuroplasticidad se presenten con mayor intensidad entre más joven sea el paciente; el grado de dependencia previo al EVC, ya que no se puede esperar una mejoría más allá del estado anterior al ictus; el entorno familiar del paciente, dado que, de manera temporal o definitiva, el paciente será dependiente de una persona. Lograr establecer un vínculo con el equipo terapéutico también es muy importante, pues de ello dependerá la adherencia al tratamiento y su aprovechamiento (Verheyden & Ashburn, 2013).

Respecto a la función motora, en general, los déficits graves a las 3 semanas tienden a mantenerse a los 6 meses. La evolución de la fuerza del miembro inferior suele ser mejor que la del superior. El reinicio del movimiento proximal del miembro superior no se asocia con la recuperación de la funcionalidad distal, pero la recuperación distal dentro del primer mes posterior al evento, especialmente si existe prensión manual voluntaria, prevé una recuperación rudimentaria de la motricidad fina al quinto o sexto mes (Jones, 2017). Tanto la flacidez como la espasticidad intensas y mantenidas son un factor de mal pronóstico.

Los mejores indicadores que permiten pronosticar una recuperación de la capacidad ambulatoria a los 6 meses de ocurrido el evento son (Arias Cuadrado, 2009):

1. La capacidad de iniciar la marcha a las 2 semanas posteriores al EVC.
2. El grado de motricidad de la pierna parética.
3. La hemianopsia homónima⁶.
4. La capacidad de mantener el equilibrio estando sentado.
5. La incontinencia urinaria.
6. La funcionalidad cotidiana en las 2 semanas siguientes al EVC.

⁶ Condición en la que una persona sólo puede ver un lado del campo visual, derecho o izquierdo.

3.4 Principios generales de rehabilitación

Los siguientes son algunos principios que se mencionan comúnmente en la literatura (Dobkin, 2004; Figueroa et al., 2014; Hylin et al., 2017; Krakauer & Cortés, 2018; Verheyden & Ashburn, 2013) sobre rehabilitación por su gran impacto que tienen en este proceso:

1. **Equipo multidisciplinario.** Como se ha mencionado anteriormente, debido a la gran variedad de déficits que se pueden desencadenar a causa de un EVC, principalmente motores, cognitivos y emocionales, será necesario atender a los pacientes de forma integral por un profesional de diversas áreas, como son personal de fisioterapia, enfermería, terapia ocupacional, neuropsicología, trabajo social, logopedia, ortopedia, entre otras. Este equipo idealmente deberá estar coordinado por un médico especializado en rehabilitación neurológica con el que se deberá planear una estrategia de rehabilitación con base en la evaluación inicial, el pronóstico, la evolución del paciente y sus necesidades
2. **Inicio precoz.** El tratamiento de rehabilitación debe comenzar tan pronto como se haya asegurado el estado vital del paciente y se haya establecido su diagnóstico, considerando que está contraindicado comenzar durante las primeras 24 horas después de ocurrido el evento, ya que podría resultar perjudicial para el paciente. Muchas de las complicaciones a corto plazo que pueden desarrollarse están relacionadas con la inmovilidad (neumonías, embolia pulmonar, infecciones por orina, contracturas, úlceras, etc.), por lo que la movilización del paciente debe comenzar tan pronto como la situación lo permita. Es de particular importancia iniciar la sedestación lo antes posible para evitar complicaciones respiratorias. Aquellos pacientes que comienzan su rehabilitación dentro de la primera semana después del ictus tienen un menor grado de discapacidad y logran una mejora calidad de vida a largo plazo que aquellos que comienzan más tarde.
3. **Continuidad.** La rehabilitación es un proceso que requiere un seguimiento constante. Los objetivos a alcanzar deben estar coordinados a lo largo de las diferentes fases de evolución y los distintos ámbitos de atención. Se debe evitar fragmentar el programa de rehabilitación, así como reiniciar tratamientos innecesarios. Los objetivos terapéuticos deben revisarse periódicamente, así como las necesidades del paciente y cualquier alteración que se llegara a presentar en el área cognitiva, emocional, en la comunicación, el dolor, la espasticidad, la disfagia y las adaptaciones personales.
4. **Intensidad.** Existe evidencia de que una alta intensidad del tratamiento rehabilitador, sobre todo en la fase subaguda, influye de manera positiva en el resultado funcional (Dobkin, 2004; Hylin et al., 2017; Krakauer & Cortés, 2018; Sleimen-Malkoun et al., 2011). Sin embargo, no existe un claro consenso de cuál es la intensidad ideal, además de que periodos excesivamente prolongados pueden perjudicar al paciente (Lang et al., 2015). Al aumentar la intensidad en los primeros 6 meses tras el ictus, siempre que el paciente lo tolere, se consigue disminuir el grado de discapacidad.

5. Duración. Durante los primeros 6 meses tras el ictus, si hay discapacidad, debe mantenerse el tratamiento de rehabilitación y prolongarse durante el primer año hasta la estabilización del proceso, siempre que persistan objetivos funcionales claros. Los programas de rehabilitación convencionales suelen mostrar mejoras poco significativas en pacientes de más de 1 año, ya que procesos biológicos, como la neuroplasticidad aumentada, se han detenido, aunque el uso de técnicas como dispositivos electromecánicos o la rehabilitación virtual han mostrado resultados alentadores en pacientes crónicos (M. Cameirão et al., 2012). No obstante, los pacientes deben seguir teniendo acceso a los servicios de rehabilitación, ya que un paciente puede presentar un deterioro a causa de una caída, fractura, depresión, alteraciones de deglución o el propio envejecimiento, y requerirá tratamientos concretos de corta duración para regresar al nivel funcional previo.

6. Retroalimentación. La retroalimentación intrínseca es aquella que nosotros recibimos naturalmente a través de nuestro sistema sensorio-perceptivo, como la visión, la propiocepción⁷ y la audición. Todo aquello que se provea adicionalmente es una retroalimentación extrínseca, como las opiniones verbales acerca de la forma en la que realizamos una tarea. La retroalimentación es esencial para el aprendizaje (o reaprendizaje en el caso de la rehabilitación), pero la retroalimentación extrínseca se vuelve aún más crucial cuando se ve afectada la retroalimentación intrínseca (Carr et al., 2011). Si no se proporciona la adecuada retroalimentación los pacientes podrían continuar usando patrones de movimiento no deseables. Kelly Carr (Carr et al., 2011) propone 3 tipos de retroalimentación extrínseca que tienen el potencial de afectar el aprendizaje motor:
 - a. Conocimiento de resultados. Proporciona información sobre el resultado del movimiento y tiene como objetivo ayudar a que el paciente desarrolle sus capacidades de detección y corrección de errores respecto al movimiento específico que se está aprendiendo. De tal forma que en los siguientes intentos el paciente genere el patrón apropiado de movimiento comparando el resultado esperado con su referencia interna de exactitud y realizando los ajustes necesarios.
 - b. Conocimiento del desempeño. Proporciona información respecto al patrón de movimiento para promover cambios en la coordinación. Se enfoca en la ejecución del movimiento y en mejorar la técnica para alcanzar el mejor desempeño posible. En este caso al paciente se le proporciona la solución de cómo coordinar sus miembros, mientras que en conocimiento de resultados sólo se indica lo que se desea lograr y el paciente debe activamente tratar de encontrar la solución. La redundancia cinemática es aquella en la que a través de diferentes sinergias motoras se puede llegar a la misma meta, aunque en pacientes EVC la solución a

⁷ Capacidad que nos permite percibir la ubicación, movimiento y acciones de las partes de nuestro cuerpo.

menudo incluye movimientos compensatorios o patrones anómalos no deseados, al proporcionar este tipo de retroalimentación se busca evitar su aprendizaje.

- c. Retroalimentación motivacional. Mensajes como “¡Estás haciendo un gran trabajo!” son positivos, solidarios, gratificantes y/o relajantes, lo que permite que el participante mantenga interés en la tarea. Sin embargo, no se recomienda apoyarse demasiado en este tipo de retroalimentación porque puede resultar ambigua, ya que no permite que el paciente detecte entre movimientos correctos e incorrectos. Si la información de la retroalimentación es limitada se reduce la habilidad del paciente de auto-corregir errores y adquirir la destreza deseada.

7. Evaluación periódica. El uso de escalas válidas y confiables, que permitan evaluar de forma objetiva las diferentes funciones del paciente, son la piedra angular para identificar problemas, establecer objetivos, determinar las intervenciones, controlar su eficacia y evaluar los resultados.
8. Participación de los pacientes y cuidadores. Tanto el paciente, como los familiares y los cuidadores deben involucrarse en el proceso de rehabilitación desde el comienzo, para ello se les debe informar adecuadamente todo lo necesario respecto al manejo y los cuidados que requerirá el enfermo. El hecho de que el paciente tenga mejor conocimiento acerca del EVC puede reducir moderadamente la depresión. Por otra parte, los cuidadores y familiares que participan en programas de soporte específico muestran más habilidades psicosociales en el manejo de pacientes, menos depresión y mejor percepción de su salud (Figuroa et al., 2014).

3.5 Objetivos en las fases de rehabilitación

Desde un punto de vista temporal, es posible estructurar la rehabilitación en 3 periodos. El programa terapéutico deberá adaptarse en función de la situación del paciente, por lo que en algunos casos no será factible avanzar en todas las fases de tratamiento (Arias Cuadrado, 2009). A continuación se mencionan los objetivos más comunes que se persiguen en cada fase.

3.5.1 Periodo agudo

Corresponde al periodo en el que el paciente ha sido estabilizado y su característica principal es la hipotonía. Suele ser el periodo en el que el paciente permanece encamado. La rehabilitación en esta etapa debe enfocarse en lo siguiente:

- Evitar trastornos cutáneos y respiratorios a través de ejercicios respiratorios y cambios frecuentes de posturas.
- Prevenir actitudes viciosas como la hipertonia postural, hombro doloroso y equinismo, a través de ortesis o posturas protectoras.
- Se debe valorar la seguridad de la deglución antes de iniciar la alimentación o la hidratación oral. De ser necesario se debe remitir al paciente a un tratamiento logopédico.

- Corregir trastornos vesicoesfinterianos y gastrointestinales.
- Movilizaciones pasivas lentas de amplitud máxima, estiramientos musculares prolongados, seguidas de movimientos funcionales normales en ambos hemisferios.
- Enseñarle al paciente estrategias de auto movilización y transferencias.
- Iniciar equilibrio de tronco y sedestación.
- Realizar estimulación sensorial en el hemisferio afectado.
- Comenzar con terapia ocupacional para adquirir autonomía elemental en cama.

3.5.2 Periodo subagudo

Este periodo se caracteriza por la aparición de la espasticidad e hiperreflexia, así como la recuperación motora en los casos favorables, por lo que marcará el inicio de la fase de trabajo activo por parte del paciente para la recuperación de fuerza y coordinación. Se considera la fase más importante de la rehabilitación, pues en este periodo se debe aprovechar la presencia de neuroplasticidad aumentada, con lo que se logrará obtener un porcentaje importante de la máxima recuperación posible del paciente (Krakauer & Cortés, 2018). La duración habitual es entre 3 y 6 meses, pero cada caso es diferente y la estrategia de rehabilitación debe obedecer una evaluación de cambios más que un criterio temporal estricto. Durante esta etapa se debe enfocar en lo siguiente:

- Corrección de deformidades ortopédicas ya instauradas a través de ortesis.
- Trabajar en técnicas de regulación de la espasticidad.
- Continuar con movilizaciones pasivas, progresando a movimientos activos-asistidos del lado parético y trabajar en el aumento de fuerza muscular.
- Reeducación propioceptiva y de la coordinación.
- Reentrenar el control postural, normalizar el tono y continuar con la reeducación del equilibrio en bipedestación.
- Se recomienda iniciar la reeducación de la marcha cuando haya un buen equilibrio en bipedestación y un control selectivo de los movimientos de miembros inferiores.
- Continuar con la estimulación sensorial en el hemisferio afectado.
- Comenzar con ejercicios para la parálisis facial.
- Iniciar terapia ocupacional dirigida a AVD básicas e instrumentales.

3.5.3 Periodo crónico o de estado

La fase crónica comienza aproximadamente después de los 6 meses desde que ocurrió el ictus. En este punto se considera que se ha alcanzado la estabilidad del cuadro y la recuperación que mostrará el paciente será relativa, pues los procesos biológicos que están presentes durante la fase subaguda para potenciar la neuroplasticidad se ven considerablemente reducidos, por lo que el esfuerzo terapéutico no suele ir encaminado a la recuperación del déficit perdido, sino a la adaptación funcional que resta. En esta fase cobran especial relevancia las ortesis y las ayudas técnicas, que son instrumentos diseñados específicamente para prevenir, compensar, mitigar o neutralizar una deficiencia o discapacidad. Por lo general las ayudas técnicas se aplican al campo

de las AVD para resolver problemas individuales muy específicos, la más popular se conoce como andador (Arias Cuadrado, 2009; Figueroa et al., 2014).

Estos son los aspectos en los que se suele enfocar la rehabilitación en este estado:

- Continuar con el plan terapéutico con mayor énfasis en la potenciación muscular.
- Continuar con técnicas para recuperar la marcha y el equilibrio, pues esta función puede seguir mejorando hasta transcurridos 2 años.
- Continuar con terapia del lenguaje, ya que es una de las funciones que puede seguir mejorando hasta transcurridos 2 años.
- Evaluar la aplicación de estrategias de rehabilitación con un objetivo concreto, por ejemplo, un programa de 4 semanas que permita aumentar la velocidad y la resistencia de la deambulación.
- Entrenamiento en técnicas compensatorias para realizar diversas AVD.
- Valoración del uso de ayudas técnicas y ortesis.

En el momento en el que la evolución del paciente se ha estabilizado y se ha logrado una adaptación a su situación tanto para él, como para sus cuidadores, se puede dar por finalizada la rehabilitación. Este momento puede ser muy temprano, como en los casos de peor pronóstico donde no se observa una mejora en los primeros 3 meses, o ser más tardío, según la evolución que presente el paciente. Una vez que el médico rehabilitador considere que la rehabilitación ha terminado podrá requerir revisiones de control para dar un seguimiento a las capacidades funcionales del paciente, quien será remitido al área de rehabilitación sólo en caso de que se presente una pérdida o empeoramiento de su capacidad funcional con respecto al último registro (Verheyden & Ashburn, 2013).

3.6 Tratamientos de rehabilitación motora

La rehabilitación motora es un elemento fundamental y de mayor esfuerzo para el paciente. Entre mayor sea la recuperación de las funciones mejor calidad de vida podrá alcanzar el paciente. Como ya se ha mencionado anteriormente, según la localización del EVC y el tamaño de la lesión, se pueden ver afectadas la fuerza, el equilibrio y la coordinación, como en el caso de las ataxias y las apraxias, de modo que el enfoque terapéutico se suele hacer de manera conjunta. Se debe poner especial atención en los casos donde se presenten afecciones a la propiocepción o heminegligencia cuando se diseñe el programa terapéutico, ya que estos déficits impactarán de forma importante en la recuperación motriz. Otro caso especial es el de la espasticidad, la cual debe tratar de prevenirse desde las fases más tempranas de evolución a través de movilizaciones precoces, posturas de inhibición antiespástica, el uso de fármacos sistémicos o locales, como la Toxina Botulínica, y con ortesis y férulas.

3.6.1 Terapia física

El objetivo de la fisioterapia o terapia física después de un EVC es maximizar la recuperación del movimiento y la independencia en la vida cotidiana para minimizar complicaciones secundarias,

para ello se realizan ejercicios de movilización pasiva y activa del hemicuerpo afectado para la recuperación de fuerza y tono muscular, así como ejercicios para la recuperación del balance y el reentrenamiento de la marcha (Verheyden & Ashburn, 2013). Dentro de la fisioterapia existen varias escuelas que proponen diferentes métodos de tratamientos específicos basándose en el trabajo del sistema neuromuscular y la estimulación de los propioceptores. Estas técnicas de reeducación o reprogramación sensoriomotriz han gozado de menor o mayor aceptación según las épocas. La base de muchas de ellas es la existencia de una respuesta refleja inhibitoria o facilitadora ante un estímulo. Hoy en día sigue existiendo una opinión dividida entre los especialistas de rehabilitación respecto a la eficacia de cada una, por lo que su elección dependerá de la experiencia del equipo terapéutico. Por otra parte, una de las críticas más razonables a estos métodos es que sus modelos teóricos son empíricos, con insuficiente información objetiva, lo que favorece la pasividad de los pacientes y las ganancias que obtienen resultan de escasa utilidad para la vida real. A continuación se presentan de forma breve algunas de las escuelas más conocidas:

3.6.1.1 Escuelas terapéuticas

Técnica de Bobath o del Neurodesarrollo (Michielsen et al., 2019): Es un método que persigue combatir mal adaptaciones (sinergias no deseadas) e incrementar la fuerza muscular con base en repeticiones de ejercicios, con el fin de que eventualmente se tornen automáticos. Combina técnicas que inhiben movimientos anormales con técnicas que faciliten mantener una postura correcta y el equilibrio. Esta escuela clasifica a la hemiplejía en tres fases: flácida, espástica y de recuperación.

Método de Kabat o de Facilitación Neuromuscular Propioceptiva (Westwater-Wood et al., 2010): Intenta evocar movimientos voluntarios a través de estímulos simultáneos y sincronizados. Se basa en la aplicación de esquemas de movimientos facilitadores de carácter espinal y diagonal, aprovechando los músculos agonistas para favorecer a los más débiles; estos esquemas se asocian a otras técnicas de facilitación como reflejos de estiramiento, de flexión, resistencia máxima, etcétera.

Método de Perfetti o Ejercicio Terapéutico Cognoscitivo (Albiol et al., 2014): Es un modelo que se basa en la neuromotricidad y en la psicomotricidad, trata de programar movimientos a través de estímulos táctiles y cinestésicos, prescindiendo de la vista. En esta escuela, el punto de partida de la reeducación motriz ha de pasar inicialmente por la reducción de la sensibilidad y le da especial relevancia a la mano.

Método de Brunnstrom (Naghdi et al., 2010): Contrario al método de Bobath, este modelo no combate las sinergias, sino que las aprovecha. Utiliza estímulos aferentes para despertar respuestas reflejas con el fin de producir movimiento y después ejercita el control voluntario de las mismas. En las fases iniciales utilizaría las sinergias y en las fases posteriores se asemejaría más a métodos como Kabat o Perfeti. Esta escuela maneja una escala de 6 fases de recuperación:

- Fase 1. Se presenta flacidez y no se pueden iniciar movimientos de los miembros.

- Fase 2. Pueden aparecer sinergias básicas o algunos de sus componentes en miembros básicos. Pueden presentarse movimientos voluntarios mínimos. La espasticidad comienza a desarrollarse.
- Fase 3. El paciente gana control voluntario de las sinergias de movimiento, aunque no necesariamente se desarrolla un rango completo de movimiento en todos los componentes de las sinergias. La espasticidad es severa.
- Fase 4. Se dominan algunas combinaciones de movimientos que no pertenecen a sinergias y la espasticidad comienza a declinar.
- Fase 5. Se pueden lograr combinaciones de movimientos más complejos conforme las sinergias de miembros básicos pierden su dominancia sobre los actos motores.
- Fase 6. La espasticidad desaparece y es posible realizar movimientos articulares individuales.

Método de Vojta (Epple et al., 2020): El aspecto más importante de este modelo es que busca facilitar el reflejo de arrastre para distintos segmentos corporales como respuesta activa a estímulos sensoriales de presión, tacto, estiramiento y actividad muscular contra resistencia. También es conocido por mejorar el control postural automático en fases muy tempranas del EVC y suele utilizarse en niños.

Los defectos que pudieran tener estas técnicas no implican que se tengan que abandonar completamente en el plan terapéutico, pues pueden integrarse en un programa que persiga la participación activa del paciente y la ejecución de un tratamiento orientado hacia un fin concreto. Es común que se apliquen técnicas de cinesiterapia (movilización pasiva, asistida o activa encaminada a incrementar la fuerza muscular, relajación y estiramiento de la musculatura espástica), de reeducación del equilibrio y de la coordinación, tomando estrategias de alguna de las escuelas anteriores según lo requiera el caso. Actualmente se suelen combinar estas técnicas e inclusive cada terapeuta suele añadir nuevas técnicas, modificaciones y enfoques de acuerdo con su experiencia.

3.6.2 Compensación y recuperación

En términos de comportamiento, la recuperación se refiere a la restauración de una función que se perdió o deterioró debido al EVC, o a un estado lo más cercano posible antes de la lesión cerebral, mientras que la compensación trata de una sustitución de la función, de aprender estrategias con las cuales lograr el mismo resultado. Aunque en la literatura algunas veces estos dos términos se usan de forma indistinta; en muchas ocasiones el término recuperación se emplea para describir mejoras en el desempeño de una tarea y no se contempla la forma en la que se realiza, por lo que es difícil saber si se trata de una compensación o una recuperación.

En términos de mecanismos neuronales, la diferencia entre recuperación y compensación no queda muy clara. Algunos autores sostienen que la recuperación neuronal debería referirse únicamente al proceso de recuperación espontánea mencionado anteriormente, mientras que los mecanismos en los que ocurre una reorganización cortical y se reentrena a las neuronas

sobrevivientes para sustituir las funciones de las neuronas que se perdieron son mecanismos de compensación (Hyllin et al., 2017). Sin embargo, la práctica de habilidades motrices puede promover un cierto grado de recuperación, volviendo los movimientos más cercanos a lo normal, incluso mucho tiempo después de ocurrido el EVC. Para que eso suceda tiene que ocurrir una reorganización en la conectividad neuronal que permita un cierto grado de “invasión” o “reescritura” sobre las funciones originales de esas regiones (Jones, 2017).

Es muy importante prestar atención a la compensación durante la fase subaguda, ya que puede ser un mecanismo que ayude a mejorar la función, pero si no se trata adecuadamente también puede contribuir a perpetuar el daño. Los patrones de movimientos compensatorios se pueden auto-enseñar o pueden ser producto de una intervención terapéutica. Durante la rehabilitación se buscará mejorar la habilidad funcional y esto puede lograrse desarrollando estrategias compensatorias o recuperando patrones de movimiento lo más cercanos posible a los patrones normales; la forma en la que ambos contribuirán en la rehabilitación dependerá del grado de daño que presente el paciente, es decir, la compensación tendrá un mayor impacto en la mejora de la habilidad funcional en aquellos pacientes que presenten mayor discapacidad. También se debe tener en cuenta el tipo de rehabilitación que se aplicará, pues como se describió anteriormente, algunas escuelas buscan entrenar y refinar la compensación, mientras que otras la tratan de inhibir, y también existen otros métodos donde lo más importante es completar una tarea, lo que resulta permisivo para la compensación.

Comenzar la rehabilitación con ejercicios que favorezcan el uso del miembro no parético, o que promuevan estrategias compensatorias, apoyándose de movimientos de miembros proximales del brazo en vez de distales, propiciará el desuso de cualquier capacidad residual que pudiera mejorar el rango de movilidad del miembro parético a través de la práctica (Jones, 2017). En pacientes subagudos y crónicos, lo anterior generalmente es la mejor opción cuando el daño es muy severo y no hay posibilidades de una recuperación, pero no sería lo más conveniente para pacientes subagudos que presenten un daño moderado, donde existe, gracias a la ventana de neuroplasticidad aumentada, la oportunidad de lograr que la calidad del movimiento sea cercana a como era previa al EVC. Por esta misma razón, corregir adaptaciones anómalas resulta muy complicado en pacientes crónicos, pues los mismos mecanismos de neuroplasticidad que ayudan al reaprendizaje de sinergias deseadas favorecerán de igual forma que se aprendan patrones no deseados, como consecuencia de la práctica de éstos. Lograr que los circuitos neuronales sobrevivientes cambien para mejorar la función del miembro parético es una labor complicada, pero es aún más complicado cambiar esos mismos circuitos que ya han sido modificados hacia una actividad que va en contra de cualquier cambio futuro.

3.6.3 Estrategias complementarias para la rehabilitación motora

Aunado a los métodos de fisioterapia de las escuelas tradicionales, en el campo de la rehabilitación han surgido diversas estrategias, en gran parte gracias al avance de la tecnología, que buscan maximizar la recuperación de los déficits mediante técnicas cuya base está sustentada en la observación de cambios favorables en la conducta, medidos a través de diferentes escalas,

así como de cambios fisiológicos en el cerebro cuantificados a través de estudios de imagen como tomografías, resonancias magnéticas y tractografías. A continuación se mencionan algunas de las estrategias que suelen emplearse como complemento en los planes terapéuticos.

3.6.3.1 Terapia ocupacional

La terapia ocupacional es un tipo de intervención que busca que el paciente logre maximizar su independencia, esto se logra a través de ejercicios que desarrollen habilidades para realizar AVD. Éstas se clasifican en: básicas o de cuidados personales, instrumentales, y avanzadas (Arias Cuadrado, 2009). Las AVD básicas son todas aquellas que se deben realizar para sobrevivir, tales como la alimentación, vestido, higiene, transferencia, continencia y baño. Algunas de las AVD instrumentales son el uso de sistemas de comunicación, transporte público o particular, manejo del dinero, cuidado del hogar, ir de compras, así como aquellas actividades relacionadas al trabajo. Las AVD avanzadas son complejas y requieren un mayor grado de organización, pero no comprometen la independencia del individuo, en general están destinadas al desarrollo dentro de la sociedad, e.g. la educación, el trabajo y jugar.

Durante la terapia se debe evaluar la necesidad del entrenamiento para cada actividad específica, de forma que se pueda elaborar un plan terapéutico que tome en cuenta si se desarrollará una adaptación, si será necesaria la asistencia de tecnología auxiliar, como el caso de una ortesis, así como entrenamiento en el uso de dicha tecnología.

Estudios experimentales sobre neuroplasticidad muestran que el aprendizaje de habilidades motoras, en vez de actuar sobre músculos o patrones de movimiento deben de orientarse hacia actividades que tengan un significado práctico en la vida de la persona (Hoyer & Celnik, 2011). Para lograr el mejor aprendizaje, estas actividades se deben realizar en un entrenamiento repetitivo, intenso y progresivo en su dificultad, con retroalimentación sobre lo que se realiza y empleando estrategias de motivación. Hoy en día este tipo de terapia se considera fundamental en cualquier tipo de plan terapéutico para pacientes EVC.

3.6.3.2 Entrenamiento bimanual

Se trata de técnicas de entrenamiento en donde se requiere el uso de ambos miembros superiores de forma simultánea. Su objetivo principal es la recuperación y mejora del miembro parético. Este tipo de tareas se asemejan más a AVD, donde es más frecuente que usemos ambas manos de forma coordinada que una sola. Los principales argumentos detrás de este tipo de técnicas son la existencia de dependencias mediadas neuronalmente entre las extremidades, la interacción entre ambos hemisferios cerebrales entre redes neuronales similares al realizar actividades bimanuales, y la evidencia de la relación entre el entrenamiento intensivo y la neuroplasticidad. A pesar de que en teoría este tipo de ejercicios deberían maximizar la recuperación funcional mediante el restablecimiento del control del miembro parético y tareas bimanuales coordinadas, aún no está claro el tipo de población que realmente se pueda beneficiar de este entrenamiento, ya que su eficacia se ve altamente afectada por la severidad de la parálisis o espasticidad del miembro afectado y el tiempo de intervención post EVC (Sleimen-Malkoun et al., 2011).

3.6.3.3 Terapia de movimiento inducido por restricción (TMIR)

Este tipo de terapia involucra una práctica intensa de ejercicios orientados a tareas con el miembro afectado mientras se restringe el movimiento del miembro no afectado, generalmente empleando un cabestrillo o un guante, lo que significa que durante las actividades se obliga a los pacientes a emplear su miembro parético. Esta estrategia fue pensada para ser usada principalmente sobre miembros superiores. La base de esta técnica es que, tras un EVC, el brazo y la mano afectados no se usan lo suficiente, inclusive cuando existe cierto grado de actividad funcional, debido principalmente a que los movimientos causan dolor, se realizan de forma lenta y requieren mucho esfuerzo, lo que causa un “aprendizaje de no uso” del miembro afectado, es decir, el paciente aprende a utilizar exclusivamente su miembro no afectado para todas sus actividades, ignorando por completo el miembro afectado.

La TMIR ha sido asociada con cambios plásticos cerebrales tanto funcionales como estructurales; se ha demostrado que el uso del miembro parético aumenta la activación de la corteza somatosensorial y otras estructuras relacionadas con las salidas motoras del hemisferio ipsilesional o contralesional (Ordoñez Mora et al., 2017).

El tratamiento original con TMIR es ser aplicado por 6 horas al día durante 2 semanas. Aunque también existen versiones modificadas menos intensivas y más tolerables donde se aplica durante 2 horas al día, 3 días a la semana. Cabe resaltar que este no es un tratamiento aplicable a toda la población EVC, pues no se aconseja su uso en caso de que ya se hayan instalado sinergias o si el movimiento voluntario de extensión de dedos y muñeca no alcanza al menos 10 grados (Dobkin, 2004).

3.6.3.4 Terapia en espejo y práctica mental

La práctica mental se refiere a un estado en el que el individuo simula mentalmente alguna acción; se debe imaginar que la acción se repite varias veces para promover el aprendizaje. Por otro lado, en la terapia en espejo se coloca un espejo en el plano medio-sagital, de tal forma que el paciente vea reflejado el movimiento del miembro no afectado, mientras que el lado parético se encuentra detrás del espejo. De esta manera se “engaña” al cerebro para que piense que el movimiento lo realiza el miembro parético. El propósito original de esta técnica era reducir el dolor del “miembro fantasma”, una afección que sucede después de la amputación de un brazo, sin embargo, estudios en pacientes EVC han encontrado que este tipo de ejercicios activan la corteza motora primaria del hemisferio ipsilateral a la lesión (Figueroa et al., 2014). Por otro lado, la práctica mental se ha asociado con la activación de las áreas involucradas en la planeación y el control de movimientos, como la corteza motora primaria, la corteza parietal superior y el área premotora (De Almeida Oliveira et al., 2014). La hipótesis actual detrás del funcionamiento de ambas técnicas es la activación de neuronas espejo durante el entrenamiento.

Las neuronas espejo fueron descubiertas por Rizzolatti, quien las nombró de esa manera por su propiedad de “reflejar” el acto motor observado. El sistema de neuronas espejo se activa cuando un individuo observa que se realiza una acción por otra persona e inconscientemente imita lo que

vemos, escuchamos o percibimos. De hecho, se ha encontrado que algunas neuronas espejo responden a sonidos que pertenecen a acciones específicas y se les han llamado neuronas espejo audio-visuales. Estas neuronas se activan en las áreas del giro frontal inferior y la corteza premotora ventral, por lo que un entrenamiento con una estimulación multi-sensorial podría ser muy benéfico para la rehabilitación. Tanto la terapia en espejo como la práctica mental han demostrado buenos resultados en pacientes EVC cuando se combinan con la fisioterapia y la terapia ocupacional, ya que ambas activan el sistema de neuronas espejo, el cual promueve la reorganización cortical y por consecuencia mejora la recuperación funcional (Carvalho et al., 2013).

3.6.3.5 Estimulación transcraneal

La estimulación magnética transcraneal (EMT) es una técnica no invasiva que se usa para investigar la funcionalidad del cerebro en sujetos sanos y en pacientes EVC. Esta técnica se basa en las leyes de Faraday de la inducción electromagnética para crear pulsos de campo magnéticos, los cuales pueden inducir la actividad eléctrica en áreas focales del cerebro. Aplicar EMT en la corteza motora primaria (M1) activa neuronas corticoespinales a través de las sinapsis, provocando descargas de salidas neurales en forma de potenciales evocados motores. Como herramienta terapéutica, la EMT ha sido usado para mejorar el desempeño motriz, así como para estudiar la forma en que otras técnicas de rehabilitación facilitan los procesos adaptivos de neuroplasticidad del cerebro, ya que se ha encontrado que la aplicación de trenes de pulso de baja frecuencia repetitivos de EMT puede disminuir significativamente la excitación cortical, mientras que altas frecuencias logran el efecto contrario, de esta forma la EMT podría modular las redes motoras a través de mecanismos similares a la potenciación a largo plazo y la depresión a largo plazo (Hoyer & Celnik, 2011). El uso de EMT como complemento a otro tipo de terapias es prometedor, sin embargo, aún existen varias limitaciones con este tipo de tecnología y se sigue estudiando la mejor forma en la que se le puede sacar provecho, pues no se ha llegado a un consenso entre las frecuencias, dosis y las mejores ubicaciones del cerebro donde se puede aplicar, debido a que esto dependerá del tipo de EVC, la ubicación y el grado de la lesión, así como el tiempo de evolución del paciente.

3.6.3.6 Estimulación eléctrica

La estimulación eléctrica (EE) es una de las terapias más usadas y entre los beneficios que se han reportado sobre su uso se encuentran la reducción de espasticidad, mejoras en el rango de movimiento, disminución del dolor y mejoras en la sensibilidad. Sin embargo, estos beneficios no se han demostrado de forma adecuada y aún se considera que no existe suficiente evidencia en estudios con grupos control. (Guo et al., 2018) Entre los diferentes tipos de EE que se aplican en pacientes EVC se encuentran la estimulación eléctrica neuromuscular (EENM), la estimulación eléctrica funcional (EEF) y la estimulación eléctrica neuromuscular transcutánea (EENT), cada una de ellas se suele combinar con ejercicios físicos para lograr mejores resultados.

La EENM se administra a través de electrodos que se colocan sobre los músculos que se desean entrenar y a los cuales se les aplicarán pulsos estimulantes. EEF es una técnica que utiliza pulsos eléctricos de baja energía para inducir el movimiento en un grupo de músculos débiles de un miembro parético. Un estudio reportó que esta técnica es útil para recuperar el movimiento voluntario de funciones como el agarre, alcanzar y recuperar, y marcha en pacientes subagudos, sin embargo, no encontraron beneficios en pacientes crónicos. (Eraifej et al., 2017) EENT es una técnica analgésica simple y no invasiva que se puede auto administrar para aliviar el dolor asociado con la rigidez y contracturas de miembros hemipléjicos.

3.6.3.7 Hidroterapia

La hidroterapia es uno de los métodos terapéuticos más longevos, pues ya desde la época de la antigua Grecia se usaba como remedio para tratar distintos malestares. Dependiendo de la afección a tratar será la temperatura y estado de agregación del agua (sólida, líquida o gaseosa). Los atributos físicos de la hidrodinámica, como la densidad, flotabilidad, presión hidrostática y viscosidad, proveen las condiciones para mejorar los procesos fisiológicos y la actividad motora (Paizan et al., 2019). Algunos estudios han demostrado que programas de ejercicios acuáticos de entre 4 y 8 semanas en pacientes EVC crónicos realizados en albercas o tinas a una temperatura de 33-34 °C, han incrementado de forma significativa la resistencia cardiovascular, la cantidad de trabajo máxima soportada, la velocidad de la marcha y la fuerza muscular de miembros inferiores, las ganancias también ha resultado ser significativamente mayores al compararlas con las obtenidas mediante terapia física pura (Zhu et al., 2016). Entre las razones que podrían explicar esta mejora se encuentran que el agua, a la temperatura previamente indicada, incrementa la temperatura de la piel, expande los vasos sanguíneos de la piel periférica, incrementa el suministro de sangre, acelera la relajación muscular, disminuye la sensibilidad al dolor y los espasmos musculares, finalmente, las estimulaciones motoras y sensitivas durante los ejercicios acuáticos pueden potencialmente inducir neuroplasticidad (Figuerola et al., 2014).

3.6.3.8 Asistencia electromecánica

Actualmente existe una gran variedad de dispositivos robóticos como elementos auxiliares para la rehabilitación que son empleados en diferentes patologías que afectan la motricidad de miembros superiores, inferiores y marcha. Generalmente se dividen en 2 tipos: exoesqueletos y efectores finales. Además, estos sistemas pueden realizar un movimiento pasivo del miembro afectado, únicamente ayudar a soportar el peso de la extremidad (soporte anti gravitatorio) o proveer una fuerza como retroalimentación háptica para incrementar la dificultad de los ejercicios, enriquecer la información de entrada que recibe el sistema nervioso y promover el fortalecimiento muscular. Los sistemas más complejos pueden incluso adaptarse de acuerdo con las necesidades y evolución del paciente ofreciendo estos tres tipos de movimiento (pasivo, activo e interactivo). Otra de las grandes ventajas de los robots es que cuentan con una alta precisión para el rastreo de movimientos.

Los dispositivos de efector final generalmente se fijan a una superficie, varían entre 2 y 6 grados de libertad y se centran en la motricidad de la mano y muñeca. Algunos de los más populares son:

PHANTOM Omni: Manufacturado por Geomatic, de 6 grados de libertad, usa un paradigma de control de impedancia de tal forma que el dispositivo provee una fuerza de retroalimentación cuando colisiona con un objeto virtual, su área de trabajo es de 160mm de ancho x 120 mm de alto x 70mm de profundidad (Piggott et al., 2016).

PERCRO GRAB: Desarrollado en Italia, es un conjunto de dos brazos robóticos de 6 grados de libertad cada uno; el usuario debe colocar sus dedos en dedales ubicados en sendos brazos robóticos. Los dedales proveen una fuerza de retroalimentación a los dedos pulgar e índice, permitiendo múltiples puntos de contacto para una exploración háptica (M. Cameirão et al., 2012).

MIT-Manus: Desarrollado en el MIT, fue diseñado para ser usado en clínicas de rehabilitación como una interfaz segura y portable. Se trata de un robot que permite movimientos del hombro y codo en el plano horizontal gracias a su baja fricción e inercia. El dispositivo se compone de 4 módulos, el primero sólo permite el movimiento de un brazo robótico de 2 grados de libertad para ejercitar el codo y el hombro. En el segundo módulo se permite el movimiento de la manija del dispositivo, que tiene 3 grados de libertad para ejercitar la muñeca, adicionalmente el sistema cuenta con un módulo vertical de movimiento activo de 1 grado de libertad y un módulo de movimiento pasivo para la función de prensión de 1 grado de libertad (Krebs et al., 1999).

Los sistemas electromecánicos de tipo exoesqueleto pueden dividirse a su vez en aquellos que necesitan o no estar fijos a una superficie. Los dispositivos que no están fijos se enfocan a trabajar funciones específicas de la mano, aunque también se pueden combinar con exoesqueletos fijos para trabajar con el brazo completo. Los dispositivos más populares de este tipo son los guantes CyberGlove III, CyberTouch II y CyberGrasp, desarrollados por la compañía CyberGlove Systems (Piggott et al., 2016), estos guantes se ilustran en la figura 3.2.

Cyber Glove III: Es un guante inalámbrico con 22 sensores que permiten capturar los movimientos de los dedos con una frecuencia de hasta 120Hz. CyberTouch II es un guante con 6 estimuladores vibrotáctiles colocados en la palma y dedos que pueden ejercer una fuerza de hasta 1.2N, así como crear pulsos o vibraciones sostenidas para simular diferentes texturas.

CyberGrasp: Provee una fuerza de retroalimentación con 5 actuadores ajustables que ejercen una fuerza de hasta 12N en cada dedo con el fin de que el usuario sienta como si en verdad estuviera sujetando un objeto.

CyberForce: Es una mejora de CyberGrasp, se trata de un exoesqueleto que soporta el brazo completo, provee 6 grados de libertad y una fuerza de hasta 8.8N, que junto con CyberGrasp, le permite al usuario no sólo sentir que tiene un objeto en la mano, sino que además tiene peso.



Figura 3.2. Principales guantes desarrollados por CyberGlove Systems. (Piggott et al., 2016).

Los exoesqueletos fijos a una superficie tienen la ventaja de ser adaptables al usuario en términos de tamaño y confort. Se ha reportado que estos robots son muy efectivos en la rehabilitación de miembros superiores con parálisis, ya que soportan todo el miembro superior afectado y trabajan múltiples articulaciones, lo que permite ejercitar diferentes rangos de movimiento, desde aquellos muy gruesos hasta motricidad más fina.

Armeo Spring: Es un sistema que abraza el miembro superior de proximal a distal, magnificando cualquier movimiento activo residual del miembro hemiparético en un espacio tridimensional. También cuenta con un mecanismo que detecta la presión de empuñadura y se puede ajustar su sensibilidad de acuerdo a la situación del paciente. El exoesqueleto permite calibrar el espacio de trabajo en función de la movilidad activa que tenga el paciente y es capaz de medir varios parámetros de movimiento como resistencia, fuerza, rango y coordinación. Este dispositivo ha sido probado con pacientes crónicos con sesiones intensas de entrenamiento durante 4 meses, obteniendo mejoras motoras significativas (Colomer et al., 2013).



Figura 3.3. Exoesqueleto ARMin. (Piggott et al., 2016).

ARMin: Es un exoesqueleto de 6 grados de libertad, 4 activos y 2 pasivos, que se enfoca en la rehabilitación del brazo entero, ya sea el derecho o el izquierdo. Su diseño permite ajustarse al tamaño de brazo del usuario, así como a su rango de movimiento (figura 3.3). ARMin puede ser usado para ayudar al movimiento del brazo soportando únicamente su peso o proporcionar un

movimiento pasivo a una posición específica en caso de que el usuario no pueda hacerlo voluntariamente. Este exoesqueleto se ha empleado en múltiples estudios con pacientes EVC entrenándolos para realizar AVD (Guidali et al., 2011).

Es una práctica común utilizar estos dispositivos electromecánicos en combinación con ambientes virtuales y aplicaciones interactivas, lo que se conoce como rehabilitación virtual. Varios estudios han reportado una recuperación significativa en pacientes EVC con este tipo de estrategias respecto a los métodos tradicionales de rehabilitación (Guidali et al., 2011; Krakauer & Cortés, 2018; Piggott et al., 2016). Sin embargo, la mayor limitante de esta tecnología es su accesibilidad, pues su costo de adquisición y mantenimiento suele ser muy elevado y en el caso de los exoesqueletos también ocupan mucho espacio, además, en algunos casos suelen ser muy rígidos y restringen el movimiento, lo que puede llegar a impedir movimientos compensatorios y disminuir la eficacia de la rehabilitación.

3.6.3.9 Rehabilitación virtual

En los últimos años se ha incrementado el uso de sistemas de detección de movimiento humano en conjunto con videojuegos y simulaciones aplicados a la práctica y estudios clínicos, sobretodo enfocados en la recuperación de la función de miembros superiores en pacientes EVC. Este tipo de intervenciones pueden incrementar la intensidad del entrenamiento y la adherencia de los pacientes a su tratamiento al añadir un elemento lúdico a las terapias. Además, uno de sus mayores atractivos es el potencial de poder aplicarse a través de la tele rehabilitación. Sin embargo, aún se requiere de mucha investigación sobre este tema. En el siguiente capítulo se abordarán más detalles sobre la rehabilitación virtual.

Como se ha mencionado, uno de los principios fundamentales de la rehabilitación es el tiempo en el que se realiza la intervención, pues la rehabilitación es un proceso limitado en el tiempo, no obstante, algunas de las estrategias mencionadas anteriormente, como la rehabilitación virtual o la hidroterapia, han demostrado ser eficaces en casos incluso más allá de 1 año después del EVC (M. S. Cameirão et al., 2010; Ögün et al., 2019; Zhu et al., 2016), algo que no es muy común de observar en las terapias convencionales. Todas las estrategias alternativas deben tener un objetivo concreto y suelen mostrar los mejores resultados cuando se llevan a cabo de forma intensiva.

Capítulo 4

Rehabilitación virtual

Las prácticas actuales en las terapias de rehabilitación para la recuperación de la motricidad de miembros superiores se basan en principios para mejorar la neuroplasticidad, tales como la repetición de movimientos aislados, ya sea de forma pasiva o activa, que promuevan el incremento del rango de movilidad articular y el tono muscular, así como actividades que involucren la activación de diferentes músculos y articulaciones para realizar una tarea específica. Sin embargo, algunas de las limitantes que se presentan en los sistemas de salud son el enfoque en la mejora de la movilidad de miembros inferiores y el tiempo de tratamiento que recibe el paciente. Además, esta repetición intensa de movimientos provoca que los pacientes encuentren los ejercicios monótonos y aburridos, reduciendo su motivación y compromiso con su programa de rehabilitación (Demers et al., 2019). Debido a lo anterior, los objetivos de movilidad de miembros superiores no se llegan a alcanzar. Este problema se vuelve aún más crítico en los países en vías de desarrollo como México, donde los sobrevivientes de un EVC rara vez reciben terapia con la oportunidad y la frecuencia necesarias debido a la saturación de los centros de rehabilitación, carencia de personal especializado en el área, así como la situación socioeconómica de los pacientes, que dificulta su traslado para poder llegar a sus citas. Ante este escenario, la rehabilitación virtual luce como una alternativa atractiva que puede ayudar a los pacientes EVC en su recuperación.

De acuerdo con Klinger, la rehabilitación virtual se refiere a cualquier interacción física o cognitiva en tiempo real en un ambiente artificial generado por una computadora o un dispositivo móvil, que parezca o se sienta similar a eventos y objetos del mundo real (Klinger et al., 2010). De forma complementaria, Burdea propone que la rehabilitación virtual representa una intervención terapéutica, local o a distancia, que emplea dispositivos de RV y simulaciones (Burdea, 2003). Actualmente es aceptable considerar que la rehabilitación virtual es un tipo de tratamiento terapéutico que generalmente se aplica a través de juegos serios, los cuales pueden o no estar basados en RV, en conjunto con sensores de posición y movimiento.

El término Realidad Virtual ha recibido numerosas definiciones desde que fue utilizado por primera vez por Jaron Lanier en 1989 (Steuer, 1992), una de las más aceptadas considera a la RV como un espacio virtual 3D generado por computadora al que una persona puede acceder mediante un dispositivo colocado sobre la cabeza, tipo visor, donde se despliegue la imagen. El visor le permite al usuario controlar una cámara con una libertad de 360° para poder observar el mundo virtual y podrá interactuar en tiempo real con objetos virtuales y eventos que parecerán, sonarán, y en algunos casos se sentirán, similares a aquellos del mundo real, dándole la sensación al usuario de estar inmerso en el ambiente virtual (Wilson et al., 1997). Otros autores son más

flexibles cuando definen la RV, pues consideran que el entorno virtual puede ser similar al mundo real o estar basado en un mundo de fantasía (Charles et al., 2020), así como también es posible que la imagen se despliegue sobre una pantalla o tableta, en cuyo caso se suele hablar de una RV no inmersiva (Saposnik et al., 2010).

Por otra parte, existen otros términos para referirse a aquellas aplicaciones interactivas cuyo propósito va más allá del entretenimiento. Los juegos serios son videojuegos que le proporcionan al jugador una habilidad que puede ser aplicada en el mundo real, su objetivo es la educación y/o el entrenamiento en cualquier forma (Broeren et al., 2008). Los juegos serios incluyen juegos para el aprendizaje, para la salud, la política y el cambio social. Dentro de esta clasificación existen los exergames, definidos como juegos que usan el ejercicio para promover la salud y una buena condición física, y los videojuegos activos, que son todos aquellos videojuegos empleados para la rehabilitación sensomotora y cognitiva a través de dominios como la función, estructura, actividad y participación del cuerpo, este es uno de los términos más recientes y fue propuesto en la Conferencia Internacional de Rehabilitación Virtual (ICVR) 2021. Una de las razones por las que la rehabilitación virtual se ha vuelto una herramienta terapéutica popular es que propicia la práctica repetida de ejercicios gracias a que las actividades que se realizan han sido “gamificadas”, es decir, los ejercicios se convierten en juegos, específicamente juegos serios.

La idea de que se utilicen juegos para la rehabilitación virtual tiene su base en el reaprendizaje por el que tiene que pasar un paciente que ha sufrido un EVC. En un juego, el aprendizaje es central para poder avanzar, así como también es crucial que el paciente reaprenda habilidades motrices para realizar tareas adecuadamente y mejorar. Los videojuegos buscan introducir al jugador en un mundo temporal, donde aumenten sus habilidades y eso les permita superar retos y alcanzar metas específicas, lo que es básicamente un proceso de aprendizaje (Charles et al., 2020). Las mismas técnicas que se usan en los videojuegos convencionales para mantener la atención de los jugadores (logrando que las personas pasen horas tratando de dominar un juego), pueden emplearse para aumentar el compromiso en los juegos serios. Estas características de motivación y de apoyo para el aprendizaje sugieren que los videojuegos tienen el potencial de incrementar el compromiso de los pacientes con su terapia, aumentar la frecuencia e intensidad de sus ejercicios, y de mejorar su aprendizaje motor. Otro aspecto que vale la pena destacar de la rehabilitación virtual es que, a pesar de que nace con la idea de simular en ambiente virtuales los mismos ejercicios que se realizan en la terapia convencional, este tipo de tratamiento abre la posibilidad de crear nuevos ejercicios que aprovechen la tecnología y contribuyan en el proceso de mejora del paciente de formas que no serían posibles en otro tipo de estrategias.

Durante las últimas dos décadas, a la par del desarrollo de nuevas tecnologías interactivas, se han realizado numerosos estudios que investigan la forma en la que la RV y los videojuegos serios pueden ayudar a las personas que han sufrido un EVC a incrementar las mejoras motrices, su motivación y el compromiso con su rehabilitación (Burdea et al., 2020; Saposnik & Levin, 2011).

4.1 Clasificación

Una forma en la que se puede clasificar a la rehabilitación virtual es considerando la población a la que se dirige, que pueden ser:

4.1.1 Rehabilitación musculo-esquelética

Se enfoca en la rehabilitación de personas que sufrieron una lesión muscular u ósea, aunque también puede ocuparse en el proceso de recuperación de fuerza muscular después de una cirugía. Los trastornos ortopédicos son una de las principales causas de discapacidad crónica en el mundo. Este tipo de sistemas busca la restauración de los rangos funcionales de movimientos, la recuperación de la fuerza muscular y la reducción del dolor.

La rehabilitación ortopédica debe centrarse en la movilización de articulaciones individuales, con incrementos moderados, su práctica debe ser intensiva y comenzar tan pronto como sea posible. Este campo no ha sido muy explorado, pero uno de los más recientes y significativos desarrollos en el área es el sistema PERCRO-BRANDO (Padilla-Castañeda et al., 2018), del cual se ahondará más adelante.

4.1.2 Rehabilitación cognitiva

Son programas dirigidos a individuos con desórdenes psicológicos, que van desde déficits de atención hasta desórdenes alimenticios, trastornos de estrés postraumático y fobias.

Un ejemplo de la aplicación de este tipo de rehabilitación virtual es en el tratamiento de personas con miedo a volar. Tradicionalmente se les pide a los pacientes tomar un vuelo en compañía de un terapeuta para desensibilizarlos. Sin embargo, la privacidad de la persona se ve comprometida, ya que todas las manifestaciones de su padecimiento (temblores, ataques de pánico, desmayos, etc.) son observadas por todos los pasajeros. Una simulación de vuelo empleando el visor HMD V6 fue desarrollada por el Instituto de Tecnología y Ciencias de la Salud de Georgia, donde el tratamiento se aplica en total privacidad. Un año después de haber sido expuestos al ambiente virtual 92% de los pacientes mantuvieron sus ganancias y tomaban vuelos sin inconvenientes (Baños et al., 2002).

La plataforma comercial CogniFit ofrece un programa de entrenamiento cognitivo en línea que ha sido validado en diferentes grupos poblacionales, como es el caso de personas con discapacidades intelectuales y de desarrollo (Siberski et al., 2014). Se trata de una plataforma en Internet que alberga diversos videojuegos en 2D, los cuales se pueden jugar desde la computadora utilizando el teclado y el mouse. Cada uno de los juegos cuenta con 3 niveles de dificultad (fácil, intermedia y difícil). En cada sesión de entrenamiento se asignan automáticamente juegos enfocados a diferentes dominios como memoria de trabajo, atención focalizada, percepción espacial, comprensión lectora, inhibición, etc.; los resultados que se obtengan en la sesión determinarán los juegos y el nivel de dificultad que se asignarán la próxima sesión.

4.1.3 Neuro-rehabilitación

En este caso la población está compuesta principalmente por personas que han sufrido un EVC, aunque este tipo de desarrollos también puede aplicarse en personas con Parkinson, parálisis cerebral u otro tipo de padecimientos neurológicos. La investigación para pacientes EVC con este tipo de terapia casi siempre se centra en la recuperación de la función motora de miembros superiores (Shin et al., 2014), ya que, como se mencionó en el capítulo anterior, una de las prioridades antes de dar de alta a un paciente del hospital es tratar de recuperar el control de tronco, el equilibrio y la marcha. También es posible que un juego serio trabaje simultáneamente déficits cognitivos y motrices, aun cuando la aplicación no esté explícitamente enfocada en ello. Algunos estudios donde se emplean juegos serios han reportado mejoras cognitivas gracias a la misma naturaleza de los juegos, lo cuales requieren de una constante atención, habilidades para resolver problemas y de razonamiento perceptual (Trabucco et al., 2019).

Los primeros estudios hace ya más de una década acerca de la rehabilitación virtual para pacientes EVC empleaban videojuegos y dispositivos comerciales, logrando resultados favorables (Saposnik et al., 2010), sin embargo, los mejores resultados se han alcanzado empleando sistemas desarrollados específicamente para la rehabilitación de pacientes EVC (Maier et al., 2019).

Existen diversos estudios que comparan la efectividad de la rehabilitación virtual contra la terapia convencional en pacientes EVC (Henderson et al., 2007; Hung et al., 2019; Laffont et al., 2019). Sin embargo, la mayoría concluye que los pacientes logran obtener, en el mejor de los casos, las mismas ganancias que las obtenidas con una intervención tradicional, por esta razón siempre se recomienda el uso de la rehabilitación virtual como un complemento terapéutico, nunca como un sustituto de la terapia física u ocupacional. En un estudio reciente en Francia (Laffont et al., 2019), donde participaron 46 pacientes EVC subagudos divididos aleatoriamente en dos grupos, uno que realizó terapia ocupacional junto con fisioterapia y otro que realizó fisioterapia junto con rehabilitación virtual, se observó que después de 6 semanas de entrenamiento no hubo diferencia significativa en la recuperación motriz de ambos grupos ($p > 0.005$), lo que indica que, en general, todos los pacientes obtuvieron el mismo grado de recuperación, sin importar el tipo de intervención. Un hallazgo interesante de la investigación fue que aquellos pacientes que recibieron rehabilitación virtual dentro de los 30 días después del EVC mostraron mejoras en la escala de Fugl-Meyer notablemente mayores (aunque no significativamente diferentes estadísticamente) a las de pacientes que no utilizaron los videojuegos, o de aquellos que si los emplearon, pero después de los primeros 30 días, lo que podría sugerir una “ventana de oportunidad” para este tipo de intervenciones.

4.2 Interfaces sensoriales y motoras en la rehabilitación virtual

La forma en la que las personas experimentan la rehabilitación virtual es a través de interfaces sensoriales, que transmiten información del sistema al participante, y de interfaces motoras, que son las que le proporcionan información al sistema acerca de las acciones del paciente. La elección de ambas es importante, porque impactará la forma en la que el usuario responderá al sistema,

determinará su nivel de interacción e inmersión en el ambiente virtual, así como los riesgos a los que estará expuesto. Una interfaz se considerará intuitiva si su uso resulta natural para el usuario y prácticamente no hay necesidad de aprender a utilizarla (Klinger et al., 2010).

Las interfaces sensoriales más comunes en la rehabilitación virtual para desplegar la imagen de la simulación o del juego que vaya a realizar son las tablets y los monitores, donde también se proporciona retroalimentación visual a través el movimiento de un avatar, mostrando mensajes de ánimo y/o de advertencia. En algunos estudios con pacientes EVC se han utilizado visores comerciales de RV como el Oculus Rift (Ögün et al., 2019), así como el visor HTC Vive (Burdea et al., 2020), que ofrecen una experiencia más inmersiva. Proyectos más ambiciosos, como el programa CAREN (*Computer Assisted Rehabilitation Environment*) del Instituto Neurológico de Cleveland (Isaacson et al., 2013), donde la imagen se despliega en un domo y se coloca al paciente sobre una caminadora a través de un arnés que sujeta su peso (figura 4.1), logran un grado de inmersión aún mayor, sin embargo, debido a la enorme infraestructura que requieren, este tipo de interfaces son únicas en su tipo. Adicionalmente, se han explorado distintos dispositivos hápticos para enriquecer las señales que recibe el usuario, algunos ya se mencionaron anteriormente como los exoesqueletos Armeo y ARMin, así como los guantes de CyberGlove Systems, pero también existen otros dispositivos como el HapticMaster (Van der Linde et al., 2002), y más recientemente UltraHaptics (UHDK5), una mejora del dispositivo Leap Motion que emplea sensores de ultrasonido para proveer retroalimentación háptica de forma no invasiva (Georgiou et al., 2018).

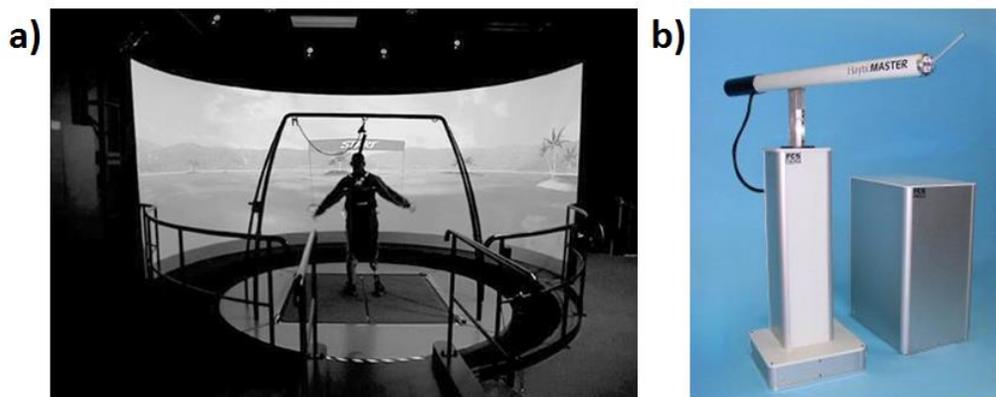


Figura 4.1. a) Interfaz sensorial del proyecto CAREN. (Isaacson et al., 2013). b) Dispositivo HapticMaster. (Van der Linde et al., 2002).

Las primeras interfaces motoras que se utilizaron en estudios de rehabilitación virtual para pacientes EVC fueron dispositivos de consolas de videojuegos y sus juegos asociados, como el Nintendo Wii (Saposnik et al., 2010) y el sensor Kinect de Microsoft (Chang et al., 2011), con los que no se lograron buenos resultados en la recuperación motora, principalmente porque la dificultad de los juegos, los cuales no estaban diseñados para la rehabilitación, era demasiado elevada para los pacientes y la espasticidad o flacidez en las manos le dificultaba a los pacientes el uso de los controles del Wii. No obstante, el Kinect se volvió en una de las interfaces motoras más populares en los estudios sobre rehabilitación virtual.

4.2.1 Sistemas de seguimiento de movimiento humano

En un artículo de la Revista Mexicana de Ingeniería Biomédica (Bravo et al., 2016) se propone una clasificación de las interfaces motoras, mejor conocidas como sistemas de captura de movimiento, según la tecnología que utilicen en los siguientes grupos:

- a) **Electromecánicos.** La captura de movimiento se realiza utilizando sensores mecánicos, como potenciómetros o servomotores, que recopilan información acerca del grado de rotación de las articulaciones del participante. Algunos de estos dispositivos que se han empleado en estudios de rehabilitación de pacientes EVC son el Armeo (Colomer et al., 2013), ARMin (Guidali et al., 2011), MIT-MANUS (Krebs et al., 1999), InMotion ARM (Hsieh et al., 2018) y PERCRO GRAB (Cameirão et al., 2012), los cuales, al proveer también retroalimentación háptica, son al mismo tiempo interfaces motores e interfaces sensoriales.
- b) **Electromagnéticos.** En este caso se cuenta con varios sensores electromagnéticos que miden la relación espacial con ayuda de un transmisor cercano. El transmisor genera un campo electromagnético de baja frecuencia que detectan los sensores, los cuales miden el flujo magnético para determinar su posición y orientación en el espacio y transmiten esta información a una unidad central. Este tipo de dispositivos no ha sido muy explorado y el sensor más representativo que ha sido usado en estudios con pacientes EVC es el Polhemus (Bohannon et al., 2009).
- c) **Inerciales.** Para detectar el movimiento del participante se deben colocar sensores IMU en la parte del cuerpo que se desee rastrear. Los IMU (Inertial Movement Units) son una combinación de giróscopos, que miden las velocidades angulares alrededor de un sistema de coordenadas de un sensor, acelerómetros, que miden la aceleración gravitacional y traslacional, y en algunos casos magnetómetros, que miden el campo magnético de la tierra. Las ventajas de estos sensores es que no presentan problemas de oclusión, son compactos y muy fáciles de transportar.



Figura 4.2. Dispositivo inercial Pablo de la empresa Tyromotion. (Tramontano et al., 2020)

Existen trajes como el Shadow Motion Capture System y el Xsens (Zhang et al., 2013) que utilizan este tipo de tecnología y se usan más que nada para la producción de películas y

videojuegos, además de que su precio es bastante elevado. Por otra parte, el dispositivo inercial Pablo de la empresa Tyromotion fue diseñado específicamente para pacientes EVC (Tramontano et al., 2020), este sensor ilustrado en la figura 4.2 se puede colocar sobre otros dispositivos como un volante, una balanza, una muñequera o una pelota, mientras que una aplicación externa recibe los datos del sensor para medir los movimientos que el participante realiza con el objeto.

- d) **Ópticos.** Los sistemas ópticos utilizan datos recogidos por sensores de imagen para inferir la posición de un elemento en el espacio, pueden utilizar una o varias cámaras y pueden o no utilizar marcadores pegados al cuerpo del participante. Los marcadores pueden ser activos (si emiten alguna luz) o pasivos, y se colocarán en todos los puntos de interés del cuerpo para ayudar al sistema a discriminar del resto de los objetos de la imagen. El sistema VICON se considera en este momento el estado del arte por su alta precisión para capturar movimientos humanos (Rastegarpanah et al., 2018), sin embargo, sus principales desventajas son que requiere mucho tiempo para adecuar el espacio de trabajo y los marcadores, su costo es muy elevado y, como todos los sistemas ópticos, sufren de la oclusión y dependen de una correcta iluminación. Un sensor que permite rastrear el movimiento ocular es el Tobii T120, este dispositivo se usó en un estudio para medir en pacientes EVC la latencia entre la observación y la acción entre el brazo parético y el no parético (Alves et al., 2014).

A continuación se mencionarán los 2 sistemas ópticos sin marcadores más populares en estudios de rehabilitación virtual.

4.2.1.1 Kinect

Es un sensor óptico desarrollado por Microsoft (figura 4.3), fue lanzado en el 2010 para la consola de videojuegos Xbox 360 y más tarde se habilitó su compatibilidad con computadoras personales a través de un adaptador USB. Posee un arreglo de 4 micrófonos para el reconocimiento de voz, una cámara RGB con una resolución espacial de 640x480 pixeles y una resolución temporal de 30 fps, un emisor infrarrojo y una cámara infrarroja monocromática CMOS que, en conjunto, permiten detectar la profundidad de los objetos hasta una distancia de 4m. Estos componentes proveen una captura de movimiento de cuerpo completo, capaz de identificar 20 articulaciones, midiendo 2 grados de libertad del hombro y 1 grado de libertad en el codo (Galna et al., 2014).

Es uno de los sensores con los que más investigaciones acerca del impacto de la rehabilitación virtual se han realizado, esto gracias a que su Software Development Kit (SDK) permite realizar desarrollos de aplicaciones para computadora. Su principal ventaja es que, al ser un sensor óptico que no usa marcadores no es invasivo, y permite detectar las articulaciones de las personas mientras están sentadas o paradas, por lo que puede emplearse en individuos con un amplio rango de déficits motores, incluyendo aquellos en silla de ruedas. Gracias a su detección de cuerpo completo se pueden desarrollar aplicaciones que involucren el movimiento de manos, pies, brazos o inclusive del cuerpo completo, lo que ofrece una amplia gama para el desarrollo.

En su época de mayor popularidad conseguir este sensor era muy sencillo, pues su costo era accesible y su oferta era alta, pero después de que fue discontinuado en 2017 por Microsoft cada vez es más difícil encontrarlo y los sensores que podrían funcionar como alternativa son significativamente más costosos (e.g. Intel RealSense, Orbbec y VicoVR).

En un estudio que midió la precisión del sensor para detectar movimientos clínicamente relevantes, comparándolo con el sistema de captura de movimiento Vicon, se concluyó que es poco confiable para detectar movimientos finos de las manos, como entrecruzar los dedos ($ICC=0.012$), su mayor correlación con los movimientos reales se obtiene con movimientos gruesos de los brazos y movimientos de sentado a parado ($ICC=0.989$) (Galna et al., 2014). Una investigación que recopila los resultados de múltiples investigaciones que evalúan la precisión y confiabilidad del Kinect concluye que es un sensor aceptable y accesible para ser usado con propósitos de rehabilitación (Mousavi Hondori & Khademi, 2014), sin embargo se debe prestar atención a sus limitaciones, ya que, como sucede con cualquier sensor óptico, es susceptible a la oclusión y depende de la iluminación donde se use. Además, el Kinect no se puede usar en exteriores, ya que los rayos UV interfieren con el sensor de profundidad y se requiere de un espacio relativamente amplio para ser usado, pues funciona mejor cuando la persona se coloca a 2 metros de distancia del sensor.



Figura 4.3. Sensor Kinect. (Galna et al., 2014).

La siguiente versión del Kinect, el Kinect V2, posee una cámara RGB con una resolución HD (1920x1080) y además es capaz de identificar hasta 25 articulaciones del cuerpo humano, sin embargo, las características mínimas que deben tener las computadoras para poder funcionar con este sensor son tener puertos USB 3.0, una tarjeta de video dedicada y Windows 8 o mayor, lo que limita mucho su accesibilidad.

Gracias a todas las características que se han expuesto, el Kinect ha sido uno de los sensores más utilizados en estudios sobre rehabilitación virtual, desde uno de los primeros estudios piloto de rehabilitación virtual, donde se observó el potencial que tenía este sensor junto con el uso de videojuegos en la rehabilitación de pacientes EVC (Chang et al., 2011), pasando por la plataforma comercial Jintronix (Archambault et al., 2014) mostrado en la figura 4.4, y más recientemente en juegos como Kinect2Scratch (Hung et al., 2019). Estas últimas dos investigaciones se llevaron a cabo en pacientes EVC crónicos con hemiplejía en miembros superiores y obtuvieron resultados

muy favorables en la recuperación motora, por lo que recomiendan su uso como una estrategia complementaria a la terapia convencional para disminuir la carga de trabajo de los terapeutas.



Figura 4.4. Ejemplo de un juego de la plataforma Jintronix junto el sensor Kinect. (Archambault et al., 2014).

Finalmente, en un estudio con pacientes subagudos (Demers et al., 2019) se utilizó el Kinect para simular la actividad de comprar la despensa en un supermercado (figura 4.5). El participante debe tomar los objetos desde los anaqueles y procurar realizar la cuenta correctamente para pagar, entrenando habilidades motoras y cognitivas simultáneamente.

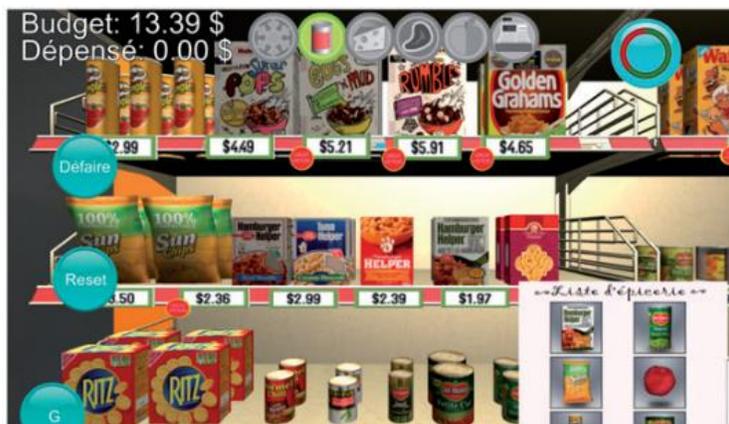


Figura 4.5. En esta aplicación se simula que el paciente compra diferentes artículos en un supermercado, los movimientos del participante se detectan con el sensor Kinect. (Demers et al., 2019).

4.2.1.2 Leap Motion

Es un dispositivo óptico de 3''x1.2''x0.5'' diseñado para colocarse enfrente de la pantalla sobre una mesa. Cuenta con 3 LEDs infrarrojos y 2 cámaras infrarrojas. Su campo de visión es una hemiesfera sobre el dispositivo con un radio entre 25 y 600 mm (Hung et al., 2019). Este sensor fue creado para reconocer movimientos de la mano y los dedos en tiempo real, y es capaz de detectar dos manos al mismo tiempo. Es común que se use con el Oculus Rift a través de un adaptador que permite montar el Leap Motion sobre el visor, de esta forma el participante puede

interactuar con el ambiente virtual con sus manos virtuales, volviendo la experiencia aún más inmersiva que si únicamente se usara el visor. El sensor proporciona una gran precisión para movimientos gruesos de la mano, sin embargo, dependiendo de su orientación y de la orientación de la mano, llega a presentar errores en movimientos finos como la pinza terminal-terminal.

Uno de los primeros estudios que utilizó este sensor en pacientes EVC (Khademi et al., 2014) se enfocó en la factibilidad de emplear el Leap Motion para la rehabilitación de miembros superiores. En su estudio piloto participaron 14 pacientes, quienes utilizaron el sensor para interactuar con un videojuego donde se requieren diferentes combinaciones de movimientos de la mano y correlacionaron sus marcadores con las pruebas estandarizadas Fugl-Meyer y Box and Block Test. Concluyeron que las lecturas del sensor son suficientemente precisas para realizar ejercicios terapéuticos, encontraron una buena correlación entre el juego y las pruebas estandarizadas, y los participantes estuvieron totalmente de acuerdo en que el juego era útil, divertido y promovió movimientos de AVD.

En un estudio en Turquía (Ögün et al., 2019), 65 pacientes EVC crónicos fueron divididos aleatoriamente en un grupo control (n=32) y en un grupo experimental (n=32), el grupo control recibió terapia ocupacional convencional y el grupo experimental recibió rehabilitación virtual utilizando el sensor Leap Motion junto con el Oculus Rift (figura 4.6) en ambientes virtuales donde se simulaba que el paciente realizaba las mismas actividades que en la terapia ocupacional, como cocinar y mover objetos de un lado a otro, las sesiones duraban 1 hora. La intervención duró 6 semanas y al finalizar se encontró que ambos grupos mostraron una mejora motriz significativa. Sin embargo, el grupo experimental logró una mejora significativamente mayor al grupo control ($p < 0.001$), mostrando mayor funcionalidad en ambos miembros superiores, independencia y autocuidado.

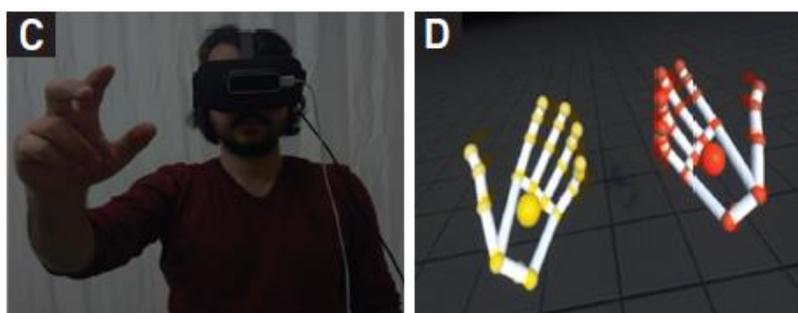


Figura 4.6. C) Muestra el sensor Leap Motion acoplado al visor Oculus Rift. D) Gracias a esta combinación, el participante puede observar los movimientos de sus manos en el ambiente virtual. (Ögün et al., 2019).

A pesar de esta clasificación de los sistemas de captura de movimiento, también existen dispositivos “híbridos”, que implementan diferentes métodos para lograr medir los movimientos humanos. Ejemplo de ello es el guante Smart Glove de la empresa Neofect mostrado en la figura 4.7, también conocido como Rapael (Kim et al., 2018). Este dispositivo fue diseñado específicamente para pacientes EVC, posee acelerómetros y sensores de flexión para medir movimientos del antebrazo, muñeca y motricidad fina de la mano. Por supuesto, tampoco se debe

olvidar que la interacción con un sistema de rehabilitación virtual también se puede realizar con el mouse y teclado tradicionales de la computadora.



Figura 4.7. Guante Rapael y la plataforma de videojuegos asociados a él. (Kim et al., 2018).

4.3 Beneficios de la rehabilitación virtual

Es probable que se obtengan mejores resultados de rehabilitación al integrar principios de control motor y aprendizaje motor involucrando retroalimentación multimodal relevante y procesos cognitivos. Dicho de otra forma, practicar tareas en ambientes virtuales puede mejorar el aprendizaje de habilidades motoras a través de la integración de múltiples procesos sensoriales como la propiocepción, la audición, la visión y la información vestibular, junto con procesos cognitivos como la planeación, resolución de problemas, la atención, la memoria y la orientación viso espacial (Levin et al., 2015). El enriquecimiento de los estímulos que proporcionen estos ambientes virtuales motiva a los pacientes a realizar sesiones de ejercicios más largas, más intensas y con mayor frecuencia (Henderson et al., 2007).

A través del uso de plataformas ad hoc para rehabilitación, es posible incrementar el número de repeticiones, duración e intensidad de los ejercicios, ya sea realizando tareas enfocadas a un objetivo o ejercicios repetitivos en un ambiente interactivo único con un contexto funcional, útil y motivacional, que le permite al terapeuta adaptar los ejercicios que realizará el paciente, así como modular la retroalimentación visual, auditiva y en algunos casos háptica que recibirá el paciente para mejorar su aprendizaje (Laffont et al., 2019).

Cuando el usuario realiza actividades divertidas y disfrutables en un ambiente virtual es más probable que se genere un apego al tratamiento, propiciando que el paciente realice sus ejercicios con una mayor frecuencia e intensidad. Estos efectos favorables se han encontrado en las investigaciones de Broeren (Broeren et al., 2008), y Yao-Jen Chang (Chang et al., 2011). Por su parte, Pallesen (Pallesen et al., 2018) y Ho-Shin (Shin et al., 2014) encontraron que la naturaleza del entrenamiento basado en RV pareció tener una influencia significativa en los estado de ánimo y el compromiso de los pacientes con su rehabilitación.

En los ambientes virtuales se tiene control sobre lo que el paciente realiza, de tal forma que no se expone al paciente a situaciones de riesgo, además se puede modular el tipo de tarea a realizar y su dificultad (Burdea, 2003).

En cada sesión de ejercicios existe la oportunidad de realizar diferentes tipos de mediciones en tiempo real, dependiendo del tipo de interfaz que se utilice se podrá capturar información acerca de la cinemática de los movimientos, tiempos de reacción, fuerza, y otros parámetros que ayudarían a detectar si el paciente está trabajando a su máxima capacidad o si está realizando movimientos compensatorios, todo ello con el fin de evaluar de una mejor forma el desempeño del paciente y poder tomar mejores decisiones respecto al programa terapéutico.

Uno de los más importantes beneficios de la rehabilitación virtual es la posibilidad de ofrecer tele rehabilitación, que es suministrar de forma remota los servicios de rehabilitación, sustituyendo el acercamiento tradicional cara a cara entre paciente y rehabilitador (Peretti et al., 2017). La tele rehabilitación nace con la idea de alcanzar los mismos resultados en casa que los que se lograrían en un hospital en presencia de un terapeuta, las dos motivaciones principales son 1) cubrir situaciones donde tener acceso a las infraestructuras tradicionales de rehabilitación resulta complicado para los pacientes y 2) comenzar el tratamiento del padecimientos lo más pronto y con la mayor intensidad posible, ya que en las fases iniciales es donde se puede alcanzar la mejor recuperación, como es el caso de los pacientes EVC en la fase subaguda.

Como se mencionó en el capítulo 2, de acuerdo con las tendencias actuales y a la posible incidencia del virus SARS-Cov2 en los infartos cerebrales, se pronostica que en los próximos años el número de casos de EVC aumente, lo que amenazaría la sustentabilidad de los sistemas de salud al no contar con la capacidad de atender a tantas personas. La tele rehabilitación puede ser una herramienta que ayude a paliar este problema, facilitando que se pueda atender a un mayor número de pacientes.

Usar sensores de posición y movimiento en conjunto con juegos serios en casa, que almacenen información relevante acerca de las tareas que realiza el paciente, brindan la opción de que un mismo terapeuta supervise de forma remota el progreso de varios pacientes, e inclusive pueda asignarles ejercicios personalizados.

En México los servicios de rehabilitación no suelen brindarse con la oportunidad y frecuencia que los pacientes necesitan debido a la falta de recursos humanos, materiales y de infraestructura en los centros de salud. Aunado a esto, los déficits que llega a presentar una persona que ha sufrido un EVC pueden complicarle mucho el traslado a la clínica, el paciente requerirá el apoyo de un familiar y usar frecuentemente transporte particular lo que, dependiendo de su situación socioeconómica, puede no ser viable. La tele rehabilitación permite que el paciente se vuelva más independiente, ya que al realizar sus ejercicios desde casa no necesitaría trasladarse tan seguido al hospital y no dependería de la disponibilidad de citas. Además, podría realizar sus ejercicios de forma oportuna, con la intensidad y frecuencia que se requiera e inclusive continuar con su terapia más tiempo de lo habitual.

4.4 Retos y limitaciones de la rehabilitación virtual

Los dispositivos y los juegos serios asociados a ellos sólo permiten practicar un número limitado de movimientos. La solución a este problema puede ser contar con distintos sensores, o

desarrollarlos desde cero para una tarea específica, así como diseñar juegos para cada uno de estos periféricos donde se tenga la oportunidad de entrenar diversas funciones motoras y así abarcar una mayor cantidad de movimientos. Estos dispositivos deben ser creados teniendo en cuenta los diferentes tamaños poblacionales de quien podría llegar a usarlos y ser lo más ergonómicos posibles.

Se corre el riesgo de que el paciente realice movimientos no óptimos o compensatorios durante sus ejercicios y además se refuerce este comportamiento, esto puede suceder si los sensores de detección de movimiento no tienen suficiente precisión y/o si no se brinda una adecuada retroalimentación respecto al desempeño de la tarea. Este es uno de los principales problemas de usar interfaces de videojuegos comerciales, como en el caso del Wii, ya que estos juegos y dispositivos están pensados para ser usados por personas sin ningún tipo de déficit motor o cognitivo.

El costo de las interfaces motoras, especialmente los dispositivos hápticos y los exoesqueletos, llegan a ser muy altos y requieren mucho espacio. En México, en términos de reportes estadísticos, las cifras del número de personas rehabilitadas no suelen despertar mucho interés en las instituciones de salud, por lo que es poco probable que realicen una inversión fuerte en esta área al no considerarla prioritaria. En ese sentido, los futuros desarrollos tecnológicos de este tipo en México deben tratar de ser lo más accesibles posible.

A pesar de que ya se han realizado diversas investigaciones de tratamientos con rehabilitación virtual en pacientes EVC, aún se desconoce el tipo o las combinaciones de retroalimentación que impacte de manera significativa en la recuperación de la función motora. En un estudio (Cameirão et al., 2012) se intentó averiguar la forma en la que la tecnología podía maximizar la explotación de los mecanismos neuronales de recuperación y compensación. Para ello emplearon 3 grupos de pacientes EVC crónicos. Un grupo (n=17) fue intervenido únicamente a través del *Rehabilitation Gaming System* (RGS), del cual se ahondará más adelante, otro grupo (n=16) utilizó el RGS junto con la interfaz háptica GRAB, y el último grupo (n=15) utilizó el RGS junto con el exoesqueleto Armeo. El tratamiento consistió en sesiones diarias de 35 minutos durante 4 semanas, los ejercicios constaban de tareas bimanuales asignadas por un Módulo de Entrenamiento Personalizado (MEP) integrado en el RGS. Se realizaron numerosas pruebas estandarizadas antes, al final y 8 semanas después de la intervención. En todos los participantes se encontraron diferencias significativas ($p < 0.001$) del impacto del tratamiento en las pruebas *Barthel Index*, *Motricity Index*, *Fugl-Meyer* y *CAHAI* al comparar los puntajes iniciales con el puntaje de la etapa de seguimiento. No se encontraron diferencias significativas en los puntajes entre grupos ($p > 0.001$), sin embargo, se observó que el grupo que trabajó con el Armeo no retuvo las ganancias obtenidas por el tratamiento en contraste con los otros 2 grupos. Se piensa que una causa de esto es que el exoesqueleto, no sólo apoya y brinda soporte anti gravitatorio, sino también restringe la cinemática de los movimientos de los brazos y no facilita estrategias compensatorias. Como se mencionó en el capítulo anterior, en el caso de pacientes crónicos la compensación no necesariamente es negativa. Por otra parte, el grupo que utilizó la interfaz háptica fue el único que logró retener todas las ganancias obtenidas por el tratamiento, esto les llevó a formular la

hipótesis de que la retroalimentación háptica junto con un videojuego serio promueve que haya una mayor cantidad de estimulación de modalidades sensoriales en el cerebro al realizar ejercicios y esto contribuya a una mejor recuperación a largo plazo. El estudio concluye reconociendo que se requieren más estudios al respecto con una mayor cantidad de participantes.

Respecto a la tele rehabilitación mediante interfaces de RV y juegos serios existen pocos estudios (Peretti et al., 2017), por lo que aún falta más investigación que compruebe su efectividad. Algunas limitaciones para su implementación es que el paciente debe de contar con una computadora y con Internet en su hogar, también debe tener espacio suficiente para poder realizar sus ejercicios y de preferencia contar con la asistencia de un familiar. Además, quien brinde este servicio debe contar con personal que pueda ofrecer asistencia técnica en caso de que ocurra algún problema con los periféricos o los programas interactivos, y también debe tener una logística de seguimiento que asegure que el paciente no sólo está realizando sus ejercicios con la frecuencia adecuada, sino también de la forma correcta. Otro punto a considerar es que algunos pacientes podrían desmotivarse con la tele rehabilitación al realizar con menos frecuencia sus ejercicios debido a la falta de atención humana.

Existe la creencia errónea que la rehabilitación virtual es un sustituto de la rehabilitación convencional, cuando en realidad toda la investigación que se ha realizado al respecto confirma que son herramientas amplificadoras para los terapeutas, que les permitirá hacer más y con más pacientes.

La integración de este tipo de desarrollos en la práctica clínica es uno de los mayores retos de la rehabilitación virtual, pues aunque existe evidencia sobre sus beneficios, al final es el personal clínico quien deberá administrar el programa de tratamiento para tomar ventaja de la tecnología. Algunos motivos por los que se puede existir una falta de aceptación son la aversión a la tecnología, dificultad para adaptarse a las interfaces, falta de interés, y escepticismo sobre su efectividad, este último punto tiene su justificación en que muchos de los estudios que existen sobre la rehabilitación virtual no llegan a conclusiones firmes principalmente debido a debilidades metodológicas o tamaños de cohorte muy pequeños.

4.5 Factores en el diseño de juegos serios para rehabilitación

De acuerdo con Schell (Schell, 2008), el objetivo central al diseñar cualquier juego debe ser crear una experiencia estética para los jugadores que inculque alguna forma de respuesta emocional. El modelo MDA (*Mechanics, Dynamics, Aesthetics*) (Hunicke et al., 2004) es uno de los más comunes para el diseño de juegos. Este marco de trabajo resulta útil para representar la estructura cíclica en la que residirá el jugador.

La base del modelo MDA es la tecnología, ya que la interacción del jugador y las mecánicas de juego dependerán tanto del tipo de sensores que se ocupen (e.g. exoesqueletos, Kinect, mouse y teclado), como de la plataforma donde se desplegará el juego (e.g. visor de RV, pantalla de computadora, tableta). Dentro del modelo MDA, las mecánicas de juego son la pieza central de todo, son las que definen las reglas del juego, dictan qué es lo que el jugador puede o no hacer

dentro del ambiente virtual, la forma en la que cambiará la dificultad y los momentos en los que sucederá, definen el propósito del juego y el alcance que tendrá. Los juegos como tal son herramientas inherentes de enseñanza, y uno de sus propósitos debe ser ayudar al jugador a que aprenda cómo jugar e incrementar su habilidad dentro del juego. Las dinámicas del juego se refieren a lo que sucederá en el ambiente virtual cuando el jugador interactúe con él, y estarán supeditadas a las mecánicas. Finalmente, la estética del juego tiene una relación directa con la experiencia del jugador y su respuesta emocional o intelectual a la retroalimentación del juego, en este componente se debe decidir el tipo de música y el aspecto visual que tendrá el juego, la interfaz del usuario, así como todas las señales que se le deseen proporcionar al jugador (visuales, auditivas o hápticas), tratando de armonizar todos los elementos con el fin de evocar una respuesta emocional. A través de las mecánicas del juego se controlará el tipo de dinámicas que podrán ocurrir en el ambiente virtual, dependiendo de las dinámicas el jugador experimentará diferentes emociones como conflicto, incertidumbre, satisfacción, frustración, etcétera. Este estado de ánimo también dependerá del componente estético, que incluye toda la retroalimentación que recibirá el jugador mientras experimenta las dinámicas del juego. En general, el diseño se debe enfocar en construir una interfaz de usuario intuitiva, fácil de usar, proporcionar una experiencia de juego disfrutable, brindar mecanismos que faciliten el aprendizaje y proporcionar retroalimentación informativa. El correcto diseño de todo el sistema en su conjunto es lo que permitirá que el jugador aprenda y se divierta.

En el caso de los juegos serios enfocados a neuro-rehabilitación, específicamente para la recuperación de déficits motores, el objetivo principal es que la aplicación sea efectiva, que su uso contribuya a la mejora motriz de tal forma que el paciente se pueda desempeñar mejor en AVD, por esta razón también se deben considerar varios factores como: el objetivo terapéutico del juego, el tipo de movimientos que se entrenarán, la exactitud que deberán tener los movimientos del paciente para poder jugar, las capacidades motrices mínimas necesarias para poder jugar, la interfaz que se empleará para que el paciente interactúe con el juego, es decir, el sensor que detectará los movimientos del paciente, la duración de las sesiones, el tiempo que se le dará al paciente para reaccionar, el número de repeticiones de movimientos que realizará en cada sesión de juego, el grado de reto (la dificultad), las habilidades cognitivas que serán necesarias para jugar y las que se entrenarán de forma directa o indirecta, el grado de autonomía que deberá tener el paciente, el funcionamiento del sistema de recompensas que tendrá el juego, las formas de retroalimentación que existirán (mensajes, sonidos, vibraciones, etc.) y los momentos en los que se proveerán (cuando se complete una tarea, cuando se realice un movimiento incorrecto, etc.). También se deben contemplar las formas en las que el paciente podría hacer trampa al jugar y, finalmente, si la estética y mecánicas de juego se alinean con la personalidad del paciente y sus preferencias.

Otros factores externos a tomar en cuenta son la relación costo-efectividad que tendrá el juego para el paciente, el grado de transferencia de las habilidades que adquiridas al jugar a AVD, si el desarrollo se alinea con estructuras clínicas existentes, si el juego abarca diferentes perfiles

psicológicos, preferencias de juego y capacidades, y si existe un correcto balance entre los beneficios de entretenimiento y de recuperación.

En el caso de aplicaciones de RV, conforme la representación virtual se acerca a la fidelidad del mundo real se puede volver más perturbador para la percepción de una persona. El nivel de realismo que reduce significativamente el confort del observador se conoce como el valle inquietante o *uncanny valley* (Mori et al., 2012). Se prefiere que el diseño sea creíble y no tan realista, esto reducirá el problema del *uncanny valley*, ya que la gente se siente cómoda aceptando un contexto irreal si el contenido es consistente y significativo dentro del contexto, además de que recrear un ambiente virtual menos realista requiere menos poder computacional y vuelve la aplicación más accesible. Es mucho más importante que la calidad de la interacción sea alta a que se trate de lograr una representación visual exacta de la realidad, porque entre mejor sea la calidad de la interacción proporcionará un mayor sentido de inmersión, para lograr eso se recomienda usar dispositivos hápticos y disminuir el tiempo que se tarda en representar en pantalla las acciones del participante (latencia).

Cuando se usan visores de RV, el tiempo que tarda el usuario en mover su cabeza y observar ese cambio debe ser menor a 15 ms (Charles et al., 2020), de lo contrario pueden ocurrir mareos. Este problema también puede aparecer, incluso si no se usa algún tipo de visor, cuando existe un retraso al procesar la respuesta visual a una interacción del usuario, especialmente en juegos con vista en primera persona.

La propiocepción es el sentido que tiene uno mismo sobre su cuerpo en movimiento y su posición en el espacio. Este sentido se requiere para inducir un sentimiento de presencia, a través de un avatar, en un ambiente virtual. Los efectos de la propiocepción pueden ser importantes en los juegos para neuro-rehabilitación, por ejemplo, en los casos donde se intente crear ilusiones como en las de la terapia en espejo, aunque sus efectos aún necesitan ser más investigados. En ese respecto, un estudio (Trabucco et al., 2019) comparó las perspectivas en primera y tercera persona para un mismo conjunto de ejercicios. En sus resultados encontraron que, en el caso de movimientos de miembro superior, no se encontraron diferencias significativas entre ambas perspectivas ($p > 0.001$). En los ejercicios donde los usuarios debían mover la mano para alcanzar objetos en un ambiente virtual, ninguno realizaba trayectorias más cortas. El desempeño en las tareas tampoco fue significativamente diferente entre las perspectivas. Los investigadores encontraron que, en una perspectiva en tercera persona, los sujetos pueden observar inmediatamente cuando realizan acciones incorrectas a través de un avatar y realizar las correcciones necesarias, mientras que, en una perspectiva en primera persona, el campo de visión es más restringido y no facilita esta retroalimentación inmediata. Sin embargo, en una perspectiva en primera persona es mucho más rápido atraer la atención en lo que realizan las manos y esto puede ser de ayuda particularmente en pacientes con negligencia espacial. En el estudio concluyen recomendando usar juegos con perspectiva en tercera persona, donde un avatar refleje los movimientos que realiza el participante, cuando se busque realizar tareas de precisión con una alta carga cognitiva.

La percepción de la profundidad es un factor importante en los juegos donde se trabajen tareas de alcanzar y recuperar. Para ayudar a dar una sensación de profundidad se deben emplear una correcta perspectiva de la cámara en el ambiente virtual, escalando los objetos apropiadamente respecto a su distancia del observador, utilizar oclusión de objetos, así como técnicas de iluminación y sombreado. Al mismo tiempo, también se debe cuidar que no se le presenten al jugador demasiados estímulos (visuales, cognitivos o táctiles), debido a que los pacientes EVC pueden presentar problemas de atención, aversión a la luminosidad y otros problemas sensoriales.

4.6 Retroalimentación positiva o afectiva

Los videojuegos siempre han recompensado a los jugadores por superar retos. Cuando se implementa un buen sistema de recompensas se pueden incrementar los niveles de retención, la motivación del jugador y mejorar su progreso. El diseño de este sistema debe estar basado en lo que es valioso para el jugador dentro del contexto del juego (Schell, 2008). Algunas estructuras de recompensas son:

- a) **Tangibles/intangibles.** Las recompensas intangibles son trofeos, medallas, o símbolos de estatus. Las recompensas tangibles son aquellas que tienen un efecto en el juego que le ayudan al jugador a alcanzar sus objetivos más fácilmente, por esta razón este tipo de recompensas pueden ser más motivadoras.
- b) **Esperadas/inesperadas.** Recibir una recompensa sorpresivamente puede ayudar a “enganchar” al jugador, pues agregan variedad al juego, pero también es importante establecer una estructura de recompensas esperadas donde se tengan objetivos claros hacia los cuales el jugador pueda progresar.
- c) **De desempeño y de terminación.** Las recompensas por completar un nivel o finalizar una tarea son muy importantes, aunque también se pueden modular la calidad o cantidad de las recompensas en relación al desempeño de la tarea. Por ejemplo, completar un ejercicio sin haber realizado ningún movimiento incorrecto o de compensación otorgaría mejores recompensas que las que obtendría el paciente de otra forma.
- d) **De intervalo fijo/variable.** Otorgar recompensas en intervalos fijos de tiempo, por ejemplo, cada 5 minutos. Este tipo de recompensas son apropiadas cuando el objetivo es que el paciente continúe realizando la tarea el mayor tiempo posible. Las recompensas que se otorgan en intervalos aleatorios de tiempo también pueden resultar motivadoras.
- e) **De proporción fija/variable.** Se otorgan recompensas después de que se hayan completado un número fijo de tareas o después de un número impredecible de acciones, éste último esquema funciona para evitar frustración en el jugador.

- f) **Retroalimentación exagerada.** Por ejemplo, si al jugador se le otorga una recompensa 10% mayor a la que en realidad merece, el jugador puede incrementar su desempeño en esa cantidad en su siguiente intento.

La inducción, orientación y entrenamiento de los pacientes con los juegos y sus periféricos asociados son de vital importancia. Una experiencia desafortunada desde el comienzo puede hacer que el paciente sienta que no es capaz de utilizar la tecnología y consecuentemente adoptar una actitud negativa y reacia a volver a intentar jugar.

Tabla 4.1 Principios del control motor y la forma en la que se pueden implementar en un ambiente virtual. (Levin et al., 2015)

Principios del control motor que pueden ser incorporados en la práctica de una tarea	Diseño del ambiente virtual que puede permitir implementar los principios de control motor
<p>La dificultad de una tarea toma en consideración la velocidad y la precisión del movimiento.</p> <p>La organización del movimiento para tomar un objeto, esto es, el número de grados de libertad, los tiempos de movimiento y la coordinación de las rotaciones de las articulaciones, se relacionan con la ubicación y la distancia que existe entre el objeto y el cuerpo.</p> <p>Para un paciente EVC, los movimientos realizados en el área contralateral (cuando el brazo cruza la línea media) son más difíciles que los movimientos en el área ipsilateral.</p>	<p>A través de una interfaz amigable, el personal clínico debe ser capaz de configurar el tamaño de los objetos a alcanzar y su ubicación, otra posibilidad es que el programa se adapte automáticamente a los cambios en el desempeño del paciente.</p> <p>Las tareas en el entorno virtual deben incluir interacciones con objetos ubicados a diferentes distancias del cuerpo y en diferentes posiciones (contralateral, ipsilateral o sobre la línea media), para promover el uso coordinado de combinaciones del brazo y tronco.</p>
<p>La organización del movimiento para tomar un objeto se relaciona con la percepción de la distancia que existe entre nuestro cuerpo y el objeto, así como de las señales visuales de nuestro brazo y de la interacción con el objeto.</p>	<p>El entorno virtual debe incluir elementos como líneas de perspectiva y sombras para ayudar a una mejor percepción de profundidad.</p>
<p>La orientación y trayectoria que tendrá la mano cuando tomamos un objeto se relacionan con la ubicación, tamaño y orientación del objeto.</p>	<p>Los objetos en el entorno virtual que se vayan a tomar deben ser de varias formas y tamaños, su ubicación también debe cambiar.</p>
<p>La organización del movimiento de alcanzar y recuperar dependerá de lo que se quiera realizar con el objeto.</p>	<p>Las tareas en la aplicación interactiva deben tener un propósito, si se van a simular tareas donde se deben tomar objetos usualmente se suelen elegir AVD como limpiar, cocinar o tomar objetos de los estantes de un supermercado.</p>
<p>La retroalimentación sobre el desempeño motor (información acerca de la calidad del movimiento y los rangos articulares utilizados) es esencial para mejorar el comportamiento motor.</p>	<p>La aplicación interactiva debe contener retroalimentación visual y auditiva de alta fidelidad, incluso retroalimentación háptica si es posible. La mayoría de los programas ofrecen retroalimentación sobre si se logró o no la tarea (conocimiento de resultados), pero también es deseable agregar retroalimentación negativa si se realizan movimientos incorrectos o compensatorios (conocimiento del desempeño).</p>

Los aspectos básicos del control motor, incluyendo la planeación de movimientos, redundancia, especificidad de la tarea, solución de problemas y la experiencia, pueden ser aplicados tanto para el entrenamiento en un ambiente virtual como en un ambiente convencional. En la tabla 4.1 se

muestran las propuestas de Levin y colaboradores (Levin et al., 2015) sobre cómo implementar los principios de control motor en ambientes virtuales.

Desde una perspectiva del aprendizaje motor, el objetivo del aprendizaje es promover mecanismos de neuroplasticidad que apoyen la adquisición y retención de nuevas habilidades motoras. Ya se ha mencionado anteriormente que algunos principios que contribuyen al reaprendizaje motor son la repetición intensiva de los ejercicios, el aumento en la dificultad de la tarea, la solución de problemas o la corrección de errores, la motivación, la calidad y la frecuencia de la retroalimentación. Los juegos serios se prestan para poder aplicar estos principios y Levin (Levin et al., 2015) engloba en la tabla 4.2 la forma en la que éstos se pueden abordar desde la perspectiva del diseño del juego.

Tabla 4.2 Propuestas de cómo aplicar los principios del aprendizaje motor en juegos serios. (Levin et al., 2015).

Principios del aprendizaje motor	Aplicación de los principios en juegos serios
El aprendizaje ocurre a través de la repetición de una tarea. Las tareas deben de ser significativas y variadas.	Las tareas a realizar en el juego deben incluir múltiples repeticiones de diferentes movimientos. Idealmente las metas del juego deben ser significativas para el usuario.
El aprendizaje ocurre cuando la dificultad de la tarea aumenta progresivamente de acuerdo con la habilidad del usuario.	El juego debe contener diferentes niveles de dificultad asociados con las habilidades que puede presentar el usuario. Idealmente deben existir reglas de decisión que le ayuden al terapeuta a modificar la dificultad del juego en los momentos adecuados.
El aprendizaje debe incluir problemas a resolver para involucrar mecanismos cognitivos y de funciones ejecutivas.	Las tareas del juego deben variar para entrenar diferentes habilidades cognitivas.
La retroalimentación sensorial es necesaria para el aprendizaje.	La retroalimentación sensorial multimodal debe ser mapeada explícitamente al desempeño del paciente y su progresión a través de los niveles de dificultad.
El aprendizaje ocurre cuando el individuo recibe retroalimentación positiva respecto la calidad del movimiento (conocimiento del desempeño) y a la meta que se intenta lograr (conocimiento de resultados).	Las retroalimentaciones de tipo conocimiento de desempeño y conocimiento de resultados deben ser proporcionadas al paciente durante sus entrenamientos en el juego.
Para evitar el aprendizaje de mal adaptaciones se debe proveer retroalimentación negativa para limitar los movimientos compensatorios durante la ejecución de la tarea, principalmente se debe evitar el desplazamiento excesivo del tronco.	La forma en la que se presente la retroalimentación que le indique al paciente cuando se comentan movimientos mal adaptativos no debe interrumpir el desempeño de la tarea.
El aprendizaje ocurre cuando el sujeto es motivado a mejorar.	Algunos elementos que se deben incluir en el juego para motivar al usuario son mostrar el marcador, mostrar mensajes alentadores y emplear técnicas para aumentar el sentido de presencia en el juego.

El aprendizaje sensomotor se basa en el principio de la minimización del error para alcanzar el objetivo deseado. Se espera que el entrenamiento en ambientes virtuales propicie aprendizaje

motor porque el sujeto debe resolver las diferencias entre las señales sensoriales que recibe, lo que requiere que activamente corrija errores. Por ejemplo, el movimiento que el paciente observa en el campo visual del entorno virtual puede no corresponder con el movimiento que el paciente intenta realizar, lo que produce una diferencia entre las señales de retroalimentación visuales y vestibulares que requieren una corrección del movimiento. De hecho, a través de la rehabilitación virtual se puede magnificar la retroalimentación en donde se desee concentrar la atención del paciente, como los detalles del movimiento que sean más relevantes para la tarea a realizar.

La retroalimentación motivacional, como un simple mensaje que diga “¡Bien hecho!” o un símbolo como una cara feliz, que indique que la tarea fue realizada correctamente anima al jugador a continuar con la actividad. Sin embargo, este tipo de retroalimentaciones no proporcionan información acerca de la calidad con la que se realizó la tarea, por lo que por sí sola no es suficiente para producir un verdadero aprendizaje (Levin et al., 2015).

4.7 Transferencia de habilidades al mundo real

Para determinar si la rehabilitación virtual es una herramienta efectiva para el aprendizaje motor se necesita evidencia de que las habilidades aprendidas en los entornos virtuales pueden ser transferidas al mundo real. Pocos estudios han resaltado aspectos importantes del aprendizaje motor, como la retención de las habilidades a largo plazo, la transferencia de los efectos del entrenamiento y la generalización. Una de las escalas más populares para evaluar la transferencia de habilidades es CAHAI, así como los cuestionarios sobre AVD como el Índice de Barthel y el *Motor Activity Log* (MAL), los cuales presentan la desventaja de ser bastante subjetivos, porque están sujetos a la percepción del paciente sobre sus propias capacidades y esto en ocasiones suele diferir de la realidad. CAHAI es una prueba más confiable, pues el paciente debe realizar las tareas en presencia de un evaluador, pero explora menos AVD que los cuestionarios.

Un estudio (Juliano & Liew, 2020) intentó averiguar el grado de transferencia de las habilidades adquiridas en la rehabilitación virtual que existe cuando se usa tecnología inmersiva (visores de RV) comparados con la no inmersiva (pantalla de la computadora), así como también los factores que pueden influenciar esta transferencia de habilidades. Emplearon 70 participantes divididos en 2 grupos, ambos grupos entrenaron con el mismo juego serio y la interacción con el ambiente era a través de un dispositivo que detectaba la fuerza de presión de la mano. Ambos grupos obtuvieron puntuaciones similares en la prueba *Sequential Visual Isometric Pinch Task* (SVIPT), lo que indicaría que todos los participantes lograron adquirir las mismas habilidades, sin embargo, al final de la intervención se pidió a los participantes que intercambiaran el modo de juego. Se observó que aquellos participantes que se entrenaron sin visor de RV mejoraron su puntuación en el juego al usar el visor, mientras que en el otro grupo su desempeño disminuyó. Esta diferencia entre grupos fue significativa ($p = 0.0004$), lo que indica que la transferencia de habilidades fue mayor en el grupo que entrenó en condiciones no inmersivas. El estudio sugiere que los factores más influyentes en las mejoras individuales y en la transferencia de habilidades motoras son la edad, la familiaridad con videojuegos y visores de RV, el género y la calidad de la interfaz del juego. Admiten que sus resultados no son concluyentes y que aún existen incógnitas sobre

porqué, en términos de la transferencia de habilidades motoras, algunas personas se ven beneficiadas de un mayor sentido de presencia en los juegos y otras no.

De acuerdo con los resultados de diversos estudios sobre rehabilitación virtual para pacientes EVC (Henderson et al., 2007), las habilidades motrices adquiridas a través de la práctica, tanto en juegos serios donde se simula la realidad, como en juegos serios basados en entornos de fantasía, pueden transferirse a funciones significativas del mundo real. Por ello, no es esencial que exista una reproducción exacta de movimientos de la vida real en los ambientes virtuales para que ocurran mejoras en la función motora. Sin embargo, algunos autores como Levin (Levin et al., 2015) recomiendan que las actividades y los entornos virtuales sean similares a entornos reales, donde se tengan que realizar ADV como cocinar, limpiar, socializar, comprar, alcanzar objetos, cruzar calles, etc., para una mejor transferencia de habilidades.

4.8 Flow y ajuste dinámico de dificultad

La teoría de Flow de Csikszentmihályi (Csikszentmihalyi, 1991) habla acerca de cómo alcanzar una experiencia óptima de vida y disfrutar todas las actividades que realizamos. La idea subyacente es que podemos disfrutar el hecho de realizar una actividad a través de un proceso deliberado de volvernos más eficientes en dicha tarea, incluso aunque esta tarea sea tediosa y repetitiva.

La teoría de Flow se relaciona con teorías de aprendizaje y actitudes para superar retos al hacer énfasis en completar las tareas. Si el reto es muy alto la persona puede ponerse ansiosa y desanimarse, por otra parte, si la tarea es muy sencilla, la persona puede aburrirse y perder el interés. Csikszentmihályi denomina como *flow channel* al rango intermedio entre estos dos extremos. Conforme una persona se vuelve competente al realizar una actividad su grado de dificultad debe aumentar. De acuerdo con Koster (Koster, 2005), este ajuste para empatar las habilidades de una persona le proporcionará una frustración placentera, el cual es un estado óptimo para el aprendizaje. En términos psicológicos, el estado de *flow* en los videojuegos es aquel en el que la persona enfoca su energía, logra una inmersión total en el juego y disfruta lo que hace.

Csikszentmihályi usa la frase entropía psíquica para explicar cómo las personas ven que hay más que hacer de lo que uno en realidad puede lograr y sentirse capaz de lograr más de lo que las condiciones permiten. Una entropía psíquica alta es un estado de desorden dentro de la mente consiente, lo que puede llevar a la infelicidad. Además, una alta entropía equivale a altos niveles de incertidumbre. Koster menciona que una de las razones por las que disfrutamos los juegos es porque aprendemos a reducir el desorden de los patrones que el juego nos presenta.

Considerando lo anterior, la entropía psíquica de un paciente EVC es potencialmente alta debido a la presencia de déficits cognitivos y motrices. Por ese motivo la complejidad inicial de los juegos debe ser mucho menor que la que se asignaría para una persona sin déficits, y la dificultad debería incrementar con mayor graduación.

Como se mencionó anteriormente, empatar la dificultad del juego al grado de habilidad del usuario es un factor importante para prevenir la frustración, el aburrimiento y la fatiga. Aunque no es evidente como deben ajustarse exactamente los parámetros de cada repetición de una actividad funcional para promover un óptimo aprendizaje motor, el principio de la redundancia cinemática sugiere que los ejercicios deben ser variados para que el usuario aprenda diferentes movimientos y para mantener su atención. (Levin et al., 2015)

Idealmente, en cualquier tipo de videojuego la dificultad debería comenzar en un nivel bajo para darle oportunidad al jugador de familiarizarse con las mecánicas de juego e incrementar gradualmente conforme las habilidades del jugador mejoran, esto es precisamente lo que buscan los algoritmos de ADD, balancear automáticamente en tiempo real el reto con la habilidad para llegar a una óptima experiencia (Hocine et al., 2015). En los métodos de ajuste de dificultad predeterminada, el mismo jugador o el terapeuta, en el caso de los juegos para rehabilitación, determinan la dificultad del juego o pueden configurar manualmente distintos parámetros que modifiquen la complejidad de la tarea.

El ajuste dinámico de dificultad (ADD) basado en el desempeño del jugador permite que éste se mantenga en un estado de *flow*, al incrementar el grado de reto cuando el juego se vuelve muy sencillo o disminuirlo cuando parece ser muy difícil, ayudando a evitar el aburrimiento y la ansiedad. Mantener este balance y extender el tiempo entre la pérdida o la ganancia de un jugador es muy ventajoso para maximizar el potencial de mejoras funcionales (Barrett et al., 2016). En este mismo sentido, la ley de Yerkes y Dodson (Teigen, 1994) plantea que el mejor modo de potenciar el rendimiento consiste en aumentar la motivación para llevar a cabo las tareas objetivo. El desempeño de una tarea incrementa con la excitación fisiológica o mental (estrés), pero sólo hasta cierto punto. Cuando los niveles de excitación se vuelven demasiado altos, el rendimiento disminuye. Esto significa que la experiencia óptima se puede obtener en cualquier tarea que este perfectamente balanceada, que no sea ni muy fácil ni muy difícil.

El objetivo de adaptar la dificultad en los juegos serios no sólo es mantener la motivación del jugador, si no también asegurar que se logren los objetivos de aprendizaje y entrenamiento. En el contexto de videojuegos serios enfocados a la rehabilitación de miembros superiores, la dificultad del juego yace principalmente en el esfuerzo motor que el paciente tiene que ejercer para alcanzar los objetivos del juego. Para que el paciente recupere sus funciones perdidas este esfuerzo debe ser maximizado (Hocine et al., 2015).

4.9 Plataformas con ajuste dinámico de dificultad

A continuación se presentan algunos ejemplos de plataformas de rehabilitación virtual que implementan diversos mecanismos para ajustar dinámicamente la dificultad del juego de acuerdo al desempeño del paciente.

4.9.1 Sistema de rehabilitación PERCRO-BRANDO

Se trata de un proyecto enfocado en la rehabilitación ortopédica de miembros superiores (Padilla-Castañeda et al., 2018). Una de sus principales características es que implementa un modelo de regresión que permite predecir el marcador que obtendrá el paciente en el juego, esto ofrece muchas posibilidades, pues el terapeuta podría configurar el sistema para que asigne automáticamente los ejercicios al paciente indicándole el marcador que se desea obtener, lo que estará directamente relacionado con la dificultad de los ejercicios que se asignarán. Sin embargo, en este estudio el modelo se emplea como una ayuda para el terapeuta al momento de definir manualmente los parámetros del juego. Este sistema consta de 4 partes:

- 1) El robot BRANDO. Se trata de un brazo robótico de efector final de 6 grados de libertad, 3 de los 6 grados permite la movilización de las articulaciones de codo y hombro. El resto de los grados de libertad permiten un movimiento pasivo de la muñeca. El robot provee una compensación gravitatoria completa para ayudar al paciente a realizar movimientos sin cargar el peso de su brazo. Se cuenta con dos modalidades hápticas, movimientos libres con la muñeca proporcionando únicamente soporte del brazo, y la aplicación de una fuerza variable para movimientos de codo. En la figura 4.8 se ilustra el robot completo.

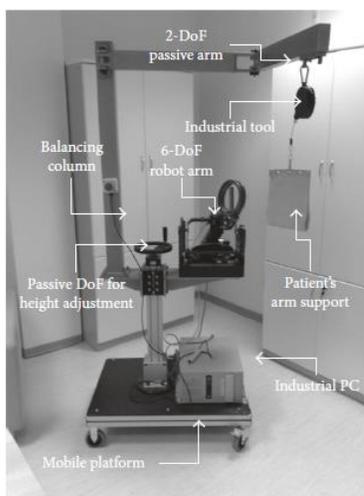


Figura 4.8. Sistema robótico de rehabilitación BRANDO, que cuenta con 6 grados de libertad, el dispositivo está montado sobre una plataforma móvil, con una manija pasiva adjunta al efector final.(Padilla-Castañeda et al., 2018).

- 2) Un conjunto de 3 videojuegos creados explícitamente para ejercitar la flexión y extensión (FE) del codo, y la pronación y supinación (PS) de la muñeca. En un juego se trabaja la FE del codo de arriba hacia abajo con el brazo en abducción, en otro escenario virtual se trabaja la FE del codo de tipo lateral/medial con el brazo en abducción, y en el último juego se ejercita la muñeca con el brazo en aducción. Durante las partidas el sistema monitorea el rango de movimiento, la velocidad angular de las articulaciones y las intensidades de las fuerzas hápticas toleradas, así como el marcador, duración de la partida y nivel de dificultad alcanzado.

- 3) Una interfaz gráfica (figura 4.9) donde el terapeuta puede evaluar los movimientos que realizó el paciente durante las partidas, modificar la dificultad de los juegos, y monitorear el progreso durante las sesiones.
- 4) Un módulo de adaptación de la dificultad, que ayudará a los terapeutas a ajustar la demanda física de los ejercicios basándose en la información sobre la cinemática del participante. Este módulo estimará el marcador que obtendrá el paciente, dada la configuración de los parámetros del juego que elija el terapeuta, a través de un modelo de regresión, previamente calibrado experimentalmente.

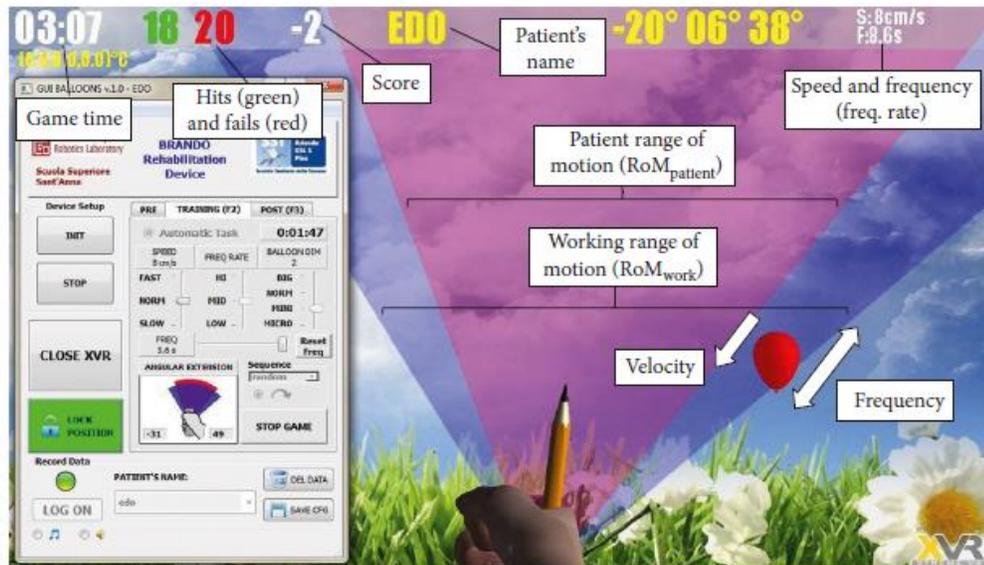


Figura 4.9. Interfaz que se le presenta al participante durante el juego y la forma en la que se representan en el juego los parámetros que determinan el nivel de dificultad. (Padilla-Castañeda et al., 2018).

Con el fin de ayudar al terapeuta a configurar de forma manual los parámetros del juego de forma personalizada, incrementando sistemáticamente el nivel de demanda del entrenamiento, se implementó un modelo de desempeño, para ello se aplicó un análisis de componentes principales (PCA) a la información kinesiológica observada de los pacientes. El modelo resultó en una combinación lineal del rango de movimiento, velocidad y la fuerza resistiva tolerada para los ejercicios de FE del codo, mientras que para los ejercicios de PS la combinación lineal estaba conformada por los parámetros de rango de movimiento y velocidad. Después, para proporcionar una estimación a priori del desempeño que tendría el paciente, se desarrollaron modelos predictivos de regresión de los resultados normalizados del rango de movimiento y velocidad para FE y PS, y para la fuerza tolerada en FE de la siguiente forma:

$$Resultado = b_0 + \sum_{i=0}^p b_i x_i + \sum_{j \neq k}^q b_{jk} x_j x_k \quad (4.1)$$

Donde x_i es el i -ésimo parámetro de entrada entre los p variables más significativas para el resultado correspondiente, x_j y x_k corresponden a los factores significativos que interactúan

entre sí, y b_n son valores constantes. La optimización de los modelos de regresión se realizó con los datos de entrenamiento de 7 pacientes para PS y 6 pacientes para FE. En la tabla 4.3 se muestran los valores de los componentes principales para predecir el desempeño en las tareas de FE y PS.

Outcome	PC ₁	PC ₂	Performance equation
<i>Flexion/Extension</i>			
RoM _{performed}	0.285	0.4424	Performance _{FE} = 0.6238PC ₁ + 0.2957PC ₂
Velocity _{performed}	0.618	0.5929	
Force _{performed}	0.7323	0.6729	
<i>Pronation/supination</i>			
RoM _{performed}	0.4899	-0.8718	Performance _{PS} = 0.6844PC ₁ + 0.3156PC ₂
Velocity _{performed}	0.8718	0.4899	

Tabla 4.3. Resultados del PCA sobre los datos kinesiológicos observados que explican el primer y el segundo componente de las ecuaciones de predicción del desempeño, los cuales son una combinación lineal del rango de movimiento, la velocidad y, en el caso de los ejercicios tipo FE, de la fuerza ejercida por el brazo robótico. (Padilla-Castañeda et al., 2018).

Consecuentemente, se obtuvieron los modelos de regresión, los cuales se muestran en las ecuaciones de la tabla 4.4, como una función de los parámetros de dificultad estimados sobre los datos de entrenamiento correspondientes.

Movement	Regression predictive model	No significant effect for the model ($p > 0.05$)
FE	$RoM^*_{performed} = b_0 + b_1 * RoM_{work}$	$t_{task} * RoM_{work}$
FE	$Velocity^*_{performed} = b_0 + b_1 * t_{task} + b_2 * RoM_{work}$	—
FE	$Force^*_{performed} = b_0 + b_1 * Force_{work} + b_2 * t_{task} * RoM_{work}$	t_{task}
PS	$RoM^*_{performed} = b_0 + b_1 * RoM_{work}$	—
PS	$Velocity^*_{performed} = b_0 + b_1 * sequence + b_2 * balls_{speed} + b_3 * sequence * balls_{speed} + b_4 * RoM_{work} * balls_{frequency}$	$balls_{frequency}$, $sequence * balls_{number}$

Tabla 4.4. Modelos de regresión para predecir la información kinesiológica del rango de movilidad logrado, la velocidad de los movimientos y las fuerzas opuestas ejercidas, como función de los parámetros de dificultad. (Padilla-Castañeda et al., 2018).

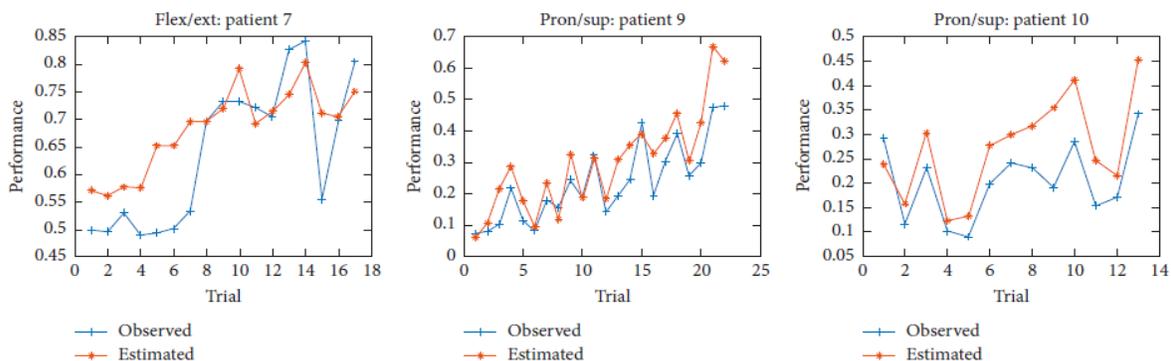


Figura 4.10. Comparación del desempeño entre el valor predicho del marcador y su valor observado. Cada uno de los 3 pacientes realizó rondas de 10 ejercicios. (Padilla-Castañeda et al., 2018).

Los modelos de regresión presentaron una correlación significativa de $r=0.837$ (coeficiente de Pearson) al comparar el desempeño predicho con el observado sobre los datos de un paciente para ejercicios tipo FE y una $r=0.917$ sobre los datos de dos pacientes que realizaron ejercicios de

PS. Esta correlación indica un buen ajuste de los datos entre el desempeño estimado durante la configuración de los parámetros juego y el desempeño real que tuvo el paciente al jugar bajo dichos parámetros, como se muestra en la figura 4.10.

4.9.2 Rehabilitation Gaming System (RGS)

El objetivo de esta plataforma basada en RV para neuro-rehabilitación es el tratamiento de déficits motores como resultado de una lesión al sistema nervioso central, que explote los procesos cognitivos que intervienen entre la percepción y la acción (Cameirão et al., 2010). Fue desarrollada con el motor *Torque Game Engine* (TGE), utiliza una cámara CCD KE-240CV colocada encima del monitor, junto con cuatro parches de diferentes colores, dos sobre las muñecas y dos sobre los antebrazos para rastrear los movimientos de miembros superiores a través del sistema de captura AnTS, así como dos guantes de datos de la marca 5DT (*Fifth Dimension Technologies*) para rastrear el movimiento de los dedos.

El RGS mapea los movimientos de brazos y dedos a un ambiente virtual, de esta forma el usuario controla los movimientos de dos miembros virtuales que se observan desde una perspectiva en primera persona. La mecánica de juego consiste en interceptar y capturar esferas que se dirigen hacia el usuario. La hipótesis de este planteamiento se basa en que la ejecución de actividades bimanuales, en combinación con el hecho de observar que los miembros virtuales imiten los movimientos realizados (o la intención del movimiento), crea condiciones que facilitan la reorganización funcional de los sistemas motor y pre-motor afectados por el EVC.

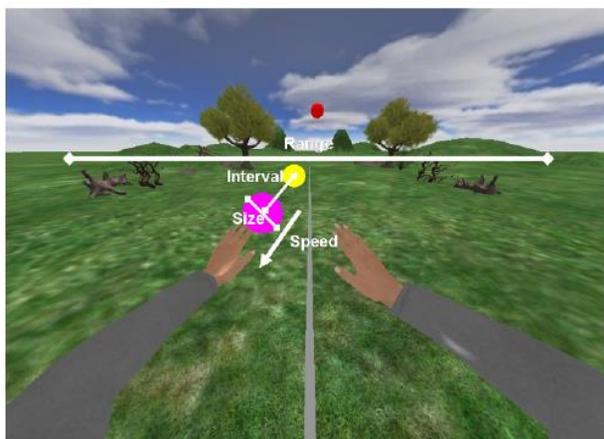


Figura 4.11. Se muestra el juego Spheroids y el ambiente virtual que observa el usuario al jugar, así como los parámetros configurables del juego y cómo se observa su impacto. (Cameirão et al., 2010).

El diseño del juego busca involucrar el sistema de neuronas espejo, pues el que el equipo de desarrollo del RGS sostiene que, comparado con una vista en tercera persona, una perspectiva en primera persona promovería de forma mucho más significativa la activación de las neuronas espejo a través de estímulos multi-modales, ya que esta es la perspectiva a la que el sistema de neuronas espejo está más frecuentemente expuesto.

Durante el juego bautizado como *Spheroids* (figura 4.11) aparecerán esferas en la línea de horizonte que avanzarán hacia el jugador, quien tendrá que interceptar tantas como le sea posible. Para el ADD, después de cada ronda de ejercicios se modifican los parámetros de velocidad de movimiento de las esferas, el intervalo de tiempo en el que aparecen y el rango de dispersión horizontal de las esferas en el campo de visión.

Antes de cada sesión de juego se realiza una calibración. En una mesa se colocan 4 marcas a donde el usuario debe mover sus manos. Durante esta tarea se extrae información para cada brazo acerca de la velocidad de movimiento, el rango de movilidad y el tiempo que tarda en iniciar cada movimiento. Esta información se usa para calcular la línea base de los parámetros de *Spheroids* y de esta manera se asigna una dificultad inicial para la sesión de entrenamiento con el RGS.

El RGS emplea un Módulo de Entrenamiento Personalizado (PTM) que de forma automática ajusta la dificultad de las sesiones de juego con base en la información histórica recabada. Su funcionamiento se muestra en la figura 4.12 y es el siguiente: Antes de que el usuario comience a jugar se realiza una calibración que definirá la dificultad con la que comenzará el juego. Después de cada bloque de 10 ejercicios (después de que han aparecido 10 esferas), el PTM ajustará la dificultad del juego basándose en el desempeño del usuario. Para cada nuevo valor de dificultad los parámetros correspondientes se calculan tomando en consideración la respuesta previa del usuario a los parámetros individuales y el modelo psicométrico de *Spheroids*. La dificultad aumenta en un 10% cuando el usuario alcanza más del 70% de las esferas, hasta llegar a un nivel de dificultad del 100%. La dificultad disminuye en un 5% si el usuario intercepta menos del 50% de las esferas. Es así que ocurre una continua adaptación de los parámetros del juego al desempeño del usuario. Cabe resaltar que la dificultad es asignada individualmente para cada brazo.

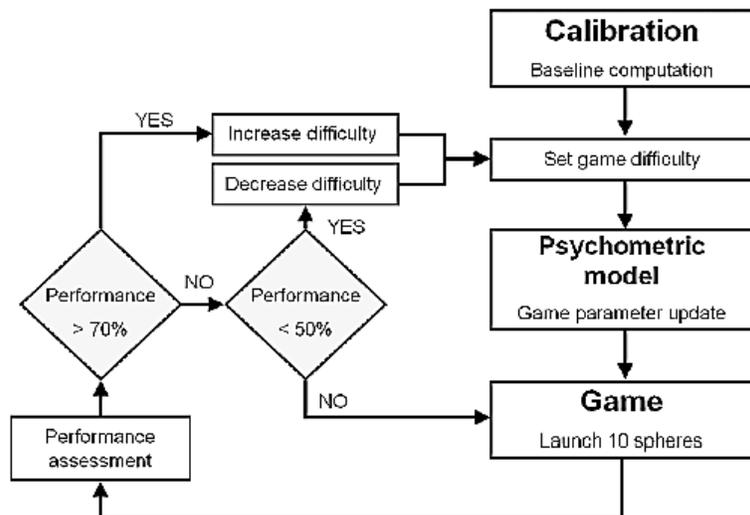


Figura 4.12. Diagrama de flujo que muestra la forma en que el PTM actualiza automáticamente los parámetros del juego con base en el desempeño del participante.(M. S. Cameirão et al., 2010).

En el contexto del PTM, el desempeño del jugador se mide con la siguiente función:

$$\text{Desempeño} = f(\text{Velocidad}, \text{Intervalo}, \text{Rango}, \text{Tamaño}) \quad (4.2)$$

El equipo de desarrollo investigó el efecto de cada uno de estos parámetros en el desempeño de forma individual y establecieron una relación cuantitativa entre múltiples variables independientes (parámetros) y una sola variable dependiente (dificultad). Emplearon un modelo cuadrático con términos de primer orden, interacciones (productos entre términos) y términos de segundo orden. Para tres variables de entrada (x_1, x_2, x_3) el modelo sería el siguiente:

$$y = m_0 + m_1x_1 + m_2x_2 + m_3x_3 + m_{12}x_1x_2 + m_{13}x_1x_3 + m_{23}x_2x_3 + m_{11}x_1^2 + m_{22}x_2^2 + m_{33}x_3^2 \quad (4.3)$$

Al ajustar el modelo a los datos de interés se extraen los parámetros de regresión (coeficientes m), que mejor describen la contribución de cada variable independiente a la variable dependiente.

Para encontrar los parámetros (coeficientes m) más relevantes del modelo psicométrico que permitieran el ADD se reclutaron 12 pacientes EVC y 10 sujetos control, a cada uno de los participantes se les asignó una combinación con parámetros fijos de juego; cada parámetro (velocidad, intervalo, rango y tamaño) podía tener 1 de 4 posibles valores, lo que significa que se podían asignar hasta 256 combinaciones diferentes. La combinación asignada cambiaba después de cada ronda de 10 esferas, cada sesión de la intervención duraba 20 minutos, en ese tiempo se asignó un promedio de 82 combinaciones a cada participante, aunque en ocasiones se repetían. Los datos de los sujetos control se usaron para cuantificar la relación entre el desempeño y los parámetros del juego. El modelo se ajustó a los datos de desempeño de los pacientes mediante un método de regresión, y una vez que se obtuvo el modelo final se volvió a evaluar con un nuevo grupo de pacientes ($n=9$) y controles ($n=10$), nuevamente cada participante realizó una sesión de 20 minutos con el PTM ya integrado.

En el estudio se encontró que el tamaño de las esferas tenía un bajo efecto en el desempeño del jugador, mientras que el intervalo, rango y velocidad modulaban significativamente el desempeño, con los datos obtenidos en sus experimentos el modelo cuadrático que obtuvieron fue el siguiente:

$$\begin{aligned} \text{Dificultad} = m_0 + m_1\text{Intervalo} + m_2\text{Velocidad} + m_3\text{Rango} + m_4\text{Intervalo} \cdot \text{Velocidad} \\ + m_5\text{Intervalo} \cdot \text{Rango} + m_6\text{Velocidad} \cdot \text{Rango} + m_7\text{Intervalo}^2 \\ + m_8\text{Velocidad}^2 + m_9\text{Rango}^2 \end{aligned} \quad (4.4)$$

Donde la dificultad es inversamente proporcional al marcador del juego. En este modelo valores positivos de dificultad corresponde a un desempeño superior al promedio, mientras que una dificultad negativa corresponde a un desempeño inferior al promedio.

Como se muestra en la figura 4.13, tras evaluar su modelo con nuevos grupos, encontraron que a) la proporción en la dificultad entre el brazo parético y no parético para pacientes, y entre el brazo dominante y el no dominante para los sujetos control fue significativa ($p < 0.05$) y b), tanto para los sujetos control, como para los pacientes, el sistema se ajustó para lograr un buen desempeño (valor del marcador mayor al 80%) con ambos grupos.

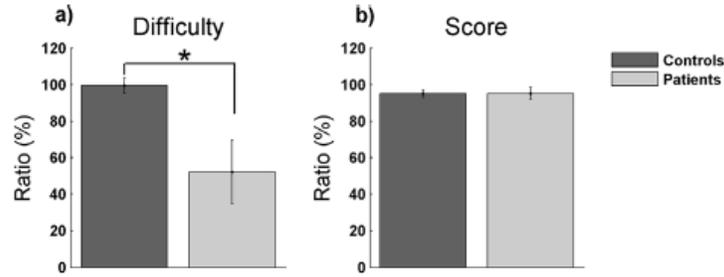


Figura 4.13. Resultados obtenidos en la dificultad y el score con los diferentes grupos empleando el modelo psicométrico final. (M. S. Cameirão et al., 2010).

4.9.3 Videojuego de rehabilitación con ajuste dinámico de dificultad con MPRL

Dentro de las numerosas plataformas de aplicaciones interactivas para rehabilitación que se han desarrollado, el único trabajo publicado donde se emplean métodos de aprendizaje por reforzamiento para el ADD es el propuesto por Yoones (Sekhvat, 2017), que desarrolla un algoritmo apoyándose en la teoría del aprendizaje por reforzamiento de múltiples periodos (MPRL). El videojuego para rehabilitación se enfoca en la recuperación de la función motora de los miembros superiores, utiliza el sensor Kinect para detectar los movimientos de los brazos y mapearlos a brazos virtuales, donde el jugador controla los brazos de un personaje en una vista en tercera persona. La mecánica de juego consiste en pegarle a pelotas que se sitúan en un arco lleno de pelotas. El jugador debe alzar su mano para golpear la pelota más brillante de todas las pelotas en el arco para acumular un punto en el marcador final. En este escenario, la ejecución de una acción orientada a una tarea se combina con la observación de una mano virtual para crear las condiciones que resulten en una reorganización funcional del paciente.

Para este juego los parámetros que se pueden manipular para modificar la dificultad son la velocidad del personaje, el tamaño de las pelotas y la distancia entre los arcos de pelotas. Consecuentemente, el desempeño del usuario será una función de estos parámetros.

En el juego, los posibles resultados que se pueden obtener cada vez que se presenta un arco de pelotas son que el jugador gane y le pegue a la pelota correcta, que el jugador pierda y le pegue a una pelota incorrecta, o que el jugador no le pegue a ninguna pelota, que también se contabiliza como una pérdida. En este trabajo se propone que el margen en cada partida entre éxitos y fracasos debe ser pequeño. Con la hipótesis de que, dados W y L , como el número total de éxitos y fracasos respectivamente. Es importante mencionar que en este estudio consideraron que un jugador se considerará satisfecho cuando el marcador en una partida resulte $|W - L| = 0$. Más específicamente, se definieron tres metas para medir la satisfacción del jugador. La primera (g_1) es minimizar $|W - L|$ para que el número de éxitos sea igual al número de fracasos. Dados los marcadores r_i y r_{i+1} que obtuvo el jugador en dos rondas consecutivas i e $i+1$, para evitar un gran progreso o retroceso después de cada ronda se debe minimizar $|r_i - r_{i+1}|$. Cuando se observan pequeñas diferencias entre los marcadores de rondas consecutivas significaría que el jugador está inmerso en el juego y trata de obtener un mejor marcador al jugar más juegos (g_2). Finalmente, el jugador debe percibir un progreso en el juego y conseguir un marcador ligeramente mejor que en

la ronda anterior, para ello se debe maximizar el número de progresos en n rondas consecutivas (g_3). Cada una de las rondas de juego puede considerarse como un episodio, un concepto del aprendizaje por reforzamiento que define un periodo de tiempo con un inicio y un fin en el que ocurre una interacción entre un agente y el ambiente, para este planteamiento cada una de las metas se evalúan en distintos episodios. La meta g_1 se evalúa con base en la información de cada episodio, g_2 se evaluará cada 2 episodios y g_3 sólo se podrán evaluar hasta tener la información de n episodios. Por esta razón en este trabajo se usa un algoritmo de múltiples periodos y así evaluar las funciones de objetivos diferentes en periodos separados. Cuando llegue el fin de un episodio en el que se deba evaluar más de una meta se realizará una fusión donde se tomarán en cuenta todas las metas asignándole a cada una un componente de severidad λ_k .

Cada ronda de juego, o episodio, está conformada por 10 ejercicios. En particular, como $W + L = 10$, el resultado que se puede esperar en cada episodio, que en términos del algoritmo MPRL nos indicará un estado del agente para g_1 , es un valor en $S_1 = \{-10, -8, \dots, 0, \dots, 8, 10\}$. El marcador que se puede obtener en cada partida es un número entre 0 y 10, por lo tanto, la diferencia entre dos episodios consecutivos será un valor en $S_2 = \{-10, -9, \dots, 9, 10\}$, los cuáles serán los posibles estados en los que se podrá encontrar el agente para g_2 . Para indicar los estados correspondientes a g_3 se considera la historia de 6 episodios consecutivos. Dada la función $\sum_{t=0}^n (r_{t+1} - r_t > 0)$ los posibles estados serán $S_3 = \{5, 4, 3, 2, 1, 0\}$. De esta forma el estado del agente, quien será el encargado de modificar la dificultad del juego, estará definido por la tripleta $(s_1, s_2, s_3) s_t \in S_t$.

Una de las propuestas que se realizan en este trabajo es que, a diferencia de otros métodos MPRL tradicionales, el agente ejecuta acciones con base en la probabilidad de que dicha acción ocurra. En su algoritmo manejan 2 tipos de acciones, las acciones indirectas, que incrementan o disminuyen la probabilidad de que se ejecuten las acciones directas, y las acciones directas, que son las que se ejecutan sobre el ambiente y se refieren al incremento o disminución de los parámetros del juego, que son la velocidad, el tamaño y la distancia de las pelotas.

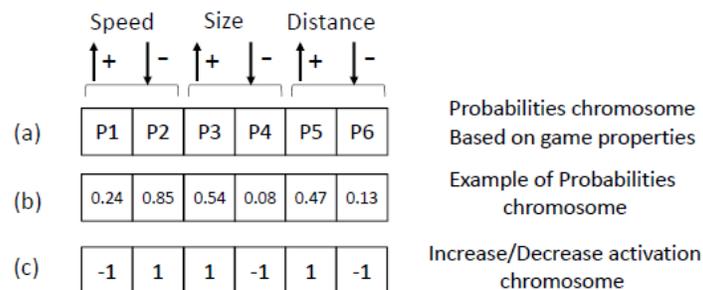


Figura 4.14. (a) Controlador de un cromosoma para indicar las probabilidades de activar o desactivar las propiedades del juego. (b), un ejemplo del cromosoma almacenando las probabilidades de cada comportamiento (c), la activación o desactivación de los comportamientos con base en las probabilidades del cromosoma. (Sekhavat, 2017).

Se emplea un controlador adaptivo de un cromosoma para indicar cómo se pueden modificar los parámetros del juego, este cromosoma es en realidad un arreglo de 6 números reales. Cada número real representa la probabilidad de activar un comportamiento en el juego. Las primeras 2

posiciones del arreglo corresponden a la velocidad (P1 incrementa la velocidad, P2 la disminuye), las siguientes 2 posiciones corresponden al tamaño (P3 incrementa el tamaño, P4 lo disminuye), y las últimas 2 posiciones corresponden a la distancia (P5 incrementa la distancia y P6 la disminuye). Como se ilustra en la figura 4.14, en cada ronda, cualquier incremento o disminución de algún parámetro se ejecutará en tiempo real con base en su probabilidad de activación, un valor de 1 activa el comportamiento y un valor de -1 lo desactiva.

Al inicio del juego los valores del cromosoma se inicializan aleatoriamente. Al final de cada ronda se actualizan los valores del cromosoma con base en el siguiente algoritmo:

Algoritmo MPRL

Sea C_i el i -ésimo lugar en el controlador de un cromosoma.

Sea $G = \{g_1, \dots, g_k\}$ el conjunto de k metas

g_i está asociado con el conjunto de estados $S_i = \{s_i^1, \dots, s_i^n\}$

Sea P_i la representación del periodo en el que la meta g_i se evalúa

Sea t el contador de rounds

Inicia el cromosoma C arbitrariamente

Itera hasta finalizar el juego:

Sea $G = null$ (metas que se evalúan en la ronda actual)

$rw = 0$ (representa la recompensa acumulada)

Desde $i = 1$ hasta k realizar:

Si $t \bmod p_i == 0$ entonces:

$G = G \cup g_i$

Para cada g_i en G realizar:

Si $g_i == g_1$ entonces:

$rw_1 = s_1 + (|t.score| - 5)$

$rw = rw + g_1 * \lambda_i * rw_1$

Si $g_i == g_2$ entonces

$rw_2 = s_2 + (|t.score| - |(t-1).score|)$

$rw = rw + g_2 * \lambda_i * rw_2$

Si $g_i == g_3$ entonces

$rw_3 = s_3 + \sum_{i=t}^{t-5} (r_{i+1} - r_i > 0)$

$rw = rw + g_3 * \lambda_i * rw_3$

$rnd = \text{Random}(0,1)$

Si ($rnd \leq 0.3$) entonces

Para $p = 1$ hasta 6 realizar:

Si p es par entonces:

$C_p = (1 - \alpha) * C_p + \alpha * (sign(p) + rw)$

Si no:

$C_p = (1 - \alpha) * C_p - \alpha * (sign(p) + rw)$

Para $p = 1$ hasta 6 realizar:

Con la probabilidad de C_i en C , ejecutar la acción directa correspondiente

Actualizar los estados (s_1, s_2, s_3)

En el algoritmo, $sign(p)$ denota el estado del comportamiento del cromosoma antes de una actualización. Si el comportamiento ya estaba activado entonces $sign(p)=1$, en caso contrario $sign(p)=-1$. La activación de un parámetro representa un aumento del 10% de dicho parámetro, así mismo, la desactivación de un parámetro indica una reducción del 10%.

En MPRL en vez de actualizar los valores de una tabla Q y seleccionar acciones con base en estos valores, se utiliza un enfoque probabilístico en donde todas las acciones tienen la oportunidad de ser ejecutadas en diferentes periodos. De acuerdo con las probabilidades de cada acción podrán o no ejecutarse. En este algoritmo la tasa de aprendizaje α indica hasta qué punto el algoritmo de explorar nuevas condiciones o explotar los aprendizajes previos. Cada meta tiene su propio impacto en la recompensa agregada rw . El parámetro λ_i asociado a la meta g_i determina la severidad y la prioridad de la meta, es decir, la proporción con la que afectará a rw .

El algoritmo ya implementado en el videojuego se probó con 8 pacientes subagudos para medir su efecto en la movilidad del brazo. Los participantes realizaron sesiones de 20 minutos de juego, 1 vez al día, durante 2 semanas en el laboratorio de Realidad Aumentada Cognitiva en Irán. De acuerdo con la meta g_1 , minimizar la diferencia entre éxitos y fracasos con $|W - L|$ resulta en una satisfacción y entretenimiento del jugador. En este caso, realizó un histograma de $|W - L|$, el cual se muestra en la figura 4.15, obteniendo valores muy bajos en un número considerable de casos, lo que, de acuerdo con su hipótesis, indica que el algoritmo fue capaz de ajustarse a las capacidades de los pacientes en la mayoría de los casos. El histograma es casi simétrico, y los valores promedio son muy cercanos a cero. En el trabajo de investigación se reporta que en más del 70% de los casos se obtuvo una diferencia en el marcador de 4 o menor, lo que implica que los jugadores se mantuvieron entretenidos en al menos el 70% de los casos.

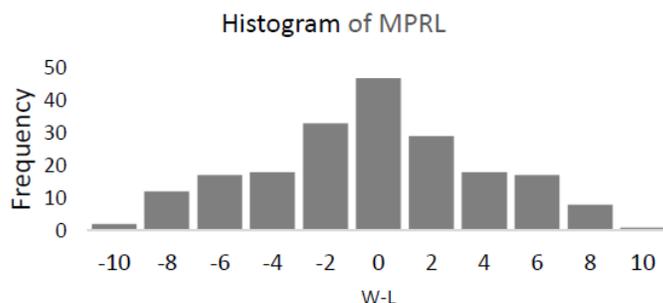


Figura 4.15. El histograma de las diferencias en el marcador $|W - L|$ utilizando MPRL de todas las partidas jugadas por 8 pacientes subagudos durante 2 semanas. (Sekhavat, 2017).

Capítulo 5

Aprendizaje por Reforzamiento

Una de las formas más naturales en las que los seres vivos aprenden es a través de la interacción con el ambiente que los rodea. Cuando nosotros realizamos una acción, sin que nadie explícitamente nos indique lo que debemos hacer, establecemos una conexión de causa y efecto, la información obtenida de las consecuencias de las acciones que realizamos nos ayudará en el futuro para saber qué hacer cuando queramos lograr algo. Aprender a través de la interacción es la idea fundamental que subyace todas las teorías de inteligencia y aprendizaje.

El aprendizaje por reforzamiento es aprender qué se debe realizar, mediante una relación entre situaciones y acciones, para maximizar una señal de recompensa (Sutton & Barto, 2005). Quien aprende debe descubrir por sí mismo las acciones que otorgan mayores recompensas mediante la ejecución de las acciones. En los escenarios más desafiantes y complejos, las acciones ejercidas no sólo determinarán las recompensas que se obtendrán de forma inmediata, sino también la situación que se presentará a continuación, dando pie a una nueva toma de decisiones y, por consiguiente, afectando a todas las recompensas subsecuentes. Ambos componentes, el ensayo y error de efectuar las acciones, y el efecto a largo plazo que puede existir en las recompensas, son las características más importantes en el aprendizaje por reforzamiento.

El aprendizaje por reforzamiento o *reinforcement learning* (RL), al igual que otras ramas del aprendizaje automático o *machine learning* (ML), es un concepto que engloba tanto un problema, como un conjunto de métodos que solucionan el problema, y también incluye al campo que estudia el problema y sus métodos para solucionarlos.

RL es diferente del aprendizaje supervisado, una de las ramas más populares de ML, porque en esta última se aprende de un conjunto de ejemplos ya etiquetados que provee un supervisor externo, idealmente un experto. Cada ejemplo es una situación que podría suceder junto con la acción que el sistema debería tomar en esa situación; el objetivo es que el sistema pueda extrapolar o generalizar sus respuestas ante situaciones que no se presentaron en el conjunto de entrenamiento. Cuando se presentan problemas en los que el sistema debe interactuar, es a menudo poco práctico generar un conjunto de entrenamiento donde se presenten tanto las situaciones más representativas con las que el sistema podría encontrarse, como las acciones a realizar en cada caso. En este tipo de problemas, donde realmente se desconoce el tipo de situaciones que podrían presentarse, lo mejor es que el sistema aprenda con base en su experiencia.

RL también es diferente del aprendizaje no supervisado, donde el objetivo es encontrar una estructura escondida en un conjunto de datos no etiquetados. Se podría pensar que RL es un caso

especial de aprendizaje no supervisado, porque ninguno de los dos trabaja con ejemplos que indiquen las acciones correctas a tomar dada una situación, la diferencia es que RL busca maximizar una señal de recompensa y el aprendizaje no supervisado busca encontrar una estructura oculta. Existe además otra rama conocida como aprendizaje semi-supervisado, que usualmente utiliza dos conjuntos de datos para entrenar, una pequeña cantidad de datos etiquetados y una gran cantidad de datos no etiquetados. Sin embargo, esto también es muy diferente de RL, que se basa esencialmente en la interacción entre el sistema con el ambiente.

Por estas razones, RL es considerado un paradigma en sí mismo en el campo de ML.

Los pasos que seguiría un algoritmo típico de RL son los siguientes (Sudharsan, 2018) :

1. El agente interactúa con el ambiente al realizar una acción.
2. Cuando el agente realiza una acción se mueve de un estado, el estado actual en el que tomó la decisión, a otro.
3. El agente recibe una recompensa dependiendo de la acción que haya realizado.
4. Con base en la recompensa, el agente ponderará si la acción que tomó fue buena o mala.
5. La acción habrá sido buena si la recompensa que se obtuvo fue positiva, cuando el agente vuelva a encontrarse en la misma situación preferirá realizar esa misma acción sobre cualquier otra. En caso de que el agente haya recibido una recompensa negativa, cuando el agente se vuelva a encontrar en esa situación elegirá alguna otra acción tratando de obtener una recompensa positiva. Es por eso que este proceso de aprendizaje es de ensayo y error.

5.1 Elementos del aprendizaje por reforzamiento

A continuación se describen los diferentes elementos que conforman a este paradigma.

Agente: En términos de RL es quien tomará las decisiones de las acciones a tomar cuando interactúe con el ambiente; como consecuencia de sus acciones recibirá una recompensa y, con base en la experiencia, aprenderá a tomar mejores decisiones.

Ambiente: Todo aquello con lo que el agente interactúe es el ambiente. El ambiente condensa todo aquello que es externo al agente, sin embargo, el límite entre ambiente y agente no necesariamente es igual que el límite físico que existe con el cuerpo del agente. Por regla general, cualquier cosa que el agente no pueda modificar de forma arbitraria es parte del ambiente. Esto es independiente del conocimiento que tenga el agente acerca del ambiente, el agente podría conocer todo acerca del ambiente, o desconocer muchas cosas. Existen diferentes tipos de ambientes no excluyentes entre ellos (Sudharsan, 2018) :

- a) **Ambiente determinista:** Un ambiente es determinista cuando se conoce el resultado con base en el estado actual. Un ejemplo de este tipo de ambiente es una partida de ajedrez, pues conocemos el resultado exacto de mover cualquier pieza.

- b) **Ambiente estocástico:** Un ambiente es estocástico cuando no se puede determinar el resultado con base en el estado actual. Un ejemplo de este ambiente es lanzar un dado, nunca se sabe qué número saldrá, aunque se conozca el número que resultó en el lanzamiento anterior. Tomar la misma acción en el mismo estado puede dar resultados diferentes.
- c) **Ambiente completamente observable:** Cuando el agente puede determinar el estado del ambiente en todo momento es un ambiente completamente observable. Nuevamente un ejemplo de este tipo de ambiente sería el ajedrez, en este caso cualquier estado del ambiente estaría determinado por la posición de cada una de las piezas en el tablero, esta información está disponible en todo momento y puede ser usada por el agente para tomar una decisión.
- d) **Ambiente parcialmente observable:** En este tipo de ambiente el agente no puede determinar su estado en todo momento. Un ejemplo sería un juego de póker, ya que el agente no tiene información acerca de las cartas de los otros jugadores.
- e) **Ambiente discreto:** Cuando, para todos los posibles estados del ambiente, sólo existe un número finito de acciones disponibles para cambiar de un estado a otro. Por ejemplo, en el videojuego Pac-man, el jugador sólo tiene 4 acciones disponibles, moverse hacia arriba, hacia abajo, hacia la izquierda o a la derecha, sin importar el estado en el que se encuentre.
- f) **Ambiente continuo:** Cuando el conjunto de acciones a tomar en un estado es infinito. El problema de un auto que se maneja automáticamente es un ejemplo de un ambiente continuo, en todo momento el agente debe controlar el ángulo de rotación del volante y la velocidad del auto, los valores de ambas acciones pertenecen a un espacio continuo.
- g) **Ambiente episódico:** En este tipo de ambiente las acciones del agente quedan confinadas al episodio actual y no se verán afectadas por acciones de episodios anteriores ni tampoco afectarán las acciones de futuros episodios. Por ejemplo, en un juego donde en cada episodio el agente debe adivinar dónde se encuentra el premio detrás de alguna de 3 puertas, la acción que el agente tome en cada episodio será independiente en cada episodio.
- h) **Ambiente secuencial:** Contrario al ambiente episódico, en este caso las acciones del agente se verán afectadas por las acciones que haya tomado en el pasado, así como también las acciones que tome en el estado actual afectarán las acciones en el futuro. Por ejemplo, en una partida de damas chinas, la acción que el agente tome en cualquier momento de la partida determinará cuál será el siguiente estado, y la acción que deberá tomar en ese momento dependerá del historial de acciones pasadas.

- i) **Ambientes de un solo agente:** En este ambiente un único agente interactúa con el ambiente, por ejemplo, un ambiente donde un agente conduzca un carro de un punto B a un punto C.
- j) **Ambientes multi-agente:** Cuando existe más de un agente interactuando con el ambiente. Por ejemplo, en un ambiente donde existan múltiples autos y cada uno de ellos sea manejado por su propio agente.

Política: Dicta la manera en la que el agente debe comportarse en una situación determinada. A grandes rasgos, la política es una función que mapea los estados percibidos del ambiente a las acciones que deben de ser tomadas cuando el agente se encuentre en esos estados. La política puede ser desde algo muy sencillo como una tabla, hasta una función que involucre un complejo proceso de búsqueda. La política es el núcleo de un agente, en el sentido de que por sí sola es suficiente para determinar su comportamiento.

Señal de recompensa: También se suele referir a ella como función de recompensa. Esta función define cuál es la meta del problema de RL. Dependiendo de la acción que realice el agente el ambiente le retorna un único número llamado recompensa. El objetivo del agente es maximizar la recompensa total que recibirá a largo plazo. La señal de recompensa definirá entonces cuando la acción que ha realizado el agente en ese momento ha sido acertada o no. El hecho de usar una señal de recompensa para formalizar la idea de una meta es una de las características más distintivas del aprendizaje por reforzamiento.

Función de valor: Mientras que la función de recompensa determina si la acción que tomó el agente fue buena de manera inmediata, la función de valor determina si la acción tomada fue acertada a largo plazo. De forma general, la función de valor de un estado es la cantidad total de recompensa que el agente puede esperar acumular en el futuro, empezando por el estado actual. De esta forma, esta función puede determinar que la función de valor de un estado es muy grande, a pesar de que la recompensa inmediata del estado es muy baja, sin embargo, los estados a los que se mueve el agente otorgan recompensas muy grandes. El caso contrario también puede cumplirse.

El proceso de elegir la mejor acción a tomar se basa en valores. Sin las recompensas no podría haber valores, y el único propósito de estimar los valores es obtener mayores recompensas. Podría decirse entonces que el agente busca tomar acciones que conduzcan a estados de mucho valor, en lugar de buscar estados que otorguen las mayores recompensas, pues estas acciones nos permitirán obtener la mayor cantidad de recompensa a largo plazo. Sin embargo, es significativamente más complicado determinar los valores que determinar recompensas. Las recompensas se otorgan directamente por el ambiente, pero los valores deben estimarse continuamente a partir de las observaciones del agente a lo largo de su ciclo de funcionamiento. Es por eso que el componente más importante de casi cualquier algoritmo de RL es un método para estimar valores de forma eficiente.

Modelo del ambiente: Se trata de un elemento que debe permitir realizar inferencias de cómo se comportaría el ambiente real. De esta forma, dado un estado y una acción, el modelo podría

predecir el siguiente estado y la siguiente recompensa. Generalmente los modelos se emplean para un proceso conocido como *planeación*, que es una forma de decidir el curso de una acción considerando todas las posibles situaciones futuras que podrían ocurrir a consecuencia de esa acción, sin que el agente tenga que experimentarlas realmente. Los métodos de RL que utilizan modelos y planeación se conocen como *métodos basados en modelos*, mientras que aquellos que sólo se basan en ensayo y error se conocen como *métodos libres de modelo*.

5.2 El problema del bandido multibrazo

En este problema tenemos k diferentes acciones a elegir. Después de elegir una acción se recibe una recompensa numérica, elegida de una distribución de probabilidad estacionaria que depende de la acción que se realizó. El objetivo es maximizar la recompensa total esperada en un número determinado de pasos de tiempo. Se conoce como el problema del bandido multibrazo en alusión a la máquina “bandido” que se encuentra en los casinos, donde se inserta una moneda y se jala una palanca, en este caso existen k máquinas en vez de una. Cada acción que se realiza simboliza jalar la palanca de alguna de las máquinas, siendo la recompensa el hecho de si ganamos el premio mayor o no. A través de una repetición de acciones debemos maximizar nuestras ganancias al elegir sólo las mejores máquinas.

Cada k acción tiene una recompensa esperada dado que esa acción haya sido efectuada, es decir, cada acción tiene un valor. Se denota a una acción de un paso de tiempo t como A_t y su correspondiente recompensa como R_t . El valor exacto de una acción arbitraria a se denota:

$$q_*(a) = E[R_t | A_t = a] \quad (5.1)$$

Si conociéramos el valor exacto de cada acción el problema sería trivial, porque siempre elegiríamos la acción de mayor valor, pero en este caso sólo se conoce un estimado del valor de cada acción. El valor estimado de una acción a en el tiempo t se denota como $Q_t(a)$, buscamos que $Q_t(a)$ sea cercano a $q_*(a)$.

Dado un conjunto de acciones en cualquier paso de tiempo, existe al menos una acción cuyo valor estimado es el mayor de todos. A esas acciones se les llama *avariciosas*. Al seleccionar una de esas acciones estamos *explotando* el conocimiento actual de los valores de las acciones. Si se selecciona cualquier otra acción entonces estamos *explorando*, de esta manera mejoramos la estimación del valor de las acciones *no-avariciosas*. La *explotación* es la mejor forma de obtener la mayor recompensa inmediata, pero la *exploración* puede ofrecer una mayor recompensa a largo plazo. Este conflicto es otra de las principales características de los problemas de RL, ya que no es posible realizar ambas, explorar y explotar, al seleccionar una acción. Lograr un balance entre *explotación* y *exploración* es un problema aún no resuelto. Los métodos actuales realizan fuertes suposiciones respecto a un conocimiento a priori o estacionario que pueden no siempre cumplirse o que son imposibles de verificar en la práctica.

El verdadero valor de una acción es el promedio de la recompensa cuando la acción es seleccionada (Sutton & Barto, 2005) :

$$Q_t(a) = \frac{\text{suma de las recompensas de elegir } a \text{ antes de } t}{\text{número de veces que se ha elegido } a \text{ antes de } t} = \frac{\sum_{i=1}^{t-1} R_i \cdot 1_{A_t=a}}{\sum_{i=1}^{t-1} 1_{A_t=a}} \quad (5.2)$$

Donde $1_{\text{predicado}}$ denota la variable aleatoria que toma el valor de 1 si el predicado es verdadero y 0 de otra forma. En caso de que el denominador sea 0 se le asigna un valor default a $Q_t(a)$, que usualmente es 0. Conforme el valor del denominador tienda al infinito, $Q_t(a)$ convergerá a $q_*(a)$. A este método de estimación del valor de la acción se conoce como *método del promedio de muestras*.

La regla más simple para seleccionar una acción es seleccionar aquella con el valor estimado más alto, es decir, una *acción avariciosa*. En caso de que varias acciones tuvieran el mismo valor una opción es seleccionar alguna de ellas de forma aleatoria.

$$A_t = \operatorname{argmax}_a Q_t(a) \quad (5.3)$$

Este método de selección de acciones es avaricioso, pues siempre explotará al conocimiento actual para maximizar la recompensa inmediata, sin explorar otras opciones que quizá podrían ser mejores. Una alternativa es un método que se comporte avaricioso la mayoría del tiempo, pero en algunas ocasiones, con una probabilidad pequeña ϵ , seleccionará una acción de forma aleatoria; todas las acciones tendrán la misma probabilidad de ser seleccionadas, independientemente de su valor. A este método de selección de acciones se le conoce como ϵ -*avaricioso* (Sudharsan, 2018). Una ventaja de este método es que, conforme el número de pasos incrementa, todas las acciones serían elegidas un número infinito de veces, asegurándonos que todos los valores de las acciones $Q_t(a)$ convergerán a $q_*(a)$.

Considerando únicamente una acción a . Sea R_i la recompensa recibida después de la i -ésima selección de a . Sea Q_n la estimación del valor de a después de haber sido seleccionada $n-1$ veces, entonces:

$$Q_n = \frac{R_1 + R_2 + \dots + R_{n-1}}{n-1} \quad (5.4)$$

Se podría guardar un registro de todas las recompensas que se han recibido al seleccionar a y después calcular Q_n cuando se necesite. Sin embargo, la memoria y otros requerimientos computacionales aumentarían conforme pase el tiempo y se observen más recompensas. Entonces, dado Q_n y la n -ésima recompensa R_n , el cálculo del nuevo promedio de todas las n recompensas puede ser:

$$\begin{aligned} Q_{n+1} &= \frac{1}{n} \sum_{i=1}^n R_i \\ &= \frac{1}{n} \left(R_n + \sum_{i=1}^{n-1} R_i \right) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{n} \left(R_n + (n-1) \frac{1}{n-1} \sum_{i=1}^{n-1} R_i \right) \\
&= \frac{1}{n} (R_n + (n-1)Q_n) \\
&= \frac{1}{n} (R_n + nQ_n - Q_n) \\
&= Q_n + \frac{1}{n} [R_n - Q_n] \tag{5.5}
\end{aligned}$$

Si $n=1$, entonces $Q_2 = R_1$, siendo Q_1 un valor arbitrario, usualmente 0. Esta implementación sólo requiere memoria para Q_n y para n , realizando operaciones pequeñas para cada nueva recompensa. A este método para estimar el valor de una acción se le conoce como *método incremental* (Sutton & Barto, 2005).

La importancia de este método es que la forma en que ocurre la actualización de un nuevo valor es muy similar en varios algoritmos de RL:

$$NuevaEstimación \leftarrow EstimaciónAnterior + TasaAprendizaje [Meta - EstimaciónAnterior] \tag{5.6}$$

La expresión $[Meta - EstimaciónAnterior]$ es el error que puede existir en la estimación. El error se reducirá al acercarnos al valor *Meta*. La *Meta* debe indicar la dirección hacia la que se desea avanzar. En este caso, la *Meta* es la n -ésima recompensa R_n .

El parámetro *TasaAprendizaje* cambia de un paso de tiempo a otro. En este caso, la tasa de aprendizaje es $\frac{1}{n}$, es común en el campo del aprendizaje de máquina, que este parámetro se represente con el símbolo α , o de forma más general $\alpha_t(a)$.

A continuación se presenta el pseudocódigo de un algoritmo bandido empleando un *método incremental* para la estimación de los valores de las acciones y un *método ϵ -avaricioso* para seleccionar acciones. Se asume que la función *bandido(A)* recibe como argumento una acción y devuelve la recompensa correspondiente.

Algoritmo bandido

Inicializa para $a=1$ hasta k :

$$Q(a) \leftarrow 0$$

$$N(a) \leftarrow 0$$

Itera por siempre:

$$A \leftarrow \begin{cases} \operatorname{argmax}_a Q(a) & \text{con probabilidad } 1 - \epsilon \\ \text{acción aleatoria } a & \text{con probabilidad } \epsilon \end{cases}$$

$$R \leftarrow \text{bandido}(A)$$

$$N(A) \leftarrow N(A) + 1$$

$$Q(A) \leftarrow Q(A) + \frac{1}{N(A)} [R - Q(A)]$$

5.3 Procesos de Decisión de Markov

El problema del bandido multibrazo no aborda el hecho de que situaciones diferentes requieren acciones diferentes. Tampoco toma en cuenta el impacto de las acciones a largo plazo, es decir, las acciones que el agente toma no solo afecta la recompensa inmediata a recibir, sino también las situaciones subsecuentes que se desencadenarán junto con sus respectivas recompensas. Para abordar los puntos antes mencionados, se enmarca el problema del aprendizaje mediante la interacción con el ambiente para alcanzar una meta a través de procesos de decisión de Markov o MDPs (Markov Decision Processes) (Sudharsan, 2018).

El agente y el ambiente interactúan a lo largo de una secuencia discreta de pasos de tiempo $t=0,1,2,3,\dots$, en un momento del tiempo t el agente recibe un estado $S_t \in S$ del ambiente, donde S es el conjunto de todos los posibles estados. Basado en este estado, el agente selecciona una acción $A_t \in A(S_t)$ de un conjunto de posibles acciones. Donde $A(S_t)$ es el conjunto de acciones válidas en el estado S_t . En el siguiente paso de tiempo de haber seleccionado la acción, el ambiente le retorna al agente un nuevo estado $S_{t+1} \in S$ y una recompensa por haber elegido esa acción $R_{t+1} \in R$. Donde R es el conjunto de todas las posibles recompensas que puede recibir el agente. Tal y como se muestra en la figura 5.1.

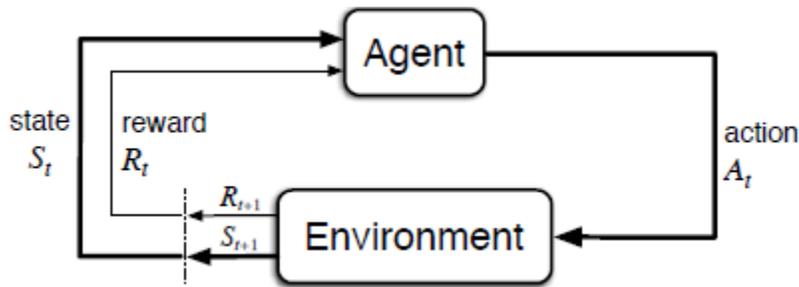


Figura 5.1. La interacción que ocurre entre el agente y el ambiente en un proceso de decisión de Markov. (Sutton & Barto, 2005).

Considerando lo anterior, El MDP y el agente producirían una trayectoria que comenzaría de la siguiente forma:

$$S_0, A_0, R_1, S_1, A_1, R_2, S_2, A_2, R_3, \dots \quad (5.7)$$

En un MDP finito los conjuntos de estados, acciones y recompensas tienen un número finito de elementos. En este caso, las variables aleatorias R_t y S_t están definidas por una distribución de probabilidad discreta que depende del estado anterior y la acción que se tomó en ese estado. Es decir, dado un estado $s \in S$ y una acción $a \in A(s)$, se calcula la probabilidad conjunta de que el siguiente estado sea $s' \in S$ y se reciba una recompensa $r \in R$.

$$p(s', r | s, a) = P\{S_t = s', R_t = r | S_{t-1} = s, A_{t-1} = a\} \quad (5.8)$$

Al ser una distribución de probabilidad debe cumplirse que:

$$\sum_{s' \in S} \sum_{r \in R} p(s', r | s, a) = 1, \forall s \in S, a \in A(s) \quad (5.9)$$

Se dice que los estados tienen la *propiedad de Markov* cuando incluyen información de todos los aspectos de la última interacción agente-ambiente que afecten las futuras decisiones del agente.

El marco de trabajo MDP es abstracto y flexible, lo que permite que pueda ser aplicado a diferentes problemas de varias formas. Por ejemplo, los pasos de tiempo no necesariamente pueden ser usados como intervalos fijos de tiempo real, también pueden referirse a etapas sucesivas de tomas de decisión. Las acciones pueden ser controles, como aplicar diferente cantidad de voltaje a un brazo robótico, girar un volante, activar o desactivar una alarma. Los estados también pueden tomar una gran variedad de formas, pueden estar compuestos por lecturas de un sensor, inclusive pueden ser completamente abstractos o subjetivos, como una sensación o la descripción simbólica de objetos en un cuarto.

El agente siempre aprenderá a maximizar la recompensa. Si buscamos que el agente haga algo por nosotros, debemos proveerle recompensas de tal forma que, cuando se maximicen, al mismo tiempo el agente estará cumpliendo nuestras metas. Por esta razón es crucial que las recompensas que nosotros definamos realmente representen lo que queremos lograr. Hay que tener especial cuidado, pues la señal de recompensa no es el lugar donde se debe indicar cómo es que queremos que se logren las metas. La señal de recompensa es nuestra forma de comunicarle al agente qué queremos lograr, no cómo queremos que se logre.

El marco de trabajo MDP permite una abstracción del problema. Propone que, sin importar los detalles del aparato de control, o recepción o almacenamiento de información, y sin importar el objetivo que se esté tratando de alcanzar, cualquier problema de aprendizaje orientado a una meta puede reducirse a tres señales que serán intercambiadas entre el agente y el ambiente. La primera señal representa las elecciones del agente (acciones), la segunda señal representa las bases sobre las cuales se toma la elección (los estados), y la última señal representa la meta del agente (las recompensas). Este marco de trabajo puede no ser suficiente para lograr representar todos los problemas de aprendizaje en donde se tomen decisiones, pero ha demostrado ser ampliamente útil y aplicable (Sudharsan, 2018).

5.4 Retornos esperados en los episodios

En general, buscamos maximizar el retorno esperado G_t , que es una función específica de la secuencia de recompensas. En el caso más sencillo el retorno es la suma de las recompensas:

$$G_t = R_{t+1} + R_{t+2} + R_{t+3} + \dots + R_T \quad (5.10)$$

Donde T es el paso de tiempo final. Este acercamiento tiene sentido cuando la interacción entre el agente y el ambiente puede dividirse en secuencias llamadas episodios, como lo serían las jugadas de un juego, número de viajes en un laberinto, o cualquier tipo de interacción repetitiva. Cada episodio termina en un estado especial llamado *estado terminal*, después de ello, todo reinicia con

un nuevo episodio que comienza en un estado inicial ya definido o en una muestra de una distribución de estados iniciales. Los episodios son independientes entre sí, ninguno se ve afectado por cómo termino el episodio anterior. Aquellos problemas que pueden dividirse en episodios de este tipo se les denominan *tareas episódicas*. Un ejemplo muy claro de este tipo de tareas son las partidas de ajedrez, siendo cada partida un episodio, el episodio termina cuando alguno de los dos jugadores gana o cuando se llega a un empate, una vez que termina el episodio el nuevo episodio comienza colocando todas las piezas en su lugar correspondiente.

Se denota como S al conjunto de todos los estados no terminales, S^+ al conjunto de estados terminales, y T , como el paso de tiempo en el que termina un episodio, siendo una variable aleatoria, pues la duración de cada episodio puede ser diferente.

Cuando la interacción entre agente y ambiente no puede dividirse en episodios identificables, sino que se extienden continuamente sin límite, se les denomina *tareas continuas*. Un ejemplo de este tipo de tareas es un termostato que regula la temperatura de un edificio, la tarea del termostato nunca se detiene, los estados pueden estar conformados por información como la temperatura actual y detalles como la hora del día. En cada estado sólo hay 2 acciones posibles, prender el calentador o apagarlo, la recompensa es -1 si alguien tiene que ajustar el calentador de forma manual y 0 de otra forma.

El retorno esperado G_t para tareas continuas, donde $T = \infty$ podría fácilmente ser también infinito, por ejemplo, si en cada paso de tiempo el agente recibiera +1 de recompensa. Una forma de calcular G_t es empleando un parámetro conocido como *tasa de descuento* γ , que hará que las recompensas del futuro tengan menor valor que las recompensas del presente, volviendo una suma infinita en una suma que convergerá por tratarse de una serie geométrica. Es así que el agente tratará de seleccionar las acciones donde la suma de las recompensas con descuento que recibirá en el futuro sea máxima. Es decir, elegirá A_t para maximizar el retorno con descuento esperado:

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \dots + \gamma^{k-1} R_{t+k} + \dots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \quad (5.11)$$

Donde la *tasa de descuento* γ es un parámetro $0 \leq \gamma \leq 1$.

Una recompensa recibida después de k pasos de tiempo en el futuro valdrá γ^{k-1} veces lo que valdría si se recibiera de forma inmediata. Si $\gamma < 1$, la suma infinita de G_t tendría un valor finito mientras la secuencia de recompensas estuviera acotada. Si $\gamma = 0$, el agente sería "miope", ya que sólo le importarían las recompensas inmediatas, sólo aprendería a maximizar R_{t+1} . Conforme el valor de la tasa de descuento se acerque a 1, el agente se interesará más por las recompensas que se recibirán en el futuro.

$$\begin{aligned} G_t &= R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \gamma^3 R_{t+4} + \dots \\ &= R_{t+1} + \gamma(R_{t+2} + \gamma R_{t+3} + \gamma^2 R_{t+4} + \dots) \\ &= R_{t+1} + \gamma(G_{t+1}) \end{aligned} \quad (5.12)$$

Aunque para obtener G_t se requiere de la suma de un número infinito de recompensas para tareas continuas, G_t será finito si la recompensa en cada paso de tiempo es diferente de 0 y es constante con $\gamma < 1$. En caso de que la recompensa siempre sea 1 en cada paso de tiempo entonces G_t se calcularía de la siguiente forma:

$$G_t = \sum_{k=0}^{\infty} \gamma^k = \frac{1}{1 - \gamma} \quad (5.13)$$

Los MDPs formalizan el problema de una gente interactuando con el ambiente, ambos interactúan en tiempos discretos. En cada paso de tiempo el agente observa el estado actual del ambiente, con base en ello el agente selecciona una acción, a continuación, el ambiente cambia a un nuevo estado y proporciona una recompensa. La meta del agente es maximizar la recompensa total del futuro, esto frecuentemente significa balancear el hecho de obtener recompensas inmediatas con las consecuencias a largo plazo de las acciones. Esta meta se formaliza con el retorno esperado G_t , que es la suma de las recompensas futuras con descuento. Con la tasa de descuento γ se garantiza que el retorno permanece finito. El valor de γ determina cuánta importancia le dará el agente a las recompensas inmediatas o a las de largo plazo.

5.5 Políticas y funciones de valor

Formalmente, una política es una función que mapea los estados del ambiente en los que el agente se puede encontrar, con las probabilidades de seleccionar cada posible acción en dichos estados (Sutton & Barto, 2005). Si el agente sigue una política π en un paso de tiempo t , entonces $\pi(a|s)$ es la probabilidad de que se seleccione la acción $A_t = a$ dado que el estado actual es $S_t = s$. Las políticas sólo dependen del estado actual, no de otros factores como el tiempo o estados previos, por esta razón, el estado debe proveer toda la información necesaria para que se seleccione la mejor acción posible. Los métodos de RL especificarán cómo cambiará la política del agente con base en la experiencia que adquiera.

La *función del valor de un estado* s bajo una política π se denota como $v_{\pi}(s)$, el valor de esta función podría interpretarse como el valor del retorno que se esperaría que obtenga el agente a partir del estado s siguiendo la política π . En el caso de los MDPs se puede definir formalmente v_{π} como:

$$v_{\pi}(s) = E_{\pi}[G_t | S_t = s] = E_{\pi} \left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} | S_t = s \right], \forall s \in S \quad (5.14)$$

Cabe resaltar que la función de valor de un estado terminal, si existe, siempre será 0.

De forma análoga se define el valor de tomar una acción a en un estado s bajo una política π como $q_{\pi}(s, a)$, que se puede interpretar como el valor de retorno que se esperaría que obtenga el agente a partir del estado s , tomando la acción a , siguiendo una política π :

$$q_{\pi}(s, a) = E_{\pi}[G_t | S_t = s, A_t = a] = E_{\pi} \left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} | S_t = s, A_t = a \right] \quad (5.15)$$

A q_{π} se le conoce como la *función de valor de una acción* para una política π .

Una propiedad fundamental de las funciones de valor es que satisfacen una relación recursiva, tal y como se observó con la función para calcular el retorno esperado G_t .

$$\begin{aligned} v_{\pi}(s) &= E_{\pi}[G_t | S_t = s] \\ &= E_{\pi}[R_{t+1} + \gamma G_{t+1} | S_t = s] \\ &= \sum_a \pi(a|s) \sum_{s'} \sum_r p(s', r|s, a) [r + \gamma E_{\pi}[G_{t+1} | S_{t+1} = s']] \\ &= \sum_a \pi(a|s) \sum_{s'} \sum_r p(s', r|s, a) [r + \gamma v_{\pi}(s')], \forall s \in S \end{aligned} \quad (5.16)$$

Donde las acciones a , son tomadas del conjunto $A(s)$, los estados siguientes, s' , son tomados del conjunto S (o de S^+ en el caso de tareas episódicas) y las recompensas r son tomadas del conjunto R . La ecuación final de v_{π} se conoce como ecuación de Bellman, donde se expresa la relación que existe entre el valor del estado actual y el valor de los estados sucesores. En el primer término de la ecuación de Bellman se denota que las acciones que el agente tiene disponible sólo dependen del estado actual, con $\sum_a \pi(a|s)$. El segundo término indica que el siguiente estado y la recompensa que obtendrá el agente dependen del estado actual y la acción que el agente tome, con $\sum_{s'} \sum_r p(s', r|s, a)$. Después se toma en cuenta la recompensa inmediata que se recibirá con r y el término final de la ecuación determinará qué tanto le interesan al agente las recompensas esperadas futuras del siguiente estado, con $\gamma v_{\pi}(s')$. De forma un poco más intuitiva, a partir de un estado en el que se encuentre el agente, éste podrá elegir distintas acciones, la política determinará cuál es la mejor opción, no obstante, cada acción puede conducir al agente a diferentes estados, cada uno con sus respectivas recompensas, es por ello que, para obtener el valor del estado en el que se encuentra el agente, la ecuación de Bellman realiza un promedio entre todas las posibilidades, ponderando cada una por su probabilidad de ocurrir. Entre mayor sea el valor de un estado, significa que será mejor para el agente llegar ahí, ya que obtendrá mayores recompensas.

La misma lógica que se empleó para obtener la ecuación de Bellman para v_{π} puede aplicarse sobre la función del valor de una acción y de esa forma obtener la ecuación de Bellman para q_{π} :

$$\begin{aligned} q_{\pi}(s, a) &= E_{\pi}[G_t | S_t = s, A_t = a] \\ &= E_{\pi}[R_{t+1} + \gamma G_{t+1} | S_t = s, A_t = a] \\ &= \sum_{s'} \sum_r p(s', r|s, a) [r + \gamma \sum_{a'} \pi(a'|s') E_{\pi}[G_{t+1} | S_{t+1} = s', A_{t+1} = a']] \end{aligned}$$

$$= \sum_{s'} \sum_r p(s', r|s, a) [r + \gamma \sum_{a'} \pi(a'|s') q_\pi(s', a')], \forall s \in S, \forall a \in A(s) \quad (5.17)$$

Existen problemas con el marco MDP donde no es posible obtener las funciones de valor de los estados empleando la ecuación de Bellman, por ejemplo, en una partida de ajedrez, donde es virtualmente imposible enlistar todos los posibles estados, pues existen alrededor de 10^{45} .

5.6 La ecuación de Bellman y funciones óptimas

La ecuación de Bellman, nombrada así por el matemático Richard Bellman (Bellman, 1952), nos ayuda a resolver MDPs, volviendo sumas infinitas inmanejables en problemas de álgebra lineal. Cuando se trata de resolver un problema de RL sobre un marco MDP, en realidad lo que se busca es encontrar una política que retorne las mayores recompensas a lo largo del tiempo, es decir, una política óptima. Cuando se tratan problemas sobre un marco MDP, se dice que una política π es mejor que una política π' , si el retorno esperado de π es mayor que el de π' para todos los estados. Esto es $\pi \geq \pi'$ si y sólo si $v_\pi(s) \geq v_{\pi'}(s) \forall s \in S$. Siempre existe una política que es igual o mejor que el resto, y se le conoce como *política óptima* π_* . De hecho, pueden existir varias políticas óptimas, pero todas ellas compartirán la misma *función óptima de valor de estado*.

La *función óptima de valor de estado* se denota como $v_*(s)$. Esta función retorna el máximo valor comparado con todas las otras funciones de valor de estado:

$$v_*(s) = \max_\pi V_\pi(s) \forall s \in S \quad (5.18)$$

De forma similar, las políticas óptimas también comparten la misma *función óptima de valor de una acción*, denominada $q_*(s, a)$:

$$q_*(s, a) = \max_\pi q_\pi(s, a) \forall s \in S \text{ y } \forall a \in A \quad (5.19)$$

Con base en lo anterior, se puede definir la *ecuación óptima de Bellman para $v_*(s)$* de la siguiente forma:

$$v_\pi(s) = \sum_a \pi(a|s) \sum_{s'} \sum_r p(s', r|s, a) [r + \gamma v_\pi(s')] \quad (5.20)$$

$$v_*(s) = \sum_a \pi_*(a|s) \sum_{s'} \sum_r p(s', r|s, a) [r + \gamma v_*(s')] \quad (5.21)$$

Una política óptima π_* elige la acción con mayor valor en cada estado. Esta política asignará una probabilidad de 1 a dicha acción y una probabilidad de 0 a todas las demás acciones, es así que reemplazamos $\sum_a \pi_*(a|s)$ con \max_a obteniendo:

$$v_*(s) = \max_a \sum_{s'} \sum_r p(s', r|s, a) [r + \gamma v_*(s')] \quad (5.22)$$

Mientras que la *ecuación óptima de Bellman para $q_*(s, a)$* resulta:

$$q_*(s, a) = \sum_{s'} \sum_r p(s', r|s, a) [r + \gamma \max_{a'} q_*(s', a')] \quad (5.23)$$

Una vez que se ha encontrado v_* se vuelve relativamente sencillo encontrar π_* . Porque para cada estado s , existirán una o más acciones con las que se alcance el valor máximo en la ecuación óptima de Bellman para $v_*(s)$. Cualquier política que le asigne una probabilidad diferente de cero únicamente a estas acciones es una política óptima. Para un MDP finito, en cada estado se debe hacer una búsqueda de un paso a futuro de aquellas acciones que maximicen $v_*(s)$ para obtener las acciones óptimas (las acciones que forman parte de π_*), es decir, para cada acción a que se tome en el estado s , se debe calcular el valor de $v_*(s)$ respecto a a considerando $v_*(s')$, el valor del estado s' al que llegará el agente a consecuencia de haber elegido a . Cualquier política que sea avariciosa con respecto a v_* es una política óptima. El término avaricioso hace alusión a una búsqueda o proceso de selección basado únicamente en consideraciones locales o inmediatas, sin considerar la posibilidad de que la opción elegida pueda prevenir el acceso en el futuro a mejores alternativas. Sin embargo, al evaluar las consecuencias inmediatas de las acciones en cada estado con v_* , en realidad también se están evaluando sus consecuencias a largo plazo, porque v_* ya está considerando las recompensas de todos los comportamientos futuros. A través de v_* , es posible acceder a los retornos esperados a largo plazo en un estado gracias a que se representan con una cifra que es accesible de forma inmediata.

Una vez que se ha encontrado q_* se vuelve aún más fácil encontrar π_* , porque entonces ya no es siquiera necesario realizar una búsqueda de un paso a futuro ni realizar cálculo alguno, simplemente se debe elegir aquella acción con la que $q_*(s, a)$ obtenga su máximo valor. La función de valor de acción captura los resultados de todas las búsquedas de un paso a futuro, proporcionando el retorno óptimo esperado a largo plazo como un valor que es accesible de forma inmediata para cada par estado-acción del MDP. De esta forma, q_* permite que se seleccionen las acciones óptimas sin tener que conocer nada acerca de los posibles estados futuros y sus valores, e inclusive sin tener conocimiento acerca de las dinámicas del ambiente.

A pesar de que resolver directamente las ecuaciones óptimas de Bellman, es en sí mismo resolver el problema de RL, este enfoque rara vez es usado en la realidad debido a que deben cumplirse tres cosas: 1) conocer perfectamente las dinámicas del ambiente, 2) tener suficiente poder de cómputo para realizar búsquedas exhaustivas, así como observar todas las posibilidades y calcular todas las probabilidades de los futuros estados en términos de las recompensas esperadas, y 3) se debe cumplir la propiedad de Markov. En la mayoría de problemas se conocen las dinámicas del ambiente y se cumple la propiedad de Markov, pero el número de estados llega a ser muy grande y resolver las ecuaciones de Bellman para obtener v_* o q_* puede tardar años, por esa razón se han desarrollado métodos que más bien tratan de aproximarse a los valores reales de v_* o q_* , de hecho, la mayoría de los algoritmos de RL pueden verse como una aproximación a la solución de las ecuaciones óptimas de Bellman, los cuales emplean la experiencia del agente de haber

transitado entre estados en lugar de un conocimiento a priori acerca de transiciones que se esperarían entre estados.

En tareas con un número pequeño de estados es posible realizar una aproximación a la solución de las ecuaciones óptimas de Bellman utilizando arreglos o tablas, de tal forma que se pueda obtener una tabla que relacione los estados con las acciones que se pueden realizar en dichos estados, y cada celda de la tabla almacene el valor de un par estado-acción, a estas tareas se les conoce como *casos tabulares*, y sus algoritmos se conocen como *métodos tabulares* (e.g. Q-learning, SARSA y Dyna-Q).

5.7 Aprendizaje por diferencias temporales

El *aprendizaje por diferencias temporales* (TD) es una de las ideas que más impacto ha tenido en el campo de RL. En este método, en lugar de tratar de calcular la recompensa total esperada, TD trata de predecir la combinación de recompensas inmediatas y su propia predicción de la recompensa que se recibirá en el siguiente paso de tiempo. Cuando el siguiente paso de tiempo llega, entonces la nueva predicción se compara contra la predicción anterior, si son diferentes el algoritmo calcula qué tan diferentes son y emplea esta “diferencia temporal” para ajustar la antigua predicción a la nueva predicción. Al tratar siempre de acercar estas dos predicciones en cada paso de tiempo, empatando expectativa versus realidad, la cadena completa de predicción se vuelve más precisa (Sutton & Barto, 2005).

Recordando del *método incremental* de (5.6):

$$\text{NuevaEstimación} \leftarrow \text{EstimaciónAnterior} + \text{TasaAprendizaje} [\text{Meta} - \text{EstimaciónAnterior}] \quad (5.6)$$

Se puede emplear este mismo método para estimar el valor de cada estado $V(S_t)$ siguiendo una política π , con una tasa de aprendizaje α :

$$V(S_t) \leftarrow V(S_t) + \alpha [G_t - V(S_t)] \quad (5.24)$$

Recordando que G_t es el retorno esperado con descuento:

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \gamma^3 R_{t+4} + \dots = R_{t+1} + \gamma(G_{t+1}) \quad (5.25)$$

Entonces, al sustituir G_t de (5.24) en (5.23) se obtiene:

$$V(S_t) \leftarrow V(S_t) + \alpha [R_{t+1} + \gamma V(S_{t+1}) - V(S_t)] \quad (5.26)$$

El método planteado en (5.24) para calcular el valor de los estados en un MDP se conoce como Monte Carlo. Los métodos tipo Monte Carlo se pueden aplicar a problemas episódicos, donde existe una secuencia de estados, acciones y recompensas. La gran desventaja de este tipo de métodos es que la actualización de los valores de los estados sólo se realiza hasta que el episodio termina, porque en ese momento se conoce el valor de G_t , esto quiere decir que este tipo de

métodos no se puede aplicar en problemas continuos, además de que los requerimientos de memoria dependen del número de estados de cada episodio y de la longitud de los episodios.

El método presentado en (5.25) realiza una predicción de diferencias temporales, donde en el siguiente paso de tiempo $t + 1$ se realiza una actualización útil de $V(S_t)$ utilizando la recompensa obtenida R_{t+1} y la estimación $V(S_{t+1})$. De esta forma, las actualizaciones de $V(S_t)$ no requieren que el episodio termine y por lo tanto se puede aplicar en problemas continuos. El algoritmo completo, conocido como $TD(0)$, es un método tabular para estimar v_π .

Algoritmo tabular $TD(0)$

Inicializa $V(s) = 0 \forall s \in S$

Itera para cada episodio:

Elige un estado inicial s

Itera para cada paso del episodio, hasta que s sea terminal:

$a \leftarrow$ acción siguiendo π para s

Ejecuta la acción a , observa la recompensa obtenida r y el siguiente estado s'

$V(s) \leftarrow V(s) + \alpha [r + \gamma V(s') - V(s)]$

$s \leftarrow s'$

$TD(0)$ basa parte de su actualización de $V(S_t)$ en una estimación ya existente $V(S_{t+1})$, es decir, actualizan sus estimaciones con base en las estimaciones de estados sucesivos, por eso se dice que el método es de tipo *bootstrapping*.

En (5.25) al término $[R_{t+1} + \gamma V(S_{t+1}) - V(S_t)]$ se le conoce como el *error TD*, donde se mide la diferencia entre el valor estimado de S_t y la nueva estimación que se basa en la experiencia real del agente $R_{t+1} + \gamma V(S_{t+1})$.

5.8 Q-learning

Este algoritmo (Watkins, 1989) libre de política (Off-policy) de tipo TD es uno de los más populares en la historia del aprendizaje por reforzamiento, la idea central se presenta en la siguiente ecuación:

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha [R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t)] \quad (5.27)$$

En este caso se aprende la función de valor de una acción $Q(S_t, A_t)$, es decir, el valor de tomar una acción a en un estado s siguiendo una política π , que directamente aproxima q_* , independientemente de la política que se siga, formando una tabla de valores Q que almacena el valor de cada acción en cada estado del MDP. Entre mayor sea el valor del par estado-acción significa que el agente obtendrá una mayor recompensa. La política π que sigue el agente aún mantiene relevancia, en el sentido de que determina los pares estado-acción que son visitados y actualizados. Sin embargo, lo único que se requiere para la convergencia del algoritmo es que todos los pares continúen siendo actualizados.

Este algoritmo se considera libre de política porque actualiza los valores de $Q(S_t, A_t)$ utilizando los valores Q del siguiente estado S_{t+1} y de aquella acción a del estado S_{t+1} que maximice el valor de

$Q(S_{t+1}, a)$, es decir, se selecciona de forma avariciosa la acción de S_{t+1} , sin importar cuál sea la política π que el agente este siguiendo para seleccionar acciones. También podría verse como si el agente tuviera dos políticas, una para elegir la acción de acuerdo con el estado actual, por ejemplo ε -avariciosa, como se mostró en el algoritmo del bandido multibrazo, y otra que sería una política avariciosa para estimar el retorno para cada par estado-acción. Cabe mencionar que también existen métodos de tipo TD que no son libres de política (On-policy) como SARSA, donde se actualizan los valores de $Q(S_t, A_t)$ utilizando los valores Q del siguiente estado S_{t+1} y de aquella acción A_t que ha sido elegida siguiendo la política actual del agente. Claro que, en el caso de que la política actual sea avariciosa, entonces no existiría diferencia entre Q-learning y SARSA al momento de actualizar los valores de Q, sin embargo, no es muy recomendable usar una política avariciosa fuera del paso de actualización de Q, pues el agente nunca exploraría, en su lugar suele usarse una política ε -avariciosa o de Boltzmann.

A continuación se muestra el pseudocódigo del algoritmo:

Algoritmo Q-learning

Inicializa $Q(s, a) = 0 \quad \forall s \in S \text{ y } \forall a \in A(s)$

Itera para cada episodio:

Elige un estado inicial s

Itera para cada paso del episodio, hasta que s sea terminal:

$a \leftarrow$ acción siguiendo π (política ε – avariciosa o Boltzmann) para s

Ejecuta la acción a , observa la recompensa obtenida r y el siguiente estado s'

$Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma \max_a Q(s', a) - Q(s, a)]$

$s \leftarrow s'$

Una de las desventajas de los métodos TD que involucran operaciones de maximización para actualizar los valores de Q, es que se corre el riesgo de que puedan ser engañados, en este tipo de algoritmos se utiliza el máximo valor de una estimación ya sea a través de una política avariciosa o ε -avariciosa, lo que puede provocar un sesgo positivo. Por ejemplo, en algún problema podría suceder que para un estado s con diversas acciones a tengan como valor verdadero 0 para la función $q(s, a)$, para el agente estos valores son desconocidos y al realizar una estimación de cada uno de ellos se obtendría una distribución en donde algunos valores serían menores a cero y otros serían mayores. El máximo de los verdaderos valores de cada par estado-acción es cero, pero el máximo de los valores estimados es positivo, a esto se le conoce como un *sesgo de maximización*. Algunos algoritmos que intentan evitar este tipo de sesgos son Double Q-learning, Double SARSA y Double Expected SARSA. Como su nombre indica, utilizan 2 tablas Q(S,A) y deciden de forma aleatoria cuál de ellas actualizan en cada paso de tiempo, las dos funciones de valor se trabajan de forma simétrica y existen diferentes formas en las que la política del algoritmo puede trabajar con ambas, por ejemplo promediándolas.

5.9 Planeación con métodos tabulares

Un modelo del ambiente le ayuda al agente a predecir la forma en la que el ambiente responderá ante sus acciones. Dado un estado y una acción, el modelo produce una predicción del siguiente estado y de la recompensa que se obtendrían. Cuando el ambiente es estocástico, para una misma acción pueden llegar a ocurrir diferentes estados y sus respectivas recompensas asociadas a ellos. Los modelos que producen una descripción de todas las posibilidades y sus probabilidades se conocen como *modelos de distribución*, mientras que aquellos modelos que sólo consideran una de las posibilidades son *modelos de muestreo*. Los modelos de distribución siempre pueden ser usados para producir muestras, sin embargo, en muchas aplicaciones es más fácil obtener modelos de muestreo.

Los modelos pueden ser usados para imitar o simular experiencia. Dado un estado inicial y una acción, un modelo de muestreo produciría una posible transición, mientras que un modelo de distribución generaría todas las posibles transiciones y las probabilidades de que ocurrieran. De hecho, dado un estado inicial y una política, un modelo de muestreo podría producir un episodio completo, mientras que un modelo de distribución podría generar todos los posibles episodios y sus probabilidades, pero en cualquier caso, el modelo se emplearía para simular al ambiente y producir experiencia simulada.

En el campo de RL, la palabra *planeación* hace alusión a cualquier proceso computacional que recibe como entrada un modelo y produce o mejora una política para interactuar con el ambiente modelado. Los métodos de RL sólo requieren la experiencia (i.e. el estado siguiente y la recompensa) del agente y en muchos casos estos mismos métodos se pueden usar con experiencia simulada tan bien como con experiencia real.

Cuando la planeación se realiza al mismo tiempo que el agente interactúa con el ambiente se obtiene nueva información que puede cambiar el modelo, lo que a su vez modificaría la interacción que se realiza en la planeación.

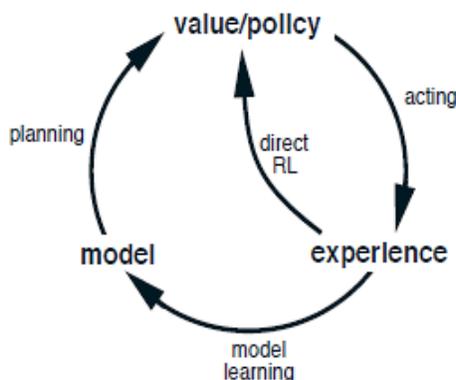


Figura 5.2. Se puede mejorar la política directamente con la experiencia e indirectamente a través de la planeación. (Sutton & Barto, 2005).

En un agente que se usa para la planeación, existen al menos 2 usos que se le pueden dar a la experiencia real, emplearla para mejorar el modelo, de manera que represente de forma más precisa el ambiente real y también puede usarse para mejorar directamente las funciones de valor y la política actual, como en el algoritmo Q-learning. A la primera opción se le conoce como *model-learning* y la segunda se le conoce como *direct RL* (Sutton & Barto, 2005). Como se muestra en la figura 5.2, la experiencia puede mejorar las funciones de valor y la política de forma directa (*direct RL*) o indirecta (*model learning*), por esa razón a los métodos que emplean la planeación para aprender se les conoce como *indirect RL*.

Los métodos indirectos a menudo emplean completamente una cantidad limitada de experiencia y con ello logran obtener una mejor política con una menor cantidad de interacciones con el ambiente. Por otra parte, los métodos directos son más simples y no se ven afectados por sesgos en el diseño del modelo.

Dyna-Q (Sutton, 1991) es un algoritmo que incluye todos los procesos de la figura 5.2, planear, actuar, *model-learning* y *direct RL*, todos ocurriendo continuamente. El proceso de planeación consiste en un paso tomado del algoritmo original de Q-learning, sólo que el estado y la acción a tomar por el agente son muestras aleatorias. El proceso de *direct RL* es prácticamente el algoritmo Q-learning. Ambos procesos son tabulares y asumen que el ambiente es determinista. Después de cada transición donde $S_t, A_t \rightarrow S_{t+1}, R_{t+1}$, el modelo almacena en una tabla, con datos de entrada S_t y A_t , los valores S_{t+1} y R_{t+1} , de esta forma, cuando al modelo se le pregunte a la hora de simular cómo respondería el ambiente dado que el agente ha elegido la acción A_t en el estado S_t , una búsqueda en la tabla le permitirá realizar esta predicción, retornando el estado siguiente y la recompensa que obtuvo el agente la última vez que experimentó lo mismo. Cabe resaltar que durante la planeación, las muestras aleatorias se realizarán sólo de pares estado-acción que el agente haya experimentado previamente, así que al modelo nunca se le preguntará acerca de un par del que no tenga información.

El modelo del ambiente mejorará conforme el agente adquiera experiencia real, lo que dará lugar a que se obtenga mayor experiencia simulada. Se emplea el término *control de búsqueda* al proceso que selecciona los estados y acciones iniciales para las simulaciones. Finalmente, la planeación se alcanza cuando se aplica el aprendizaje por reforzamiento a las experiencias simuladas como si en realidad hubieran ocurrido. Es común que el mismo método de aprendizaje se aplique tanto para aprender de la experiencia real, como para la planeación con la experiencia simulada. Aprender y planear son como dos maquinarias profundamente integradas que comparten los mismos engranajes, su única diferencia es el origen de su materia prima, la experiencia.

Conceptualmente los procesos de planear, actuar, *model learning* y *direct RL* ocurren de forma simultánea y en paralelo, pero en la práctica se debe especificar el orden en el que ocurren dentro de un paso de tiempo.

En Dyna-Q el aprendizaje y la planeación son logrados por el mismo algoritmo, empleando experiencia real para el aprendizaje y obteniendo experiencia simulada con la planeación.

Conforme el agente obtiene nueva información, el modelo se actualiza, conforme el modelo cambia el proceso de planeación calculará diferentes maneras de manejarse para empatar con el nuevo modelo.

En la vida real, los modelos construidos por los algoritmos pueden llegar a ser incorrectos, pues en muchos casos los ambientes son estocásticos y sólo se alcanzan a observar un número limitado de muestras, o también puede deberse a que el modelo fue aprendido empleando funciones de aproximación que generalizan de forma imperfecta, o simplemente porque el ambiente ha cambiado y su nuevo comportamiento no ha sido observado. En los casos en los que el modelo es incorrecto, se espera que el proceso de planeación calcule una política subóptima. Sin embargo, en muchas ocasiones esta política subóptima conduce rápidamente el descubrimiento y corrección de errores en el modelo. Esto suele suceder cuando el modelo es muy optimista, prediciendo mejores recompensas o transiciones que las que en verdad son posibles. La política que se aprendió con ayuda de la planeación tratará de explotar estas oportunidades en el ambiente real, y al hacerlo descubrirá que en realidad no existen.

A continuación se presenta el algoritmo Dyna-Q.

Algoritmo Dyna-Q

Inicializa $Q(s, a) = 0$ y $Model(s, a) = 0 \forall s \in S$ y $\forall a \in A(s)$

Itera para siempre:

- a) Elige un estado inicial s
- b) $a \leftarrow$ acción siguiendo π (política ϵ – avariciosa) para s
- c) Ejecuta la acción a , obtén la recompensa r y el siguiente estado s'
- d) $Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma \max_a Q(s', a) - Q(s, a)]$
- e) $Model(s, a) \leftarrow r, s'$ (Asumiendo un ambiente determinista)
- f) Repite n veces:
 - $s \leftarrow$ estado aleatorio previamente observado
 - $a \leftarrow$ acción aleatoria previamente elegida en S
 - $r, s' \leftarrow Model(s, a)$
 - $Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma \max_a Q(s', a) - Q(s, a)]$

Donde $Model(s, a)$ denota el contenido de la predicción del siguiente estado s' y la recompensa para el par estado-acción (s, a) . El *direct RL* se implementa en el paso d), *model learning* se ejecuta en e), que es donde se actualiza el modelo con base en la experiencia obtenida por el agente, mientras que la planeación está implementada en los pasos correspondientes al ciclo en f). Como se puede observar, el paso con el que aprende el agente actualizando la tabla Q que almacena los valores para cada par estado-acción, es el mismo tanto para *direct RL* (utilizando experiencia real), como para la planeación (utilizando experiencia simulada). Si se omitieran los pasos e) y f) entonces el algoritmo que quedaría sería Q-learning.

En la figura 5.3, se compara el comportamiento de un agente que emplea Dyna-Q con 50 pasos de planeación contra un agente que también emplea Dyna-Q pero con 5 pasos de planeación y

también contra un agente que sólo usa *direct RL* con Q-learning. El agente debe llegar de un estado inicial S a un estado final G en un laberinto. Cada una de las casillas del laberinto representan un estado donde se puede realizar una de 4 posibles acciones: moverse hacia arriba, hacia abajo, a la derecha o a la izquierda. Estos movimientos llevarán al agente de forma determinística al estado vecino correspondiente, excepto cuando el movimiento esté bloqueado por un obstáculo o se quiera salir del laberinto, en cuyo caso el agente permanecerá en el mismo estado. La recompensa que recibirá el agente será de 0 para todas las transiciones, excepto para la transición que conduzca al estado G , en ese caso se recibirá una recompensa de +1. Cuando el agente llegue a G entonces se moverá al agente al estado S para comenzar un nuevo episodio. En esta tarea episódica el factor de descuento fue $\gamma=0.95$ y la tasa de aprendizaje fue $\alpha=0.1$. Todos los valores de los pares estado-acción de Q comenzaron con 0. Se empleó una política ϵ avariciosa, donde $\epsilon=0.1$. Después de cada paso real en el laberinto el agente realizaba n pasos de planeación, en el caso del agente que sólo usó Q-learning $n=0$.

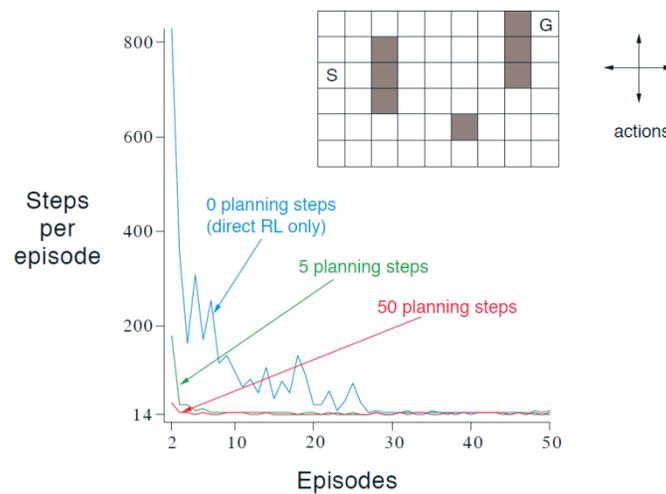


Figura 5.3. Curvas de aprendizaje de cada agente para encontrar el menor número de pasos para llegar de S a G en un laberinto. (Sutton & Barto, 2005).

En la figura 5.3 se observa que después del primer episodio el desempeño mejoró rápidamente para los agentes que utilizaron planeación. El agente que sólo usó Q-learning fue por mucho el que peor desempeño tuvo, ya que le tomó alrededor de 25 episodios llegar a una solución óptima, mientras que al agente con $n=5$ le tomó 5 episodios y al agente con $n=50$ sólo 3 episodios. Este ejemplo ilustra la forma en la que la planeación logra mejorar el desempeño de un algoritmo tabular como Q-learning cuando se trabaja en un ambiente determinístico.

Capítulo 6

Desarrollo

6.1 Aplicación interactiva Penal Madness

Desarrollé una aplicación interactiva para computadora con SO Windows con el motor de videojuegos Unity 5.6.0f3 en el Laboratorio de Investigación y Desarrollo de Aplicaciones Interactivas para la Neuro-Rehabilitación (LANR) del Instituto de Fisiología Celular (IFC) de la UNAM, con la asesoría de una especialista en neuro-rehabilitación. El videojuego lleva el nombre de Penal Madness y su objetivo terapéutico es incrementar el rango de movilidad articular y fuerza muscular del hombro en el plano frontal y sagital. Para la detección de movimientos del jugador se utiliza el sensor Kinect.

6.1.1 Mecánica de juego

En esta aplicación el jugador asume el rol de un portero y debe tratar de detener la mayor cantidad de penaltis posibles tratando de no flexionar el codo durante sus movimientos. Durante el juego aparecerá una diana indicando la dirección en la que se tirará el penal, el jugador tendrá 3 segundos a partir de ese momento para realizar el movimiento correspondiente, ya sea de abducción o de flexión de uno o ambos hombros. Cuando pasen los 3 segundos se lanzará un balón hacia la posición de la diana, dando la ilusión de que lo han pateado con fuerza, un segundo después tanto el balón como la diana desaparecerán. Cuando el jugador mueva sus brazos el avatar que se muestra en pantalla imitará sus movimientos.



Figura 6.1 Se muestran todas las posibles posiciones donde puede aparecer una diana durante el juego. En esta interfaz el terapeuta puede definir la secuencia en la que aparecerán las dianas en cada set.

Unity contiene una herramienta para crear cajas y esferas invisibles que permiten detectar y simular colisiones entre objetos en el ambiente virtual. Los avatares, el balón y la portería contienen estos elementos para ofrecer una experiencia creíble y al mismo tiempo poder detectar si el balón colisionó con una de las manos del avatar, con la red de la portería o con alguna otra parte del cuerpo del avatar.

Una vez que el jugador haya realizado el ejercicio debe regresar a su posición inicial, que consiste en colocar los brazos a los costados, en caso de que el jugador no regrese a esa posición pasados 5 segundos desde el último ejercicio no aparecerá la siguiente diana y se mostrará un mensaje indicándole al jugador que debe colocar sus brazos a los costados para que el juego continúe.

En la figura 6.1 se presentan todas las posiciones en las que puede aparecer una diana en el juego. La captura de pantalla pertenece a la sección donde un terapeuta puede crear una rutina personalizada, cada diana es en realidad un botón; el aspecto visual de las dianas 11 a 25 es diferente al de las dianas 1 a 10 para tratar de que los botones no se superpongan entre sí, pero durante el juego todas las dianas que aparezcan tendrán las mismas dimensiones.

Las dianas 1 a 5 requieren que el paciente realice movimientos de abducción con el hombro izquierdo para detener el balón; las dianas 6 a 10 demandan realizar movimientos de abducción con el hombro derecho; las dianas 11 a 15 se traducen en movimientos de flexión con el hombro izquierdo y las dianas 16 a 20 con el hombro derecho; finalmente, las dianas 21 a 25 también requieren movimientos de flexión de hombro, pero en este caso la tarea debe realizarse con ambos hombros.

Para alcanzar las dianas 1 y 10 se requiere que el jugador ejecute entre 20 y 40 grados de abducción, para las dianas 9 y 2 entre 50 y 70 grados, para detener el balón dirigido a las dianas 3 y 8 se requiere que entre 80 y 100 grados de abducción con el hombro, para las dianas 4 y 7 entre 120 y 140 grados, y para las dianas 5 y 6 entre 150 y 170 grados de abducción. Para las dianas 11, 21 y 16 se necesita una flexión de hombro, o de ambos hombros al mismo tiempo, según el caso, de entre 10 y 40 grados; para las dianas 12, 22 y 17 una flexión entre 50 y 70 grados; para las dianas 13, 23 y 18 una flexión entre 80 y 100 grados, para alcanzar las dianas 14, 24 y 19 se requiere una flexión entre 120 y 140 grados y para las dianas 15, 25 y 20 una flexión entre 150 y 180 grados.

Los movimientos que se practicarán durante el juego son muy importantes para la realización de AVD y son la base para el entrenamiento de movimientos más complejos. Se eligieron ejercicios para el hombro porque el enfoque más habitual en rehabilitación motora de pacientes EVC es comenzar por la recuperación de los miembros proximales antes que los distales (Figuroa et al., 2014), de hecho, se podría considerar que la base de este tipo de ejercicios siguen los principios del método de Bobath (Michielsen et al., 2019), que persigue la recuperación e incremento de la fuerza muscular de un miembro con base en la repetición de ejercicios y combatiendo mal adaptaciones. En la mecánica de juego se incluyen tareas bimanuales, ya que el entrenamiento de este tipo de tareas promoverá que se puedan volver a realizar AVD más pronto, esto debido a que las actividades bimanuales propicia una interacción entre ambos hemisferios cerebrales entre

redes neuronales similares, lo cual en teoría debería maximizar la recuperación funcional (Sleimen-Malkoun et al., 2011).

El avatar imitará los movimientos que realice el paciente con sus brazos y tronco, pero no imitará los movimientos de las piernas ni se moverá de lugar durante el juego, esto se implementó principalmente porque el enfoque de la aplicación es sobre los miembros superiores. Las ventajas del Kinect en este tipo de desarrollos se mencionan con más detalle en el capítulo 4 (Rehabilitación virtual), en este caso se eligió este sensor para mapear las acciones del paciente al avatar debido a su precisión para la detección de movimientos gruesos con los brazos, al hecho de ser un sensor óptico no invasivo (no requiere marcadores) y de su fácil integración con Unity.

Durante el juego se escuchará el sonido de una audiencia de un estadio, para incrementar el grado de inmersión. Se reproducirá un sonido cuando se dispare el balón y tres diferentes tipos de sonido para cada acción que realice el paciente. Un movimiento se considerará correcto cuando la mano correspondiente del avatar colisione con el balón, deteniendo el penalti, cuando esto suceda el brazo con el que se realizó el movimiento no debe estar flexionado, en el caso de las tareas bimanuales el balón debe colisionar con ambas manos al mismo tiempo y ambos brazos deben estar completamente extendidos. Si lo anterior se cumple la audiencia gritará una porra de alegría. Si el paciente logra detener el tiro, pero lo realiza flexionando el codo de uno o ambos brazos, se considerará como un falso acierto y se reproducirá un sonido como si la audiencia aplaudiera. Finalmente, se considerará como un fallo cuando el paciente no logre detener el tiro y el balón colisione con la red de la portería o con alguna parte del cuerpo del avatar que no sean las manos, en ese caso el marcador del fondo aumentará indicando que han anotado un gol y la audiencia reproducirá un sonido similar al que hacen cuando un jugador está cerca de meter gol pero falla. En el caso de los ejercicios bimanuales, se considerará como un fallo si alguna de las manos del avatar no colisiona con el balón. El sonido de la audiencia se puede desactivar a través de un botón, sin embargo, los sonidos de retroalimentación para los éxitos o fracasos no se pueden omitir, ya que se consideran una parte importante para ofrecer una retroalimentación multimodal.



Figura 6.2 Se ilustra la forma correcta en la que el paciente debe realizar el ejercicio de abducción para detener el penal. Constantemente se indica el marcador de la partida, cuántos ejercicios se han realizado y cuántos restan.

El sensor Kinect proporciona la orientación y la posición de todas las articulaciones que detecta con una frecuencia máxima de 30Hz. Se considera que el paciente flexiona el codo cuando el ángulo formado entre los segmentos codo-muñeca y codo-hombro es mayor a 15 grados. Cuando el juego detecta que el paciente está doblando el brazo se muestra un mensaje grande en pantalla indicando que se está realizando un movimiento incorrecto. Flexionar el codo se considera un movimiento incorrecto porque no se está cumpliendo cabalmente el objetivo de ejercitar el hombro, si esto se permitiera la mitad de los ejercicios se podrían realizar sin movilizar el hombro.

Utilizar una perspectiva en tercera persona en este juego permite una retroalimentación inmediata, para el participante cuando realiza movimientos gruesos de miembros superiores, como el movimiento de los brazos, lo que facilita la auto-corrección de errores durante la ejecución de los ejercicios (Trabucco et al., 2019). Además, el hecho de que el participante observe a un personaje humanoide imitando sus movimientos puede influir en la activación de las neuronas espejo (Carvalho et al., 2013), favoreciendo la reorganización cortical y por consecuencia apoyando en la recuperación funcional.



Figura 6.3 Un mensaje especial aparece cuando el paciente realiza un movimiento incorrecto. Detener un balón con el codo flexionado se considera un falso acierto. Adicional a este mensaje se reproducirá un sonido característico.

Este videojuego está dirigido tanto a pacientes EVC subagudos, como a pacientes crónicos que presenten una espasticidad en el codo con puntaje de máximo de 1+ en la Escala de Ashworth Modificada o bien, que tengan la capacidad de realizar al menos 20 grados de abducción y flexión de hombro, pues esto les permitiría realizar los ejercicios más básicos del juego y eventualmente progresar a aquellos que requieren mayor esfuerzo físico. Es preferible que el paciente tenga la suficiente fuerza en las piernas y control de tronco para mantenerse de pie mientras juega, pero si esto no fuera posible el Kinect puede detectar de forma aceptable los movimientos de los brazos incluso si el paciente se encuentra sentado.

Se cuidó que los modelos, la perspectiva de la cámara y la iluminación logran recrear un ambiente virtual creíble, pero sin ser demasiado realista para no entrar en una zona aversiva como el *uncanny valley* (Mori et al., 2012). Cabe resaltar que todos los elementos gráficos están optimizados, los modelos de personas que simulan a la audiencia contienen menos de 1000 polígonos, y toda la escena de juego en su conjunto tiene 108,800 polígonos y 87,300 vértices, además, no se cuenta con ningún efecto de iluminación o *shader* que sean computacionalmente

demandantes, por lo que este programa puede ejecutarse incluso en computadoras que no cuenten con tarjeta gráfica dedicada. Los requerimientos mínimos que necesita tener la computadora son 4 GB de memoria RAM, procesador Intel Core i3 de 3.6 GHz y 5 GB de espacio libre en el disco duro.

El juego ofrece los 3 tipos de motivación que se mencionaron en el capítulo 3 (Rehabilitación posterior a un EVC). Se ofrece conocimiento de resultados a través de una opción donde se puede acceder a una gráfica en pantalla que muestra el progreso del paciente en el juego y a los detalles de cada partida que se ha jugado (figura 6.4), se proporciona conocimiento del desempeño de los movimientos durante el juego al ofrecer retroalimentación auditiva y visual que indica si el movimiento es correcto o no, y se brinda retroalimentación motivacional mediante mensajes de ánimo que aparecen después de que termina cada set y cuando termina la partida (figura 6.5).

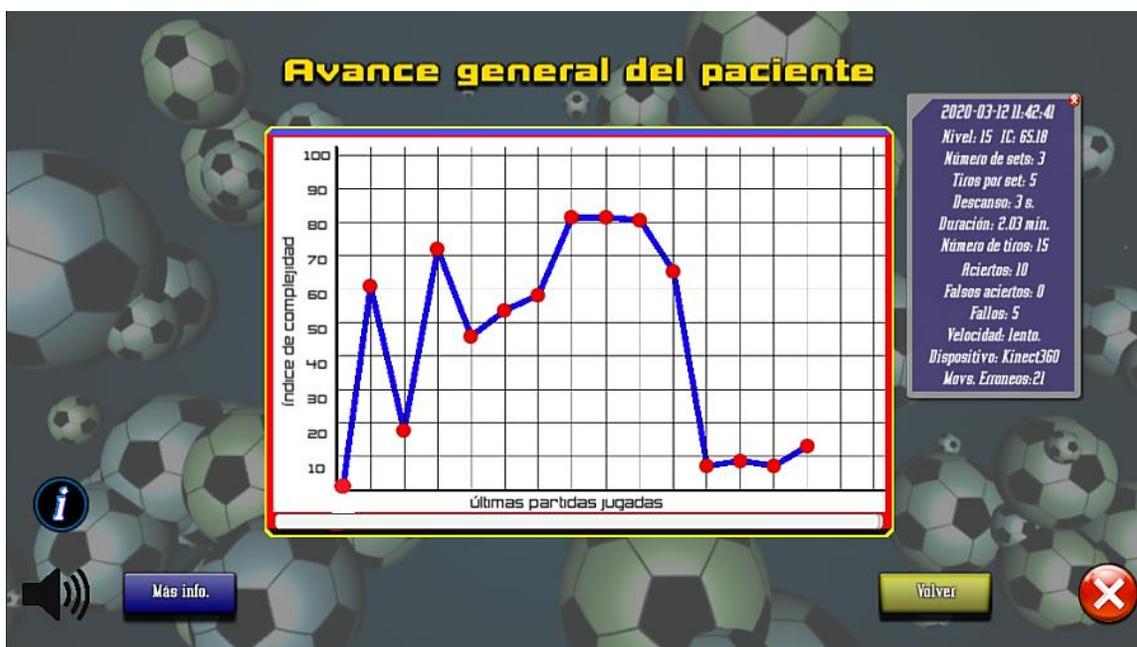


Figura 6.4 Pantalla que muestra la forma en la que se puede consultar el progreso del paciente. La gráfica muestra las últimas 15 partidas que se han jugado, representadas por botones rojos, los cuales al darles clic muestran un panel con información acerca de la partida.



Figura 6.5 Una vez que la partida termina aparecerá un efecto de confeti para celebrar que el paciente ha concluido los ejercicios.

Antes de poder usar la aplicación se deben registrar los datos del paciente en una base de datos a través de un portal llamado CITAN (<https://lanr.ifc.unam.mx/citan/public/login>), en donde se le asignará un número de identificación (ID). Con el fin de poder ofrecer tele-rehabilitación a través de este juego serio, la aplicación cuenta con tres interfaces diferentes: a) una pensada para que el terapeuta le asigne ejercicios al paciente en un hospital o clínica, b) una dirigida únicamente a los terapeutas y c) una interfaz pensada para que el paciente pueda usar en casa. Las credenciales que se ingresen en la pantalla de inicio (figura 6.6) determinarán la interfaz del juego a la que se accederá.



Figura 6.6 Pantalla de inicio de sesión del juego Penal Madness, desde donde se podrán ingresar las credenciales del paciente o terapeuta para iniciar sesión.

6.1.2 Interfaz de clínica

Para usar esta interfaz se asume que el paciente está siendo acompañado por un terapeuta para asignarle y asistirle en la realización de sus ejercicios. Para ingresar a la interfaz de clínica se debe ingresar el ID del paciente y una contraseña compuesta por su fecha de nacimiento (YYYYMMDD). Para verificar que los datos ingresados son correctos la computadora debe contar con conexión a Internet. Una vez que se verifiquen los datos aparecerá un panel que le preguntará al paciente si se encuentra acompañado por un terapeuta, a lo que debe responder afirmativamente. Posteriormente aparecerá un panel donde el terapeuta debe ingresar su contraseña y correo tal cual y como se registraron en la plataforma CITAN (figura 6.7). De esta forma se asegura que quien asignará los ejercicios será un terapeuta registrado y no el mismo paciente o alguna persona sin conocimientos sobre el juego. En caso de que la computadora no cuente con conexión a Internet el terapeuta deberá ingresar unas credenciales especiales para este tipo de eventos que le permitirán acceder al menú principal del juego.



Figura 6.7 Paneles que se muestran cuando se desea ingresar a la interfaz de clínica.

Una vez que se ha accedido al menú principal de la interfaz de clínica (figura 6.8) se podrán elegir varias opciones:

- Crear una rutina de ejercicios personalizada.
- Cargar una rutina de ejercicios.
- Jugar empleando un módulo de ADD.
- Elegir un personaje
- Revisar un tutorial animado que le enseñará al paciente a jugar.
- Consultar los datos estadísticos de las partidas que el paciente ha jugado.

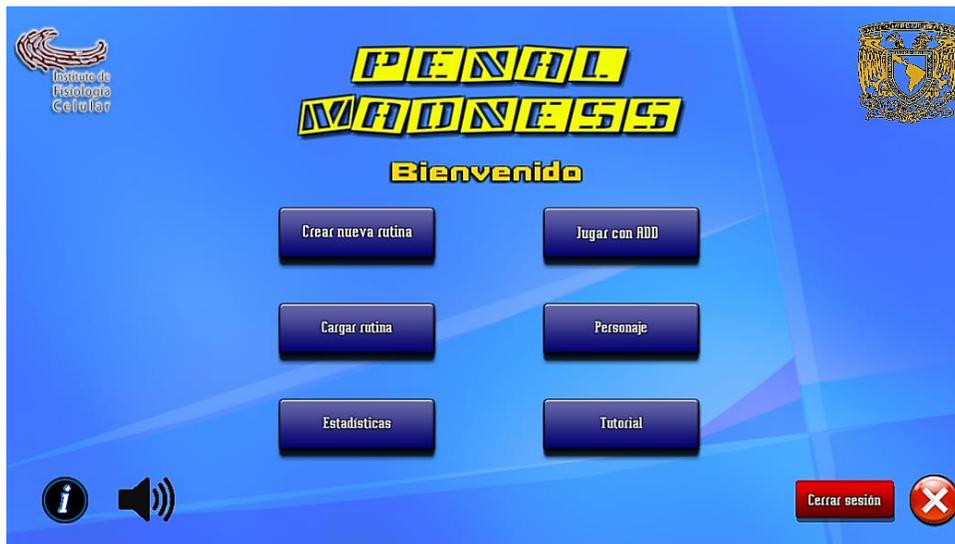


Figura 6.8 Menú principal de la interfaz de clínica.

La opción para crear una rutina personalizada le permitirá al terapeuta diseñar una secuencia de movimientos específicos, de tal forma que la rutina cubra las capacidades y necesidades que el terapeuta haya identificado en el paciente. Para crear una nueva rutina se debe asignar un nombre que no se haya utilizado antes, así como definir el número de sets que tendrá la rutina y el tiempo de descanso que existirá entre cada set (figura 6.9).



Figura 6.9 Pantalla inicial para la creación de una rutina personalizada.

El siguiente paso en la creación de la rutina es definir la secuencia de movimientos que deberá realizar el paciente en cada set de ejercicios. Para ello se presentará una interfaz en donde cada una de las dianas se representará con un botón, al dar clic sobre la diana aparecerá su número correspondiente sobre un cuadro de texto. Conforme se seleccionen las dianas se visualizará en el cuadro de texto el orden en la que aparecerán las dianas durante el juego, como se muestra en la figura 6.1.



Figura 6.10 Interfaz donde el terapeuta puede elegir una rutina personalizada creada previamente para que juegue el paciente.

El límite de movimientos que se pueden asignar para un solo set es 20 y el número máximo de sets que puede tener una rutina son 5. Una vez que se finalice con la creación de la rutina el terapeuta podrá opcionalmente agregar un mensaje adicional que ayude a identificar su propósito, al terminar el proceso la rutina se almacenará sólo de forma local como un archivo en formato XML.

Desde el menú principal, al seleccionar la opción “Cargar rutina” se mostrará una interfaz con una lista desplegable que contendrá los nombres de todas las rutinas que se han creado en esa computadora, el terapeuta podrá elegir alguna de ellas y presionar el botón “Siguiente” para que el paciente pueda jugar, lo anterior se ilustra en la figura 6.10.

La opción “Personaje” permite elegir al portero que se mostrará durante el juego para ayudar a que el jugador se identifique con el avatar en pantalla. Los modelos se obtuvieron del programa de acceso libre MakeHuman (Bastioni, 2000), una aplicación para la creación de personajes humanoides.



Figura 6.11 Interfaz donde se puede seleccionar un avatar de un hombre o mujer.

En la opción de “Tutorial” se le mostrará una animación tipo película dividida en secciones, que se podrá controlar a través de los botones “Atrás” y “Siguiente”. El objetivo de este tutorial es que el paciente aprenda cómo se debe jugar, el tipo de retroalimentación que recibirá, los diferentes movimientos que se deben realizar dependiendo de la posición donde aparezca la diana, las condiciones ideales que debe tener el espacio donde se jugará, incluyendo la distancia a la que se debe de colocar el jugador del Kinect y las condiciones de iluminación (figura 6.12).



Figura 6.12 Se muestran algunas animaciones del tutorial, que indican las diferentes dianas que aparecerán durante el juego y el tipo de movimiento que se debe realizar para detener el balón. Además de las condiciones del espacio donde se realizarán los movimientos que se deben cumplir para poder jugar.

En la opción de “Estadísticas” se podrá consultar el progreso del paciente. Se muestra una gráfica en pantalla con información sobre las últimas 15 partidas que ha jugado el paciente (figura 6.4). El índice de complejidad de cada partida está determinado por la proporción entre aciertos y el número total de ejercicios. Al dar clic sobre alguno de los puntos rojos de la gráfica (los cuales

representan las partidas), aparecerá un panel que mostrará información más específica sobre la partida que se jugó como: fecha, número de sets, número de ejercicios en cada set, duración total de la partida, tiempo de descanso entre sets, la secuencia de ejercicios realizada en cada set, número de aciertos, número de fallos, número de movimientos incorrectos y número de falsos aciertos. De esta forma se pueden mostrar varios paneles al mismo tiempo y comparar la información de las partidas.

Después de que se haya jugado una partida, toda la información respecto a ella se almacena en un archivo XML que contendrá los resultados de todas las partidas que haya jugado el paciente, funcionando como una especie de historial; asimismo, en caso de contar con conexión a Internet, la información también se almacena en una base de datos en el servidor CITAN.

6.1.3 Interfaz del terapeuta

Esta interfaz está pensada para que, en conjunto con la interfaz del paciente, se pueda realizar tele-rehabilitación. Desde esta interfaz el terapeuta puede crear una rutina personalizada, asignársela de forma remota a algún paciente y consultar el progreso de los pacientes a su cargo. Para poder acceder el terapeuta debe ingresar en la pantalla de inicio de sesión su correo electrónico y la contraseña con las que fue dado de alta en el sistema CITAN. Esta es la única interfaz en la que es obligatorio contar con conexión a Internet para poder ingresar.

Una vez que el terapeuta ha iniciado sesión (figura 6.13) se descargarán automáticamente desde el servidor CITAN todas las rutinas personalizadas que el terapeuta haya creado y todos los registros de las partidas que han jugado los pacientes que se encuentran bajo su supervisión.

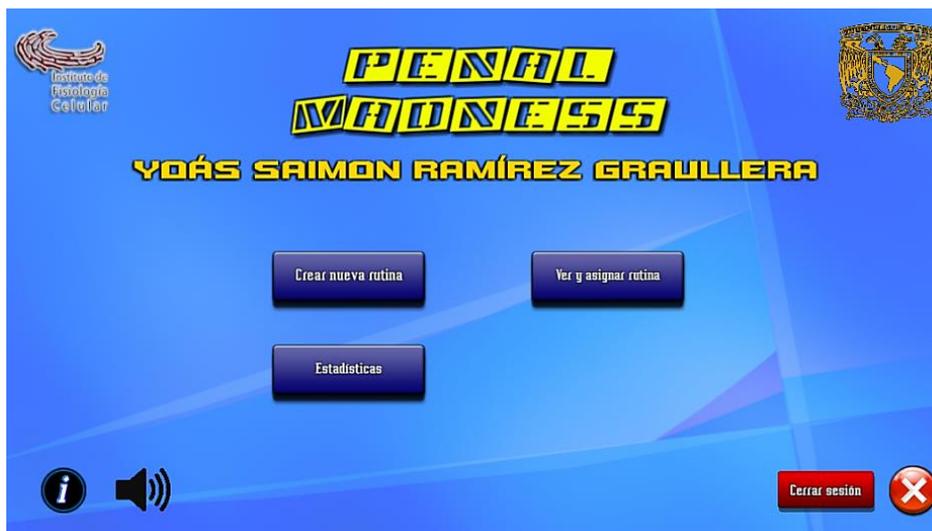


Figura 6.13 Menú principal de la interfaz de terapeuta.

La opción “Ver y asignar rutina” le permite al terapeuta seleccionar alguna de las rutinas que ha creado, visualizar esa rutina y asignársela a algún paciente (figura 6.14 y figura 6.15).

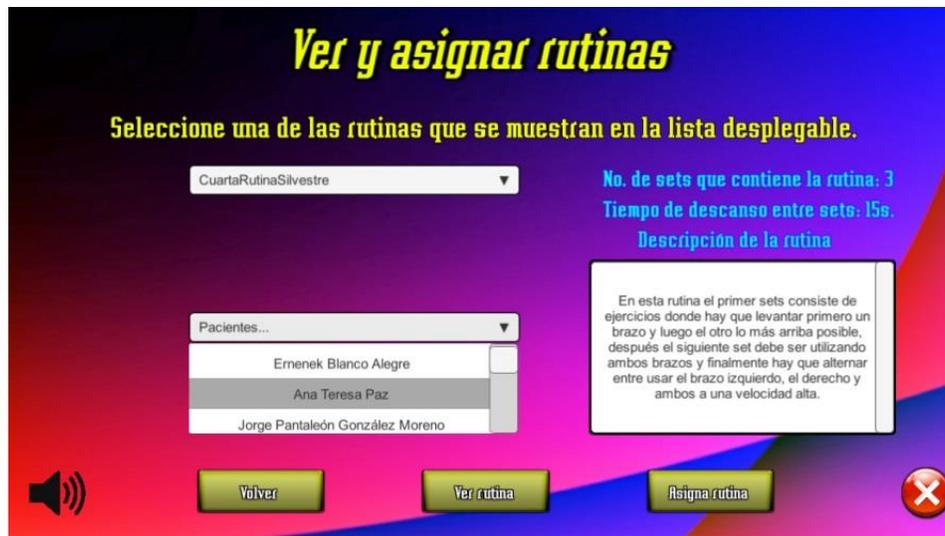


Figura 6.14 Pantalla que muestra una lista desplegable de todas las rutinas que ha creado el terapeuta, así como una lista de todos los pacientes que tiene a su cargo. Al seleccionar una rutina del lado derecho se muestra una descripción sobre ella.



Figura 6.15 Pantalla que aparece una vez que se ha asignado con éxito una rutina.

La opción para crear una rutina personalizada en la interfaz del terapeuta es la misma que la mostrada en la interfaz de clínica.

Cuando se selecciona la opción “Ver rutina”, se cargará una escena donde se reproducirá toda la rutina como si alguien estuviera jugando, aunque el personaje que se muestra en pantalla no se moverá (figura 6.16). La velocidad a la que se puede visualizar la rutina puede aumentar hasta 5 veces respecto a la velocidad normal.



Figura 6.16 Interfaz que se muestra al seleccionar la opción “Ver rutina”.

En la opción “Estadísticas”, el terapeuta puede elegir en una lista desplegable a alguno de los pacientes a su cargo y visualizar la misma pantalla de resultados que se mostraría en la interfaz de clínica (figura 6.17).



Figura 6.17 Se muestra como el terapeuta puede seleccionar al paciente del cuál desea consultar su progreso, cuando se haya elegido al paciente se mostrará una gráfica para ilustrar su progreso.

6.1.4 Interfaz del paciente

Esta interfaz está pensada para que el paciente pueda realizar sus ejercicios desde casa, para ello deberá contar con una computadora con el programa Penal Madness instalado y con el sensor Kinect. Para ingresar a esta interfaz el paciente debe ingresar su ID y una contraseña que se conforma su fecha de nacimiento (YYYYMMDD), posteriormente se debe seleccionar la opción “No” cuando aparece el panel preguntado si se encuentra asistido por un especialista. Para verificar que los datos ingresados de inicio de sesión son correctos la computadora debe contar con conexión a Internet, en caso contrario se permitirá acceder al menú de la interfaz del paciente, pero todas sus partidas se guardarán en un archivo temporal; cuando el programa detecte que existe conexión a Internet se le pedirá al paciente ingresar su contraseña, de lo contrario el archivo temporal se eliminará. En cuanto el paciente inicia sesión se descargará desde el servidor CITAN la última rutina que su terapeuta le haya asignado. Si el paciente ya contaba con una rutina asignada el archivo de la rutina más reciente sobre escribirá al antiguo. Las opciones que se podrán elegir en esta interfaz son consultar los resultados de las partidas jugadas, acceder

al tutorial del juego, seleccionar un personaje y jugar la rutina asignada por el terapeuta (figura 6.18).



Figura 6.18 Cuando el paciente cuenta con una rutina asignada se habilitará el botón jugar. Al seleccionar esta opción se mostrará una pantalla donde aparecerán algunos detalles sobre la rutina y el botón para jugar.

En caso de que el paciente no cuente con alguna rutina asignada no podrá jugar (figura 6.19). Sin embargo, se planea integrar en esta interfaz el módulo de ADD para que el paciente pueda jugar, aunque no cuente con una rutina.



Figura 6.19 Cuando el paciente no cuenta con una rutina asignada aparecerá un mensaje de advertencia y se deshabilitará el botón para jugar.

6.2 Penal Madness con aprendizaje por reforzamiento

La implementación de un algoritmo de ADD en Penal Madness es un elemento muy importante para el juego, ya que de esta manera la aplicación interactiva podría adaptarse a la evolución de cada paciente, de forma personalizada y con la oportunidad que se requiere, sin la necesidad de que un terapeuta esté pendiente de asignar la dificultad adecuada del juego, por ejemplo, en la modalidad de tele-rehabilitación. No obstante, la participación del terapeuta aún será necesaria para supervisar el progreso del paciente, corregir movimientos no deseados y también cuando se busque trabajar con ejercicios muy específicos, en cuyo caso se tendría que asignar la rutina particular correspondiente.

Diseñé un algoritmo basado en Q-learning y Dyna-Q para implementar un ADD en el juego. El algoritmo, al igual que Q-learning, será un método tabular que tratará de optimizar la función de valor de acción $Q(S, A)$ para obtener indirectamente la política óptima, π^* . Para alcanzar más

rápido esta meta se realizará un proceso de planeación empleando un modelo del ambiente construido con base en la experiencia del agente.

6.2.1 Definición de la función de valor estado-acción $Q(S,A)$

El primer paso para diseñar el algoritmo de RL fue abordar el problema bajo el marco MDP, una partida del juego se puede ver como un episodio que contiene un conjunto finito de estados, un único estado inicial y un conjunto finito de estados terminales, que corresponden al fin de la partida. En este MDP la acción que realizará el agente será la elección del ejercicio que deberá realizar el jugador, la manera en la que el jugador realice el ejercicio determinarán la recompensa y el siguiente estado que recibirá el agente, es decir, en este MDP el jugador será el ambiente. Como el jugador será un paciente con secuelas motoras se considera que el ambiente es estocástico hasta cierto punto, pues algunos ejercicios los podrá realizar siempre de la misma forma, pero existirán otros que en algunas ocasiones realizará correctamente y en otras no. Adicionalmente, se trata de un ambiente secuencial, ya que las acciones de los episodios anteriores afectarán la forma en la que se tomen las acciones en futuros episodios. A continuación se mencionarán el conjunto de acciones y estados que pueden existir en un episodio del juego.

Una de las secuelas más comunes del EVC es la hemiparesia, por esta razón se propuso que cada partida esté conformada de 10 ejercicios, de los cuáles 6 deberán realizarse del lado parético, 2 del lado no parético y 2 deberán ser tareas bimanuales. Por ejemplo, si el hemicuerpo afectado de un paciente fuera el lado izquierdo, entonces el juego le asignaría 6 ejercicios del lado izquierdo, 2 ejercicios bimanuales y 2 ejercicios dirigidos al brazo derecho. Es importante enfocarse en el hemicuerpo afectado, pero también se deben practicar tareas bimanuales por los beneficios que aporta a la recuperación motora, por otra parte, se decidió realizar 2 ejercicios del lado no afectado para añadir variabilidad al juego y ejercitar también ese miembro en caso de que tuviera algún grado de afección, aunque no es común.

Como se mostró en la parte de la mecánica del juego, se cuenta en total con 25 diferentes posiciones en las que puede aparecer una diana, que es el punto hacia donde el paciente tiene que mover su brazo para detener el balón.

Para identificar la acción de que aparezca una diana, su posición y el movimiento que exige, se empleó una nomenclatura tal que I se refiere a un movimiento de abducción con el brazo izquierdo, IF a un movimiento de flexión con el brazo izquierdo, B como un movimiento de flexión con ambos brazos hacia la línea media, D como un movimiento de abducción con el brazo derecho y DF como un movimiento de flexión con el brazo derecho. El número asociado a cada acción indica la altura a la que aparecerá la diana, siendo 1 la más baja y 5 la más alta. Las dianas de menor altura son más sencillas de alcanzar, pues requieren una menor activación de músculos, y las más altas las más difíciles, ya que requieren el mayor esfuerzo físico.



Figura 6.20 Se muestran las diferentes posiciones en las que puede aparecer una diana y su correspondiente valor asignado, todas las dianas de una misma altura reciben el mismo valor. Entre mayor rango de movimiento se necesite para alcanzar las dianas, mayor será su valor.

La acción que tomará el agente será la diana a mostrar. Como se muestra en la figura 6.20, el número total de acciones será de 25. Sea A el conjunto de acciones que el agente puede tomar en una partida entonces:

$$A = \{D1, D2, D3, D4, D5, DF1, DF2, DF3, DF4, DF5, B1, B2, B3, B4, B5, I1, I2, I3, I4, I5, IF1, IF2, IF3, IF4, IF5\} \quad (6.1)$$

Cada episodio del algoritmo será una partida del juego y estará compuesto por 10 ejercicios, por lo tanto, el episodio estará dividido en 10 tiempos. El valor del tiempo actual de la partida se representará con $t \in \{0, 1, \dots, 9\}$

La primera acción a que tomará el agente en cada episodio siempre será aleatoria, pero con la condición de que $Q(s, a) > 0$, de tal forma que siempre se comience con un ejercicio que el paciente pueda realizar correctamente, además, esta mecánica ayuda a que la secuencia de dianas cambie y no siempre sea la misma, de esta forma se evitará la monotonía y la habituación. Si ninguna acción cumple la condición anterior entonces se elegirá de forma aleatoria.

Durante el juego se tendrá un registro del número de movimientos exitosos y el número de fallos. Sea $score$ el conjunto de valores que se pueden obtener al calcular la diferencia entre éxitos y fallos en cada momento t de la partida.

$$score = \{-10, \dots, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} \quad (6.2)$$

Sea $score_t \in score$ el valor en el tiempo t de la operación *Exitos* – *Fallos*

El estado proporcionará información sobre cuántos tiros se han realizado del lado menos afectado, del lado parético y en el centro. También indicará el valor de $score_t$, con esta información el agente podrá tomar una acción, de esta forma se busca cumplir la *propiedad de Markov*. En caso de que en un tiempo t de un episodio el agente ya haya agotado el número

máximo de acciones permitidas para una tarea bimanual, del lado parético o del lado menos afectado, en los episodios subsecuentes ya no podrá asignar ese tipo de ejercicios.

Para el final del episodio se deben haber cumplido 3 condiciones:

- 2 movimientos bimanuales.
- 2 movimientos del lado no parético.
- 6 movimientos del lado parético.

Sin importar el valor $score_t$, alguna de estas 3 condiciones se han cumplido o no, entonces, las posibles combinaciones en las que se pueden cumplir son 8, por lo tanto, si se incluyeran los 21 posibles valores de $score_t$ se tendrían 168 estados diferentes ($21 \times 2 \times 2 \times 2$). Sin embargo, los estados terminales no son de interés, es decir, cuando las 3 condiciones se han cumplido y $score_t = 10$, ya que en ese momento se realizará ninguna acción. Además, se considerará que es tan malo que el paciente tenga más de 6 aciertos, como que tenga más de 5 fallos durante el episodio, por ello de los 21 valores de $score_t$ sólo será de interés si $score_t = 0$, $score_t = 1$, $score_t = 2$, $score_t > 2$ ó $score_t < 0$. Por lo tanto, el número de estados que se considerarán para definir la tabla Q serán 35. Los primeros 12 estados de Q se muestran en la tabla 6.1.

Tabla 6.1 Se muestran los primeros 12 estados de la tabla Q para el algoritmo de RL. Cada una de las acciones en cada estado tendrán su propio valor que podrá ser accedido a través de la función $Q(S, A)$.

score	Límite alcanzado Parético	Límite alcanzado Centro	Límite alcanzado Sano	$Q(S,A)$
0	False	False	False	{1:0,2:0,3:0,4:0,5:0,D1:0,D2:0,D3:0,D4:0,D5:0,B1:0,B2:0,B3:0,B4:0,B5:0,IF1:0,IF2:0,IF3:0,IF4:0,IF5:0,DF1:0,DF2:0,DF3:0,DF4:0,DF5:0}
0	False	False	True	{1:0,2:0,3:0,4:0,5:0,IF1:0,IF2:0,IF3:0,IF4:0,IF5:0, B1:0, B2:0, B3:0, B4:0, B5:0}
0	False	True	False	{1:0,2:0,3:0,4:0,5:0,D1:0,D2:0,D3:0,D4:0,D5:0,IF1:0,IF2:0,IF3:0,IF4:0,IF5:0,DF1:0,DF2:0,DF3:0,DF4:0,DF5:0}
0	True	False	False	{D1:0,D2:0,D3:0,D4:0,D5:0,DF1:0,DF2:0,DF3:0,DF4:0,DF5:0,B1:0,B2:0,B3:0,B4:0, B5:0}
0	False	True	True	{1:0,2:0,3:0,4:0,5:0, IF1:0,IF2:0,IF3:0,IF4:0,IF5:0}
0	True	True	False	{D1:0,D2:0,D3:0,D4:0,D5:0, DF1:0,DF2:0,DF3:0,DF4:0,DF5:0}
0	True	False	True	{B1:0,B2:0,B3:0,B4:0,B5:0}
0	True	True	True	Estado terminal, no se incluye en Q
1	False	False	False	{1:0,2:0,3:0,4:0,5:0,D1:0,D2:0,D3:0,D4:0,D5:0,B1:0,B2:0,B3:0}
1	False	False	True	{1:0,2:0,3:0,4:0,5:0,D1:0,D2:0,D3:0,D4:0,D5:0,B1:0,B2:0,B3:0,B4:0,B5:0,IF1:0,IF2:0,IF3:0,IF4:0,IF5:0,DF1:0,DF2:0,DF3:0,DF4:0,DF5:0}
1	False	True	False	{1:0,2:0,3:0,4:0,5:0,IF1:0,IF2:0,IF3:0,IF4:0,IF5:0,B1:0,B2:0,B3:0,B4:0,B5:0}
1	True	False	False	{1:0,2:0,3:0,4:0,5:0,D1:0,D2:0,D3:0,D4:0,D5:0,IF1:0,IF2:0,IF3:0,IF4:0,IF5:0,DF1:0,DF2:0,DF3:0,DF4:0,DF5:0}

La función $Q(S,A)$ se representa computacionalmente como una matriz, donde los renglones representan todos los posibles estados S en los que se puede encontrar el agente y las columnas son todas las posibles acciones A que puede tomar el agente en ese estado. Entre mayor sea el valor de la celda significa que en un plazo determinado esa acción otorgará mayores recompensas. La longitud de este plazo de tiempo se modula con el parámetro γ , el factor de descuento; si esta constante tiene un valor de 0 al agente sólo le interesarán las recompensas inmediatas que otorguen las acciones A , entre mayor sea el valor del factor de descuento el agente considerará las recompensas de la acción a más largo plazo.

6.2.2 Definición de la función de recompensa

Con el fin de que el juego no resulte demasiado fácil y aburra al jugador, ni tan difícil, de forma que lo frustre, se propuso que el marcador ideal para que el paciente entre en un “*flow channel*” (Csikszentmihalyi, 1991) sea conseguir 6 aciertos de 10 ejercicios. De acuerdo con un estudio donde se realizó un experimento con un videojuego al cual se le modificaba la dificultad (Hagelbäck & Johansson, 2009), se encontró que el margen entre éxitos y fallos debe ser pequeño para que el jugador disfrute el juego. Este hallazgo se alinea con las propuestas de otros estudios que se presentaron en el capítulo 4 (Rehabilitación virtual), como el de Yoones (Sekhavat, 2017), en donde una de las metas de su algoritmo MPRL para alcanzar el mayor grado de disfrute es conseguir $|Exitos - Fallos| = 0$; mientras que en el proyecto RGS de Cameirão (Cameirão et al., 2010), el algoritmo incrementaba la dificultad del juego sólo si el paciente lograba más del 70% de aciertos y la disminuía si obtenía menos del 50%, en cualquier otro caso se asumía que la dificultad era óptima y el paciente estaba en un estado de máximo rendimiento. Considerando lo anterior, para este juego se optó por que el algoritmo RL buscara obtener una diferencia de 2 puntos positivos en la operación $Exitos - Fallos$, con el fin de no desmotivar a los pacientes y tener al menos un acierto más que un fallo. Las dianas que el paciente no pueda alcanzar (fallos), deben ubicarse en una posición muy cercana a la de aquellas dianas donde sí puede detener el balón (aciertos), de esta forma, el algoritmo de RL lograría que el paciente alcance un grado de “frustración placentera” (Koster, 2005) y de máximo esfuerzo (Hocine et al., 2015). Cabe resaltar que el algoritmo también considerará los “falsos aciertos” como fallos.

El agente siempre va a tratar de obtener la mayor recompensa posible. Por esta razón es muy importante definir una función de recompensa que mejor represente el comportamiento que buscamos tenga el agente:

- Al final de la partida el jugador debió conseguir 6 aciertos de 10 ejercicios, es decir, se debe cumplir que $Exitos - Fallos = 2$.
- Los ejercicios asignados con la intención de que el jugador falle deben ubicarse al límite de sus capacidades.

Sea $a_t \in A$ la acción que el agente ha tomado (la diana que ha mostrado en pantalla) en el tiempo t del episodio.

Sea $score_{t+1}$ la diferencia entre éxitos y fallos resultado de haber elegido la acción a_t .

La función de recompensa r estará definida a su vez por las funciones f_1 y f_2 :

$$r_{t+1}(a_t, score_{t+1}) = f_1(a_t) + f_2(score_{t+1}) \quad (6.3)$$

La función f_1 se enfocará en que la asignación de las dianas sea variada y se encuentre dentro de las capacidades del paciente, mientras que la función f_2 se buscará lograr que al final de la partida el paciente haya logrado 6 aciertos y 4 fallos.

Sea la constante V_a el valor de la acción a asociado al esfuerzo que necesita realizar el paciente para alcanzar la diana a a causa de los grados articulares que se debe rotar el hombro.

$$V_a \in \{10,20,30,40,50\} \quad (6.4)$$

Entre más esfuerzo anti gravitatorio requiera el movimiento que debe realizar el paciente para alcanzar la diana a mayor será el valor de V_a .

$$V_{D1} = V_{DF1} = V_{B1} = V_{I1} = V_{IF1} = 10 \quad (6.5)$$

$$V_{D2} = V_{DF2} = V_{B2} = V_{I2} = V_{IF2} = 20 \quad (6.6)$$

$$V_{D3} = V_{DF3} = V_{B3} = V_{I3} = V_{IF3} = 30 \quad (6.7)$$

$$V_{D4} = V_{DF4} = V_{B4} = V_{I4} = V_{IF4} = 40 \quad (6.8)$$

$$V_{D5} = V_{DF5} = V_{B5} = V_{I5} = V_{IF5} = 50 \quad (6.9)$$

En la figura 6.20 se ilustran los valores de las diferentes acciones (dianas).

Sea n_a el número de veces que el agente ha asignado la diana a durante la partida y c una constante de penalización, se define la función $f_1(a_t)$ como:

$$f_1(a_t) = \begin{cases} \frac{V_{a_t}}{n_{a_t}}, & \text{Logró detener el tiro} \\ -V_{a_t} * c, & \text{Falló al detener el tiro} \end{cases} \quad (6.10)$$

Si el jugador logró detener correctamente el penal dirigido a la diana a_t , entonces $f_1(a_t)$ estará disminuido en razón de la cantidad de veces que el agente ha asignado la diana a_t , de manera que el agente no elija siempre la misma diana con mayor valor que sabe que el paciente puede detener. Para evitar que el agente asigne ejercicios que le resulten imposibles al paciente se aplica una penalización del siguiente modo, si el paciente no realizó correctamente el ejercicio correspondiente a la diana asignada a_t , entonces $f_1(a_t)$ tomará un valor negativo, de tal forma que un fallo asociado a una diana con un alto valor se volverá una gran pérdida para el agente. Con esto se busca que el agente sea más cauteloso al momento de elegir los ejercicios en los que el paciente debe fallar, procurando elegir aquellos que produzcan la menor pérdida, es decir, de todo el universo de acciones que el paciente no puede realizar, se tratarán de asignar aquellas de menor altura.

Cada episodio que juegue el paciente estará compuesto por 10 ejercicios. El objetivo del agente es regular el juego para que no sea ni muy fácil, ni muy difícil. La hipótesis es que si $score_{10} = 2$

entonces el jugador tuvo el suficiente grado de reto durante la partida. Por ello, el mayor valor de $f_2(score_{t+1})$ se dará cuando $score_{t+1}=2$, de tal forma que se le incentive al agente a alternar entre una diana que sabe que el paciente no podrá alcanzar y una que sabe que si podrá. Poner muchas dianas consecutivas de fallo o éxito provocará que $score_{t+1}$ se aleje del valor de 2.

$$f_2(score_{t+1}) = \begin{cases} 200, & score_{t+1} = 2 \\ 100, & score_{t+1} = 1 \\ 0, & score_{t+1} = 0 \\ -1000, & otherwise \end{cases} \quad (6.11)$$

En las tablas 6.2 y 6.3 se muestran ejemplos de cómo se comportaría la función f_2 de acuerdo con el comportamiento del paciente. Como se puede observar, los valores más altos se obtendrán cuando el agente alterne entre acciones asequibles para el paciente y acciones que no podrá realizar. Cabe resaltar que será f_1 quien ponderará si la acción destinada a fallar fue justa o no.

Tabla 6.2 Ejemplo que muestra el comportamiento de f_2 en caso de que el paciente falle en todos sus ejercicios.

t	Respuesta del jugador	$score_{t+1}$	$f_2(score_{t+1})$	$\sum_{i=1}^{t+1} f_2(score_i)$
0	Fallo	-1	-1000	-1000
1	Fallo	-2	-1000	-2000
2	Fallo	-3	-1000	-3000
3	Fallo	-4	-1000	-4000
4	Fallo	-5	-1000	-5000
5	Fallo	-6	-1000	-6000
6	Fallo	-7	-1000	-7000
7	Fallo	-8	-1000	-8000
8	Fallo	-9	-1000	-9000
9	Fallo	-10	-1000	-10000

Tabla 6.3 Ejemplo que muestra un caso ideal del comportamiento del algoritmo, ya que se logra que el paciente concluya el episodio con 6 aciertos y 4 fallos. La secuencia en la que ocurren los éxitos y los fallos permiten obtener el mayor valor acumulativo posible para f_2 .

t	Respuesta del jugador	$score_{t+1}$	$f_2(score_{t+1})$	$\sum_{i=1}^{t+1} f_2(score_i)$
0	Éxito	1	100	100
1	Éxito	2	200	300
2	Fallo	1	100	400
3	Éxito	2	200	600
4	Fallo	1	100	700
5	Éxito	2	200	900
6	Fallo	1	100	1000
7	Éxito	2	200	1200
8	Fallo	1	100	1300
9	Éxito	2	200	1500

6.2.3 Planeación

La primera vez que el paciente juegue se realizará un “episodio exploratorio”. El objetivo será determinar cuál es el brazo parético y cuáles son las dianas que el paciente logra alcanzar. El episodio exploratorio consta de 10 ejercicios que siguen la lógica de una búsqueda binaria. Primero se asignan ejercicios que requieren una flexión o abducción del hombro a 90 grados, si se supera el ejercicio se asume que el paciente es capaz de realizar ejercicios de una altura menor y se aumenta la altura de la diana para el siguiente ejercicio, por el contrario, si no se logra alcanzar la diana se asignará una de menor altura para el siguiente ejercicio. El lado donde el paciente logre detener menos balones se considerará el lado parético para el algoritmo.

Posteriormente se creará una lista M , la cual contendrá una estimación de los movimientos que el paciente puede realizar. Esta lista en realidad se trata del modelo del ambiente con el que el agente planificará para que pueda determinar más rápidamente qué acciones le debe exigir al paciente. El modelo M contendrá 25 valores, cada uno asociado a una acción del conjunto A . El valor de cada acción del modelo M se calcula de la siguiente forma:

$$M_a = \frac{\text{número de veces que el paciente logró realizar } a}{\text{número de veces que } a \text{ ha sido asignada}}, \quad a \in A \quad (6.12)$$

De esta manera, M trata de capturar la probabilidad que tiene el paciente de realizar correctamente el ejercicio que se le asigne. Cada vez que el paciente logre realizar una acción o falle se actualizará el modelo M de forma correspondiente.

La planeación se realizará después de que haya ocurrido un número determinado ER (episodios reales) de episodios. Cuando eso ocurra se ejecutará una subrutina en donde se simularán ES (episodios simulados) episodios, la subrutina es muy similar al algoritmo principal, para la planeación el agente asignará una acción aleatoria $a' \in A$ en cada estado $s' \in S$, pero en este caso el paciente no va a interactuar directamente con el juego, lo que sucederá es que la subrutina obtendrá un número aleatorio m' entre el rango $(0,1)$, si $m' > M_a$, significa que el paciente no habría sido capaz de realizar el ejercicio correctamente, de lo contrario se asume que el paciente habría realizado el ejercicio correctamente, en cualquier caso se actualiza $Q(s', a')$ como si el paciente en realidad hubiera interactuado con el juego. De esta forma se utiliza la experiencia que el agente ha acumulado hasta el momento y con base en ello se intenta acelerar el proceso de aprendizaje mediante una alta exploración, pues en cada estado se elige una acción aleatoria, no la mejor acción posible, lo anterior le permitirá al agente adaptarse mucho más rápido a los cambios de comportamiento que presente el paciente.

Cabe mencionar que el contenido del modelo y de la función de valor de acción nunca se volverá a reiniciar a cero una vez que el paciente comience a jugar, ya que tanto el modelo M , como la tabla $Q(S,A)$ se almacenan en un mismo archivo XML en la computadora donde se ejecute la aplicación, así como en el servidor CITAN. Cada vez que el paciente inicia sesión en el videojuego, una de las primeras acciones que se realiza es buscar este archivo y cargar sus respectivos valores en M y

$Q(S,A)$. En caso de que no se encuentre el archivo, por ejemplo, si el paciente juega en una computadora diferente, se intentará descargarlo desde la nube.

Se podría considerar que este algoritmo realiza una exploración en 4 de 10 ejercicios y una explotación en 6 de 10 ejercicios, una vez que el paciente logre realizar las acciones que antes no podía, entonces el valor de esa acción comenzará a disminuir si se asigna demasiadas veces, hasta que en un punto el valor de acciones no exploradas sea mayor que las otras, en ese momento el agente determinará cuál es el nuevo límite del paciente. No es necesario que el algoritmo explore cada cierto tiempo eligiendo acciones aleatorias, ya que, si el paciente no es capaz de alcanzar las dianas de una cierta altura, se asume que tampoco logrará alcanzar dianas de una altura aún mayor.

Con base en todo anterior, el algoritmo final, bajo el cual funciona el agente que aprende para realizar un ADD es el siguiente:

Algoritmo de Ajuste Dinámico de Dificultad

Inicializa valores para la tasa de aprendizaje α , tasa de descuento γ , frecuencia de planeación d y constante de penalización c .

Inicializa $V(A)$, donde $V(A) \in \{10,20,30,40,50\}$

Inicializa $Q(S, A)$ a cero.

Crea una lista vacía M , que representa el modelo del ambiente.

Si es la primera vez que el paciente juega:

Realiza episodio de exploración para identificar el hemicuerpo afectado.

Actualiza los valores de M

Si no:

$Q(S, A) \leftarrow$ Carga los valores de $Q(S, A)$ de la última sesión

$M \leftarrow$ Carga los valores de M de la última sesión

$episodios \leftarrow 0$

Repetir para cada episodio:

1. Si ($módulo(episodios, d) = 0$):

1.1. $Q(S, A) \leftarrow Planifica(M)$

2. Inicializa $N(A)$ a cero.

3. $score \leftarrow 0$

4. $tiros \leftarrow 0$

5. Mientras ($tiros < 9$):

5.1. Si ($tiros = 0$):

5.1.1. $s \leftarrow S_0$

5.1.2. $a \leftarrow$ una acción aleatoria $\tilde{a} \in A(s)$ tal que $Q(s, \tilde{a}) > 0$

5.2. Si no:

5.2.1. $a \leftarrow \max_a Q(s', a)$

5.2.2. $Q(s, a) \leftarrow Q(s, a) + \alpha [r + \gamma * \max_a Q(s', A) - Q(s, a)]$

5.2.3. $s \leftarrow s'$

5.3. $N(a) \leftarrow N(a) + 1$

5.4. Si el jugador logró detener el tiro a :

5.4.1. $f_1 \leftarrow \frac{V(a)}{N(a)}$

5.4.2. $score \leftarrow score + 1$

5.5. Si no:

5.5.1. $f_1 \leftarrow -V(a) * c$

5.5.2. $score \leftarrow score - 1$

5.6 Si el tiro fue del lado derecho:

5.6.1. $tirosDer \leftarrow tirosDer + 1$

5.7. Si el tiro fue del lado izquierdo:

5.7.1. $tirosIzq \leftarrow tirosIzq + 1$

5.8. Si el tiro fue en el centro:

5.8.1. $tirosCen \leftarrow tirosCen + 1$

5.9. Determinar el sig. estado s' con base en $tirosDer$, $tirosIzq$, $tirosCen$ y $score$

5.10. Si ($score = 2$):

5.10.1. $f_2 \leftarrow 200$

5.11. Si ($score = 1$):

5.11.1. $f_2 \leftarrow 100$

5.12. Si ($score = 0$):

8.5.12.1. $f_2 \leftarrow 0$

5.13. En otro caso:

5.13.1. $f_2 \leftarrow -1000$

5.14. $r \leftarrow f_1 + f_2$

5.15. $tiros \leftarrow tiros + 1$

5.16. Actualiza M_a

5.17. Si $tiros = 9$:

5.17.1. $Q(s, a) \leftarrow Q(s, a) + \alpha [r - Q(s, a)]$

6. $episodios \leftarrow episodios + 1$

6.3 Simulaciones

Para poder comprobar la efectividad del algoritmo en ausencia de pacientes reales, se realizaron varias simulaciones en el entorno Google Colab, con el lenguaje de programación Python, estas simulaciones tratan de imitar las diferentes limitaciones que podría presentar un paciente, con el fin de observar cómo reaccionaría el ADD ante diferentes circunstancias.

Para todas las simulaciones se empleó una tasa de aprendizaje $\alpha=0.05$ y un factor de descuento $\gamma=0.7$. En todas las simulaciones se evalúan 2 aspectos, el número de veces que se logra la meta de que el paciente obtenga 6 aciertos y 4 fallos en una partida, y el número de veces que ocurrió una transgresión, es decir, cuando se asignan dianas que sobrepasan el límite de las capacidades del paciente. También se registró el número de veces que el agente asignó cada una de las dianas.

El proceso de simulación fue básicamente igual al de planeación, descrito en la sección anterior. Se creó una estructura de datos P que almacenaba listas que contenían la probabilidad que tenía el paciente de realizar cada una de las 25 diferentes acciones, $P_{acciones}$. El número de listas i correspondía al número de cambios que iba a presentar el paciente después de un determinado número de episodios, $P_{episodios}$. Mientras que $P[i]_{episodios} > episodios$ la lista con la que se evaluaban las acciones que tomaba el agente sería $P[i]_{acciones}$, una vez que esto no se cumplía se incrementaba el índice i .

En las tablas 6.4 y 6.5 se muestran ejemplos de cómo se asignan las probabilidades a priori para diferentes acciones, así como el cambio en estas probabilidades para cada acción dependiendo del número de episodios que han transcurrido, intentando simular diferentes comportamientos que podría presentar un paciente EVC.

Tabla 6.4 Ejemplo que muestra sólo 5 acciones de cómo se simuló un paciente que fue mejorando su rango de abducción izquierdo. En este caso el paciente “mejoraba” cada 100 episodios, pues se observa que la probabilidad de alcanzar las dianas I3, I4 y I5 incrementa.

Episodio a partir del cual el paciente presentaría un cambio de conducta.	Probabilidades asociadas a cada acción
100	I1: 1.0, I2:0.8, I3:0.4, I4:0.2 ,I5:0.0
200	I1: 1.0, I2:1.0, I3:0.6, I4:0.4, I5:0.1
300	I1: 1.0, I2:1.0, I3:0.8, I4:0.6, I5:0.2

Tabla 6.5 Ejemplo que muestra sólo 5 acciones de cómo se simuló un paciente que fue “empeorando” en su rango de flexión del hombro derecho. En este caso el paciente “empeoraba” cada 50 episodios, pues se observa que la probabilidad de alcanzar las dianas DF3, DF4 y DF5 disminuye.

Episodio a partir del cual el paciente presentaría un cambio de conducta.	Probabilidades asociadas a cada acción
50	DF1: 1.0, DF2:1.0, DF3:1.0, DF4:0.8 ,DF5:0.6
100	DF1: 1.0, DF2:1.0, DF3:0.8, DF4:0.6 ,DF5:0.4
150	DF1: 1.0, DF2:1.0, DF3:0.5, DF4:0.4 ,DF5:0.1

6.4 Transgresiones

Como uno de los objetivos del algoritmo es que los ejercicios destinados a fallar se encuentren dentro del límite de las capacidades del paciente, se integró una métrica dentro de las simulaciones para determinar el número de veces que el algoritmo exigía más de lo que debía. Una de las ventajas de la forma en que se realizó la simulación del comportamiento de un paciente es que se tiene acceso en todo momento a la probabilidad que tiene de realizar cada ejercicio, de esta forma se propone que, si el algoritmo llegara a asignar una diana, donde la probabilidad de alcanzar la diana inmediata inferior tuviera una probabilidad menor del 60%, se considera una transgresión, porque se consideraría que el ejercicio asignado está superando el límite de las capacidades del paciente. Por ejemplo, si el paciente tuviera una probabilidad del 10% de detener el balón correctamente dirigido a la diana I3 y una probabilidad del 40% de realizar correctamente el ejercicio correspondiente a la diana I2, y el algoritmo tomara la acción de asignar la diana I3, entonces estaría cometiendo una transgresión.

De esta forma se podrá evaluar no sólo si el algoritmo cumple con tratar de que el paciente obtenga 6 aciertos y 4 fallos en cada partida, sino también si la asignación de los ejercicios es adecuada.

Capítulo 7

Resultados

7.1 Experimentos exploratorios

7.1.1 Experimento 1: Comportamiento del algoritmo sin planeación.

Para el primer experimento se simuló a un paciente con hemiparesia del lado izquierdo. Es decir, es capaz de realizar todos los ejercicios del lado derecho sin problema, pero sólo puede alcanzar las dianas I1, IF1 y B1 (figura 7.1), las dianas de una altura mayor le son imposibles de alcanzar. El experimento simula la interacción entre el paciente y el juego durante 100 episodios. No se realizó ningún paso de planeación.



Figura 7.1 Se muestra en color blanco los ejercicios que el paciente simulado es capaz de realizar y en color azul los que no.

El objetivo del experimento fue observar el cambio de efectividad del algoritmo de aprendizaje por reforzamiento cuando interactúa con un ambiente determinista (el paciente siempre realiza las mismas acciones) o con un ambiente estocástico (el paciente no siempre realiza la misma acción). Cuando se presenta un ambiente estocástico es necesario repetir el mismo experimento varias veces, entre mayor sea la muestra, mejor será la descripción de la distribución de las medias.

Las dos variables que se modificaron en este experimento fueron:

- El número de veces que se repite el experimento de 100 episodios.
- La probabilidad con la que el paciente puede alcanzar las dianas I2, IF2 y B2.

No se reporta cómo el agente asignó las dianas del lado derecho durante los 100 episodios porque se asume que, al tratarse del lado sano del paciente, alcanzaría todas las dianas.

Como se observa en la tabla 7.1, el algoritmo logra un ajuste prácticamente perfecto cuando se presenta un ambiente determinista (probabilidad del 100% de que el paciente alcance las dianas I2, IF2, B2). En las 100 partidas se logra el primer objetivo del algoritmo: el paciente obtiene 6 aciertos y 4 fallos al final de la partida. También se logra el segundo objetivo: todas las dianas asignadas para que el paciente falle se encontraron dentro de sus límites, es decir, el juego asignó ejercicios con una altura igual o menor a las de las dianas I3, IF3 y B3, nunca dianas con una altura mayor, pues le resultarían imposibles de alcanzar al paciente. Ambos objetivos se lograron incluso con una muestra de tamaño 50.

Tabla 7.1. Evaluación del algoritmo ante ambientes deterministas y estocásticos con diferentes tamaños de muestra, para cada valor se reporta el promedio obtenido de las muestras \pm la desviación estándar.

Probabilidad de alcanzar I2, IF2 y B2	1.00			0.60			0.30		
	Tamaño de muestra	1	25	50	1	25	50	1	25
Partidas con 6 aciertos	100 \pm 0	100 \pm 0	100 \pm 0	96 \pm 0	97.3 \pm 0.01	97.0 \pm 0.14	96 \pm 0	98.6 \pm 0.14	98.2 \pm 0.16
Porcentaje de transgresiones	0 \pm 0	0 \pm 0	0 \pm 0	0 \pm 0	0 \pm 0	0 \pm 0	33.2 \pm 0	33.7 \pm 3.92	33.6 \pm 2.69
Porcentaje de aparición de las dianas									
I1	1.09 \pm 0	1.64 \pm 0.82	1.74 \pm 0.81	12.9 \pm 0	14.39 \pm 2.34	14.13 \pm 2.08	14.39 \pm 0	13.9 \pm 1.12	13.91 \pm 1.25
I2	12.5 \pm 0	13.13 \pm 2.15	13.81 \pm 2.18	2.19 \pm 0	1.96 \pm 0.5	2.09 \pm 0.56	3.59 \pm 0	3.07 \pm 1.21	3.41 \pm 1.02
I3	11.89 \pm 0	15.41 \pm 2.34	14.67 \pm 1.87	15.8 \pm 0	14.14 \pm 1.89	13.88 \pm 1.63	13.1 \pm 0	12.9 \pm 1.35	12.42 \pm 1.29
I4	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.1 \pm 0	0.07 \pm 0.14	0.16 \pm 0.29
I5	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.03 \pm 0.06	0.01 \pm 0.03
IF1	1.7 \pm 0	1.65 \pm 0.911	1.92 \pm 0.93	15.5 \pm 0	13.8 \pm 2.36	14.32 \pm 2.09	12.9 \pm 0	13.95 \pm 1.18	13.95 \pm 1.17
IF2	14.7 \pm 0	13.5 \pm 2.19	12.42 \pm 2.26	1.2 \pm 0	2.19 \pm 0.54	2.03 \pm 0.44	5.3 \pm 0	3.63 \pm 1.06	3.23 \pm 0.96
IF3	18.0 \pm 0	14.5 \pm 2.38	15.39 \pm 1.92	12.4 \pm 0	13.49 \pm 1.56	13.53 \pm 1.57	10.5 \pm 0	12.18 \pm 0.46	12.69 \pm 1.42
IF4	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.10 \pm 0.24	0.12 \pm 0.20
IF5	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.1 \pm 0	0.04 \pm 0.08	0.05 \pm 0.11
B1	0.70 \pm 0	0.66 \pm 0.26	0.73 \pm 0.27	8.9 \pm 0	8.08 \pm 0.8	7.78 \pm 1.01	9.6 \pm 0	9.54 \pm 0.62	9.66 \pm 0.56
B2	9.3 \pm 0	9.35 \pm 0.61	9.34 \pm 0.41	1.3 \pm 0	1.44 \pm 0.32	1.50 \pm 0.45	2 \pm 0	2.02 \pm 0.71	2.14 \pm 0.54
B3	10.0 \pm 0	9.98 \pm 0.64	9.92 \pm 0.44	9.8 \pm 0	10.47 \pm 0.81	10.71 \pm 0.96	7.9 \pm 0	8.25 \pm 0.66	7.99 \pm 0.66
B4	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.136 \pm 0.20	0.13 \pm 0.17
B5	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.0 \pm 0	0.5 \pm 0	0.04 \pm 0.11	0.06 \pm 0.14

Sin embargo, en un escenario real un paciente no se comportará como un robot repitiendo siempre las mismas acciones a la perfección, algunas veces el paciente realizará un esfuerzo considerable para superar sus límites, pero no se esperaría que pudiera mantener ese comportamiento durante toda la partida, también puede ocurrir que en algún punto reaccione un poco más lento o inclusive que, debido al agotamiento, en ocasiones no logre alcanzar dianas que normalmente si podría.

Cuando se simuló un paciente que logra alcanzar las dianas I2, IF2 y B2 en un 60% de las ocasiones, la efectividad del algoritmo disminuye un poco. Con un solo paciente se cumple el objetivo de obtener 6 aciertos y 4 errores en 96 partidas de 100, aunque un valor más cercano al real es cuando el tamaño de muestra es de 50, donde, en promedio, en 97.08 de 100 partidas se logra esta meta. En este escenario el algoritmo logra adaptarse muy bien, a pesar de que algunas veces

el paciente no reaccionaba de la forma que el agente esperaba, al fallar dianas que normalmente alcanzaría. Por esta razón se puede observar, a diferencia del escenario anterior, que ante la incertidumbre de lo que podría ocurrir con las dianas I2, IF2 y B2, el agente asignó de forma mucho más frecuente las dianas I1, IF1 y B1, pues con ellas se aseguraba el éxito cuando fuera necesario para lograr cumplir con la meta de 6 aciertos y 4 fallos. Al mismo tiempo, la demanda de las dianas I3, IF3 y B3 no disminuyó, pues el agente sabe que el paciente no logra alcanzar esas dianas bajo ninguna circunstancia. Es decir, el agente solía probar algunas veces lo que ocurría con las dianas I2, IF2 y B2, dependiendo de los resultados el agente se apoyaba en las dianas I1, IF1 y B1 para éxitos seguros y en las dianas I3, IF3 y B3 para fallos seguros. En este tipo de condiciones también se logró el segundo objetivo del algoritmo, ya que no se cometió ninguna transgresión, ni siquiera cuando el mismo experimento se repitió 50 veces.

En el último escenario, cuando el paciente presenta un 30% de probabilidad de alcanzar las dianas I2, IF2 y B2, el comportamiento del agente resultó similar al escenario anterior en términos de desempeño del agente y las dianas asignadas, aunque en estas simulaciones si ocurrieron transgresiones, pues el agente asignó dianas que se encontraban muy por encima de las capacidades del paciente, por ejemplo, las dianas I5, IF5 y B5. Además, en este escenario tampoco es muy deseable que se asignen con tanta frecuencia las dianas I3, IF3 y B3, pues con un 30% de probabilidad de alcanzar las dianas anteriores significa que al paciente le cuesta mucho trabajo realizar las acciones I2, IF2 y B2. Recordando el criterio mencionado anteriormente, el ejercicio asignado se considera una transgresión si la diana inmediata de altura inferior cuenta con una probabilidad menor al 60% de ser alcanzada, por esta razón cualquier acción en la que se haya asignado las dianas I3, IF3 o B3 es una transgresión por parte del agente. A pesar de que el algoritmo tuvo en promedio 98.24 partidas donde logró 6 aciertos y 4 fallos, también se cometieron en promedio 33.6 transgresiones de un total de 100 ejercicios, lo que indica que el agente tuvo dificultades para determinar los límites de las capacidades del paciente.

A continuación se muestra cómo la implementación en el algoritmo de un proceso de planeación le ayuda al agente a responder de mejor forma ante ambientes estocásticos.

7.1.2 Experimento 2: Comportamiento del algoritmo con planeación

En este experimento se vuelve a simular el mismo tipo de paciente que en el experimento 1. Su brazo parético es el izquierdo, la probabilidad de que alcance las dianas I2, IF2 y B2 es de 30%, las dianas de altura inferior las alcanza con un 100% de probabilidad y las dianas de mayor altura con un 0% de probabilidad. Cada experimento consta de 100 episodios, el cual se repitió 50 veces y se obtuvo el promedio para tener un valor más representativo de la distribución de los valores obtenidos. En esta ocasión se incluyó un proceso de planificación en el algoritmo. Cada determinado número de episodios el algoritmo “planifica” con una alta exploración un cierto número de episodios con base en lo que ha experimentado (modelo del ambiente) y actualiza los valores de $Q(S, A)$.

Las dos variables que se modificaron en este experimento fueron:

- La frecuencia con la que el algoritmo planifica, es decir, cuántos episodios tiene que transcurrir para que el algoritmo vuelva a planificar.
- El número de pasos de planificación, es decir, durante cuántos episodios el agente va a explorar las diferentes posibilidades empleando el modelo que ha construido con base en la experiencia.

Es importante mencionar que los episodios de planificación únicamente modifican los valores de $Q(S,A)$, en una aplicación real este proceso es imperceptible para el usuario, pues el agente sólo “imagina” lo que sucedería. Los resultados de este experimento se muestran en la tabla 7.2 y se compararon con los del experimento anterior (última columna), que se realizó bajo las mismas condiciones, pero sin realizar ningún proceso de planeación.

Tabla 7.2. Evaluación del algoritmo ante un ambiente estocástico combinando diferentes parámetros de planeación. Se reportan los valores promedio obtenidos de las muestras \pm la desviación estándar.

Cada cuántos episodios se planifica	5	10	25	50	5	10	25	50	0
Número de pasos de planificación	3000				5000				0
Partidas con 6 aciertos	88.5 \pm 0.4	92.5 \pm 2.6	95.5 \pm 0.3	96.9 \pm 0.1	80.3 \pm 0.4	87.3 \pm 0.5	92.9 \pm 0.3	95.2 \pm 0.2	98.2 \pm 0.1
Porcentaje de transgresiones	15.6 \pm 1.3	22.2 \pm 2.9	28.7 \pm 3.5	31.4 \pm 3.2	12.3 \pm 1.0	17.8 \pm 2.0	26.72 \pm 2.3	29.92 \pm 3.9	33.6 \pm 2.6
Porcentaje de aparición de las dianas									
I1	10.2 \pm 1.4	11.7 \pm 2.1	12.24 \pm 1.5	13.1 \pm 1.1	8.82 \pm 1.4	9.89 \pm 1.1	11.4 \pm 1.16	13.0 \pm 1.23	13.91 \pm 1.2
I2	14.3 \pm 2.8	9.43 \pm 2.7	5.89 \pm 1.7	4.56 \pm 1.3	17.03 \pm 2.5	12.82 \pm 2.7	7.63 \pm 1.39	5.47 \pm 1.58	3.41 \pm 1.0
I3	5.3 \pm 1.4	8.70 \pm 1.7	11.0 \pm 2.1	11.98 \pm 2.0	3.97 \pm 1.0	7.32 \pm 1.5	10.8 \pm 1.51	11.7 \pm 1.29	12.42 \pm 1.2
I4	0.3 \pm 0.2	0.19 \pm 0.2	0.16 \pm 0.2	0.12 \pm 0.1	0.29 \pm 0.2	0.21 \pm 0.2	0.14 \pm 0.19	0.11 \pm 0.20	0.16 \pm 0.2
I5	0.04 \pm 0.1	0.01 \pm 0.0	0.03 \pm 0.0	0.01 \pm 0.0	0.03 \pm 0.0	0.02 \pm 0.0	0.04 \pm 0.12	0.03 \pm 0.08	0.01 \pm 0.0
IF1	10.31 \pm 1.2	11.1 \pm 2.0	12.2 \pm 1.1	13.53 \pm 1.3	9.06 \pm 1.0	9.91 \pm 1.2	11.84 \pm 0.93	12.8 \pm 0.76	13.95 \pm 1.1
IF2	13.5 \pm 2.6	9.66 \pm 2.7	6.62 \pm 1.9	4.47 \pm 1.3	16.11 \pm 2.3	12.01 \pm 2.5	7.0 \pm 1.44	5.29 \pm 1.41	3.23 \pm 0.9
IF3	5.41 \pm 1.2	8.93 \pm 1.6	11.54 \pm 2.1	12.0 \pm 1.5	4.36 \pm 1.1	7.56 \pm 1.5	10.8 \pm 1.28	11.5 \pm 1.10	12.7 \pm 1.4
IF4	0.35 \pm 0.2	0.19 \pm 0.2	0.17 \pm 0.2	0.07 \pm 0.1	0.28 \pm 0.2	0.02 \pm 0.2	0.12 \pm 0.18	0.08 \pm 0.17	0.12 \pm 0.2
IF5	0.03 \pm 0.7	0.02 \pm 0.0	0.03 \pm 0.0	0.05 \pm 0.1	0.01 \pm 0.0	0.03 \pm 0.0	0.02 \pm 0.05	0.03 \pm 0.08	0.05 \pm 0.1
B1	7.85 \pm 1.3	9.26 \pm 1.9	10.2 \pm 1.0	9.64 \pm 0.7	7.72 \pm 0.9	9.83 \pm 1.1	10.53 \pm 0.88	9.76 \pm 6.53	9.66 \pm 0.5
B2	7.94 \pm 1.6	6.47 \pm 2.1	4.26 \pm 1.3	3.13 \pm 0.9	8.92 \pm 1.1	7.44 \pm 1.5	4.65 \pm 1.3	3.69 \pm 0.8	2.14 \pm 0.5
B3	3.91 \pm 0.9	4.02 \pm 1.7	5.4 \pm 1.2	7.06 \pm 1.0	3.08 \pm 0.9	2.55 \pm 0.9	4.67 \pm 1.1	6.39 \pm 0.9	7.99 \pm 0.6
B4	0.23 \pm 0.1	0.2 \pm 0.2	0.10 \pm 0.1	0.11 \pm 0.1	0.25 \pm 0.2	0.13 \pm 0.1	0.11 \pm 0.1	0.08 \pm 0.1	0.13 \pm 0.1
B5	0.05 \pm 0.0	0.02 \pm 0.0	0.03 \pm 0.0	0.04 \pm 0.0	0.03 \pm 0.0	0.02 \pm 0.0	0.02 \pm 0.0	0.05 \pm 0.0	0.06 \pm 0.1

En la tabla 7.2 se puede apreciar que la planeación le ayuda bastante al agente a identificar que las dianas I2, IF2 y B3 presentan un comportamiento errático, lo que significa que se encuentran dentro del límite de las capacidades del paciente y por lo tanto son esas dianas las que se deben de asignar con mayor frecuencia. Realizar una planeación cada 25 o 50 episodios tiene un impacto negativo, pues el agente se comporta de forma similar a que si no hubiera planeación, en el sentido de que se cometen casi el mismo número transgresiones, por lo tanto, este proceso se debe realizar máximo cada 10 episodios para conservar sus beneficios sobre al agente.

El número de pasos de planeación también es importante, entre más pasos se realicen al agente le será más difícil lograr la meta de obtener 6 aciertos y 4 fallos, al mismo tiempo asignará en mayor medida los ejercicios adecuados para el paciente durante la partida y cometerá menos transgresiones. Los resultados de planificar cada 5 episodios con 3000 pasos son muy similares a los obtenidos por planificar cada 10 episodios con 5000 pasos, obteniendo 6 aciertos en alrededor de 88% de las partidas y cometiendo transgresiones entre el 12 y 15% del total de los ejercicios (1000 para este experimento). La asignación de las dianas en ambos escenarios es mucho mejor que cuando no se realiza ninguna planeación (transgresiones en promedio en 336 ejercicios de 1000), por lo que es deseable tener este tipo de comportamiento a costa de una disminución en el número de partidas donde se logra la meta de obtener 6 aciertos y 4 fallos.

En los experimentos anteriores se han presentado ambientes estocásticos, pero de alguna forma estables, en el sentido de que las probabilidades asignadas que tiene el paciente de realizar correctamente el ejercicio asociado a cada diana se mantienen igual a lo largo de todos los episodios. Los siguientes experimentos exploran el comportamiento del algoritmo cuando esto cambia.

7.1.3 Experimento 3: Planeación en ambientes estocásticos

El siguiente experimento consta de 300 episodios, con una muestra de tamaño 50. El paciente simulado presenta hemiparesia en el brazo izquierdo y “mejora” cada 100 episodios.

- De los episodios 0 a 100 la probabilidad de que el paciente logre realizar correctamente los ejercicios de las dianas I3, IF3 y B3 es de 0%. Cualquier diana de altura inferior tiene una probabilidad asociada del 100% y cualquier diana de altura superior tiene una probabilidad de 0%.
- De los episodios 100 a 200 la probabilidad de que el paciente logre alcanzar las dianas I3, IF3 y B3 es 30%. El resto de las probabilidades se mantienen sin cambios.
- A partir del episodio 200 la probabilidad de que el paciente logre alcanzar las dianas I3, IF3 y B3 es del 60%. El resto de las probabilidades se mantienen sin cambios.

Los resultados de este experimento se dividieron en dos tablas para comparar el efecto de cambiar el número de pasos de planificación y la frecuencia con la que se planifica.

En las tablas 7.3 y 7.4 se observa que la mayor cantidad de transgresiones se cometen cuando el agente no realiza ninguna planeación, a pesar de que en estas condiciones es donde se logra una mayor cantidad de partidas con 6 aciertos y 4 fallos. Como ya se había observado en el experimento 2 (tabla 2.2), la distribución en la que el agente asigna las dianas cuando no existe planeación no es muy buena, ya que se asigna con mucha frecuencia las dianas I4, IF4 y B4 entre los episodios 100 y 200, con un porcentaje promedio de 11.65, 12.70 y 10.14 respectivamente, cuando el paciente apenas puede alcanzar I3, IF3 y B3 con un 30% de probabilidad. En un caso ideal, el porcentaje promedio de aparición de las dianas I4, IF4 y B4 debería de ser 0% cuando el

paciente simulado tiene una probabilidad del 30% de realizar correctamente los ejercicios de las dianas I3, IF3 y B3, pues esto significaría que el algoritmo no le está asignando ejercicios al jugador más allá de sus capacidades motrices.

Tabla 7.3. Evaluación del algoritmo con diferentes parámetros de planeación ante un paciente que evoluciona. Se reportan los valores promedio obtenidos de las muestras \pm la desviación estándar.

Cada cuántos episodios se planifica	5			10 *			0		
Número de pasos de planificación	3000			3000			0		
Episodios transcurridos	100	200	300	100	200	300	100	200	300
Probabilidad de realizar acciones I3,IF3 y B3	0.0	0.30	0.6	0.0	0.30	0.6	0.0	0.30	0.6
Partidas con 6 aciertos	100 \pm 0.0	88.3 \pm 0.42	77.08 \pm 0.7	100 \pm 0.0	90.5 \pm 0.33	86.46 \pm 0.5	100 \pm 0.0	94.64 \pm 3.9	95.62 \pm 3.1
Porcentaje de transgresiones	0.0 \pm 0.0	11.2 \pm 1.6	1.75 \pm 0.1	0.0 \pm 0.0	17.7 \pm 3.3	1.0 \pm 0.67	0.0 \pm 0.0	35 \pm 3.04	0.45 \pm 0.3
Porcentaje de aparición de las dianas									
I1	1.61 \pm 0.7	1.26 \pm 0.4	1.37 \pm 0.5	1.8 \pm 0.93	2.18 \pm 0.6	2.75 \pm 0.8	1.68 \pm 0.86	4.55 \pm 1.44	4.44 \pm 1.8
I2	14.6 \pm 1.7	8.15 \pm 1.3	6.48 \pm 1.0	14.3 \pm 1.8	9.19 \pm 1.1	8.22 \pm 1.0	12.89 \pm 1.8	10.9 \pm 2.17	11.4 \pm 2.41
I3	13.9 \pm 2.0	16.2 \pm 2.4	13.9 \pm 2.0	13.9 \pm 1.9	12.54 \pm 2.6	8.43 \pm 1.5	15.36 \pm 1.9	2.63 \pm 1.56	0.67 \pm 0.40
I4	0.0 \pm 0.0	3.92 \pm 1.2	7.57 \pm 1.5	0.0 \pm 0.0	5.96 \pm 1.43	10.27 \pm 1.44	0.0 \pm 0.0	11.65 \pm 1.76	13.0 \pm 3.81
I5	0.0 \pm 0.0	0.43 \pm 0.33	0.58 \pm 0.39	0.0 \pm 0.0	0.31 \pm 0.31	0.33 \pm 0.33	0.0 \pm 0.0	0.24 \pm 0.26	0.38 \pm 0.20
IF1	1.62 \pm 0.6	1.34 \pm 0.46	1.42 \pm 0.59	1.8 \pm 0.74	2.16 \pm 0.59	2.79 \pm 0.84	1.67 \pm 0.84	4.8 \pm 1.96	4.53 \pm 1.64
IF2	14.5 \pm 1.7	8.25 \pm 1.6	6.0 \pm 1.02	14.1 \pm 1.8	8.84 \pm 1.2	8.23 \pm 1.1	13.94 \pm 7.4	10.62 \pm 2.3	10.96 \pm 2.1
IF3	13.6 \pm 2.2	16.1 \pm 2.8	13.9 \pm 2.4	14.0 \pm 1.8	12.3 \pm 2.9	8.26 \pm 1.27	14.67 \pm 1.9	2.49 \pm 1.36	0.76 \pm 0.58
IF4	0.0 \pm 0.0	3.85 \pm 1.0	7.95 \pm 1.1	0.0 \pm 0.0	6.09 \pm 1.4	10.23 \pm 1.6	0.0 \pm 0.0	12.70 \pm 2.0	14.0 \pm 3.78
IF5	0.0 \pm 0.0	0.43 \pm 0.2	0.72 \pm 0.4	0.0 \pm 0.0	0.36 \pm 0.3	0.44 \pm 0.35	0.0 \pm 0.0	0.19 \pm 0.17	0.04 \pm 0.06
B1	0.71 \pm 0.2	0.66 \pm 0.3	0.66 \pm 0.2	0.68 \pm 0.2	0.75 \pm 0.3	0.67 \pm 0.3	0.73 \pm 0.26	1.91 \pm 0.58	0.6 \pm 0.21
B2	6.84 \pm 0.5	8.06 \pm 1.1	4.08 \pm 0.9	7.31 \pm 0.6	7.12 \pm 0.9	4.57 \pm 0.80	9.31 \pm 0.53	6.29 \pm 0.91	6.98 \pm 0.73
B3	12.4 \pm 0.5	8.7 \pm 1.31	7.2 \pm 0.91	11.9 \pm 0.5	7.04 \pm 1.6	4.64 \pm 0.71	9.95 \pm 0.47	1.56 \pm 1.0	0.39 \pm 0.12
B4	0.0 \pm 0.0	2.39 \pm 0.8	7.5 \pm 0.97	0.0 \pm 0.0	4.89 \pm 1.3	9.83 \pm 1.11	0.0 \pm 0.0	10.14 \pm 1.1	12.0 \pm 0.63
B5	0.0 \pm 0.0	0.17 \pm 0.2	0.45 \pm 0.3	0.0 \pm 0.0	0.16 \pm 0.2	0.27 \pm 0.29	0.0 \pm 0.0	0.09 \pm 0.14	0.03 \pm 0.0

Al comparar las tablas 7.3 y 7.4 se puede observar que el mejor comportamiento del agente se obtiene al realizar 3000 episodios de planeación cada 10 episodios. En promedio, durante los 300 episodios se logra que el paciente finalice con 6 aciertos y 4 fallos en más del 80% de las partidas. También, en promedio, se logra una baja cantidad de transgresiones y una mejor distribución de la asignación de las dianas de acuerdo con la evolución del paciente. Entre los episodios 0 y 100 el agente asigna las dianas I3, IF3 y B3 en un porcentaje promedio de 13.94, 14.70 y 11.90% respectivamente, del 100% de los ejercicios asignados, lo cual es adecuado, pues se encuentran dentro del límite de las capacidades del paciente. Cabe resaltar que en este caso la distribución de la asignación de estas 3 dianas en específico es muy similar. El agente decide que al paciente, al presentar una limitación equivalente en movimientos de abducción, movimientos de flexión y movimientos bimanuales, se le deben exigir estos tres tipos de movimientos con la misma intensidad. Cuando el paciente evoluciona entre los episodios 100 y 200, el agente decide mantener en gran medida esta distribución de I3, IF3 y B3, lo cual se observa con un porcentaje

promedio de 12.54, 12.3 y 7.04 respectivamente del 100% de los ejercicios asignados. Esto ocurre porque el paciente sólo es capaz de alcanzar esas dianas con un 30% de probabilidad, por lo que se consideraría que aún se encuentra en el límite de sus capacidades. Finalmente, no es hasta el último periodo, a partir del episodio 200, que el agente decide realizar un cambio más notable en la distribución de la asignación de los ejercicios, exigiendo de igual forma las dianas I3, IF3, B3, I4, IF4 y B4 pues son las que se en ese momento se encuentran dentro del límite de las capacidades del paciente simulado.

Tabla 7.4. Evaluación del algoritmo con diferentes parámetros de planeación ante un paciente que evoluciona.

Cada cuántos episodios se planifica	5			10			0		
Número de pasos de planificación	5000			5000			0		
Episodios transcurridos	100	200	300	100	200	300	100	200	300
Probabilidad de realizar acciones I3,IF3 y B3	0.0	0.30	0.6	0.0	0.30	0.6	0.0	0.30	0.6
Partidas con 6 aciertos	100±0.0	83.6±0.4	74.6±0.6	100±0.0	88.5±0.4	84.1±0.5	100±0.0	94.6±3.9	95.6±3.1
Porcentaje de transgresiones	0.0±0.0	10.4±2.3	2.16±0.9	0.0±0.0	16.5±3.8	1.2±0.35	0.0±0.0	35±3.0	0.45±0.3
Porcentaje de aparición de las dianas									
I1	1.6±0.6	1.2±0.4	1.42±0.5	1.76±0.8	2.23±0.6	2.49±0.8	1.7±0.9	4.5±1.4	4.4±1.8
I2	14.7±1.9	7.45±1.1	5.5±0.99	14.2±1.5	8.36±1.0	7.8±1.0	12.9±1.8	10.9±2.2	11.4±2.4
I3	12.9±2.1	17.5±2.4	14.3±1.4	14±1.75	13.4±2.7	9.6±1.3	15.4±2.0	2.63±1.5	0.67±0.4
I4	0.0±0.0	3.5±0.94	8.0±1.34	0.0±0.0	5.7±1.11	9.77±1.2	0.0±0.0	11.65±1.8	13.0±3.8
I5	0.0±0.0	0.37±0.3	0.83±0.5	0.0±0.0	0.38±0.3	0.45±0.4	0.0±0.0	0.24±0.26	0.38±0.2
IF1	1.4±0.67	1.35±0.5	1.32±0.53	1.92±0.93	1.94±0.5	2.31±0.6	1.67±0.84	4.8±1.96	4.53±1.64
IF2	14.9±2.09	7.63±1.2	5.57±1.1	14.4±1.7	8.84±1.0	7.82±1.1	13.94±7.4	10.62±2.3	10.9±2.1
IF3	14.44±2.1	16.9±2.4	14.2±2.0	13.7±1.8	12.9±2.3	9.62±1.5	14.67±1.9	2.49±1.36	0.76±0.6
IF4	0.0±0.0	3.43±1.0	7.95±1.4	0.0±0.0	5.85±1.5	9.72±1.1	0.0±0.0	12.70±2.0	14.0±3.8
IF5	0.0±0.0	0.55±0.4	0.8±0.45	0.0±0.0	0.39±0.3	0.42±0.3	0.0±0.0	0.19±0.17	0.04±0.0
B1	0.73±0.26	0.66±0.2	0.6±0.23	0.73±0.27	0.87±0.29	0.67±0.25	0.73±0.26	1.91±0.58	0.6±0.21
B2	6.67±0.56	7.87±0.9	3.8±0.68	6.96±0.6	6.93±0.8	4.42±0.75	9.31±0.53	6.29±0.91	6.98±0.7
B3	12.6±0.48	8.9±0.90	8.35±0.87	12.3±0.5	8.0±1.84	5.28±0.7	9.95±0.47	1.56±1.0	0.392±0.1
B4	0.0±0.0	2.3±0.81	6.69±1.0	0.0±0.0	4.03±1.45	9.28±1.04	0.0±0.0	10.14±1.16	12.0±0.63
B5	0.0±0.0	0.26±0.2	0.53±0.3	0.0±0.0	0.13±0.2	0.33±0.3	0.0±0.0	0.09±0.14	0.03±0.0

En el experimento 2 se observó que una planeación entre 3000 o 5000 episodios era buena dependiendo de si se realizaba cada 5 o 10 episodios. Sin embargo, en el experimento 3, donde se simulan condiciones más apegadas a la realidad respecto a cómo podría responder el paciente, la opción de realizar procesos de planeación de 3000 pasos cada 10 episodios obtuvo mejores resultados y por esa razón se emplearon estos parámetros en el resto de los experimentos.

7.1.4 Experimento 4: Comportamiento del algoritmo ante diversos pacientes

En este experimento se evaluó el comportamiento del agente ante pacientes con diferentes grados de movilidad, por ejemplo, en donde a un paciente le sea más fácil realizar movimientos de

flexión que movimientos bimanuales y de abducción, mientras que a otro le resulte muy difícil realizar cualquiera de estos 3 movimientos.

En este experimento se manipularon 4 variables:

1. El brazo afectado del paciente.
2. El grado de movilidad en movimientos de flexión.
3. El grado de movilidad en movimientos de abducción.
4. El grado de movilidad en tareas bimanuales.

Una movilidad baja significa que el paciente podrá alcanzar las dianas I1, IF1, D1, DF1 o B1, dependiendo del brazo afectado y del tipo de ejercicio, con una probabilidad entre 70 y 100%, así mismo podrá alcanzar las dianas I2, IF2, D2, DF2 o B2 con un 30% de probabilidad. Cualquier diana de altura superior tendrá un 0% de probabilidad.

Una movilidad media significa que el paciente podrá alcanzar las dianas I2, IF2, D2, DF2 o B2 con una probabilidad entre 70 y 100%, así mismo podrá alcanzar las dianas I3, IF3, D3, DF3 o B3 con un 30% de probabilidad. Cualquier diana de altura superior tendrá un 0% de probabilidad.

Una movilidad alta significa que el paciente podrá alcanzar las dianas I3, IF3, D3, DF3 o B3 con una probabilidad entre 70 y 100%, así mismo podrá alcanzar las dianas I4, D4, IF4, DF4 o B4 con un 30% de probabilidad. Cualquier diana de altura superior tendrá un 0% de probabilidad.

Al tener 3 distintos valores de movilidad (alta, media y baja) ante 3 distintos tipos de ejercicio (abducción o flexión de hombro o tareas bimanuales) se simuló 27 pacientes diferentes, cada uno con una combinación distinta de capacidades motrices. Para cada paciente se simuló 300 episodios con un tamaño de muestra de 50, realizando un proceso de planeación de 3000 pasos cada 10 episodios.

En la tabla 7.5 se puede observar que el agente se logra adaptar correctamente ante las diferentes necesidades que presenta cada paciente, sin importar el brazo afectado. Es decir, cuando el paciente presenta el mismo rango de movilidad en movimientos de flexión, abducción y tareas bimanuales, entonces el agente decide presentarle al paciente el mismo número de dianas de flexión y de abducción, recordando que el número de tareas bimanuales no es modificable y se encuentra fijo en 2. Sin embargo, se puede apreciar que de los 4 fallos que busca que cometa el agente, es decir, las 4 dianas que el agente debe asignar que se encuentren al límite de las capacidades del paciente, estas se distribuyen siempre hacia el tipo de ejercicio en el que el paciente tenga menos movilidad. Por ejemplo, si un paciente presenta una movilidad alta en tareas de abducción, baja en tareas de flexión y media en movimientos bimanuales, se esperaría que el agente asigne en mayor medida ejercicios de flexión, y que asigne más ejercicios al límite de las capacidades del paciente (ejercicios destinados para que el paciente falle) en tareas bimanuales que en tareas de abducción, tal y como se muestra en la tabla 7.5, donde del 100% de los ejercicios en promedio 22.12% fueron tareas de abducción y 37.87% de flexión, el 20% restante

fueron ejercicios bimanuales, de las 4 fallas que se buscan obtener en cada partida, en promedio 1.14% fueron ejercicios de abducción, 89.24% fueron ejercicios de flexión y 9.6% fueron tareas bimanuales, lo que significa que el juego detectó que el tipo de movimiento más débil del paciente era la flexión de hombro, y fue al que se le destinó una mayor carga de trabajo, mientras que casi nunca se asignaron fallas en las tareas de abducción y existió algo de exigencia para las tareas bimanuales. Este tipo de resultados indican que el algoritmo se aproxima al comportamiento que podría seguir un terapeuta al enfocarse en trabajar en mayor medida el tipo de movimiento más afectado del hombro.

Tabla 7.5. Comportamiento del algoritmo ante pacientes con diferentes grados de movilidad.

Se reportan los porcentajes promedio de la simulación de 27 pacientes diferentes. Cada paciente realizó 300 partidas y cada simulación se repitió 50 veces, con una planeación de 3000 episodios cada 10 episodios.									
Movilidad Abducción	Movilidad Flexión	Movilidad bimanual	Brazo afectado	Episodios donde se lograron 6 aciertos	Ejercicios abducción	Ejercicios flexión	Fallos abducción	Fallos flexión	Fallos bimanual
Baja	Baja	Baja	Derecho	86.59	21.94	38.06	32.8	35.91	31.29
Baja	Baja	Media	Izquierdo	86.23	34.09	25.9	47.9	46.9	5.13
Baja	Baja	Alta	Derecho	78.73	28.34	31.65	43.04	50.32	6.62
Baja	Media	Baja	Izquierdo	86.77	31.16	28.83	63.08	17.31	19.60
Baja	Media	Media	Derecho	84.56	38.45	21.54	72.01	14.87	13.11
Baja	Media	Alta	Derecho	82.7	40.69	19.3	77.7	14.09	8.19
Baja	Alta	Baja	Izquierdo	81.74	32.59	27.4	71.92	1.71	26.36
Baja	Alta	Media	Izquierdo	94.46	37.30	22.69	88.47	1.27	10.24
Baja	Alta	Alta	Derecho	83.29	41.43	18.57	87.45	4.77	6.52
Media	Baja	Baja	Izquierdo	85.49	28.29	31.7	4.88	68.09	27.02
Media	Baja	Media	Derecho	81.98	22.68	37.31	11.98	74.84	13.17
Media	Baja	Alta	Izquierdo	89.82	20.18	39.81	4.17	92.17	3.65
Media	Media	Baja	Derecho	87.17	29.96	30.03	31.08	31.45	37.52
Media	Media	Media	Izquierdo	78.14	29.98	30.01	37.35	37.91	24.75
Media	Media	Alta	Izquierdo	87.02	30.07	29.92	43.15	42.96	13.88
Media	Alta	Baja	Derecho	82.78	31.16	28.83	47.15	12.11	40.73
Media	Alta	Media	Izquierdo	87.57	34.43	25.56	59.92	6.52	33.54
Media	Alta	Alta	Derecho	69.81	40.38	19.62	74.54	12.19	13.25
Alta	Baja	Baja	Izquierdo	82.72	28.18	31.81	3.55	62.13	34.31
Alta	Baja	Media	Derecho	94.85	22.12	37.87	1.14	89.24	9.6
Alta	Baja	Alta	Izquierdo	85.5	18.56	41.43	3.92	88.17	7.89
Alta	Media	Baja	Derecho	73.56	29.38	30.61	13.97	44.58	41.44
Alta	Media	Media	Izquierdo	72.96	24.49	35.5	13.55	56.36	30.07
Alta	Media	Alta	Derecho	69.13	20.20	39.79	13.11	71.16	15.71
Alta	Alta	Baja	Izquierdo	69.61	30.09	29.91	26.84	26.93	46.22
Alta	Alta	Media	Derecho	70.48	29.4	30.59	27.56	34.84	37.59
Alta	Alta	Alta	Izquierdo	57.15	30.20	29.79	38.60	37.88	23.50

Con la mayoría de los pacientes simulados el agente logra el objetivo en más del 80% de las 300 partidas de obtener 6 aciertos y 4 fallos al final de cada episodio. Sin embargo, un paciente especial es el que presenta una alta movilidad en todos los tipos de movimiento, en este caso al agente le resulta difícil lograr cumplir la meta de tener 6 aciertos y 4 errores, como se muestra en la tabla 7.5, en donde sólo se logra cumplir este objetivo en 57.15% de 300 partidas, no obstante,

esto no debería representar un problema, pues un paciente con estas características está prácticamente rehabilitado.

7.2 Experimentos simulando casos específicos de pacientes

En las siguientes simulaciones se trata de tener una aproximación de diferentes tipos de comportamiento que podría presentar un paciente real al usar el videojuego. Las simulaciones están diseñadas pensando que los ejercicios se realizarían por un periodo de 2 meses. Considerando que el paciente pudiera jugar 30 minutos de lunes a viernes, en un día podría realizar 15 partidas, ya que cada partida dura un poco más de un minuto, esto significaría 75 partidas en una semana, por lo tanto, las siguientes simulaciones contemplan que se realizarían 600 partidas a lo largo de 2 meses.

7.2.1 Simulación de un paciente EVC que mejora

En la tabla 7.6 se muestra la simulación de un paciente con hemiparesia del lado izquierdo que presenta una rápida recuperación, por lo tanto, podría suponerse que se trata de un paciente subagudo, cuyo EVC ocurrió hace menos de 6 meses, y por lo tanto se encuentra dentro de un periodo de neuroplasticidad aumentada, lo cual propiciará que se produzca la mayor recuperación dentro de toda la evolución del paciente, así que es de esperar que obtenga notables beneficios con los ejercicios. Tal y como se mencionó en el capítulo “Rehabilitación posterior a un EVC”, esta recuperación tan marcada es característica de los pacientes subagudos.

Para la simulación se empleó una tasa de aprendizaje $\alpha=0.05$ y un factor de descuento $\gamma=0.7$. El agente realizó una planeación de 3000 episodios por cada 10 episodios reales. La simulación completa dura 600 episodios y se considera que el paciente presentaría cambios en su comportamiento cada 80 partidas. Para tratar de acercarnos al comportamiento real que tendría el agente cada simulación de 600 episodios se repitió 50 veces, los valores mostrados en la tabla son los valores promedio.

En este experimento, en promedio, en 76.9% de las 600 partidas se lograron 6 aciertos y 4 fallos. Conforme el paciente era capaz de alcanzar dianas de mayor altura el agente se iba adaptando y asignándole las dianas acorde con los nuevos límites que iba detectando. Se puede observar que, a lo largo del tiempo existieron ocasiones en las que el agente presentó dianas que se encontraban fuera del alcance de las posibilidades del paciente, estas transgresiones llegaron a representar el 24% del total de los ejercicios entre las partidas 80 y 160, posteriormente este tipo de comportamientos no deseados no llegan a superar el 19% del total de los ejercicios. En promedio, de los 6000 ejercicios asignados a lo largo del experimento, el 13.60% de ellos sobrepasó las capacidades del paciente.

Tabla 7.6. Simulación del comportamiento del agente ante un paciente subagudo con hemiparesia del lado izquierdo, que mejora en su movilidad a lo largo de 2 meses usando el juego.

Episodios transcurridos	80	160	240	320	400	480	560	600
Probabilidad de realizar acciones de abducción con brazo izquierdo	I1:1.0	I1:1.0	I1:1.0	I1:1.0	I1:1.0	I1:1.0	I1:1.0	I1:1.0
	I2:0.2	I2:0.4	I2:0.7	I2:0.9	I2:1.0	I2:1.0	I2:1.0	I2:1.0
	I3:0.0	I3:0.1	I3:0.4	I3:0.6	I3:0.8	I3:0.9	I3:1.0	I3:1.0
	I4:0.0	I4:0.0	I4:0.0	I4:0.1	I4:0.3	I4:0.4	I4:0.5	I4:0.6
	I5:0.0	I5:0.0	I5:0.0	I5:0.0	I5:0.0	I5:0.0	I5:0.1	I5:0.1
Probabilidad de realizar acciones de flexión con brazo izquierdo	IF1:1.0	IF1:1.0	IF1:1.0	IF1:1.0	IF1:1.0	IF1:1.0	IF1:1.0	IF1:1.0
	IF2:0.2	IF2:0.3	IF2:0.6	IF2:0.8	IF2:0.9	IF2:1.0	IF2:1.0	IF2:1.0
	IF3:0.0	IF3:0.1	IF3:0.3	IF3:0.5	IF3:0.6	IF3:0.7	IF3:0.8	IF3:0.9
	IF4:0.0	IF4:0.0	IF4:0.0	IF4:0.0	IF4:0.0	IF4:0.1	IF4:0.2	IF4:0.3
	IF5:0.0	IF5:0.0	IF5:0.0	IF5:0.0	IF5:0.0	IF5:0.0	IF5:0.0	IF5:0.0
Probabilidad de realizar acciones bimanuales	B1:1.0	B1:1.0	B1:1.0	B1:1.0	B1:1.0	B1:1.0	B1:1.0	B1:1.0
	B2:0.2	B2:0.3	B2:0.6	B2:0.8	B2:0.9	B2:1.0	B2:1.0	B2:1.0
	B3:0.0	B3:0.1	B3:0.3	B3:0.5	B3:0.6	B3:0.7	B3:0.8	B3:0.9
	B4:0.0	B4:0.0	B4:0.0	B4:0.0	B4:0.0	B4:0.1	B4:0.2	B4:0.3
	B5:0.0	B5:0.0	B5:0.0	B5:0.0	B5:0.0	B5:0.0	B5:0.0	B5:0.0
Partidas con 6 aciertos	93.9±0.3	83.8±0.56	67.4±1.0	68.5±1.3	74.6±1.1	76.3±1.2	74.65±1.3	72.0±1.6
Fallos flexión	39.6±9.4	41.0±7.1	42.2±5.5	46.7±5.1	55.1±4.8	53.35±6.3	53.4±6.1	51.9±8.2
Fallos abducción	38.6±8.3	34.8±7.5	33.0±4.6	26.4±4.9	14.0±2.9	19.0±4.3	18.8±3.9	18.4±6.6
Fallos bimanuales	21.8±5.3	24.0±3.1	24.75±2.8	26.9±2.9	30.8±3.0	27.6±3.8	27.8±3.2	29.6±4.8
Porcentaje de transgresiones	13.7±5.4	24.1±2.7	13.0±1.9	18.7±2.0	1.79±1.6	7.0±2.5	14.4±2.3	16.1±3.8
Porcentaje de aparición de las dianas								
I1	11.9±1.6	9.7±1.7	5.0±1.1	3.1±0.9	2.5±1.0	1.7±0.8	1.7±0.95	1.55±1.1
I2	12.6±3.7	9.68±2.3	9.2±1.6	8.69±1.7	7.0±1.3	4.57±1.1	3.6±1.2	3.29±2.2
I3	5.34±1.6	8.59±2.91	9.23±2.28	8.27±1.39	8.88±1.66	9.31±1.84	8.80±1.39	8.13±2.5
I4	0.17±0.2	0.54±0.47	4.38±1.32	5.98±1.82	4.33±1.05	6.30±1.47	7.36±1.6	6.93±2.4
I5	0.02±0.1	0.08±0.17	0.37±0.40	0.66±0.52	0.71±0.64	2.74±1.18	3.77±1.1	4.42±2.1
IF1	11.5±2.2	10.3±1.2	5.09±1.14	2.94±0.75	2.59±0.89	1.61±0.68	1.39±0.7	1.80±1.3
IF2	12.7±4.7	11.8±3.7	10.89±2.3	9.75±2.01	8.32±2.16	5.21±1.40	3.40±1.1	3.34±1.5
IF3	5.51±1.8	8.54±2.5	10.67±2.4	9.25±2.04	8.87±1.64	9.42±2.89	9.0±2.23	8.9±3.3
IF4	0.12±0.2	0.59±0.4	4.72±1.74	10.65±2.0	16.16±2.1	16.38±3.1	13.83±3.8	11.98±4.8
IF5	0.03±0.1	0.20±0.3	0.39±0.4	0.69±0.58	0.64±0.47	2.71±1.26	7.06±1.8	9.64±3.4
B1	9.64±2.4	7.69±1.5	4.36±0.9	2.91±1.09	1.5±0.76	0.93±0.46	0.71±0.3	0.7±0.51
B2	7.79±3.0	6.73±2.3	6.48±1.25	5.16±0.77	4.29±0.89	4.02±1.03	3.23±0.9	2.11±1.2
B3	2.36±1.7	4.80±1.8	5.96±1.59	5.18±1.3	4.54±0.83	5.20±0.89	5.13±1.1	5.2±1.7
B4	0.16±0.2	0.74±0.6	2.93±1.14	6.41±1.36	9.21±1.34	8.26±2.2	7.33±1.7	5.54±1.8
B5	0.02±0.1	0.02±0.06	0.26±0.27	0.33±0.37	0.44±0.4	1.57±1.0	3.59±1.4	6.44±1.9

Es interesante observar que, como el paciente no mejoraba de forma uniforme, el algoritmo trataba de asignar siempre más fallos en aquellos movimientos donde la mejora era rezagada, de tal forma que se incentivaba a un mayor esfuerzo por parte del paciente para llegar a un equilibrio en términos de capacidades motrices.

En la figura 7.2 se puede apreciar que conforme el paciente mejoraba, en promedio, la tendencia era que obtuviera más de 6 aciertos, sobre todo después de los 320 episodios. Esto tiene sentido, pues significa que en varias ocasiones el agente asignaba una diana, pensando que el paciente no la alcanzaría y ocurría justo lo contrario, sin embargo, esta tendencia no es muy significativa, ya

que en más del 70% de las partidas el agente lograba adaptarse al cambio de comportamiento del paciente y cumplir con la meta de lograr 6 aciertos y 4 errores.

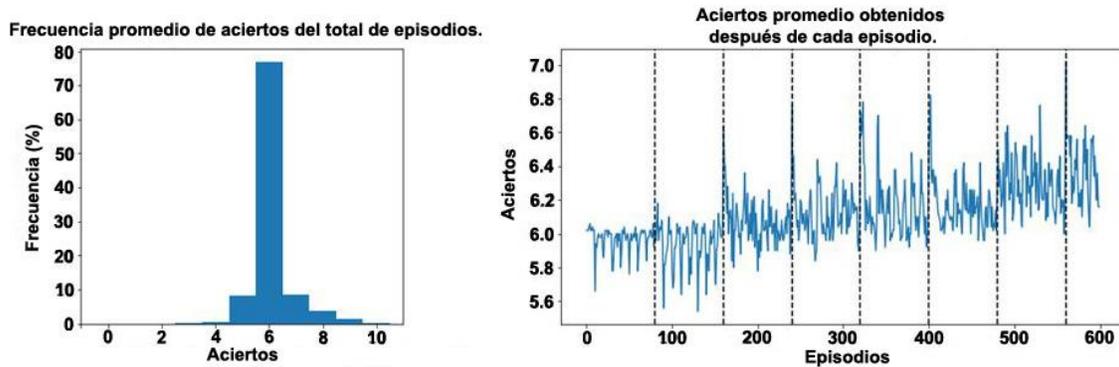


Figura 7.2. Se muestra la frecuencia promedio de los marcadores obtenidos en las 600 partidas simuladas, así como el número de aciertos promedio obtenido en cada partida simulando a un paciente EVC que presenta una mejoría en el tiempo.

En la figura 7.3 se puede apreciar la manera en la que el agente cambia la asignación de las dianas a lo largo del tiempo. En un inicio, cuando el paciente tenía una baja movilidad, el agente se enfocó en asignar dianas de baja altura, como las dianas I1, IF1 y B1, no era común asignar dianas que superaran con creces las capacidades del paciente, como las dianas I5, IF5 o B5. Sin embargo, conforme el paciente va mejorando se puede observar una disminución en la asignación de las dianas de altura más baja al mismo tiempo que aumenta la aparición de dianas de mayor altura, las cuales en ocasiones se encontraban más allá de las capacidades del paciente. Para la parte final del experimento, a partir del episodio 480, el agente exige en mayor medida las dianas I4, IF4 y B5.

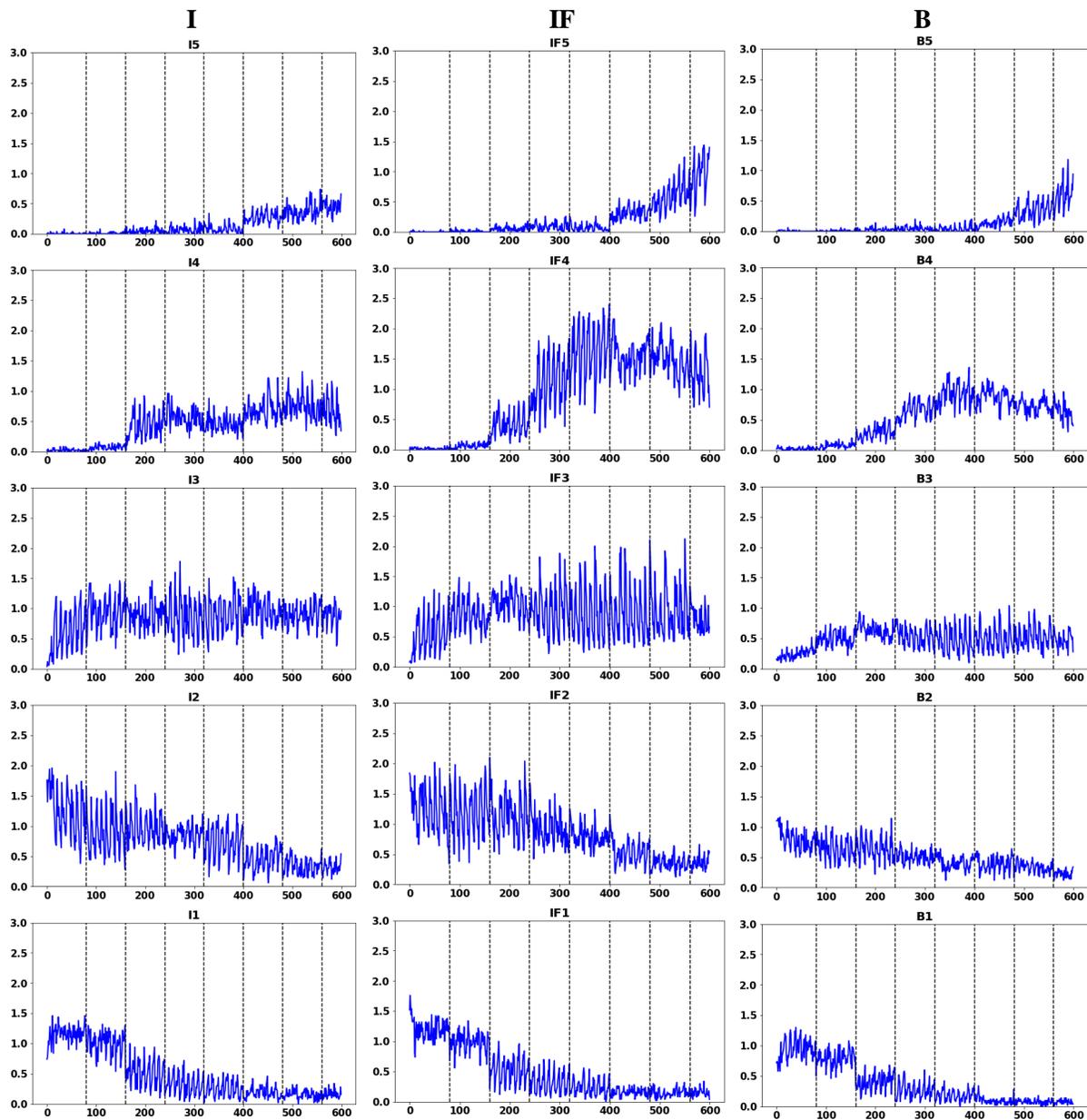


Figura 7.3. Progreso de la asignación de las dianas de abducción (columna I) y flexión (columna IF) del lado izquierdo, así como de las dianas bimanuales (columna B), de acuerdo al comportamiento que el paciente presentó a lo largo de 600 episodios. En el eje Y se mide la frecuencia promedio de aparición de las dianas y en el eje X los episodios transcurridos.

7.2.2 Simulación de un paciente EVC que no muestra cambios

En la tabla 7.7 se muestra la simulación de un paciente con hemiparesia del lado derecho que en las primeras partidas presenta cambios favorables, sin embargo, ese progreso se verá “estancado” después de cierto tiempo. Este tipo de comportamiento se esperaría de un paciente EVC crónico, en donde la neuroplasticidad se encuentra significativamente disminuida en comparación con la fase subaguda. Considerando lo anterior, con una práctica intensa de los ejercicios es probable que el paciente aún logre mejorar, pero tomará más de 2 meses y los cambios serán pequeños.

Para esta simulación se empleó una tasa de aprendizaje $\alpha=0.05$, un factor de descuento $\gamma=0.7$. El agente realizó una planeación de 3000 episodios por cada 10 episodios reales. La simulación dura 600 episodios y se considera que el paciente presenta cambios en su comportamiento cada 80 partidas, sólo durante los primeros 320 episodios. Cada simulación se repitió 50 veces, en la tabla 7.7 se muestran los valores promedio.

Tabla 7.7. Simulación del comportamiento del agente ante un paciente crónico con hemiparesia del lado derecho, deje de presentar mejoría después de una evolución inicial favorable a lo largo de 2 meses usando el juego.

Episodios transcurridos	80	160	240	320	400	480	560	600
Probabilidad de realizar acciones de abducción con brazo derecho	D1:1.0	D1:1.0	D1:1.0	D1:1.0	D1:1.0	D1:1.0	D1:1.0	D1:1.0
	D2:0.7	D2:0.8	D2:0.9	D2:1.0	D2:1.0	D2:1.0	D2:1.0	D2:1.0
	D3:0.4	D3:0.5	D3:0.6	D3:0.7	D3:0.7	D3:0.7	D3:0.7	D3:0.7
	D4:0.0	D4:0.1	D4:0.2	D4:0.3	D4:0.3	D4:0.3	D4:0.3	D4:0.3
	D5:0.0	D5:0.0	D5:0.0	D5:0.0	D5:0.0	D5:0.0	D5:0.0	D5:0.0
Probabilidad de realizar acciones de flexión con brazo derecho	DF1:1.0	DF1:1.0	DF1:1.0	DF1:1.0	DF1:1.0	DF1:1.0	DF1:1.0	DF1:1.0
	DF2:0.6	DF2:0.7	DF2:0.8	DF2:0.8	DF2:0.8	DF2:0.8	DF2:0.8	DF2:0.8
	DF3:0.0	DF3:0.1	DF3:0.2	DF3:0.2	DF3:0.2	DF3:0.2	DF3:0.2	DF3:0.2
	DF4:0.0	DF4:0.0	DF4:0.0	DF4:0.0	DF4:0.0	DF4:0.0	DF4:0.0	DF4:0.0
	DF5:0.0	DF5:0.0	DF5:0.0	DF5:0.0	DF5:0.0	DF5:0.0	DF5:0.0	DF5:0.0
Probabilidad de realizar acciones bimanuales	B1:1.0	B1:1.0	B1:1.0	B1:1.0	B1:1.0	B1:1.0	B1:1.0	B1:1.0
	B2:0.6	B2:0.7	B2:0.8	B2:0.8	B2:0.8	B2:0.8	B2:0.8	B2:0.8
	B3:0.0	B3:0.1	B3:0.2	B3:0.2	B3:0.2	B3:0.2	B3:0.2	B3:0.2
	B4:0.0	B4:0.0	B4:0.0	B4:0.0	B4:0.0	B4:0.0	B4:0.0	B4:0.0
	B5:0.0	B5:0.0	B5:0.0	B5:0.0	B5:0.0	B5:0.0	B5:0.0	B5:0.0
Partidas con 6 aciertos	84.1±0.6	73.5±1.0	73.2±0.9	77.75±0.9	77.22±0.9	78.2±1.0	77.3±0.74	77.14±1.3
Fallos flexión	68.1±4.0	60.9±6.2	65.4±5.9	66.36±4.2	65.81±5.1	66.03±5.1	64.8±6.1	66.23±5.9
Fallos abducción	0.78±0.7	6.05±3.0	1.6±1.36	1.21±1.21	1.82±1.66	1.40±1.12	1.42±1.16	1.44±2.02
Fallos bimanuales	31.1±3.4	32.9±4.2	32.9±3.7	32.41±3.0	32.3±3.5	32.56±3.1	33.78±3.0	32.32±4.3
Porcentaje de transgresiones	0.8±0.05	5.93±0.8	14.4±1.1	12.44±1.4	13±0.9	13±1.4	12.54±0.6	13.12±1.2
Porcentaje de aparición de las dianas								
D1	8.8±1.5	5.36±1.1	3.16±1.1	2.32±0.85	2.41±0.8	2.48±0.86	2.55±0.84	2.37±0.9
D2	8.7±1.5	10.7±1.4	9.45±1.62	9.87±1.5	9.55±1.6	10.24±1.6	9.9±1.47	9.79±2.1
D3	4.3±1.0	6.55±1.32	7.67±1.43	8.4±1.64	8.4±1.3	8.02±1.4	8.42±1.4	8.11±1.9
D4	0.2±0.2	2.07±0.99	3.5±1.2	3.54±1.0	3.82±1.1	3.61±1.13	3.85±1.22	3.62±1.5
D5	0.04±0.1	0.28±0.36	0.52±0.4	0.42±0.4	0.63±0.6	0.48±0.38	0.49±0.4	0.49±0.7
DF1	8.85±1.4	5.65±1.20	3.45±1.0	2.38±0.7	2.36±0.7	2.53±0.79	2.41±0.75	2.47±1.0
DF2	7.95±1.4	9.79±1.7	9.44±1.6	8.3±1.8	8.30±1.8	8.48±1.77	8.64±1.6	8.9±2.2
DF3	20.7±1.5	17.58±3.1	15.72±3.4	16.5±3.1	16.01±3.2	15.72±3.8	15.57±4.2	15.43±4.0
DF4	0.3±0.2	2.24±1.47	6.5±2.14	7.62±2.2	7.99±2.33	7.91±2.44	7.48±2.3	8.30±3.3
DF5	0.07±0.1	0.27±0.35	0.56±0.49	6.0±0.55	0.50±0.42	0.48±0.42	0.64±0.6	0.51±0.8
B1	6.0±1.5	4.01±1.04	3.07±1.17	2.61±0.84	2.85±0.92	2.72±0.97	2.35±0.8	2.72±1.4
B2	4.8±0.8	5.08±1.27	5.71±1.14	5.47±1.0	5.50±0.95	5.46±0.9	5.39±1.0	5.62±1.5
B3	9.0±1.3	9.82±2.07	7.8±2.13	8.08±2.0	7.74±2.18	7.64±2.28	8.32±2.4	7.82±2.9
B4	0.12±0.2	0.93±0.74	3.09±1.3	3.56±1.74	3.63±1.75	3.97±1.42	3.69±1.6	3.6±2.17
B5	0.04±0.1	0.14±0.17	0.28±0.33	0.25±0.3	0.25±0.3	0.19±0.26	0.24±0.3	0.22±0.4

En este experimento, en promedio, en 77.49% partidas se lograron 6 aciertos y 4 fallos. El número de transgresiones se mantuvo bajo, llegando a alcanzar su máximo punto entre los episodios 160 y 240, que es cuando se presenta la última mejora significativa. En promedio, el 10.65% del total de los ejercicios asignados (6000) superaron las capacidades del paciente, en este caso se cometieron

menos transgresiones porque el paciente mostró menos cambios de comportamiento en comparación al paciente anterior. El mayor número de transgresiones suele ocurrir en las partidas en las que ocurren cambios en las probabilidades de las acciones, ya que el agente tarda algunos episodios en identificar que algunas de sus acciones ya no reciben la recompensa esperada.

Se puede observar que, al igual que en el experimento anterior, en un inicio el agente se iba adaptando a los cambios que presentaba el paciente, exigiéndole dianas cada vez más altas con mayor frecuencia. Sin embargo, en cuanto el paciente deja de presentar mejoras el agente también mantiene la misma distribución de las dianas a asignar, manteniéndose todo prácticamente igual a partir del episodio 320. El último cambio ocurre en la función de abducción del brazo derecho en el episodio 240, en ese momento se vuelve más fácil para el paciente realizar la abducción de hombro para detener el balón dirigido a las dianas D2, D3 y D4, que se refleja en un cambio muy ligero en la distribución de asignación de esas dianas del episodio 240 al 320. Al igual que en los anteriores experimentos, se sigue manteniendo el comportamiento del agente de demandar aquellos ejercicios en donde el paciente tenga menor movilidad.

En la figura 7.4 se puede apreciar que prácticamente a lo largo de todos los episodios se logró que el paciente obtuviera 5, 6 ó 7 aciertos, obteniendo la mayoría de las veces 6 aciertos y 4 fallos.

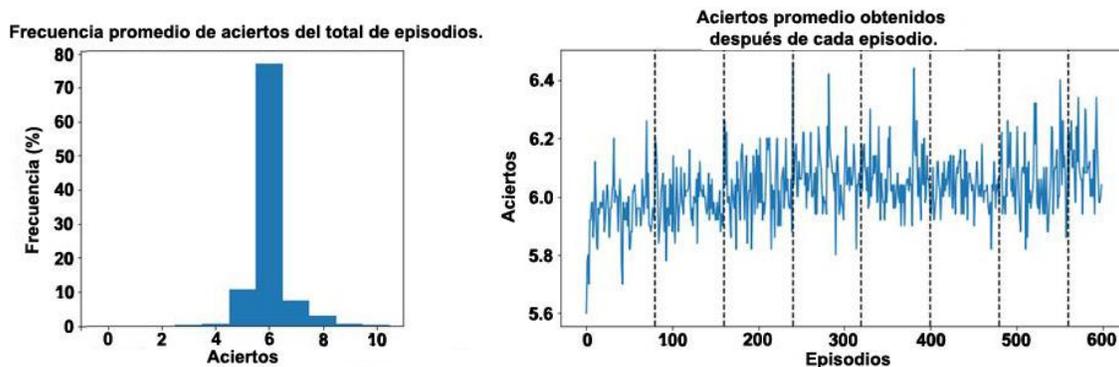


Figura 7.4. Se muestra la frecuencia promedio de los marcadores obtenidos en las 600 partidas simuladas, así como el número de aciertos promedio obtenido en cada partida simulando a un paciente EVC que presenta una mejora durante las primeras 240 partidas y posteriormente deja de presentar cambios en su función motriz.

En la figura 7.5 se puede apreciar la forma en que cambió la asignación de las dianas por parte del agente a lo largo del tiempo. En un inicio, cuando el paciente tenía una baja movilidad, el agente se enfocó en asignar dianas de baja altura, como las dianas D1, DF1 y B1, en raras ocasiones se asignaron dianas que superaran con creces las capacidades del paciente, como las dianas D5, DF5 o B5. Una vez que el paciente deja de mejorar y sus límites se vuelven las acciones asociadas a las dianas D4, DF3 Y B4, que es alrededor del episodio 240, entonces el agente decide asignar esas dianas en mayor medida durante la partida y no vuelve a ocurrir ningún cambio notable. El agente prácticamente nunca asignar las dianas D5, DF5 y B5.

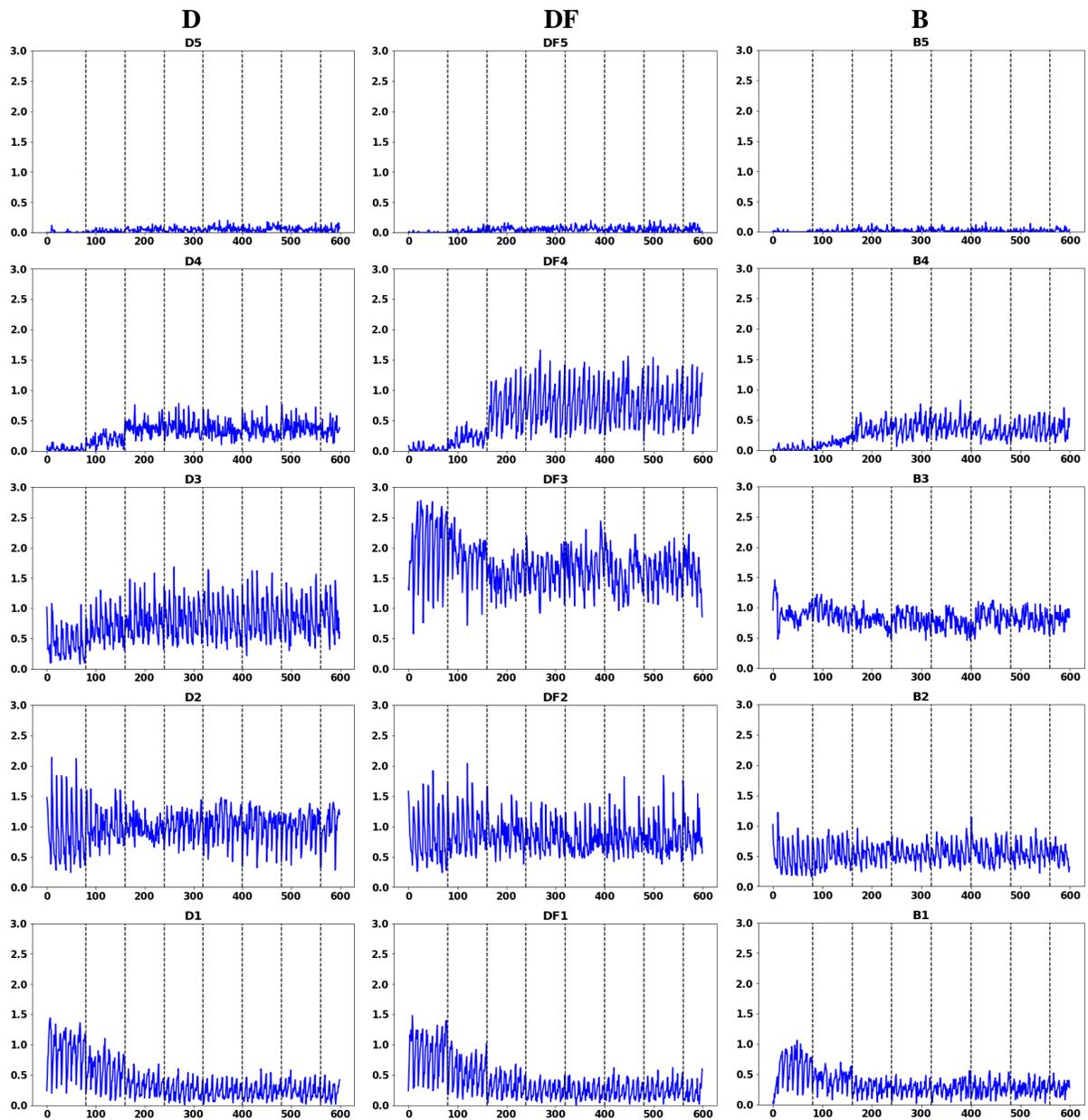


Figura 7.5. Progreso de la asignación de las dianas de abducción (columna D) y flexión (columna DF) del lado derecho, así como de las dianas bimanuales (columna B), de acuerdo a los cambios que el paciente presentaba a lo largo de 600 episodios. En el eje Y se mide la frecuencia promedio de aparición de las dianas y en el eje X los episodios transcurridos.

7.2.3 Simulación de un paciente EVC que decae

En la tabla 7.8 se muestra la simulación de un paciente EVC con hemiparesia del lado derecho. En este caso se intenta simular un escenario en el cual el paciente comience a decaer en su desempeño con el juego, este hecho podría presentarse debido a diversas causas, por ejemplo, que el paciente presente depresión, que haya sufrido alguna lesión muscular, o que a la hora en la que realiza sus ejercicios se encuentre agotado por haber realizado otra actividad.

Tabla 7.8. Simulación del comportamiento del agente ante un paciente EVC que decae en sus funciones motrices a lo largo de un periodo de 2 meses usando el juego.

Episodios transcurridos	80	160	240	320	400	480	560	600
Probabilidad de realizar acciones de abducción con brazo derecho	D1:1.0	D1:1.0	D1:1.0	D1:1.0	D1:1.0	D1:1.0	D1:1.0	D1:1.0
	D2:0.8	D2:0.8	D2:0.7	D2:0.6	D2:0.5	D2:0.4	D2:0.3	D2:0.2
	D3:0.5	D3:0.5	D3:0.4	D3:0.3	D3:0.2	D3:0.1	D3:0.0	D3:0.0
	D4:0.2	D4:0.1	D4:0.1	D4:0.0	D4:0.0	D4:0.0	D4:0.0	D4:0.0
	D5:0.0	D5:0.0	D5:0.0	D5:0.0	D5:0.0	D5:0.0	D5:0.0	D5:0.0
Probabilidad de realizar acciones de flexión con brazo derecho	DF1:1.0	DF1:1.0	DF1:1.0	DF1:1.0	DF1:1.0	DF1:1.0	DF1:1.0	DF1:1.0
	DF2:0.8	DF2:0.8	DF2:0.7	DF2:0.7	DF2:0.6	DF2:0.5	DF2:0.4	DF2:0.3
	DF3:0.6	DF3:0.6	DF3:0.5	DF3:0.4	DF3:0.3	DF3:0.2	DF3:0.1	DF3:0.0
	DF4:0.3	DF4:0.2	DF4:0.2	DF4:0.1	DF4:0.0	DF4:0.0	DF4:0.0	DF4:0.0
	DF5:0.0	DF5:0.0	DF5:0.0	DF5:0.0	DF5:0.0	DF5:0.0	DF5:0.0	DF5:0.0
Probabilidad de realizar acciones bimanuales	B1:1.0	B1:1.0	B1:1.0	B1:1.0	B1:1.0	B1:1.0	B1:1.0	B1:1.0
	B2:0.8	B2:0.8	B2:0.7	B2:0.7	B2:0.6	B2:0.5	B2:0.4	B2:0.3
	B3:0.6	B3:0.6	B3:0.5	B3:0.4	B3:0.3	B3:0.2	B3:0.1	B3:0.0
	B4:0.3	B4:0.2	B4:0.2	B4:0.1	B4:0.0	B4:0.0	B4:0.0	B4:0.0
	B5:0.0	B5:0.0	B5:0.0	B5:0.0	B5:0.0	B5:0.0	B5:0.0	B5:0.0
Partidas con 6 aciertos	52.9±1.5	57.35±1.6	56.2±1.13	64±1.08	68.57±0.9	75.12±0.6	86.7±0.4	91±0.45
Fallos flexión	39.6±6.2	39.55±6.3	44.55±6.0	44.07±6.7	38.10±7.1	40.22±7.4	34.05±5.9	42.5±10.1
Fallos abducción	32.1±6.1	32.44±8.8	25.0±7.2	23.44±8.6	35.5±6.2	34.50±8.8	43.60±6.5	31.9±12.7
Fallos bimanuales	28.2±2.1	28.0±3.1	30.4±3.1	32.47±2.6	26.39±2.7	25.27±3.0	22.33±3.0	25.55±5.8
Porcentaje de transgresiones	15.9±0.6	14.16±1.0	19.4±1.84	15.12±2.1	34.2±1.5	29.52±3.2	24.57±0.6	19.61±0.6
Porcentaje de aparición de las dianas								
D1	2.0±0.70	2.34±0.81	3.13±0.87	4.27±1.02	5.90±1.20	7.80±1.45	9.23±1.43	10.07±2.3
D2	7.12±2.3	8.2±2.31	9.57±2.31	10.37±2.6	11.92±2.8	11±3.21	11.81±2.8	14.89±5.6
D3	12.3±3.1	10.40±3.0	10.64±2.7	10.70±2..	11.13±2.8	11.45±3.2	13.71±2.0	8.99±3.6
D4	9.72±1.9	9.58±2.69	6.87±2.23	6.05±2.34	2.98±1.53	1.27±0.77	0.30±0.32	0.14±0.26
D5	1.88±1.1	1.39±0.85	1.07±0.78	0.48±0.4	0.31±0.40	0.1±0.19	0.06±0.13	0.03±0.12
DF1	2.01±0.7	2.38±0.72	3.35±0.94	4.66±1.04	6.11±1.15	7.76±1.33	9.66±1.7	10.91±2.4
DF2	6.95±2.1	8.34±2.35	9.61±2.54	9.88±1.8	9.86±1.65	10.34±2.2	9.0±2.17	8.35±2.91
DF3	8.84±2.1	8.75±2.37	9.205±2.5	9.04±1.71	8.53±2.34	8.71±2.2	5.8±1.41	6.38±2.53
DF4	6.96±1.8	7.05±1.79	5.29±1.52	3.84±1.36	2.89±1.25	1.35±0.8	0.34±0.36	0.19±0.38
DF5	2.20±1.2	1.53±0.94	1.23±0.9	0.67±0.55	0.33±0.42	0.18±0.3	0.05±0.12	0.02±0.09
B1	2.2±0.84	2.44±0.82	3.37±0.9	4.09±0.7	5.61±0.94	7.42±1.3	10.34±1.1	11.1±1.58
B2	5.42±1.0	5.56±1.01	6.21±1.17	6.19±1.09	6.31±1.12	6.10±1.15	5.33±1.16	5.03±1.46
B3	5.77±1.1	5.99±0.98	5.48±1.22	5.63±1.4	5.05±1.18	4.59±1.42	3.03±1.02	3.27±1.75
B4	4.46±1.0	4.34±1.06	3.6±1.1	3.15±1.16	2.43±1.05	1.72±0.7	1.21±0.48	0.56±0.72
B5	2.13±0.9	1.66±0.81	1.33±0.77	0.93±0.65	0.57±0.42	0.15±0.2	0.07±0.16	0.03±0.14

Para la simulación se empleó una tasa de aprendizaje $\alpha=0.05$ y un factor de descuento $\gamma=0.7$. El agente realizó una planeación de 3000 episodios por cada 10 episodios reales. La simulación dura

600 episodios y se considera que el paciente presenta cambios en su comportamiento cada 80 partidas. Cada simulación se repitió 50 veces, en la tabla se muestran los valores promedio.

En este experimento, en promedio, en 68.9% de las partidas se lograron 6 aciertos y 4 fallos, mientras que el promedio de transgresiones fue de 21.56% del total de ejercicios asignados. Conforme el paciente perdía movilidad y cada vez lograba alcanzar menos dianas de gran altura, el agente se iba adaptando, asignándole ejercicios acorde con los nuevos límites que iba detectando. A lo largo del tiempo aumenta el porcentaje de partidas en las que se logran 6 aciertos y 4 fallos, esto es porque le resultaba más fácil al agente asignar ejercicios que el paciente no lograba realizar correctamente, pues la mayoría de estos ejercicios en realidad creía que si los iba a poder realizar, pero cuando el agente notaba que esto no era así recurría a asignar dianas de menor altura, de las cuales estaba mucho más seguro sobre lo que ocurriría, esto se nota en la tabla 7.8 cuando aumenta la asignación de las dianas D2, DF2 y B2 a lo largo de los episodios.

Como durante los primeros episodios el agente detectó que el paciente era capaz de realizar correctamente los ejercicios de las dianas D3, DF3 y B3, a pesar de que el paciente fue disminuyendo su probabilidad de realizar esos ejercicios, el agente no logró disminuir completamente la asignación de esas dianas, las cuales fueron las causantes de la mayor parte de las transgresiones, esto aunado al hecho de que el paciente esporádicamente lograba alcanzar estas dianas aún después de 500 episodios ocasionaron que el agente no lograra dejar de asignar dichas dianas, es decir, al agente le toma más tiempo adaptarse a la disminución del rendimiento que a una mejora. En este experimento, en los últimos 40 episodios la probabilidad que tiene el paciente de alcanzar las dianas D3, DF3 y B3 es 0, pero ese intervalo de tiempo no es suficiente para que el agente logre aprender que el paciente no volverá a alcanzar esas dianas, aunque según se observa tanto en la figura 7.7, como en la tabla 7.8, existe una tendencia para ir disminuyendo cada vez más la aparición de esas acciones.

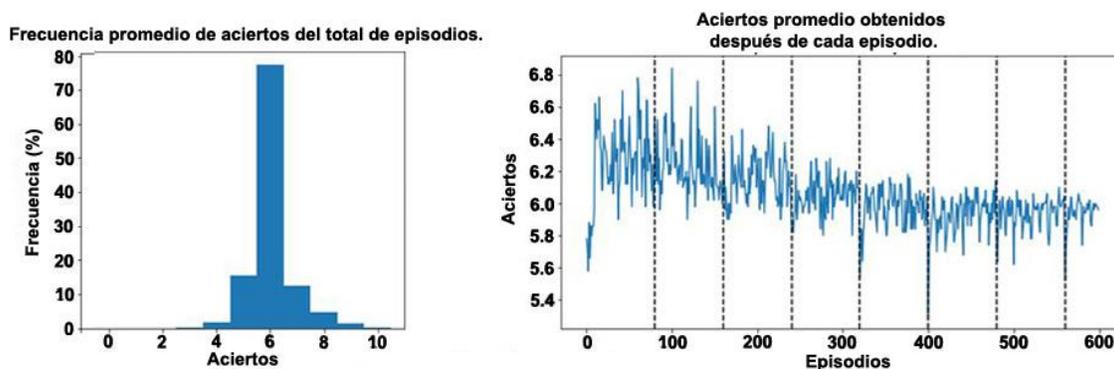


Figura 7.6. Se muestra la frecuencia promedio de los marcadores obtenidos en las 600 partidas simuladas, así como el número de aciertos promedio obtenido en cada partida simulando a un paciente EVC que presenta una disminución en sus capacidades motoras a lo largo del tiempo.

En la figura 7.7 se puede apreciar de mejor forma como cambió la asignación de las dianas por parte del agente a lo largo del tiempo. Contrario al caso en el que el paciente mejoraba, en este experimento se puede observar como las dianas de menor altura, como D1, D2, DF1, DF2, B1 y B2, comienzan a presentar un aumento en su número de apariciones, mientras que las dianas de

mayor altura disminuyen su presencia de forma marcada a lo largo del experimento. Este comportamiento responde a la disminución progresiva en las probabilidades del paciente ficticio de alcanzar las dianas de mayor altura. En un inicio las dianas D4, DF4 y B4 eran asignadas ocasionalmente, sin embargo, para el final del experimento casi nunca se presentan.

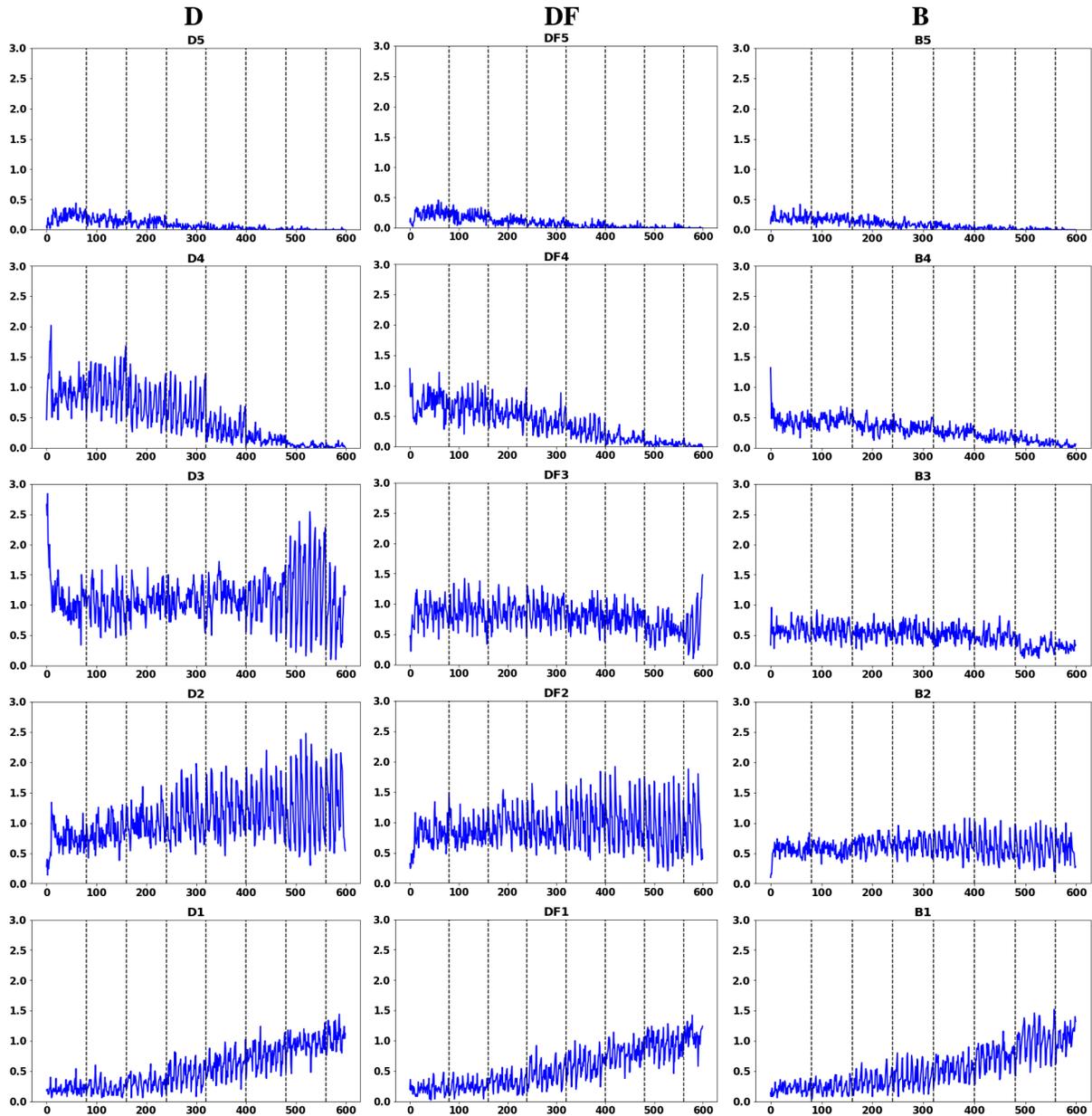


Figura 7.7. Progreso de la asignación de las dianas de abducción (columna D) y flexión (columna DF) del lado derecho, así como de las dianas bimanuales (columna B), de acuerdo a los cambios negativos que el paciente presentaba a lo largo de 600 episodios. En el eje Y se mide la frecuencia promedio de aparición de las dianas y en el eje X los episodios transcurridos.

7.3 Experimentos en un ambiente real

Todos los experimentos anteriores se realizaron a través de simulaciones, mediante la asignación de probabilidades a cada una de las acciones que podría ejecutar el paciente a lo largo de los episodios. En estos experimentos se evaluó el comportamiento del agente ya implementado en el videojuego, interactuando con un usuario real. Al igual que en los experimentos de casos específicos, en esta ocasión también se intentó simular 3 diferentes comportamientos: mejora, estabilidad y decaimiento.

Para todos los experimentos el participante usó una mancuerna de 8 libras para simular que su brazo tenía hemiparesia, de esta forma, aunque se le olvidara el comportamiento que debía actuar el peso del brazo le impediría comportarse de forma habitual.

Se conservaron los parámetros empleados en las simulaciones anteriores, es decir, una tasa de aprendizaje $\alpha=0.05$ y un factor de descuento $\gamma=0.7$, realizando un proceso de planeación de 3000 episodios, aunque en estos experimentos este proceso se realizó cada 5 episodios, en lugar de 10, debido a que el número de partidas que se jugaron en la realidad fue mucho menor que en los experimentos simulados. En ninguna tabla de este tipo de experimentos se muestra la primer partida, que es donde se crea por primera vez el modelo del ambiente M que emplea el agente para planificar.

7.3.1 Comportamiento del algoritmo ante un participante que mejora

En este primer experimento se evaluó la habilidad del algoritmo de adaptarse ante un jugador que mejora rápidamente en sus funciones motrices. El participante sujetó una mancuerna con su mano izquierda durante todo el experimento para tratar de comportarse como un paciente con hemiparesia. El participante jugó 20 partidas y cada 5 partidas se le pidió que intentará alcanzar dianas de una altura un poco mayor que las de las últimas 5 partidas.

Al inicio del experimento se le indicó al participante que las dianas más altas que debería alcanzar con el brazo izquierdo serían aquellas que requirieran tener su brazo levantado a menos de 90 grados respecto a su tronco, debía tratar de no detener cualquier diana que requiriera tener que realizar un movimiento de flexión o abducción con el hombro en un ángulo igual o mayor a 90 grados. De esta forma, de la partida 1 a la 5 las dianas más altas que lograba alcanzar el participante eran I2, IF2 y B2. De la partida 6 a la partida 10 las dianas más altas que consigue alcanzar son I3, IF2 y B2. A partir de la partida 11 hasta la partida 15, las dianas más altas que consigue alcanzar son I3, IF3 y B3. En las últimas 5 partidas las dianas más altas que logra alcanzar son I4, IF4 y B4.

Como se trata de experimentos mucho más cortos, en comparación de las simulaciones, se puede apreciar a mayor detalle el comportamiento del algoritmo.

En la tabla 7.9 se muestran los resultados de este experimento. La primera característica a destacar es que, a pesar de que el participante presentó 4 cambios de conducta en un periodo de

tiempo muy corto, el agente logró adaptarse muy bien, logrando en el 75% de las partidas que el usuario tuviera 6 aciertos y 4 fallos. De igual forma, se debe destacar que en ningún momento el agente le asignó dianas al participante que, con base en el comportamiento observado, excedieran sus capacidades, aunque probablemente haya ocurrido una coincidencia en la partida 16, ya que en ese momento el agente asignó la diana B5, justo cuando ocurre un cambio de comportamiento en el participante y ya es capaz de realizar correctamente el ejercicio asociado a esa diana.

Durante las primeras 5 partidas el participante tiene la misma movilidad en tareas de flexión, abducción y bimanuales. Por esta razón el agente asigna de forma alternada más ejercicios en tareas de abducción en algunas partidas y después más ejercicios en tareas de flexión en otras. Tampoco existe un enfoque muy marcado en algún tipo de ejercicio donde se busque que el participante cometa más errores, reforzando la idea de que el agente busca trabajar de manera equilibrada y con la misma intensidad los tres tipos de movimientos.

Tabla 7.9. Comportamiento del agente en un ambiente real con un participante que mejora en su motricidad a lo largo de 20 partidas. El color amarillo indica los episodios a partir de los cuales el participante presenta una mejora.

Episodios	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Puntaje																				
Aciertos	6	6	6	6	6	8	8	6	6	6	6	6	6	6	8	9	10	10	9	6
Ejercicios asignados al brazo izquierdo																				
Abducción	5	2	3	2	1	4	5	1	2	2	4	3	3	2	3	5	2	2	3	3
Flexión	1	3	3	4	5	2	1	5	4	4	2	3	3	4	3	1	4	4	3	3
Fallos																				
Abducción	2	1	2	1							3	2		1						2
Flexión		3	1	2	4			3	2	3	1	2	3	2	1					1
Bimanual	2		1	1		2	2	1	2	1			1	1	1	1			1	1
Transgresiones																				
Abducción																				
Flexión																				
Bimanual																				
Asignación de dianas																				
I1	1									1		1	1					2	2	
I2	2	1	1	1	1			1	1	1			1				2			
I3	2	1	2	1		4	5		1		1		1	1	3	3			1	
I4											3	2		1		2				2
I5																				1
IF1		1	1	1	1	1	1	1						1	1			3		1
IF2	1		1	1		1		1	2	1				1		1	3			
IF3		2	1	2	4			3	2	3	1	2		2	2		1			
IF4											1	1	3					1	3	1
IF5		1																		1
B1		1																		
B2			1	1	2			1		1	1		1				2			
B3	2	1	1	1		2	2	1	2	1	1	2		1	1	1				
B4													1	1	1			2	1	
B5																1			1	2

De la partida 6 a la partida 10 el participante presenta mayor movilidad de abducción que de flexión o de abducción. En ese periodo el agente tendría que haber exigido en algún momento la diana I4, que es la que se encuentra al límite de las capacidades del paciente. Sin embargo, de forma lógica, el agente decide enfocarse en las funciones motrices que se han quedado atrás y exige más movimientos de flexión. También en este periodo casi todas las dianas que requieren

movimientos bimanuales están pensadas para que el participante falle, es decir, el agente le exige al participante un esfuerzo mucho mayor para este tipo de movimientos que en las 5 primeras partidas.

De la partida 11 a la 15 el participante nuevamente presenta la misma movilidad en los tres tipos de tareas, lo que provoca que el agente nuevamente asigne dianas de flexión y de abducción de con una distribución equilibrada.

A partir de la partida 16, cuando el participante presenta un nuevo cambio de comportamiento, y hasta la partida 19, al agente le cuesta trabajo encontrar las dianas correctas para que el paciente falle, pues en este punto el participante logra realizar casi todos los ejercicios asignados y el agente no logra el objetivo de obtener 6 aciertos y 4 fallos en cada partida. Durante este periodo al participante no se le asignaron los ejercicios correctamente, pero se prefiere este tipo de comportamiento por parte de algoritmo al hecho de cometer transgresiones. En este periodo el algoritmo tarda un poco más de tiempo en reconocer qué acciones le otorgarán mayores recompensas, y es hasta la partida 19 cuando logra registrar que el participante es capaz de alcanzar las dianas IF4 y B4, entonces en la partida 20 el agente se comporta de forma eficiente y le exige al usuario las dianas I5, IF5 y B5, que son las que se encuentran al límite de sus capacidades, logrando finalmente la meta de obtener 6 aciertos y 4 fallos sin incurrir en transgresiones.

7.3.2 Comportamiento del algoritmo ante un participante que decae

En el segundo experimento se evaluó la habilidad del algoritmo de adaptarse ante un participante que empeora rápidamente en sus funciones motrices. El participante sujetó una mancuerna con su mano derecha, para tratar de comportarse como un paciente con hemiparesia. El usuario jugó 20 partidas, cada 5 partidas se le pidió al usuario que intentará alcanzar dianas de una altura un poco menor que las de las últimas 5 partidas.

De la partida 1 a la 5 las dianas más altas que lograba alcanzar el participante eran D3, DF4 y B4. De la partida 6 a la partida 10 las dianas más altas que consigue alcanzar son D3, DF3 y B3. A partir de la partida 11 hasta la partida 15, las dianas más altas que consigue alcanzar son D3, DF2 y B2. En las últimas 5 partidas las dianas más altas que logra alcanzar son D2, DF2 y B2.

En la tabla 7.10 se pueden observar los resultados de este experimento. En este caso el agente logró que el participante tuviera 6 aciertos y 4 fallos en el 80% de las partidas. Sin embargo, en esta ocasión en el 20% de los episodios ocurrió al menos 1 transgresión, aunque de forma comprensible, ya que la mayoría ocurrieron justo después de que el usuario presentara un cambio en su comportamiento, posteriormente el agente aprende cuáles son los nuevos límites del participante y no ocurren transgresiones.

Durante las primeras 5 partidas el participante presenta una mayor movilidad en tareas bimanuales y de flexión que en tareas de abducción. Por esta razón el agente asigna más ejercicios de abducción que de flexión, además de que todos los fallos se concentran en tareas de

abducción. Es interesante que en la primera partida ocurra una transgresión. El agente asigna la diana D5, cuando el límite del usuario es D4.

De la partida 6 a la partida 10 el participante presenta un decremento en su función motriz de forma homogénea en los rangos de movimiento para tareas bimanuales, de abducción y de flexión. Como consecuencia, el agente asigna dianas de flexión y abducción en la misma proporción y los ejercicios pensados para que el participante falle ya no se centran en movimientos de abducción.

Entre las partidas 11 y 15 el participante presenta una mayor movilidad en movimientos de abducción que en movimientos de flexión o en tareas bimanuales. Esto se ve inmediatamente reflejado en las dianas que asigna el agente. Ninguno de los ejercicios pensados para que el usuario falle se encuentran asociados a tareas de abducción, y durante las partidas el número de veces que se asignan dianas que requieren movimientos de flexión aumenta respecto a las dianas que requieren movimientos de abducción. En la partida 11 es donde varias transgresiones, pero son justificables, porque antes de ese momento las acciones que elegía el agente eran adecuadas.

Tabla 7.10. Comportamiento del agente en un ambiente real con un participante que disminuye progresivamente en sus funciones motrices a lo largo de 20 partidas. El color amarillo señala los episodios a partir de los cuales el participante presenta una decaída.

Episodio	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Puntaje																				
Aciertos	6	6	6	6	6	6	6	6	6	6	4	5	6	6	6	3	4	6	6	6
Ejercicios asignados al brazo derecho																				
Abducción	4	5	4	4	4	3	3	5	4	4	1	2	2	4	2	1	2	3	3	4
Flexión	2	1	2	2	2	3	3	1	2	2	5	4	4	2	4	5	4	3	3	2
Fallos																				
Abducción	3	4	4	4	4	2	3	2	2	2							1	1	1	2
Flexión							1	1	2	1	4	4	4	2	2	5	3	2	2	1
Bimanual	1					2		1		1	2	1		2	2	2	2	1	1	1
Transgresiones																				
Abducción	1																			
Flexión											2					5				
Bimanual											2					2	2			
Asignación de dianas																				
D1								1							1				1	
D2		1				1				1	1		1	2	1	1	1	2	1	2
D3								2	2	1		2		2			1	1	1	2
D4	3	4	4	4	4	2	3	2	2	2										
D5	1																			
DF1		1					1						1							1
DF2				1		2	1				1				2		1	1	1	
DF3			1	1		1				1	2	4	4	2	2		3	2	2	1
DF4	2		1		2		1	1	2	1	2					5				
DF5																				
B1										1										
B2												1	2					1	1	1
B3					1		2	1	2			1		2	2			1	1	1
B4	1	2	2	2	1	2		1		1	2					2	2			
B5	1																			

A partir de la partida 16, cuando el participante presenta un nuevo cambio de comportamiento, y hasta la partida 19, el usuario vuelve a tener la misma movilidad en los 3 tipos de tareas, por consiguiente, el agente vuelve a asignar ejercicios de flexión y de abducción de forma equilibrada, En la partida 16 es donde ocurren la mayor cantidad de transgresiones y donde el paciente obtiene la menor cantidad de aciertos, sin embargo, es justamente en este episodio donde el participante presenta un cambio de comportamiento, sin embargo, es interesante el hecho de que las transgresiones ocurrieron al signar dianas que desde hace más de 5 partidas se convirtieron en inalcanzables para el paciente.

Aunque el comportamiento del algoritmo no fue tan bueno como en el experimento anterior, ya que se cometieron transgresiones en 14 de los 200 ejercicios asignados por el agente (7% del total de ejercicios), en este experimento el agente logra adaptarse al comportamiento del participante al ir asignando dianas de altura cada vez menor, acorde con el cambio en el rendimiento observado.

7.3.3 Comportamiento del algoritmo ante un participante que no muestra cambios

En este último experimento, cuyos resultados se muestran en la tabla 7.1, se evaluó la efectividad del algoritmo ante un participante que no presenta cambios a lo largo del tiempo. El participante sujetó una mancuerna con su mano derecha para tratar de comportarse como un paciente con hemiparesia. El usuario jugó 20 partidas, en todas ellas se le pidió que las dianas más altas que debería alcanzar con el brazo izquierdo serían aquellas que requirieran tener su brazo levantado a 90 grados respecto a su tronco, debía tratar de no detener cualquier diana que requiriera tener que realizar un movimiento de flexión o abducción con el hombro en un ángulo mayor a 90 grados.

En este experimento, como se esperaba, es donde se obtuvo el mejor comportamiento del agente, obteniendo en el 95% de las partidas 6 aciertos y 4 fallos. Además, en ninguna partida se cometió transgresión alguna. En todo momento las dianas más difíciles que asignó el agente fueron D4, DF4 o B4. Como el grado de movilidad que presentaba el participante era el mismo para las tareas de flexión, abducción y bimanuales, el agente asignó este tipo de ejercicios de forma balanceada, a veces alternando entre exigir más movimientos de flexión y otras veces más de abducción, pero siempre tratando de mantener un equilibrio. Únicamente ocurre un suceso extraño en la partida 5, donde el agente no asigna ninguna diana pensada para que el participante falle, pero queda como anecdótico, pues algo similar no se vuelve a repetir.

Tabla 7.11. Comportamiento del agente en un ambiente real con un participante que no presenta ningún cambio en su función motriz a lo largo de 20 partidas.

Episodio	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Puntaje																				
Aciertos	6	6	6	6	10	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
Ejercicios asignados al brazo derecho																				
Abducción.	3	4	2	3	1	3	4	4	2	3	4	3	3	4	3	3	2	4	2	2
Flexión	3	2	4	3	5	3	2	2	4	3	2	3	3	2	3	3	4	2	4	4
Fallos																				
Abducción	2	2	1	1	0	1	2	2	2	2	2	1	2	2	1	1	2	2	2	1
Flexión	1	1	2	2	0	2	1	1	1	1	1	2	1	1	2	2	1	1	1	2
Bimanual	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Transgresiones																				
Abducción																				
Flexión																				
Bimanual																				
Asignación de dianas																				
D1																				
D2	1				1	1				1		1	1			1				
D3	1	2	1	2		1	2	2	1	1	2	1	1	2	2	1	1	2	1	1
D4	1	2	1	1		1	2	2	1	1	2	1	1	2	1	1	1	2	1	1
D5																				
DF1								1					1							
DF2	2			1	3				1	2					1		1		1	
DF3	1	2	2	1	2	1	1	1	2	1	1	1	1	2	1	1	2	2	2	2
DF4			2	1		2	1		1		1	2	1		1	2	1		1	2
DF5																				
B1							1													
B2					2	1										1				
B3	1	1	1	1				1	1	1		1		1	1		1	1	1	1
B4	1	1	1	1		1	1	1	1	1	2	1	2	1	1	1	1	1	1	1
B5																				

Finalmente, vale la pena resaltar que, incluso en escenarios donde el participante no cambia su comportamiento, el agente cumple cabalmente con una tercera meta para la cual no fue entrenado: la secuencia de dianas asignadas siempre es diferente en cada partida. En algunos momentos puede que se asignen las mismas dianas que en la partida anterior, pero el orden siempre es diferente. Es verdad que este objetivo es menos importantes que los otros 2 (lograr 6 aciertos y 4 errores en cada partida y no sobrepasar las capacidades del paciente), sin embargo, el añadir variedad al juego le suma dinamismo y le resta monotonía, lo que aporta a que la experiencia siempre se sienta diferente y contribuye a que el paciente no se aburra. Si el agente siempre repitiera la misma secuencia de dianas en el mismo orden, el paciente podría perder interés porque el juego se volvería predecible e inclusive se podría generar una habituación.

Capítulo 8

Conclusiones y trabajo a futuro

Con esta aplicación se espera motivar al paciente a que realice sus ejercicios de forma intensiva y generar un mayor apego a su tratamiento a través de una experiencia divertida en un ambiente enriquecido. Las numerosas repeticiones de estos ejercicios favorecerán la recuperación de la motricidad del hombro. Además, se esperaría que las actividades bimanuales y la retroalimentación multimodal, junto con la neuroplasticidad aumentada en el caso de pacientes subagudos, favorezcan una reorganización cortical y con ello una mejor recuperación funcional. El videojuego también permite la aplicación de una tele-rehabilitación complementaria a la terapia convencional, pues un terapeuta puede asignar desde el mismo programa una rutina de ejercicios de forma remota que el paciente podrá jugar desde casa, sin tener que trasladarse a una clínica ni depender de la disponibilidad de citas en un hospital, de esta forma el paciente podrá incrementar la frecuencia con la que realiza sus ejercicios y lograr una mejor recuperación.

Ante la imposibilidad de poder probar el desarrollo con pacientes reales debido a la pandemia, la simulación del comportamiento de pacientes resulta una alternativa interesante que se puede aplicar a otro tipo de desarrollos en donde se necesiten grandes cantidades de datos. Entre sus principales beneficios se encuentran permitir abarcar una mayor cantidad de casos con mucho menos recursos y acelerar el proceso de desarrollo.

El algoritmo basado en técnicas de aprendizaje por reforzamiento logra ajustar adecuadamente y en tiempo real la dificultad del juego Penal Madness. Se realizaron diversas pruebas que demuestran el adecuado comportamiento del algoritmo ante distintas circunstancias, ya sea un paciente que presente hemiparesia del brazo derecho o del brazo izquierdo, que tenga diferentes grados de movilidad para distintos tipos de movimiento, que presente una evolución favorable o un retroceso en su función motriz de forma uniforme o dispar en movimientos de flexión, abducción y tareas bimanuales, o incluso si el paciente no llegara a presentar ningún cambio. En la mayoría de las pruebas se logra en más del 70% de las partidas lograr un marcador final de 6 aciertos y 4 fallos, además del total de ejercicios asignados, en general, menos del 15% son ejercicios que llegan a sobrepasar las capacidades del paciente. Estos resultados son ligeramente mejores similares a los obtenidos en el trabajo de Yoones (Sekhavat, 2017), donde también se emplea aprendizaje por reforzamiento para el ajuste de la dificultad del juego para rehabilitación y en alrededor del 50% del total de las partidas se logró el objetivo de obtener 5 fallos y 5 aciertos, aunque no se debe obviar que sus pruebas fueron realizadas con pacientes EVC reales, sin embargo, a diferencia del trabajo aquí presentado, en este estudio no se consideró que los ejercicios destinados a fallar se encontrarán dentro del límite de las capacidades del paciente, por lo que a la larga puede inducir frustración al exigir ejercicios imposibles de realizar.

Durante el desarrollo del algoritmo se observó que el agente asignaba ejercicios relativamente bien sin utilizar procesos de planeación, en más del 90% de las partidas se obtenían 6 aciertos y 4 fallos, pero también el porcentaje de ejercicios que superaran las capacidades del paciente era alto cuando se presentaban cambios de comportamiento. Al añadir los procesos de planeación al algoritmo las mejoras en este punto fueron son sustanciales, a pesar de que en algoritmos basados en modelos, como Dyna-Q, este proceso es mucho más efectivo en ambientes deterministas, el impacto positivo de este mecanismo se mantuvo presente al simular diversos pacientes como ambientes estocásticos.

En un ambiente determinista el algoritmo logra un comportamiento perfecto, sin embargo, la mayoría de las simulaciones y experimentos se realizaron tratando acercarse lo más posible a la realidad a través de ambientes estocásticos, donde el programa no sólo logra que en la mayoría de las partidas el paciente alcance 6 aciertos y 4 fallos con el objetivo de mantener un suficiente grado de reto para el paciente, también logra en muchas ocasiones sólo exigir movimientos que se encuentran en el límite de las capacidades del paciente, aunque este último punto le resulta un poco más difícil de lograr cuando el paciente presenta cambios significativos en su comportamiento, sobre todo cuando ocurre una disminución en la función motriz, pues el agente tarda más en reconocer que el paciente ya no es capaz de realizar movimientos que antes si podía. El algoritmo también logra enfocarse en trabajar los aspectos de movilidad más débiles, presentando un mayor grado de exigencia en aquellos movimientos en los que el paciente tiene menor rango de movilidad, de forma similar a como lo haría un terapeuta.

Con la funcionalidad de ADD implementado en el juego, Penal Madness es una aplicación que tiene el potencial de poder ser usada de forma remota, sin necesidad de que el paciente cuente con la asistencia presencial de un terapeuta, pues el juego le asignará los ejercicios a realizar de forma automática. Sin embargo, este tipo de herramientas no están pensadas para sustituir la terapia convencional, si no para servir como complemento y como una alternativa más para aquellos pacientes que por distintas causas no puedan asistir a un hospital y, al mismo tiempo, ofrecer a los terapeutas una herramienta con la cual puedan atender a una mayor cantidad de pacientes.

Respecto al trabajo a futuro, a pesar de los buenos resultados obtenidos de las simulaciones, es claro que se deben realizar pruebas con pacientes EVC reales con distintos grados de afección con el fin de corroborar la efectividad del algoritmo; entre mayor sea el tamaño de cohorte y la diversidad de los comportamientos, el funcionamiento del algoritmo tendrá mayor validez.

Uno de los aspectos que se mencionó en los principios del desarrollo de juegos serios es el hecho de ofrecer recompensas a los jugadores, en ese sentido se planea agregar más modelos de porteros femeninos y masculinos, además de diferentes fondos que se habilitarán dependiendo del número de partidas que se hayan jugado, de esta forma se busca incrementar aún más el apego del paciente al juego y consecuentemente la frecuencia con la que realice sus ejercicios.

También se debe considerar desarrollar juegos alternativos, pues no todas las personas se verán atraídas hacia la idea de jugar como un portero, se deben crear una mayor diversidad de entornos donde las mecánicas de juego se conserven, en términos de los movimientos a realizar.

Finalmente, será conveniente explorar tecnologías alternativas para la detección de movimientos humanos, como los sensores inerciales, tomando en cuenta que el sensor Kinect ha sido discontinuado.

Bibliografía

- Albiol, S., Palacios, G., Guerrón, N., Gil-Gomez, J.-A., Quilis, J.-A., Gil-Gómez, H., & Manzano, P. (2014). The Perfetti method, a novel Virtual Fine Motor Rehabilitation system for Chronic Acquired Brain Injury. In *Proceedings - REHAB 2014*.
<https://doi.org/10.4108/icst.pervasivehealth.2014.255251>
- Alves, J., Vourvopoulos, A., Bernardino, A., & Bermúdez i Badia, S. (2014). Eye Gaze Patterns after Stroke: Correlates of a VR Action Execution and Observation Task. In *Proceedings - REHAB 2014*. <https://doi.org/10.4108/icst.pervasivehealth.2014.255288>
- Arauz, A., & Ruiz-Franco, A. (2012). Enfermedad vascular cerebral. *Revista Mexicana de Neurociencia*, 7(6), 617–621.
- Archambault, P., Norouzi-Gheidari, N., Kairy, D., Solomon, J., & Levin, M. (2014). Towards Establishing Clinical Guidelines for an Arm Rehabilitation Virtual Reality System. In *Biosystems and Biorobotics* (Vol. 7, pp. 263–270). https://doi.org/10.1007/978-3-319-08072-7_45
- Arias Cuadrado, A. (2009). Rehabilitación del ACV: evaluación, pronóstico y tratamiento. *Galicia Clinica*. [https://doi.org/10.1016/0263-8223\(91\)90022-Q](https://doi.org/10.1016/0263-8223(91)90022-Q)
- Baños, R., Botella, C., Perpiñá, C., Alcañiz Raya, M., Lozano, J., Osma, J., & Gallardo, M. (2002). Virtual reality treatment of flying phobia. *IEEE Transactions on Information Technology in Biomedicine : A Publication of the IEEE Engineering in Medicine and Biology Society*, 6, 206–212. <https://doi.org/10.1109/TITB.2002.802380>
- Barreca, S., Gowland, C. K., Stratford, P., Huijbregts, M., Griffiths, J., Torresin, W., Dunkley, M., Miller, P., & Masters, L. (2004). Development of the Chedoke Arm and Hand Activity Inventory: theoretical constructs, item generation, and selection. *Topics in Stroke Rehabilitation*, 11(4), 31–42. <https://doi.org/10.1310/JU8P-UVK6-68VW-CF3W>
- Barrett, N., Swain, I., Gatzidis, C., & Mecheraoui, C. (2016). The use and effect of video game design theory in the creation of game-based systems for upper limb stroke rehabilitation. *Journal of Rehabilitation and Assistive Technologies Engineering*, 3, 205566831664364. <https://doi.org/10.1177/2055668316643644>
- Basic-Science. (2020). *Ilustración de un EVC hemorrágico*. <https://www.orthobullets.com/basic-science/12288/stroke>
- Bastioni, M. (2000). *MakeHuman*. <http://www.makehumancommunity.org/>
- Bellman, R. (1952). On the Theory of Dynamic Programming. *Proceedings of the National Academy of Sciences of the United States of America*, 38(8), 716–719. <https://doi.org/10.1073/pnas.38.8.716>
- Benjamin, E. J., Blaha, M. J., Chiuve, S. E., Cushman, M., Das, S. R., Deo, R., de Ferranti, S. D., Floyd, J., Fornage, M., Gillespie, C., Isasi, C. R., Jiménez, M. C., Jordan, L. C., Judd, S. E., Lackland, D., Lichtman, J. H., Lisabeth, L., Liu, S., Longenecker, C. T., ... Muntner, P. (2017). Heart Disease

and Stroke Statistics-2017 Update: A Report From the American Heart Association. *Circulation*, 135(10), e146–e603. <https://doi.org/10.1161/CIR.0000000000000485>

- Berg, K., Wood-Dauphinee, S., & Williams, J. I. (1995). The Balance Scale: reliability assessment with elderly residents and patients with an acute stroke. *Scandinavian Journal of Rehabilitation Medicine*, 27(1), 27–36.
- Beyrouti, R., Adams, M. E., Benjamin, L., Cohen, H., Farmer, S. F., Goh, Y. Y., Humphries, F., Jäger, H. R., Losseff, N. A., Perry, R. J., Shah, S., Simister, R. J., Turner, D., Chandratheva, A., & Werring, D. J. (2020). Characteristics of ischaemic stroke associated with COVID-19. In *Journal of neurology, neurosurgery, and psychiatry* (Vol. 91, Issue 8, pp. 889–891). <https://doi.org/10.1136/jnnp-2020-323586>
- Bohannon, R., Harrison, S., & Kinsella-Shaw, J. (2009). Reliability and validity of pendulum test measures of spasticity obtained with the Polhemus tracking system from patients with chronic stroke. *Journal of Neuroengineering and Rehabilitation*, 6, 30. <https://doi.org/10.1186/1743-0003-6-30>
- Bohannon, R., & Smith. (1987). Interrater reliability of a modified Ashworth scale of muscle spasticity. *Physical Therapy*, 67(2), 206–207. <https://doi.org/10.1093/ptj/67.2.206>
- Bravo, M. D. A., Rengifo, R. C. F., & Agredo, R. W. (2016). Comparación de dos Sistemas de Captura de Movimiento por medio de las Trayectorias Articulares de Marcha. *Revista Mexicana de Ingenieria Biomedica*, 37(2), 149–160. <https://doi.org/10.17488/RMIB.37.2.2>
- Broeren, J., Björkdahl, A., Claesson, L., Goude, D., Lundgren-Nilsson, Å., Samuelsson, H., Blomstrand, C., Sunnerhagen, K., & Rydmark, M. (2008). Virtual rehabilitation after stroke. *Studies in Health Technology and Informatics*, 136, 77–82. <https://doi.org/10.3233/978-1-58603-864-9-77>
- Burdea, G. (2003). Virtual rehabilitation--benefits and challenges. *Yearbook of Medical Informatics*, 42(November). <https://doi.org/10.1267/METH03050519>
- Burdea, G., Grampurohit, N., Kim, N., Polistico, K., Kadaru, A., Pollack, S., Oh-Park, M., Barrett, A. M., Kaplan, E., Masmela, J., & Nori, P. (2020). Feasibility of integrative games and novel therapeutic game controller for telerehabilitation of individuals chronic post-stroke living in the community. *Topics in Stroke Rehabilitation*, 27(5), 321–336. <https://doi.org/10.1080/10749357.2019.1701178>
- Cameirão, M. S., Badia, S. B., Duarte, E., Frisoli, A., & Verschure, P. (2012). The combined impact of virtual reality neurorehabilitation and its interfaces on upper extremity functional recovery in patients with chronic stroke. *Stroke*, 43(10), 2720–2728. <https://doi.org/10.1161/STROKEAHA.112.653196>
- Cameirão, M. S., Badia, S. B., Oller, E. D., & Verschure, P. F. (2010). Neurorehabilitation using the virtual reality based Rehabilitation Gaming System: methodology, design, psychometrics, usability and validation. *Journal of Neuroengineering and Rehabilitation*, 7, 48. <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed11&NEWS=N&AN=20860808>
- Campbell, B. C. V., De Silva, D. A., Macleod, M. R., Coutts, S. B., Schwamm, L. H., Davis, S. M., &

- Donnan, G. A. (2019). Ischaemic stroke. *Nature Reviews Disease Primers*, 5(1). <https://doi.org/10.1038/s41572-019-0118-8>
- Carr, K., Zachariah, N., Weir, P., & McNevin, N. (2011). An examination of feedback use in rehabilitation settings. *Critical Reviews in Physical and Rehabilitation Medicine*, 23(1–4), 147–160. <https://doi.org/10.1615/CritRevPhysRehabilMed.v23.i1-4.110>
- Carvalho, D., Teixeira, S., Lucas, M., Yuan, T. F., Chaves, F., Peressutti, C., Machado, S., Bittencourt, J., Menéndez-González, M., Nardi, A. E., Velasques, B., Cagy, M., Piedade, R., Ribeiro, P., & Arias-Carrión, O. (2013). The mirror neuron system in post-stroke rehabilitation. *International Archives of Medicine*, 6(1), 1–7. <https://doi.org/10.1186/1755-7682-6-41>
- Chang, Y. J., Chen, S. F., & Huang, J. Da. (2011). A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities. *Research in Developmental Disabilities*, 32(6), 2566–2570. <https://doi.org/10.1016/j.ridd.2011.07.002>
- Charles, D., Holmes, D., & Charles, T. (2020). Virtual Reality Design for Stroke Rehabilitation. In *Biomedical Visualisation. Advances in Experimental Medicine and Biology*, vol 1235. Springer, Cham. https://doi.org/10.1007/978-3-030-37639-0_4
- Chiquete, E., Ruiz Sandoval, J. L., Murillo Bonilla, L. M., Arauz, A., Villarreal-Careaga, J., Barinagarrementería, F., & Cantú-Brito, C. (2011). Mortalidad por enfermedad vascular cerebral en México, 2000-2008: Una exhortación a la acción. *Revista Mexicana de Neurociencia*, 12(5), 235–241.
- Chung, J. W., Park, S. H., Kim, N., Kim, W. J., Park, J. H., Ko, Y., Yang, M. H., Jang, M. S., Han, M. K., Jung, C., Kim, J. H., Oh, C. W., & Bae, H. J. (2014). Trial of ORG 10172 in acute stroke treatment (TOAST) classification and vascular territory of ischemic stroke lesions diagnosed by diffusion-weighted imaging. *Journal of the American Heart Association*, 3(4), 1–8. <https://doi.org/10.1161/JAHA.114.001119>
- Collin, C., & Wade, D. (1990). Assessing motor impairment after stroke: a pilot reliability study. *Journal of Neurology, Neurosurgery, and Psychiatry*, 53(7), 576–579. <https://doi.org/10.1136/jnnp.53.7.576>
- Colomer, C., Baldoví, A., Torromé, S., Navarro, M. D., Moliner, B., Ferri, J., & Noé, E. (2013). Eficacia del sistema Armeo®Spring en la fase crónica del ictus. Estudio en hemiparesias leves-moderadas. *Neurología*, 28(5), 261–267. <https://doi.org/10.1016/j.nrl.2012.04.017>
- Csikszentmihalyi, M. (1991). *Flow: The Psychology of Optimal Experience*. Harper & Row.
- De Almeida Oliveira, R., Cintia Dos Santos Vieira, P., Fernanda Rodrigues Martinho Fernandes, L., Patrizzi, L., De Oliveira, S. F., & De Souza, L. A. P. S. (2014). Mental practice and mirror therapy associated with conventional physical therapy training on the hemiparetic upper limb in poststroke rehabilitation: A preliminary study. *Topics in Stroke Rehabilitation*, 21(6), 484–494. <https://doi.org/10.1310/tsr2106-484>
- De la Garza-Longoria, R. S., Maldonado-Mancillas, J. A., Mendoza-Múzquiz, P. L., & Sánchez, L. (2018). Incidencia de enfermedad cerebrovascular en un servicio de Medicina Interna. *Incidence of Cerebrovascular Disease in an Internal Medicine Service.*, 34(6), 874–880. <http://search.ebscohost.com/login.aspx?direct=true&db=asn&AN=134121264&lang=es&site>

=ehost-live

- Demers, M., Chan Chun Kong, D., & Levin, M. F. (2019). Feasibility of incorporating functionally relevant virtual rehabilitation in sub-acute stroke care: perception of patients and clinicians. *Disability and Rehabilitation: Assistive Technology*, 14(4), 361–367. <https://doi.org/10.1080/17483107.2018.1449019>
- Devesa Gutiérrez, I., María Elena Mazadiego González, D., Miguel Ángel Baldomero Hernández Hernández, T., & Héctor Antonio Mancera Cruz, T. (2014). Rehabilitación del paciente con enfermedad vascular cerebral (EVC). In *Rev Mex Med Fis Rehab* (Vol. 26, Issue 4). www.medigraphic.org.mx<http://www.medigraphic.com/medicinafisica>
- Dobkin, B. H. (2004). Strategies for stroke rehabilitation. *Lancet Neurology*, 3(9), 528–536. [https://doi.org/10.1016/S1474-4422\(04\)00851-8](https://doi.org/10.1016/S1474-4422(04)00851-8)
- Dobkin, B. H. (2011). Rehabilitation and Recovery of the Patient with Stroke. In *Stroke* (Fifth Edit, Vol. 36). Elsevier. <https://doi.org/10.1016/B978-1-4160-5478-8.10056-9>
- Epple, C., Maurer-Burkhard, B., Lichti, M.-C., & Steiner, T. (2020). Vojta therapy improves postural control in very early stroke rehabilitation: a randomised controlled pilot trial. *Neurological Research and Practice*, 2, 23. <https://doi.org/10.1186/s42466-020-00070-4>
- Eraifej, J., Clark, W., France, B., Desando, S., & Moore, D. (2017). Effectiveness of upper limb functional electrical stimulation after stroke for the improvement of activities of daily living and motor function: A systematic review and meta-analysis. *Systematic Reviews*, 6(1), 1–21. <https://doi.org/10.1186/s13643-017-0435-5>
- Figuerola, J., Villamayor, B., & Antelo, A. (2014). Rehabilitación del ictus cerebral: evaluación, pronóstico y tratamiento. In *Reeducación funcional tras un ictus* (Issue 2014, pp. 89–104). ProQuest Book Central. <https://ebookcentral.proquest.com>
- Fugl-Meyer, A. R., Jääskö, L., Leyman, I., Olsson, S., & Steglind, S. (1975). The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. *Scandinavian Journal of Rehabilitation Medicine*, 7(1), 13.
- Galna, B., Barry, G., Jackson, D., Mhiripiri, D., Olivier, P., & Rochester, L. (2014). Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease. *Gait & Posture*, 39(4), 1062–1068. <https://doi.org/10.1016/j.gaitpost.2014.01.008>
- Gauthier, L., Dehaut, F., & Joannette, Y. (1989). The Bells Test: A quantitative and qualitative test for visual neglect. *The International Journal of Clinical Neuropsychology*, 11, 49–54.
- Georgiou, O., Jeffrey, C., Chen, Z., Tong, B., Chan, S., Yang, B., Harwood, A., & Carter, T. (2018). *Touchless Haptic Feedback for VR Rhythm Games*. <https://doi.org/10.1109/VR.2018.8446619>
- Guidali, M., Duschau-Wicke, A., Broggi, S., Klamroth-Marganska, V., Nef, T., & Riener, R. (2011). A robotic system to train activities of daily living in a virtual environment. *Medical and Biological Engineering and Computing*, 49(10), 1213–1223. <https://doi.org/10.1007/s11517-011-0809-0>
- Guo, X. X., Fan, B. Y., & Mao, Y. Y. (2018). Effectiveness of neuromuscular electrical stimulation for

- wrist rehabilitation after acute ischemic stroke. *Medicine (United States)*, 97(38), 0–2.
<https://doi.org/10.1097/MD.00000000000012299>
- Hagelbäck, J., & Johansson, S. J. (2009). Measuring player experience on runtime dynamic difficulty scaling in an RTS game. *CIG2009 - 2009 IEEE Symposium on Computational Intelligence and Games, September*, 46–52. <https://doi.org/10.1109/CIG.2009.5286494>
- Henderson, A., Korner-Bitensky, N., & Levin, M. (2007). Virtual reality in stroke rehabilitation: a systematic review of its effectiveness for upper limb motor recovery. *Topics in Stroke Rehabilitation*, 14(2), 52–61. <https://doi.org/10.1310/tsr1402-52>
- Hernández-Fernández, F., Sandoval Valencia, H., Barbella-Aponte, R. A., Collado-Jiménez, R., Ayo-Martín, Ó., Barrena, C., Molina-Nuevo, J. D., García-García, J., Lozano-Setién, E., Alcahut-Rodríguez, C., Martínez-Martín, Á., Sánchez-López, A., & Segura, T. (2020). Cerebrovascular disease in patients with COVID-19: neuroimaging, histological and clinical description. *Brain*, 143(10), 3089–3103. <https://doi.org/10.1093/brain/awaa239>
- Hocine, N., Gouaïch, A., Cerri, S. A., Mottet, D., Froger, J., & Laffont, I. (2015). Adaptation in serious games for upper-limb rehabilitation: an approach to improve training outcomes. *User Modeling and User-Adapted Interaction*, 25(1), 65–98. <https://doi.org/10.1007/s11257-015-9154-6>
- Holden, M. K., Gill, K. M., & Magliozzi, M. R. (1986). Gait assessment for neurologically impaired patients. Standards for outcome assessment. *Physical Therapy*, 66(10), 1530–1539. <https://doi.org/10.1093/ptj/66.10.1530>
- Hoyer, E. H., & Celnik, P. A. (2011). Understanding and enhancing motor recovery after stroke using transcranial magnetic stimulation. *Restor Neurol Neurosci.*, 29(6), 395–409. <https://doi.org/10.3233/RNN-2011-0611>. Understanding
- Hsieh, Y.-W., Lin, K., Wu, C.-Y., Shih, T.-Y., Li, M., & Chen, C. (2018). Comparison of proximal versus distal upper-limb robotic rehabilitation on motor performance after stroke: A cluster controlled trial. *Scientific Reports*, 8. <https://doi.org/10.1038/s41598-018-20330-3>
- Hung, J. W., Chou, C. X., Chang, Y. J., Wu, C. Y., Chang, K. C., Wu, W. C., & Howell, S. (2019). Comparison of Kinect2Scratch game-based training and therapist-based training for the improvement of upper extremity functions of patients with chronic stroke: A randomized controlled single-blinded trial. *European Journal of Physical and Rehabilitation Medicine*, 55(5), 542–550. <https://doi.org/10.23736/S1973-9087.19.05598-9>
- Hunicke, R., Leblanc, M., & Zubek, R. (2004). MDA: A Formal Approach to Game Design and Game Research. *AAAI Workshop - Technical Report*, 1.
- Hylin, M. J., Kerr, A. L., & Holden, R. (2017). Understanding the mechanisms of recovery and/or compensation following injury. In *Neural Plasticity* (Vol. 2017). <https://doi.org/10.1155/2017/7125057>
- Isaacson, B. M., Swanson, T. M., & Pasquina, P. F. (2013). The use of a computer-assisted rehabilitation environment (CAREN) for enhancing wounded warrior rehabilitation regimens. *The Journal of Spinal Cord Medicine*, 36(4), 296–299. <https://doi.org/10.1179/2045772313Y.0000000119>

- ISSSTE. (2020). *Anuario estadístico de México*.
<http://www.issste.gob.mx/datosabiertos/anuarios/anuarios2019.html>
- Jones, T. A. (2017). Motor compensation and its effects on neural reorganization after stroke. *Nature Reviews Neuroscience*, 18(5), 267–280. <https://doi.org/10.1038/nrn.2017.26>
- Juliano, J. M., & Liew, S.-L. (2020). Transfer of motor skill between virtual reality viewed using a head-mounted display and conventional screen environments. *Journal of NeuroEngineering and Rehabilitation*, 17(1), 48. <https://doi.org/10.1186/s12984-020-00678-2>
- Kellor, M., Frost, J., Silberberg, N., Iversen, I., & Cummings, R. (1971). Hand strength and dexterity. *The American Journal of Occupational Therapy : Official Publication of the American Occupational Therapy Association*, 25(2), 77–83.
- Kerr, A. L., Cheng, S.-Y., & Jones, T. A. (2011). Experience-dependent neural plasticity in the adult damaged brain. *Journal of Communication Disorders*, 44(5), 538–548.
<https://doi.org/10.1016/j.jcomdis.2011.04.011>
- Khademi, M., Mousavi Hondori, H., McKenzie, A., Dodakian, L., Lopes, C., & Cramer, S. (2014). Free-hand interaction with leap motion controller for stroke rehabilitation. *Conference on Human Factors in Computing Systems - Proceedings*.
<https://doi.org/10.1145/2559206.2581203>
- Kim, H., Lee, A., Shin, Y. I., Chang, W. H., Koo, K. H., Seong, H., & Kim, Y. H. (2018). Effects of digital smart glove system on motor recovery of upper extremity in subacute stroke patients. *Annals of Physical and Rehabilitation Medicine*, 61, e28.
<https://doi.org/https://doi.org/10.1016/j.rehab.2018.05.061>
- Klinger, E., Weiss, P. L. T., & Joseph, P. (2010). Virtual reality for learning and rehabilitation Fundamental VR basic issues. In *Rethinking physical and rehabilitation medicine* (Collection). Springer, Paris. https://doi.org/10.1007/978-2-8178-0034-9_11
- Kolb, B., & Gibb, R. (2011). Brain plasticity and behaviour in the developing brain. *Journal of the Canadian Academy of Child and Adolescent Psychiatry = Journal de l'Academie Canadienne de Psychiatrie de l'enfant et de l'adolescent*, 20(4), 265–276.
<https://pubmed.ncbi.nlm.nih.gov/22114608>
- Koster, R. (2005). *A theory of fun for game design*. Scottsdale, AZ : Paraglyph Press, [2005] ©2005.
<https://search.library.wisc.edu/catalog/9910102220502121>
- Krakauer, J. W., & Cortés, J. C. (2018). A non-task-oriented approach based on high-dose playful movement exploration for rehabilitation of the upper limb early after stroke: A proposal. *NeuroRehabilitation*, 43(1), 31–40. <https://doi.org/10.3233/NRE-172411>
- Krebs, H., Hogan, N., Volpe, B., Aisen, M., Edelman, L., & Diels, C. (1999). Overview of clinical trials with MIT-MANUS: A robot-aided neuro- rehabilitation facility. *Technology and Health Care : Official Journal of the European Society for Engineering and Medicine*, 7, 419–423.
<https://doi.org/10.3233/THC-1999-7606>
- Laffont, I., Froger, J., Jourdan, C., Bakhti, K., Dokkum, L., Gouaich, V. A., Bonnin, H. Y., Armingaud, P., Jaussent, A., Picot, M. C., Bars, E. Le, Dupeyron, A., Arquizan, C., Gelis, A., & Mottet, D.

- (2019). Rehabilitation of the upper arm early after stroke: Video games versus conventional rehabilitation. A randomized controlled trial. *Disability and Rehabilitation: Assistive Technology*, 14(4), 317–324. <https://doi.org/10.1080/17483107.2018.1447608>
- Lang, C. E., Lohse, K. R., & Birkenmeier, R. L. (2015). Dose and timing in neurorehabilitation: Prescribing motor therapy after stroke. *Current Opinion in Neurology*, 28(6), 549–555. <https://doi.org/10.1097/WCO.0000000000000256>
- LANR. (2020). *¿Qué es un EVC?* https://lanr.ifc.unam.mx/pdf/01_EVC_Web.pdf
- Levin, M. F., Weiss, P. L., & Keshner, E. A. (2015). Emergence of virtual reality as a tool for upper limb rehabilitation: Incorporation of motor control and motor learning Principles. *Physical Therapy*, 95(3), 415–425. <https://doi.org/10.2522/ptj.20130579>
- López Arguelles, J., Alfonso León, D., Barboza Sanchis, S., & Pérez Manso, D. (2012). Heminegligencia y hemianopsia. Presentación de un caso. In *MediSur* (Vol. 10, pp. 318–321). scielocu.
- Maier, M., Rubio Ballester, B., Duff, A., Duarte Oller, E., & Verschure, P. F. M. J. (2019). Effect of Specific Over Nonspecific VR-Based Rehabilitation on Poststroke Motor Recovery: A Systematic Meta-analysis. *Neurorehabilitation and Neural Repair*, 33(2), 112–129. <https://doi.org/10.1177/1545968318820169>
- Michielsen, M., Vaughan-Graham, J., Holland, A., Magri, A., & Suzuki, M. (2019). The Bobath concept - a model to illustrate clinical practice. *Disability and Rehabilitation*, 41(17), 2080–2092. <https://doi.org/10.1080/09638288.2017.1417496>
- Mori, M., MacDorman, K., & Kageki, N. (2012). The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine*, 19, 98–100. <https://doi.org/10.1109/MRA.2012.2192811>
- Mousavi Hondori, H., & Khademi, M. (2014). A review on technical and clinical impact of microsoft Kinect on physical therapy and rehabilitation. In *Journal of Medical Engineering*. <https://doi.org/10.1155/2014/846514>
- Naghdi, S., Ansari, N. N., Mansouri, K., & Hasson, S. (2010). A neurophysiological and clinical study of Brunnstrom recovery stages in the upper limb following stroke. *Brain Injury*, 24(11), 1372–1378. <https://doi.org/10.3109/02699052.2010.506860>
- Ögün, M. N., Kurul, R., Yaşar, M. F., Turkoglu, S. A., Avci, Ş., & Yildiz, N. (2019). Effect of leap motion-based 3D immersive virtual reality usage on upper extremity function in ischemic stroke patients. *Archivos de Neuro-Psiquiatria*, 77(10), 681–688. <https://doi.org/10.1590/0004-282X20190129>
- Ordoñez Mora, L. T., Delgado Serna, L. J., Gutiérrez Muñoz, Y. E., Pinzón Bernal, M. Y., & Castellanos Ruiz, J. (2017). Terapia de restricción del lado sano como opción de manejo de personas con secuelas de enfermedad cerebrovascular. *Archivos de Medicina (Manizales)*, 17(1), 173–184. <https://doi.org/10.30554/archmed.17.1.1585.2017>
- Padilla-Castañeda, M. A., Sotgiu, E., Barsotti, M., Frisoli, A., Orsini, P., Martiradonna, A., Laddaga, C., & Bergamasco, M. (2018). An Orthopaedic Robotic-Assisted Rehabilitation Method of the Forearm in Virtual Reality Physiotherapy. *Journal of Healthcare Engineering*, 2018.

<https://doi.org/10.1155/2018/7438609>

- Paizan, N. L. M., Da Silva, R., & Borges, M. A. (2019). Hydrotherapy. *Revista Neurociências*, 17(4), 314–318. <https://doi.org/10.34024/rnc.2009.v17.8523>
- Pallesen, H., Andersen, M. B., Hansen, G. M., Lundquist, C. B., & Brunner, I. (2018). Patients' and Health Professionals' Experiences of Using Virtual Reality Technology for Upper Limb Training after Stroke: A Qualitative Substudy. *Rehabilitation Research and Practice*, 2018, 4318678. <https://doi.org/10.1155/2018/4318678>
- Peretti, A., Amenta, F., Tayebati, S. K., Nittari, G., & Mahdi, S. S. (2017). Telerehabilitation: Review of the State-of-the-Art and Areas of Application. *JMIR Rehabilitation and Assistive Technologies*, 4(2), e7–e7. <https://doi.org/10.2196/rehab.7511>
- Piggott, L., Wagner, S., & Ziat, M. (2016). Haptic neurorehabilitation and virtual reality for upper limb paralysis: A review. *Critical Reviews in Biomedical Engineering*, 44(1–2), 1–32. <https://doi.org/10.1615/CritRevBiomedEng.2016016046>
- Rastegarpanah, A., Scone, T., Saadat, M., Rastegarpanah, M., Taylor, S. J., & Sadeghein, N. (2018). Targeting effect on gait parameters in healthy individuals and post-stroke hemiparetic individuals. *Journal of Rehabilitation and Assistive Technologies Engineering*, 5, 2055668318766710–2055668318766710. <https://doi.org/10.1177/2055668318766710>
- Rosenich, E., Hordacre, B., Paquet, C., Koblar, S. A., & Hillier, S. L. (2020). Cognitive Reserve as an Emerging Concept in Stroke Recovery. *Neurorehabilitation and Neural Repair*, 34(3), 187–199. <https://doi.org/10.1177/1545968320907071>
- Saposnik, G., & Levin, M. (2011). Virtual reality in stroke rehabilitation: A meta-analysis and implications for clinicians. *Stroke*, 42(5), 1380–1386. <https://doi.org/10.1161/STROKEAHA.110.605451>
- Saposnik, G., Teasell, R., Mamdani, M., Hall, J., McIlroy, W., Cheung, D., Thorpe, K. E., Cohen, L. G., & Bayley, M. (2010). Effectiveness of virtual reality using wii gaming technology in stroke rehabilitation: A pilot randomized clinical trial and proof of principle. *Stroke*, 41(7), 1477–1484. <https://doi.org/10.1161/STROKEAHA.110.584979>
- Schell, J. (2008). *The art of game design : a book of lenses*. Amsterdam ; Boston : Elsevier/Morgan Kaufmann, [2008] ©2008. <https://search.library.wisc.edu/catalog/9910060875802121>
- Sekhvat, Y. A. (2017). MPRL: Multiple-Periodic Reinforcement Learning for difficulty adjustment in rehabilitation games. *2017 IEEE 5th International Conference on Serious Games and Applications for Health, SeGAH 2017*. <https://doi.org/10.1109/SeGAH.2017.7939260>
- Shin, J. H., Ryu, H., & Jang, S. H. (2014). A task-specific interactive game-based virtual reality rehabilitation system for patients with stroke: A usability test and two clinical experiments. *Journal of NeuroEngineering and Rehabilitation*, 11(1), 1–10. <https://doi.org/10.1186/1743-0003-11-32>
- Siberski, J., Shatil, E., Siberski, C., Eckroth-Bucher, M., French, A., Horton, S., Loefflad, R. F., & Rouse, P. (2014). Computer-Based Cognitive Training for Individuals With Intellectual and Developmental Disabilities: Pilot Study. *American Journal of Alzheimer's Disease & Other*

- Dementias*[®], 30(1), 41–48. <https://doi.org/10.1177/1533317514539376>
- Sleimen-Malkoun, R., Temprado, J. J., Thefenne, L., & Berton, E. (2011). Bimanual training in stroke: How do coupling and symmetry-breaking matter? *BMC Neurology*, 11. <https://doi.org/10.1186/1471-2377-11-11>
- Steuer, J. (1992). Defining Virtual Reality: Dimensions Determining Telepresence. *Journal of Communication*, 42(4), 73–93. <https://doi.org/10.1111/j.1460-2466.1992.tb00812.x>
- Stone, S. P., Ali, B., Auberleek, I., Thompsell, A., & Young, A. (1994). The Barthel index in clinical practice: use on a rehabilitation ward for elderly people. *Journal of the Royal College of Physicians of London*, 28(5), 419–423.
- Sudharsan, R. (2018). *Hands-On Reinforcement Learning with Python*.
- Sutton, R. S. (1991). Planning by Incremental Dynamic Programming. *ML*.
- Sutton, R. S., & Barto, A. (2005). Reinforcement Learning: An Introduction. In *IEEE Transactions on Neural Networks* (Vol. 16, Issue 1). <https://doi.org/10.1109/tnn.2004.842673>
- Teigen, K. H. (1994). Yerkes-Dodson: A Law for all Seasons. *Theory & Psychology*, 4(4), 525–547. <https://doi.org/10.1177/0959354394044004>
- Trabucco, J. T., Rottigni, A., Cavallo, M., Bailey, D., Patton, J., & Marai, G. E. (2019). User perspective and higher cognitive task-loads influence movement and performance in immersive training environments. *BMC Biomedical Engineering*, 1(1), 1–12. <https://doi.org/10.1186/s42490-019-0021-0>
- Tramontano, M., Morone, G., Palomba, A., De Angelis, S., Mercurio, A., Caltagirone, C., & Grasso, M. G. (2020). Effectiveness of a sensor-based technology in upper limb motor recovery in post-acute stroke neurorehabilitation: a randomized controlled trial. *Journal of Biological Regulators and Homeostatic Agents*, 34(5 Suppl. 3), 165-174. Technology in Medicine.
- Van der Linde, R., Lammertse, P., Frederiksen, E., & Ruiters, B. (2002). *The HapticMaster, a new high-performance haptic interface*.
- Verheyden, G., & Ashburn, A. (2013). Accidente cerebrovascular. In *Fisioterapia en la rehabilitación neurológica* (pp. 9–28).
- Warlow, C. P. (1998). Epidemiology of stroke. *Lancet (London, England)*, 352 Suppl, S111-4. [https://doi.org/10.1016/s0140-6736\(98\)90086-1](https://doi.org/10.1016/s0140-6736(98)90086-1)
- Watkins, C. (1989). *Learning From Delayed Rewards*.
- Westwater-Wood, S., Adams, N., & Kerry, R. (2010). The use of proprioceptive neuromuscular facilitation in physiotherapy practice. *Physical Therapy Reviews*, 15, 23–28. <https://doi.org/10.1179/174328810X12647087218677>
- Wilson, P. N., Foreman, N., & Stanton, D. (1997). Virtual reality, disability and rehabilitation. *Disability and Rehabilitation*, 19(6), 213–220. <https://doi.org/10.3109/09638289709166530>
- Wolf, S. L., Thompson, P. A., Morris, D. M., Rose, D. K., Winstein, C. J., Taub, E., Giuliani, C., &

Pearson, S. L. (2005). The EXCITE Trial: Attributes of the Wolf Motor Function Test in Patients with Subacute Stroke. *Neurorehabilitation and Neural Repair*, 19(3), 194–205. <https://doi.org/10.1177/1545968305276663>

World Stroke Organization. (2020). Annual report 2019. In *AIMS Mathematics* (Vol. 5, Issue 1). <https://doi.org/10.3934/math.2020i>

Yang, Y., Wang, A., Zhao, X., Wang, C., Liu, L., Zheng, H., Wang, Y., Cao, Y., & Wang, Y. (2016). - The Oxfordshire Community Stroke Project classification system predicts clinical. *Ther Clin Risk Manag*, 12, 1049–1056.

Zhang, J.-T., Novak, A. C., Brouwer, B., & Li, Q. (2013). Concurrent validation of Xsens MVN measurement of lower limb joint angular kinematics. *Physiological Measurement*, 34(8), N63-9. <https://doi.org/10.1088/0967-3334/34/8/N63>

Zhu, Z., Cui, L., Yin, M., Yu, Y., Zhou, X., Wang, H., & Yan, H. (2016). Hydrotherapy vs. conventional land-based exercise for improving walking and balance after stroke: A randomized controlled trial. *Clinical Rehabilitation*, 30(6), 587–593. <https://doi.org/10.1177/0269215515593392>