Introduction
The Babinski sign is a pathological response elicited by a stimulus to the sole of the foot. The resulting reflex involves dorsiflexion (upward motion) of the toes with accompanying flexion in the ankle, knee and hip (Van Gijn, 1995). The inter rater reliability of this clinically diagnostic sign has been greatly contested. The aim of the current study was to assess the inter rater reliability of neurologists and medical students of the Babinski reflex, with movements of the toes, feet and leg being objectively measured using conventional kinematics. An assessment of the aspects of the reflex used when rating the reflex was also performed.

Method
The reflex was performed by the same investigator on 15 participants. Five participants had neurological defects causing them to have a positive Babinski response. The reflex was elicited on the subjects, in a supine position, by running a sharp object along the lateral plantar. The change in big toe angle, time to maximum big toe angle, change in ankle angle, time to maximum ankle angle and movement latency were calculated from the biomechanical data. The muscle activity in the tibialis anterior, biceps femoris and gastrocnemius was concurrently recorded through surface electromyography. These reflexes were recorded on conventional video cameras and the footage was then shown to 12 neurologists and 12 medical students who were instructed to rate the given reflexes as either pathological (P) or non-pathological(NP).

The inter-rater reliability was assessed using the kappa statistic (Landis, 1977). For each rater, the 15 reflexes were grouped as either P or NP and the mean values for each biomechanical and electromyographical variable in each group were determined. ANOVAs were used to assess the change in big toe angle, time to maximum big toe angle, change in ankle angle, time to maximum ankle angle, movement latency and the maximum amplitude of the tibialis anterior, biceps femoris and gastrocnemius differences between the groups.

Results
The kappa values for neurologists and students agreements were 0.72 and 0.67 respectively. For both neurologists and students there were significant differences between P and NP for change in big toe angle, time taken to reach maximum ankle angle, movement latency and the maximum amplitude of gastrocnemius. For neurologists alone there were significant differences for time taken to reach maximum big toe angle and change in ankle angle between P and NP. For students alone maximum amplitude of biceps femoris was significantly different between P and NP.

Discussion
Substantial agreement for the inter rater reliability for both groups was shown. This is higher than most previous research which showed kappa values ranging from 0.30 to 0.75 (Lee, 2011; McCance, 1968; Miller 2005; Singerman, 2008; Vogel, 1992). This is most likely due to the raters assessing the same recorded reflex therefore removing variation caused by eliciting the reflex. Moreover the fact that the raters had to chose either P or NP and didn’t have an option for equivocal movement could also have increased the kappa values. In conclusion there was no significant difference between the inter rater reliability of students and neurologists when assessing the Babinski sign. However, the aspects of the reflex used to make the assessment did differ between the two groups.

References
Lee et al, Neurol Asia, 16(2):143–147, 2011.