

تعیین اعتبار درون و بین ارزیاب "آزمون ارزیابی گفتار افراد مبتلا به شکاف کام بر اساس سیستم پارامترهای جهانی - نسخه فارسی"

آزاده امیریان^۱، فاطمه درخشنده^{*}، احمد صالحی^۲، بهرام سلیمانی^۳

چکیده

مقدمه: شکاف لب و کام یکی از شایع‌ترین نقایص مادرزادی است. افراد مبتلا به شکاف کام ویژگی‌های گفتاری خاصی دارند که با ارزیابی ادراکی گفتار، توسط آسیب شناسان گفتار و زبان، شناسایی می‌شوند. ارزیابی ادراکی، استاندارد طلایی برای ارزیابی گفتار است. بر اساس این ویژگی‌های گفتاری و شیوع شکاف کام، لازم است از آزمونی پایا، برای ارزیابی ادراکی خطاهای گفتاری افراد مبتلا به شکاف کام استفاده نمود.

مواد و روش‌ها: در این مطالعه توصیفی-تحلیلی و مقطعی، به منظور تعیین اعتبار درون و بین ارزیاب "آزمون ارزیابی گفتار افراد مبتلا به شکاف کام بر اساس سیستم پارامترهای جهانی - نسخه فارسی"، نمونه گفتاری ۳۴ کودک ۵ تا ۹ ساله مبتلا به شکاف کام، مراجعه کننده به کلینیک شکاف کام اصفهان، توسط دو آسیب شناس گفتار و زبان (کم تجربه و باتجربه)، مورد ارزیابی قرار گرفتند. ۳۰ درصد از نمونه‌ها به طور تصادفی انتخاب و بعد از دو ماه، توسط همان دو ارزیاب ارزیابی شدند. مواد گفتاری شامل ۴۰ تک واژه و ۱۷ جمله کنترل شده و گفتار پیوسته بود. میزان اعتبار درون ارزیاب با استفاده از ضریب همبستگی درون طبقه‌ای (Intraclass correlation coefficient یا ICC) و میزان اعتبار بین ارزیاب با استفاده از ضریب همبستگی کندال و ICC برآورد گردید.

یافته‌ها: میانگین میزان اعتبار درون ارزیاب با استفاده از ICC، برای ارزیاب کم تجربه ۰/۹۶۷ و برای ارزیاب باتجربه ۰/۹۷۱، برآورد گردید. میانگین میزان اعتبار بین ارزیاب با استفاده از ضریب همبستگی کندال، ۰/۹۵۳ و با استفاده از ICC، ۰/۹۵۷ برآورد گردید.

نتیجه‌گیری: این ابزار از اعتبار درون و بین ارزیاب قابل قبولی برای ارزیابی گفتار افراد مبتلا به شکاف کام برخوردار بود. از این ابزار می‌توان به عنوان ابزاری پایا در مطالعات مربوط به گفتار کودکان ۵ تا ۹ ساله مبتلا به شکاف کام، استفاده نمود.

کلید واژه‌ها: شکاف کام، اعتبار، ارزیابی ادراکی گفتار، سیستم پارامترهای جهانی

تاریخ دریافت: ۹۰/۱۰/۲۸

تاریخ پذیرش: ۹۰/۱۲/۲۵

مقدمه

(۳، ۲). با ارزیابی ادراکی که استاندارد طلایی برای ارزیابی گفتار می‌باشد (۴)، وضعیت غیر طبیعی شناسایی و نوع و شدت اختلال تعیین می‌گردد، تا بر اساس آن اقدامات درمانی مناسب صورت گیرد (۵). با توجه به مشکلات تولیدی و تشدید گفتار و شیوع شکاف کام، وجود روش‌های مناسب

شکاف لب و کام یکی از شایع‌ترین نقایص مادرزادی است (۱، ۲). این نقص با تغییر در عملکرد درجه نرم‌کامی حلقی و همچنین ایجاد مشکلات دندانی، موجب پرخیشومی شدن گفتار، خطاهای تولیدی و کاهش کیفیت صوت می‌شود

* دانشجوی دکتری گفتاردرمانی، عضو هیأت علمی، دانشکده علوم توان‌بخشی، دانشگاه علوم پزشکی تهران، تهران، ایران

Email: derakhshandeh@rehab.mui.ac.ir

۱- دانشجوی کارشناسی ارشد گفتاردرمانی، کمیته تحقیقات دانشجویی، دانشکده علوم توان‌بخشی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران

۲- کارشناسی ارشد روان‌سنجی، عضو هیأت علمی، دانشکده علوم توان‌بخشی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران

۳- دکتری، عضو هیأت علمی، گروه آمار حیاتی، دانشگاه آزاد اسلامی، واحد نجف‌آباد، دانشکده پزشکی، اصفهان، ایران

ابزار برای فراهم کردن اطلاعاتی درباره نتایج گفتار و نیازهای مداخله بالینی در بیماران شکاف کام طراحی شد. از این پروتکل در مطالعات بررسی بین مرکزی انگلیس، ایرلند و دیگر کشورهای انگلیسی زبان و برای گزارش نتایج گفتار روش‌های جراحی استفاده می‌شود (۶).

اگر چه ارزیابی‌های ابزاری، به واسطه فراهم کردن اطلاعات عینی، در میان درمان‌گران مورد توجه هستند اما هنوز هیچ ابزاری به زبان فارسی، وجود ندارد که بتوان آن را به عنوان جایگزینی برای ارزیابی ادراکی گفتار استفاده نمود (۳).

در حال حاضر که آزمون ارزیابی گفتار افراد مبتلا به شکاف کام بر اساس سیستم پارامترهای جهانی - نسخه فارسی، در دسترس ما وجود دارد، بهتر است که نتایج ارزیابی‌های ادراکی یک ارزیاب در زمان‌های مختلف و نیز ارزیاب‌های در یک زمان، بر اساس این آزمون مقایسه شود، تا میزان اعتبار درون و بین ارزیاب این ابزار ادراکی به دست آید؛ چرا که برای اتخاذ تصمیمات درمانی تنها نمی‌توان بر قضاوت‌های درمان‌گری خاص متکی بود. پس لازم است که اعتبار بین ارزیاب برای این آزمون به دست آید. با توجه به این که در ارزیابی‌های مکرر حتی توسط یک ارزیاب نیز ممکن است نتایج متفاوتی به دست آید، نیاز است اعتبار درون ارزیاب این آزمون، به دست آید.

هدف کلی این پژوهش، تعیین اعتبار درون و بین ارزیاب "آزمون ارزیابی گفتار افراد مبتلا به شکاف کام بر اساس سیستم پارامترهای جهانی - نسخه فارسی" بود. با انجام این پژوهش به این سؤال پاسخ داده می‌شود که آیا این آزمون از اعتبار درون و بین ارزیاب قابل قبولی برخوردار می‌باشد؟

مواد و روش‌ها

پژوهش حاضر، یک مطالعه توصیفی - تحلیلی و مقطعی بود که در آن نمونه‌های گفتاری، ۳۴ کودک دارای شکاف کام با یا بدون شکاف لب (۱۴ دختر و ۲۰ پسر) بودند که به روش نمونه‌گیری در دسترس، از میان ۱۰۳ کودک ۵ تا ۹ ساله مراجعه کننده به کلینیک شکاف لب و کام دانشکده علوم

برای ارزیابی و درمان گفتار این جمعیت از اهمیت ویژه‌ای برخوردار است (۲). بنابراین، لازم است ابزاری استاندارد و دارای اعتبار برای تشخیص و ثبت دقیق خطاهای گفتاری افراد مبتلا به شکاف کام تهیه شود.

کیفیت درک شده از گفتار، اصلی‌ترین معیار برای تعیین تصمیمات درمانی است و گوش درمان‌گر، داور نهایی برای قضاوت در مورد وضعیت گفتاری یک بیمار مبتلا به شکاف کام می‌باشد (۳).

به منظور ارزیابی، به ویژه برای کودکان، نیاز است که از شیوه‌ای کاربردی و غیر تهاجمی استفاده شود. بنابراین ارزیابی‌های ادراکی هنوز هم به عنوان مناسب‌ترین استاندارد برای این امر قابل قبول هستند (۳).

هر چند نویسندگان و محققان در مورد لزوم انجام ارزیابی‌های ادراکی در تشخیص و درمان مشکلات گفتار در شکاف لب و کام اتفاق نظر دارند، ولی در مورد شیوه انجام این نوع ارزیابی، بحث‌های متنوعی مطرح است (۳).

در مطالعات به تفاوت در شیوه‌های ارزیابی ادراکی میان محققان، نظیر ثبات در نمونه‌های گفتاری مورد بررسی، میزان تجربه و تبحر ارزیاب و نیز شدت‌های مختلف مشکلات گفتاری این بیماران اشاره شده است. به نظر می‌رسد عواملی نظیر نوع مقیاس مورد استفاده در ارزیابی، نوع نمونه گفتاری مورد استفاده ارزیاب و تجربه وی در قضاوت‌های ادراکی، همگی بر اعتبار نتایج ارزیابی تأثیر گذار هستند (۳). همچنین می‌توان به تأثیر فراوانی و شدت خطاها، کیفیت ضبط نمونه گفتاری، شرایط محیطی ارزیابی نمونه‌ها، تعاریف گویا از هر پارامتر و آموزش قبل از استفاده از ابزار اشاره کرد (۶).

در مرور متون، مطالعات کمی در مورد اعتبار درون و بین ارزیاب آزمون‌های ارزیابی ادراکی گفتار شکاف کام وجود داشت. اولین تحقیق جدی در این مورد، مربوط به آزمون Cap's augmented (CAPS-A) بود که توسط Sell و همکاران تهیه شد و اعتبار درون و بین ارزیاب خوب و خیلی خوب (کاپا ۰/۶۱ تا ۱)، برای ۷ بخش از این ابزار و متوسط (کاپا ۰/۴۱ تا ۰/۶۰)، برای سه بخش از آن برآورد گردید. این

آزمودنی، مقابل رایانه به گونه‌ای روی صندلی نشست که فاصله دهان او از میکروفون رایانه حدود ۴۰ سانتی‌متر بود. قبل از شروع نمونه‌گیری، آزمون‌گر به کودک توضیح می‌داد که در هر مرحله چه کاری باید انجام دهد (نامیدن تصویر بعد از دیدن هر تصویر، تکرار جملات بعد از آزمون‌گر، خواندن شعر و غیره). برای استخراج تک واژه‌ها از نامیدن تصویر و برای استخراج جملات، از روش تکرار بعد از آزمون‌گر استفاده می‌شد.

برای شروع آزمون و در بخش نامیدن تصویر، آزمون‌گر به تصویر مورد نظر در صفحه نمایش‌گر اشاره می‌کرد و از آزمودنی می‌خواست که نام آن را بگوید. اگر آزمودنی نمی‌توانست به صورت خود انگیزه تصویر را بنامد، از سر نخ معنایی استفاده می‌شد و اگر سر نخ مؤثر نبود، از جملات حاوی تک واژه هدف و در نهایت تکرار تک واژه استفاده می‌شد (۷). سپس جمله‌ها ارزیابی می‌شدند و کودک باید هر جمله را بعد از آزمون‌گر تکرار می‌نمود.

نمونه‌ای از گفتار پیوسته آزمودنی شامل یک شعر ساده مثل "یه توپ دارم"، شمارش از ۱ تا ۲۰، روزهای هفته، خواندن یک متن در صورت امکان، ۱ تا ۲ دقیقه گفتار محاوره درباره خانواده، دوستان و زندگی روزمره نیز نمونه‌گیری شد. پس از اتمام نمونه‌گیری و ذخیره نمونه‌های صوتی و تصویری آزمودنی‌ها در یک لوح فشرده، ۳۰ درصد از آن‌ها (۱۲ نمونه صوتی و تصویری) به صورت تصادفی انتخاب و رونوشت آن در پوشه‌ای دیگر ذخیره شد (۸).

دو آسیب‌شناس گفتار و زبان به عنوان ارزیاب در پژوهش شرکت کردند. برای هماهنگ‌تر شدن ارزیاب‌ها، در یک جلسه آموزشی، نحوه ارزیابی پارامترها و آوانگاری آن‌ها در فرم نمونه گفتار و سپس نحوه درجه‌بندی آن‌ها بر اساس اصول و شرایط ذکر شده در سیستم پارامترهای جهانی، بررسی و چند نمونه صوتی و تصویری به صورت آزمایشی توسط دو ارزیاب به صورت مستقل، تجزیه و تحلیل شد. بعد از ایجاد هماهنگی، هر دو ارزیاب نمونه صوتی و تصویری آزمودنی‌ها را از طریق یک رایانه همراه دریافت و به طور مستقل ارزیابی

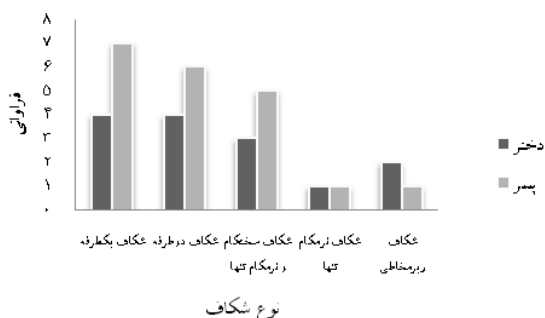
توان‌بخشی دانشگاه علوم پزشکی اصفهان انتخاب شدند. به منظور تعیین اعتبار بین ارزیاب در ارزیابی ادراکی گفتار مبتلایان به شکاف کام "آزمون ارزیابی گفتار افراد مبتلا به شکاف کام بر اساس سیستم پارامترهای جهانی - نسخه فارسی" مورد استفاده قرار گرفت. محدوده سنی ۵ تا ۹ سال (امکان همکاری بیشتر و تشکیل یک گروه متجانس)، داشتن تاریخچه‌ای از شکاف کام و لب یک‌طرفه، دوطرفه، شکاف کام تنها و شکاف زیر مخاطی (امکان بررسی شدت اختلالات گفتاری متفاوت)، زبان درکی و بیانی متناسب با سن و شنوایی در محدوده طبیعی (امکان همکاری و پاسخ‌گویی به سؤالات آزمون) معیارهای ورود به پژوهش بودند. عدم همکاری آزمودنی تا پایان اجرای آزمون و عدم تمایل والدین به شرکت در آزمون نیز از معیارهای خروج از پژوهش به شمار می‌رفتند. نمونه‌گیری در ماه‌های بهمن و اسفند سال ۱۳۸۸ و در کلینیک شکاف لب و کام دانشکده علوم توان‌بخشی دانشگاه علوم پزشکی اصفهان انجام شد.

نمونه‌گیری گفتار، در یک اتاق آکوستیک و به صورت انفرادی انجام شد. از یک رایانه همراه (مدل ۱۵۰۰ Dell vostro)، برای ضبط نمونه‌های صوتی آزمودنی‌ها به کمک نرم‌افزار AD sound recorder 4.5 و نیز نمایش تصاویر آزمون استفاده شد. همچنین از یک دوربین فیلم‌برداری دیجیتال Canon (مدل IXUS ۸۷۰ IS)، برای ضبط نمونه‌های تصویری آزمودنی‌ها استفاده شد. آزمون استفاده شده "آزمون ارزیابی گفتار افراد مبتلا به شکاف کام بر اساس سیستم پارامترهای جهانی - نسخه فارسی" بود که از روایی و اعتبار قابل قبولی برخوردار است. اعتبار و روایی این آزمون در گروه کودکان ۵ تا ۹ ساله دارای شکاف کام مراجعه کننده به کلینیک شکاف لب و کام دانشکده علوم توان‌بخشی دانشگاه علوم پزشکی اصفهان تعیین شد. میزان اعتبار آزمون (فرم درجه‌بندی)، با استفاده از ضریب Cronbach's alpha، برابر با ۰/۸۹۷ بود. این آزمون دارای یک فرم نمونه‌گیری گفتار و یک فرم درجه‌بندی است (پیوست).

ضبط شده نزد پژوهش‌گر محفوظ می‌ماند و استفاده از آن‌ها در همایش‌های عمومی با کسب اجازه از خانواده‌ها صورت می‌گیرد.

یافته‌ها

نمودار ۱ توزیع فراوانی نمونه‌ها به تفکیک نوع شکاف در هر جنس را نشان می‌دهد. بیشترین فراوانی مربوط به شکاف یک‌طرفه در جمعیت پسرها بود.



نمودار ۱. توزیع فراوانی نمونه‌ها به تفکیک نوع شکاف در هر جنس

برآورد اعتبار درون ارزیاب

به منظور برآورد اعتبار درون ارزیاب فرم درجه‌بندی و پارامترهای آن، ابتدا میزان توافق یک ارزیاب در درجه‌بندی پارامترها در دو زمان مختلف به فاصله دو ماه، با استفاده از ضریب همبستگی درون طبقه‌ای (ICC)، به شرح جدول ۱ برآورد گردید و سپس میانگین آن‌ها به عنوان اعتبار درون ارزیاب (فرم درجه‌بندی) در نظر گرفته شد.

برآورد اعتبار بین ارزیاب

به منظور برآورد اعتبار بین ارزیاب فرم درجه‌بندی پارامترهای جهانی، ابتدا میزان توافق دو ارزیاب در مورد درجه‌بندی پارامترها، با استفاده از ضریب همبستگی کندال و ضریب همبستگی درون طبقه‌ای (ICC)، به شرح جدول ۲ برآورد گردید. میانگین ضریب همبستگی پارامترها، به عنوان اعتبار بین ارزیاب (فرم درجه‌بندی) در نظر گرفته شد.

کردند. ارزیاب‌ها می‌توانستند بلندی صدا را تنظیم و در صورت لزوم، نمونه گفتاری را دوباره بشنوند (۹).

برای هر آزمودنی، ابتدا زیر آزمون تک واژه‌ها از نظر وجود یا عدم وجود خطاهای هم‌خوان (پارامتر خطاهای تولید هم‌خوان و خروج خیشومی) مورد ارزیابی قرار گرفت. زیر آزمون جملات نیز از این نظر مورد ارزیابی قرار گرفت. سپس این پارامترها بر اساس مقیاس درجه‌بندی سیستم پارامترهای جهانی و با استفاده از نمونه‌های گفتاری مربوطه، درجه‌بندی شدند. بعد از ارزیابی نمونه‌های گفتاری همه آزمودنی‌ها از نظر پارامترهای خطای تولید هم‌خوان و خروج خیشومی، بار دیگر این مراحل برای ارزیابی پارامترهای پر خیشومی و سپس کم خیشومی تکرار شد. پارامتر اختلال صوت و قابلیت پذیرش گفتار بر اساس کل نمونه‌های گفتاری آزمودنی و پارامتر قابلیت فهم گفتار بر اساس گفتار پیوسته درجه‌بندی شد. بنابراین از فرم درجه‌بندی هر ارزیاب، ۱۰ درجه مربوط به پارامترهای اختلال در قابلیت فهم گفتار، اختلال در قابلیت پذیرش گفتار، پر خیشومی در سطح تک واژه و پر خیشومی در سطح جمله (از ۰ تا ۳)، کم خیشومی و اختلال صوت (از ۰ تا ۱)، خروج خیشومی در سطح تک واژه و خروج خیشومی در سطح جمله (از ۰ تا ۲) و خطای تولید هم‌خوان در سطح تک واژه و خطای تولید هم‌خوان در سطح جمله (از ۰ تا ۸) به دست آمد (۴). همچنین ۳۰ درصد از نمونه‌ها (۱۲ نمونه) که به طور تصادفی از بین دیگر نمونه‌ها انتخاب شده بودند، دو ماه بعد مورد بررسی قرار گرفتند و از نتایج آن‌ها در برآورد اعتبار درون ارزیاب استفاده شد.

داده‌های به دست آمده با نرم‌افزار SPSS^{۱۶} تجزیه و تحلیل و اعتبار درون ارزیاب به روش ضریب همبستگی درون طبقه‌ای (Intraclass correlation coefficient یا ICC) و اعتبار بین ارزیاب، به روش ضریب همبستگی کندال و ضریب همبستگی درون طبقه‌ای (ICC)، برآورد گردید.

ملاحظات اخلاقی

هیچ گونه الزامی جهت شرکت افراد در مطالعه وجود نداشت و همکاری کاملاً اختیاری بود. اطلاعات شخصی و تصاویر

است. این آزمون در سنجش‌های مختلف یک نمونه، مقادیر یکسانی را ارائه می‌دهد و از دقت لازم برخوردار است. به عبارت دیگر این آزمون در سنجش اختلالات گفتاری افراد مبتلا به شکاف کام، از دقت بالا و خطای کمی برخوردار است. به همین دلیل کارشناسان می‌توانند نتایج گفتار گروه هدف را با اطمینان بیشتری گزارش کنند.

میزان اعتبار درون ارزیاب برای هر دو ارزیاب کم‌تجربه و باتجربه در تمام پارامترها به جز قابلیت فهم گفتار از ۰/۹۱ تا ۱ گزارش شده است که نشان دهنده اعتبار قابل توجه این ابزار ادراکی است. اعتبار درون ارزیاب در این پژوهش در مورد پارامترهای کم‌خیشومی، اختلال صوت، خروج خیشومی در سطح تک واژه و جمله در بالاترین حد ممکن یعنی ۱ بوده است. دلیل آن را می‌توان چنین ذکر کرد که در این ۴ پارامتر، از مقیاس‌های دو تایی و سه تایی برای درجه‌بندی استفاده شده است. میانگین اعتبار هر یک از ارزیاب‌ها بیشتر از ۰/۹۶ بود و بین میانگین اعتبار این دو ارزیاب، تفاوت ناچیزی وجود داشت (جدول ۱).

در مرور متون، مطالعات کمی در مورد اعتبار درون و بین ارزیاب آزمون‌های ارزیابی ادراکی گفتار شکاف کام وجود داشت. تنها تحقیق جدی در این مورد، مربوط به آزمون CAPS-A بود. این آزمون توسط Sell و همکاران در سال ۲۰۰۶ تهیه شد (۶). آن‌ها میزان اعتبار بین ارزیاب و درون ارزیاب ابزار را بر اساس نمونه‌های ۱۰ آزمودنی و با استفاده از ۷ ارزیاب باتجربه و بی‌تجربه برآورد کردند. بر اساس یافته‌های پژوهشگران، گزارش خطاها باید بر اساس قوانین و تعاریف یکسانی صورت گرفته باشد. در غیر این صورت مقایسه نتایج بی‌معنی خواهد بود (۱). با توجه به این موضوع، ابزار CAPS-A تنها ابزار ادراکی بود که با سیستم پارامترهای جهانی مشابه بود.

میانگین اعتبار درون ارزیاب در مطالعه Sell و همکاران در بخش‌های مختلف ابزار CAPS-A، از ۰/۶۲ تا ۰/۸۳ بود (۶). دلیل بالا بودن میزان اعتبار درون ارزیاب در پژوهش حاضر، شاید کمتر بودن تعداد ارزیاب‌ها و تعداد پارامترهای مورد بررسی نسبت به ابزار CAPS-A بوده است. بنابر مطالعات انجام شده، عواملی نظیر نوع مقیاس مورد استفاده در ارزیابی، نوع نمونه گفتاری مورد استفاده ارزیاب و تجربه وی در قضاوت‌های ادراکی، همگی بر

جدول ۱. اعتبار درون ارزیاب فرم درجه‌بندی آزمون فارسی

پارامترها	ارزیاب کم‌تجربه ^{ICC*}	ارزیاب باتجربه ^{ICC}
قابلیت فهم گفتار	۰/۹۱۶	۰/۸۵۶
قابلیت پذیرش گفتار	۰/۹۶۳	۰/۹۵۶
پرخیشومی در سطح تک واژه	۰/۹۵۴	۰/۹۴۹
پرخیشومی در سطح جمله	۰/۹۵۴	۰/۹۴۹
کم خیشومی	۱/۰۰۰	۱/۰۰۰
اختلال صوت	۱/۰۰۰	۱/۰۰۰
خروج خیشومی در سطح تک واژه	۱/۰۰۰	۱/۰۰۰
خروج خیشومی در سطح جمله	۱/۰۰۰	۱/۰۰۰
خطای تولید هم‌خوان در سطح تک واژه	۰/۹۵۲	۱/۰۰۰
خطای تولید هم‌خوان در سطح جمله	۰/۹۲۸	۱/۰۰۰
کل	۰/۹۶۷	۰/۹۷۱

* Intraclass correlation coefficient

جدول ۲. اعتبار بین ارزیاب فرم درجه‌بندی آزمون فارسی

پارامترها	کندال	ICC*
قابلیت فهم گفتار	۰/۹۵۸	۰/۹۷۷
قابلیت پذیرش گفتار	۰/۹۲۶	۰/۹۵۹
پرخیشومی در سطح تک واژه	۰/۹۸۵	۰/۹۹۰
پرخیشومی در سطح جمله	۱/۰۰۰	۱/۰۰۰
کم خیشومی	۱/۰۰۰	۱/۰۰۰
اختلال صوت	۰/۶۸۵	۰/۶۵۲
خروج خیشومی در سطح تک واژه	۱/۰۰۰	۱/۰۰۰
خروج خیشومی در سطح جمله	۱/۰۰۰	۱/۰۰۰
خطای تولید هم‌خوان در سطح تک واژه	۰/۹۷۹	۰/۹۹۶
خطای تولید هم‌خوان در سطح جمله	۰/۹۹۶	۰/۹۹۲
کل	۰/۹۵۳	۰/۹۵۷

* Intraclass correlation coefficient

بحث

با توجه به مقادیر اعتبار برآورد شده و دامنه اعتبار از ۰ تا ۱، می‌توان نتیجه گرفت که آزمون ارزیابی گفتار افراد مبتلا به شکاف کام- نسخه فارسی، از اعتبار قابل قبولی برخوردار

و نیازهای درمانی بعدی را نیز مشخص می‌نماید. همان طور که Sell و همکاران در مطالعه خود دریافتند (۶)، و در این مطالعه نیز نشان داده شد، استفاده از این ابزار نیاز به آموزش ویژه، از طریق سمینارهای آموزشی دارد. این آزمون با ارایه تعاریف واضح از هر پارامتر و نحوه دقیق درجه‌بندی پارامترها، امکان ایجاد هماهنگی بیشتر بین ارزیاب‌ها در نقاط مختلف کشور و در مراکز مختلف درمانی را فراهم می‌آورد.

پیشنهادات

پیشنهاد می‌شود روایی معیار این آزمون با استفاده از روش‌های ابزاری برآورد گردد. به علاوه آزمون بر روی حجم نمونه بیشتری از افراد مبتلا به شکاف کام اجرا شود. همچنین اعتبار درون و بین ارزیاب با تعداد ارزیاب‌های بیشتر برآورد گردد.

تشکر و قدردانی

از سرکار خانم هدیه هاشمی، کارشناس ارشد گفتاردرمانی و جناب آقای دکتر عبدالکریم کریمی معاون پژوهشی دانشکده علوم توان‌بخشی دانشگاه علوم پزشکی اصفهان و سرکار خانم نادری، صمیمانه سپاسگزاری می‌گردد.

اعتبار نتایج ارزیابی تأثیر گذارند (۳). اعتبار بین ارزیاب در تمام پارامترها به جز اختلال صوت، از ۰/۹۲ تا ۱ برآورد شد که نشان دهنده اعتبار بالای این ابزار ادراکی می‌باشد. میزان اعتبار بین ارزیاب پارامتر اختلال صوت کم‌تر از ۰/۷۰ (۰/۶۸۵) برآورد گردید. شاید دلیل این امر تجربه کمتر ارزیاب‌ها در ارزیابی اختلال صوت و کم بودن میزان حساسیت شنوایی آن‌ها و یا عدم هماهنگی آن‌ها در تعیین اختلال صوت به دلیل کم بودن نمونه‌های دارای اختلال صوت بود (جدول ۲).

در مطالعه Sell و همکاران اعتبار بین ارزیاب در بخش‌های مختلف ابزار CAPS-A، از ۰/۳۴ تا ۰/۹۲ برآورد شد. میزان اعتبار بین ارزیاب پارامتر اختلال صوت برابر با ۰/۵۸ و در حد متوسط بود که کمتر از بخش‌های دیگر به جز یک مورد (یک نوع خطای تولید) بود (۶). این نتیجه نیز با نتیجه حاصل از این پژوهش هم‌سو بود.

نتیجه‌گیری

بدین ترتیب آزمون فارسی، ابزاری پایا و مناسب برای استفاده در مراکز شکاف کام است که بر اساس ارزیابی ادراکی، اطلاعاتی درباره مشکلات گفتاری خاص شکاف ارایه می‌دهد.

References

1. Lohmander A, Olsson M. Methodology for perceptual assessment of speech in patients with cleft palate: a critical review of the literature. *Cleft Palate Craniofac J* 2004; 41(1): 64-70.
2. Cordero KN. Assessment of cleft palate articulation and resonance in familiar and unfamiliar languages: English, Spanish, and Hmong [Thesis]. University of Minnesota; 2008. 2012. Ref Type: Generic
3. Keuning KH, Wieneke GH, Dejonckere PH. The intrajudge reliability of the perceptual rating of cleft palate speech before and after pharyngeal flap surgery: the effect of judges and speech samples. *Cleft Palate Craniofac J* 1999; 36(4): 328-33.
4. Henningsson G, Kuehn DP, Sell D, Sweeney T, Trost-Cardamone JE, Whitehill TL. Universal parameters for reporting speech outcomes in individuals with cleft palate. *Cleft Palate Craniofac J* 2008; 45(1): 1-17.
5. Kummer AW. *Cleft Palate & Craniofacial Anomalies: Effects on Speech and Resonance*. 2nd ed ed. Delmar Cengage Learning; 2008.
6. John A, Sell D, Sweeney T, Harding-Bell A, Williams A. The cleft audit protocol for speech-augmented: A validated and reliable measure for auditing cleft speech. *Cleft Palate Craniofac J* 2006; 43(3): 272-88.
7. Lohmander A, Willadsen E, Persson C, Henningsson G, Bowden M, Hutters B. Methodology for speech assessment in the Scandcleft project--an international randomized clinical trial on palatal surgery: experiences from a pilot study. *Cleft Palate Craniofac J* 2009; 46(4): 347-62.
8. Lohmander A, Persson C. A longitudinal study of speech production in Swedish children with unilateral cleft lip and palate and two-stage palatal repair. *Cleft Palate Craniofac J* 2008; 45(1): 32-41.
9. Brunnegard K, Lohmander A. A cross-sectional study of speech in 10-year-old children with cleft palate: results and issues of rater reliability. *Cleft Palate Craniofac J* 2007; 44(1): 33-44.

Evaluating intra- and inter-rater reliability for "cleft palate speech assessment test based on universal parameters system- in Persian"

Azadeh Amirian¹, Fatemeh Derakhshandeh^{*}, Ahmad Salehi², Bahram Soleimani³

Received date: 18/01/2012

Accept date: 15/03/2012

Abstract

Introduction: Cleft palate is one of the most common congenital birth defects. Individuals with cleft palate have certain speech characteristics that can be identified perceptually by speech-language pathologists. Perceptual assessment has been regarded to as the gold standard for evaluating cleft palate speech. Considering the prevalence of cleft palate and its related speech characteristics, it is necessary for clinicians to use a reliable test for perceptual assessment of cleft palate speech errors. The aim of this study was to evaluate Intra- and inter-rater reliabilities for a recently developed assessment tool named "Cleft palate speech assessment test based on universal parameters system- Farsi version".

Materials and Methods: In this descriptive-analytic cross-sectional study, speech samples obtained from 34 five-to-nine-year-old Farsi-speaking children with cleft palate who were referred to Isfahan cleft palate clinic were rated by two speech-language pathologists (one of them was considered as a naive and the other one as an expert rater). Two months later, 30% of the same speech samples were randomly selected and rerated by the same raters. Speech materials included 40 single words and 17 sentences along with a sample of connected speech. Intra- and inter-rater reliability measures were calculated via Intraclass Correlation Coefficient (ICC).

Results: Mean Intra-rater reliability measured by ICC was found to be 0.967 and 0.971 for the naive and expert raters respectively. Mean Inter-rater reliability was 0.957 as measured by ICC.

Conclusion: This test has an acceptable Intra and inter-rater reliability to be used in assessment of cleft palate speech and can be applied in those studies aimed at evaluating different aspects of speech in Farsi-speaking children with cleft palate aged from 5 to 9 years.

Keywords: Cleft palate, Reliability, Perceptual speech assessment, Universal parameters system

* PhD Student in Speech Therapy, Academic Member, School of Rehabilitation Sciences, Tehran University of Medical Sciences, Tehran, Iran
Email: derakhshandeh@rehab.mui.ac.ir

1. MSc Student in Speech Therapy, Students Research Committee, School of Rehabilitation Sciences, Isfahan University of Medical Sciences, Isfahan, Iran

2. MSc, Academic Member, School of Rehabilitation Sciences, Isfahan University of Medical Sciences, Isfahan, Iran

3. PhD in Biostatics, Academic Member, Azad University, Najaf-Abad Branch, School of Medicine, Isfahan, Iran